

UNIVERSIDAD AUTÓNOMA DE MADRID

FACULTAD DE CIENCIAS

DEPARTAMENTO DE BIOLOGÍA



**ESTUDIO MOLECULAR DE LA ANIRIDIA Y OTRAS
MALFORMACIONES OCULARES CONGÉNITAS:
BÚSQUEDA DE NUEVOS MECANISMOS GENÉTICOS**

TESIS DOCTORAL

MARÍA TARILONTE MISAS

Directoras

MARTA CORTÓN PÉREZ Y CARMEN AYUSO GARCÍA

MADRID, 2019



Facultad de Ciencias

Departamento de Biología

**ESTUDIO MOLECULAR DE LA ANIRIDIA Y OTRAS
MALFORMACIONES OCULARES CONGÉNITAS:
BÚSQUEDA DE NUEVOS MECANISMOS GENÉTICOS**

TESIS DOCTORAL

María Tarilonte Misas

Madrid, 2019





Facultad de Ciencias

Departamento de Biología

**ESTUDIO MOLECULAR DE LA ANIRIDIA Y OTRAS
MALFORMACIONES OCULARES CONGÉNITAS:
BÚSQUEDA DE NUEVOS MECANISMOS GENÉTICOS**

Memoria presentada por **María Tarilonte Misas, Licenciada en Biología**

para optar al título de Doctor por la Universidad Autónoma de Madrid

Directoras de Tesis:

Dra. Marta Cortón Pérez

Dra. Carmen Ayuso García

Tutor de Tesis:

Dr. José Fernández Piqueras



Realizado en el Instituto de Investigación Sanitaria- Fundación Jiménez Díaz (IIS-FJD)

Dña. MARTA CORTÓN PÉREZ, investigadora Miguel Servet en el Departamento de Genética del Instituto de Investigación Sanitaria - Fundación Jiménez Díaz,

Dña. CARMEN AYUSO GARCÍA, jefa del Departamento de Genética del Hospital Universitario Fundación Jiménez Díaz y Directora Científica del Instituto de Investigación Sanitaria Fundación Jiménez Díaz,

CERTIFICAN:

Que la presente memoria titulada: *“Estudio molecular de la aniridia y otras malformaciones oculares congénitas: búsqueda de nuevos mecanismos genéticos”*, que presenta Dña. María Tarilonte Misas para obtener el grado de Doctor, ha sido realizada bajo nuestra dirección. Por ello la autorizamos para su presentación ante un Tribunal Calificador.

Madrid, 02 de octubre de 2019

Fdo. Dra. Marta Cortón Pérez

Fdo. Dra. Carmen Ayuso García

Directora de la Tesis Doctoral.....Directora de la Tesis Doctoral

Fdo. Dr. José Fernández Piqueras

Tutor de la Tesis Doctoral

A mis padres

Mina y Julián

A mis hermanos

A mis sobrinos

A mis tías

A Alejandro

**“Es el tiempo que le dedicaste a tu rosa
lo que la hace tan importante”**

El Principito

Antoine de Saint-Exupéry

(1900-1944)

Esta investigación ha sido posible gracias a las siguientes ayudas de investigación:

- Contrato predoctoral concedido por la “Fundación Conchita Rábago de Jiménez Díaz” (periodo 2016-2019).
- Proyectos de Investigación del Instituto de Salud Carlos III - (Fondo Europeo de Desarrollo Regional FEDER) (PI17/01164 y CP12/03256).
- Proyecto Retos del MINECO (SAF2013-46943-R).
- III Convocatoria de ayudas a la investigación de la Fundación FEDER para la investigación en Enfermedades Raras
- Proyecto de Investigación 2014 de la Fundación Mutua Madrileña

Es el momento de concluir esta etapa de aprendizaje de mi vida. Cuando tomé la decisión de emprender este reto no podía imaginarme todo el esfuerzo y sufrimiento que llevaba consigo, pero sobre todo la gran satisfacción que supone contribuir al avance del conocimiento de las enfermedades poco frecuentes.

En primer lugar quiero agradecer a mis directoras de Tesis el que me hayan brindado la oportunidad de formarme en el Departamento de Genética de una institución tan importante como es el Hospital Universitario Fundación Jiménez Díaz, ha sido emocionante participar en charlas, cursos y congresos científicos. A Marta Cortón por enseñarme el complejo mundo de la Genética, transmitirme sus conocimientos y ayudarme siempre que lo he necesitado. A Carmen Ayuso, por sus consejos.

A la Fundación Conchita Rábago, por financiar este proyecto durante toda mi etapa predoctoral.

También quiero dar las gracias a todos los compañeros del Servicio de Genética que me habéis hecho vivir buenos momentos, y que tuvisteis alguna palabra o gesto de apoyo en los malos. A Patri, por ser mi mano derecha y mi confidente en mi primera etapa. A Delpo, por su apoyo y por los buenos momentos compartidos fuera y dentro del laboratorio hasta el final, literalmente. A Clara e Inma, por ser buenas compañeras y por animarme siempre que lo he necesitado. A Jesús y Camilo, mis Zipi y Zape preferidos, porque se puede hablar de todo con vosotros y siempre sabéis sacarme una sonrisa. A Marta Rodríguez de Alba, por ser cercana y clara. A Miguel, por amenizar mis días con historietas. A Fernando, por su “buen carácter” y sus vaciles. A Loli, por su buen carácter y su simpatía. A Olga, por enseñarme el funcionamiento del labo. A Ruth, Carol, Rosa, Mariajo, Almudena, Ana, Choni, Berta, por sus palabras de apoyo. A los nuevas incorporaciones de nuestro grupo, las Ales, por su gran apoyo logístico y moral en mi última etapa. A los bioinfos: Pablo, Perce, Ionut, Raquel y Lorena, por amenizar mis comidas con temas curiosos y graciosos. A mis compañeros de otros grupos de investigación, Álvaro, Dani, Olaya y Leti, por poder contar con ellos.

Sobre todo quiero dedicarles este trabajo de Tesis Doctoral a mis padres, Mina y Julián, por su esfuerzo, sacrificio y dedicación para poder darme una formación académica. Sin vosotros nada de esto hubiese sido posible.

A toda mi familia, en especial a mis abuelos, por sus enseñanzas y cariño incondicional. A mis hermanos, Miguel y Jose, a mis tías, Sole, Visi, Rosalina, Ermelina y Celsa, por estar ahí siempre que os he necesitado, por vuestros consejos y vuestro apoyo desde la distancia. A mis cuatro sobrinos, Unai, Nabar, Jorge y Sergio, por hacerme ver el mundo desde otra perspectiva.

A todos mis amigos, en especial a Lurdes, por reaparecer en mi vida en el momento indicado y permanecer en las duras y en las maduras.

Por supuesto a Alejandro, mi compañero de batallas y aventuras. Tú eres el que mejor sabe el sacrificio personal que he tenido que hacer para llegar hasta aquí, y tengo que agradecerte tu apoyo, comprensión y ayuda. Por hacerme reír e ilusionarme con las pequeñas cosas de la vida. Ahora sí que empieza una nueva etapa en nuestras vidas.

A todos los pacientes con aniridia y a las familias que han participado en este estudio, y también a la Asociación de Aniridia España.

Resumen/

Abstract

Hace casi tres décadas se identificó la causa genética de la aniridia, siendo *PAX6* el principal gen implicado hasta el día de hoy. Sin embargo, los estudios rutinarios de diagnóstico genético solo permiten obtener resultados concluyentes en <75% de los pacientes de una cohorte recopilada en el Hospital Universitario Fundación Jiménez Díaz. A la vista de estos datos se sospecha que puedan existir otras causas o mecanismos genéticos todavía sin identificar entre los casos considerados “*PAX6* negativos”, lo que ha constituido el principal objetivo de esta Tesis Doctoral.

Gracias al desarrollo de nuevas tecnologías personalizadas de secuenciación masiva y *arrays* de CHG, en este trabajo hemos ampliado el estudio genético a otras regiones del *locus* de *PAX6* habitualmente no estudiadas de forma rutinaria, incluyendo UTRs, intrones y *enhancers*, así como otros genes asociados a formas fenotípicas solapantes. Se han detectado microdeleciones afectando los *enhancers* de *PAX6*, lo que ha permitido acotar una región crítica para su regulación transcripcional. Asimismo, se han detectado variantes en el 5'UTR de *PAX6*. También hemos podido detectar variantes patogénicas en otros genes, *FOXC1* y *PITX2*, en el 4% de la cohorte.

Por otro lado, hemos profundizado en la implicación de las variantes de significado clínico incierto en *PAX6* identificadas en nuestra cohorte. Nos hemos centrado en el estudio de variantes no codificantes, principalmente localizadas en el 5'UTR, en sitios no canónicos de *splicing* y variantes sinónimas. Para ello, se han realizado estudios de modelización *in vitro* utilizando minigenes para valorar posibles alteraciones del *splicing*, lo que ha permitido la reclasificación patogénicas de 10 variantes.

Por último, este trabajo ha confirmado por primera vez la existencia de mosaicismo gonosomal de origen parental como causa de un patrón de transmisión inusual y/o variabilidad fenotípica en 4 familias con aniridia u otros fenotipos asociados a *PAX6*, lo cual tiene importantes implicaciones en el consejo genético.

Finalmente, este estudio de Tesis Doctoral ha permitido la detección de nuevos mecanismos genéticos implicados en la aniridia con un consiguiente aumento del rendimiento de diagnóstico genético en un 20%, llegando así al 84% de casos en los que se ha podido identificar una variante claramente patogénica.

The genetic cause of aniridia was first described nearly three decades ago, being *PAX6* the main gene involved until the present day. However, the routine genetic diagnostic tools has only allowed us to obtain conclusive results in less than 75% of the patients in a cohort recruited in the University Hospital-Fundacion Jimenez Diaz. Considering these data, it is suspected that there are still other causes or genetic mechanisms to be unraveled in the “Negative-*PAX6*” cases. Further genetic analysis of these cases constituted the main objective of this Doctoral Thesis.

In this work, thanks to the development of new customized technologies of Next Generation Sequencing and arrays CHG, we have expanded the genetic study to other regions of the *PAX6* locus, usually not covered with the conventional approaches, including UTRs regions, introns and enhancers, as well as other overlapping phenotypes associated genes. Identification of microdeletions affecting *PAX6* enhancers has allowed us to delimit a new critical region involved in its transcriptional regulation. Additionally, variants in the 5'UTR region of *PAX6* have been identified. Moreover, pathogenic variants in other genes, as *FOXC1* and *PITX2*, have been found in the 4% of the cases of our cohort.

Furthermore, the implication of variants of unknown significance in *PAX6* identified in our cohort was deeply analyzed. This work is focused on the study of non-coding variants, mainly located in the 5'UTR region, in non-canonical splice sites and synonymous variant. For that purpose, we carried out in vitro functional studies using minigenes to evaluate potential alterations in the splicing mechanism, which has allowed us to prove the pathogenicity of 10 variants in *PAX6*.

Finally, it has been demonstrated that parental mosaicism in *PAX6* causes intra-familial variability in four families with aniridia and/or other phenotypes associated to *PAX6* mutations, which has substantial implications for genetic counseling in congenital aniridia and it has not been previously reported.

In conclusion, this Doctoral Thesis has allow us to identify new genetic mechanisms involved in aniridia, resulting in an increase of the diagnostic yield by 20%, reaching up to a 84% of the cases carrying a pathogenic variant.

Índice

Resumen/Abstract	1
Índice	7
Índice de tablas y figuras	13
Abreviaturas.....	21
1. Introducción	25
1.1. El proceso de la visión.....	27
1.2. Anatomía del globo ocular humano	28
1.3. Embriología ocular.....	31
1.4. Mecanismos moleculares asociados al desarrollo ocular	33
1.5. Características del gen <i>PAX6</i>	35
1.5.1. Estructura génica y proteica	35
1.5.2. Regulación transcripcional de <i>PAX6</i>	36
1.5.3. Expresión de <i>PAX6</i> en diferentes tejidos.....	37
1.6. Defectos en <i>PAX6</i> asociados a malformaciones oculares	39
1.7. Aniridia congénita.....	40
1.7.1. Manifestaciones oculares de la aniridia	40
1.7.2. Fenotipos extraoculares asociados a aniridia.....	42
1.8. Genética de la aniridia clásica	43
1.8.1. Espectro mutacional de <i>PAX6</i>	44
1.8.2. Correlación genotipo-fenotipo	47
1.9. Diagnóstico diferencial	48
1.10. Estudios genéticos en la aniridia	50
1.11. Casos de aniridia <i>PAX6</i> negativos	52
1.12. Manejo clínico y perspectivas terapéuticas	53
2. Objetivos	57
3. Materiales y métodos.....	61
3.1. Pacientes y muestras biológicas.....	63

3.1.1. Descripción de la cohorte del estudio	63
3.1.2. Muestras de pacientes y controles.....	64
3.1.3. Líneas celulares establecidas	65
3.2. Aislamiento de ácidos nucleicos.....	65
3.3. Análisis de CNVs mediante arrays de CGH y MLPA	66
3.3.1. aCGH	66
3.3.2. MLPA.....	67
3.4. Técnicas de secuenciación masiva.....	68
3.4.1. Panel customizado de genes para MOC	68
3.4.2. Secuenciación del exoma completo	69
3.4.3. Análisis bioinformático	69
3.5. Priorización y clasificación de las variantes	72
3.6. Secuenciación Sanger.....	75
3.7. Análisis de haplotipos	75
3.8. Construcción de minigenes	76
3.9. Análisis de expresión	78
3.9.1. Análisis de expresión de minigenes	78
3.9.2. Análisis de <i>splicing</i> en muestras de pacientes	80
3.10. Análisis de mosaicismo	80
3.11. Bases de datos y herramientas bioinformáticas utilizadas.....	82
4. Resultados.....	83
4.1. Implicación de variantes en regiones no codificantes en <i>PAX6</i> y/o variantes en otros genes en los casos de aniridia “<i>PAX6</i> negativos”	89
4.1.1. Identificación de microdeleciones afectando elementos <i>cis</i> -reguladores	92
4.1.2. Cribado de variantes puntuales en las regiones no codificantes de <i>PAX6</i>	96
4.1.3. Identificación de variantes probablemente patogénicas en el 5'UTR de <i>PAX6</i>	98
4.1.4. Caracterización funcional de la variantec.-129+1G>A en el exón 2	103
4.1.5. Caracterización funcional de la variantec.-129+1G>A en el exón 3	107

4.1.6. Identificación de variantes patogénicas en otros genes de disgenesia del segmento anterior	110
4.1.6.1. Variantes en el gen <i>FOXC1</i>	111
4.1.6.2. Variantes identificadas en el gen <i>PITX2</i>	113
4.2. Caracterización funcional de VUS que afectan al <i>splicing</i> del exón 6 de <i>PAX6</i>.....	115
4.2.1. Identificación de variantes silentes que crean donadores crípticos exónicos	123
4.2.2. Creación de un nuevo donador críptico intrónico por una variante situada en sitio no canónico de <i>splicing</i>	129
4.2.3. Una pequeña duplicación conduce a un defecto de <i>splicing</i> mediado por la activación de un ESS	131
4.2.4. La posible alteración proteica producida por dos variantes <i>missense</i>	133
4.2.5. Clasificación patogénica de variantes en función al defecto de <i>splicing</i> y posibles correlaciones fenotípicas.....	137
4.3. Estudio de la implicación del mosaicismo parental en familias con variantes en <i>PAX6</i>..	141
4.3.1. Descripción clínico-genética de las familias con mosaicismo en <i>PAX6</i>	146
4.3.1.1. Familia ANI-062.....	146
4.3.1.2. Familia ANI-100.....	147
4.3.1.3. Familia ANF-060.....	149
4.3.1.4. ANI-RyC.....	150
4.4. Actualización del espectro mutacional en nuestra cohorte de aniridia	153
5. Discusión	157
5.1. Caracterización genética de casos de aniridia “ <i>PAX6</i> negativos”	161
5.2. Caracterización funcional y reclasificación clínica de VUS en <i>PAX6</i>	170
5.3. Implicación del mosaicismo parental en la aniridia	175
5.4. Actualización del espectro mutacional en nuestra cohorte de aniridia	180
6. Conclusiones	183
7. Bibliografía	189
8. Anexos.....	211

8.1. Anexo 1. Listado de los 20 elementos cis-reguladores de PAX6 incluidos en el panel de NGS personalizado	213
8.2. Anexo 2. Lista de oligonucleótidos utilizados en el apartado 4.1	214
8.3. Anexo 3. Lista de oligonucleótidos utilizados en el apartado 4.2	216
8.4. Anexo 4. Microsatélites utilizados en el análisis de haplotipos y sus características	217
8.5. Anexo 5. Resumen de los resultados obtenidos mediante PCR semi-cuantitativa en las siete variantes estudiadas	218
8.6. Anexo 6. Resumen resultados minigenes obtenidos tras semi-cuantificación por electroforesis capilar	219
8.7. Anexo 7. Estudio de mosaicismo ligado a variantes en PAX6 mediante ddPCR	220
9. Publicaciones.....	221

Índice de tablas y figuras

Índice de tablas

Materiales y métodos

Tabla M1. Bases de datos y herramientas bioinformáticas utilizadas en este trabajo. 82

Resultados

Tabla R1. Resumen de las microdeleciones en la región cromosómica 11p13 que afectan a elementos 3' reguladores de PAX6 identificadas en nuestra cohorte..... 93

Tabla R2. Resumen de variantes no codificantes encontradas en nuestra cohorte de aniridia y en fenotipos relacionados..... 97

Tabla R3. Predicciones de *splicing* y resultados in vitro de variantes identificadas en el 5'UTR de PAX6 101

Tabla R4. Variantes patogénicas identificadas en los genes *FOXC1* y *PITX2* 110

Tabla R5. Predicciones *in silico* del potencial spliceogénico de las variantes en el exón 6 de PAX6. 119

Tabla R6. Predicciones *in silico* del potencial espliceogenico de variantes en el exón 6 de PAX6. 120

Tabla R7. Características clínicas de los individuos portadores de las variantes estudiadas en el exón 6. 139

Tabla R8. Características clínicas y genéticas de individuos pertenecientes a familias con sospecha de mosaicismo con variantes en PAX6..... 145

Índice de figuras

Introducción

Figura I1. Esquema del proceso de la visión.	27
Figura I2. Anatomía del ojo humano	28
Figura I3. Desarrollo ocular.	32
Figura I4. Procesos moleculares involucrados en el desarrollo ocular	33
Figura I5. Señalización mediada por <i>PAX6</i>	34
Figura I6. Representación del gen y la proteína <i>PAX6</i>	35
Figura I7. Representación del locus de <i>PAX6</i> y sus elementos reguladores.	37
Figura I8. Representación de la expresión de la isoforma NM_000280.4 de <i>PAX6</i> en diferentes tejidos humanos	38
Figura I9. Variedad de anomalías oculares en distintos modelos animales y en humanos.	39
Figura I10. Fenotipos asociados a defectos en <i>PAX6</i>	41
Figura I11. Representación del número de variantes publicadas para cada exón de <i>PAX6</i>	45
Figura I12. Representación de la localización de las variantes descritas en 5'UTR de <i>PAX6</i>	46
Figura I13. Algoritmo de diagnóstico genético en aniridia.	45

Materiales y métodos

Figura M5. Preparación de librerías de ADN siguiendo el método de captura de <i>HaloPlex</i>	69
Figura M6. Representación de la <i>pipeline</i> implementada en nuestro laboratorio.	71
Figura M7. Algoritmo llevado a cabo para la clasificación de las variantes.	74
Figura M4. Insertos de <i>PAX6</i> introducidos en el vector pSPL3 para transfección	77
Figura M5. Estructura del constructo silvestre " <i>PAX6_ex1-4</i> " en ensayos con minigenes.	78

Figura M6. Estructura del constructo silvestre “PAX6_ex6” en ensayos con minigenes 78

Resultados

Figura R8. Representación del espectro mutacional en la cohorte de familias españolas con sospecha de aniridia aplicando el algoritmo clásico de diagnóstico. 85

Figura R9. Segregación de las familias con deleciones 3' a PAX6 94

Figura R10. Identificación de deleciones en regiones reguladoras 3' a PAX6 en pacientes con aniridia..... 95

Figura R11. Representación esquemática de variantes en 5'UTR en el gen PAX6 98

Figura R5. Segregaciones de las tres variantes identificadas en el 5'UTR de PAX6 en cinco familias. 100

Figura R6. Análisis de splicing para la construcción silvestre mediante un ensayo *in vitro* con un midigene para el 5'UTR de PAX6..... 102

Figura R7. Análisis de splicing para la variante c.-129+1G>A mediante un ensayo *in vitro* con un midigene para el 5'UTR de PAX6..... 105

Figura R8. Diferentes isoformas resultantes del análisis de splicing de la variante c.-129+1G>A mediante PCR-semicuantitativa..... 106

Figura R9. Análisis de expresión de isoformas de PAX6 tras inhibición del sistema NMD en portadores de la mutación c.-129+1G>A. 106

Figura R10. Análisis de *splicing* para las variantes del exón 3 mediante un ensayo *in vitro* con un midigen para el 5'UTR de PAX6..... 108

Figura R11. Diferentes isoformas resultantes del análisis de splicing de las variantes del exón 3 mediante PCR-semicuantitativa..... 108

Figura R12. Representación esquemática del patrón de *splicing* obtenido para los exones del 5'UTR para cada isoforma 109

Figura R13. Identificación y segregación de variantes en FOXC1 en tres pacientes con aniridia.112

Figura R14. Segregación familiar y secuenciación Sanger de una variante en <i>PITX2</i> en la familia ANI-114.	113
Figura R15. Representación de la localización de las 10 VUS identificadas en <i>PAX6</i>	117
Figura R16. Representación esquemática del potencial espliceogénico para el donador canónico y los donadores alternativos identificados en el exón 6 de <i>PAX6</i>	121
Figura R17. Resultados estudio funcional en la construcción silvestre mediante el minigen “ <i>PAX6_ex6</i> ”.	122
Figura R18. Estudio funcional de la variante c.174C>T.....	124
Figura R19. Estudio funcional de la variante c.249T>G.....	126
Figura R20. Estudio funcional de la variante c.333C>A	128
Figura R21. Estudio funcional de la variante c.357+6T>G	130
Figura R22. Estudio funcional de la variante c.326_331dup	132
Figura R23. Estudio funcional de las variantes c.178T>C y c.356G>T.....	134
Figura R24. Modelización de las variantes <i>missense</i> Tyr60His y Ser119Ile en la estructura 3D de la proteína <i>PAX6</i> cristalizada unida a una molécula de ADN.	136
Figura R25. Hallazgos oftalmológicos en las familias ANI-062 y ANI-001.....	140
Figura R26. Información sobre la familia ANI-062.....	147
Figura R27. Información sobre la familia ANI-100.....	148
Figura R28. Información sobre la familia ANF-060	150
Figura R29. Información sobre la familia ANI-RyC.....	151
Figura R30. Gráficas de los resultados de las familias con aniridia y aniridia-like estudiadas en este trabajo.	156

Discusión

Figura D1. Nuevo algoritmo de diagnóstico genético de la aniridia propuesto tras los resultados obtenidos en el presente trabajo..... 182

Abreviaturas

aCGH	<i>Array de hibridación genómica comparativa (Comparative Genomic Hybridization array).</i>
ACMG	<i>American College of Medical Genetics.</i>
ADNc	ADN complementario.
ARNm	ARN mensajero.
ARPE19	<i>Adult Retinal Pigment Epithelial Cell.</i>
AV	Agudeza visual.
CNV	Variación en el número de copias (<i>Copy number variation</i>).
CREs	Elementos <i>cis</i> -reguladores (<i>Cis-regulatory elements</i>).
CRISPR	Repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas (<i>clustered regularly interspaced short palindromic repeats</i>).
CT	Tránsito canónico (Canonical transcript).
ddPCR	PCR digital (digital droplet PCR).
DMEM	<i>Dulbecco's Modified Eagle Medium</i>
DTT	<i>Dithiothreitol.</i>
DGP	Diagnóstico genético preimplantacional.
EBV	Virus de Epstein-Barr (<i>Epstein-Barr virus</i>).
ESE	Potenciador exónico del <i>splicing</i> (<i>Exonic Splicing Enhancer</i>).
ESS	Silenciador exónico del <i>splicing</i> (<i>Exonic Splicing Silencer</i>).
FAM	Fluoróforo 6-Carboxyfluorescein.
FBS	Suero fetal bovino (<i>Fetal bovine serum</i>).
FISH	Hibridación fluorescente <i>in situ</i> (<i>Fluorescent in situ hybridization</i>).
HD	Homeodominio (<i>Homeodomain</i>).
HEK293T	Células epiteliales de riñón (<i>Human Epithelial Kidney cells</i>).
HEPES	Tampón Ácido 4- (2-hidroxietil)-1-piperazinaetanosulfónico.
HGMD	<i>Human Gene Mutation Database.</i>
HSF	<i>Human Splicing Finder.</i>
<i>iPSCs</i>	<i>Células madre pluripotentes inducidas (induced Pluripotent Stem Cells)</i>
LCL	Líneas celulares linfoblastoides (<i>Lymphoblastoid cell lines</i>).
LOF	Pérdida de función (<i>Loss of function</i>).

MAF	Frecuencia alélica menor (<i>Minor allele frequency</i>).
MLPA	<i>Multiplex ligation-dependent probe amplification</i> .
MOC	Malformaciones oculares congénitas.
NCS	Sitio de <i>splicing</i> no canónico (<i>Non Canonical Splice Site</i>).
NGL	Núcleo geniculado lateral.
NGS	Secuenciación masiva (<i>Next generation sequencing</i>).
NMD	<i>Non- sense mediated decay</i> .
OMIM	<i>Online Mendelian Inheritance in Man</i> .
Pb	Pares de bases.
PD	Domino pareado (<i>Paired Domain</i>).
PCR	Reacción en cadena de la polimerasa (<i>Polymerase Chain Reaction</i>).
PST	Dominio de PAX6 rico en prolina, serina y treonina.
PTC	Codón de parada prematura (<i>Premature translation-termination codón</i>).
RPE	Epitelio pigmentario de la retina (<i>Retinal Pigment Epithelium</i>).
SBS	Secuenciación por síntesis (<i>sequencing by synthesis</i>).
snRNAs	ARNs nucleares pequeños (<i>Small nuclear RNAs</i>).
SNP	Polimorfismo de único nucleótido (<i>Single-nucleotide polymorphisms</i>).
SNV	Variación de único nucleótido (<i>Single nucleotide variant</i>).
SRE	Elementos reguladores del <i>splicing</i> (<i>Splicing regulatory elements</i>).
STR	Repetición corta en tandem (<i>Short tandem repeat</i>).
VIC	<i>Fluorescent dye</i> .
VUS	Variante de significado incierto (<i>Variant of uncertain significance</i>).
WAGR	<i>Wilms tumor, Aniridia, Genitourinary anomalies and mental Retardation</i> .
WAGRO	<i>Wilms tumor, Aniridia, Genitourinary anomalies, mental Retardation and Obesity</i> .
WES	Secuenciación del Exoma completo (<i>Whole exome sequencing</i>).
WGS	Secuenciación del genoma completo (<i>Whole genome sequencing</i>).
WT	Silvestre/nativo (<i>wild-type</i>).

Introducción

1.1. El proceso de la visión

En la visión están implicados tanto el cerebro como los ojos, ambos actúan coordinadamente para que podamos percibir nuestro entorno y relacionarnos con él. Este proceso consiste en una serie de sinapsis celulares que transportan la información visual del entorno para su procesamiento cerebral. En primer lugar la luz atraviesa la córnea y penetra en el ojo a través de la pupila, esta apertura del iris regula la cantidad de luz que lo atraviesa (Figuras I1 e I2). La luz llega al cristalino, donde se enfoca hacia la fovea de la retina, allí las células fotorreceptoras, los conos y los bastones, convierten la energía luminosa en nerviosa. Esta energía nerviosa pasa a las células bipolares, de ahí a las células amacrinas y finalmente a las células ganglionares, cuyos axones abandonan la retina a través del nervio óptico (Figura I1). Éste llega al cerebro y se une al del ojo contrario en el quiasma óptico, donde se cruzan las fibras nerviosas que terminarán en el lado opuesto del cerebro. A partir del quiasma las fibras se llaman tractos ópticos, estos penetran en el núcleo geniculado lateral (NGL), donde tiene lugar la siguiente sinapsis. Las fibras abandonan el NGL como radiaciones ópticas que terminan en la corteza visual de los lóbulos occipitales.

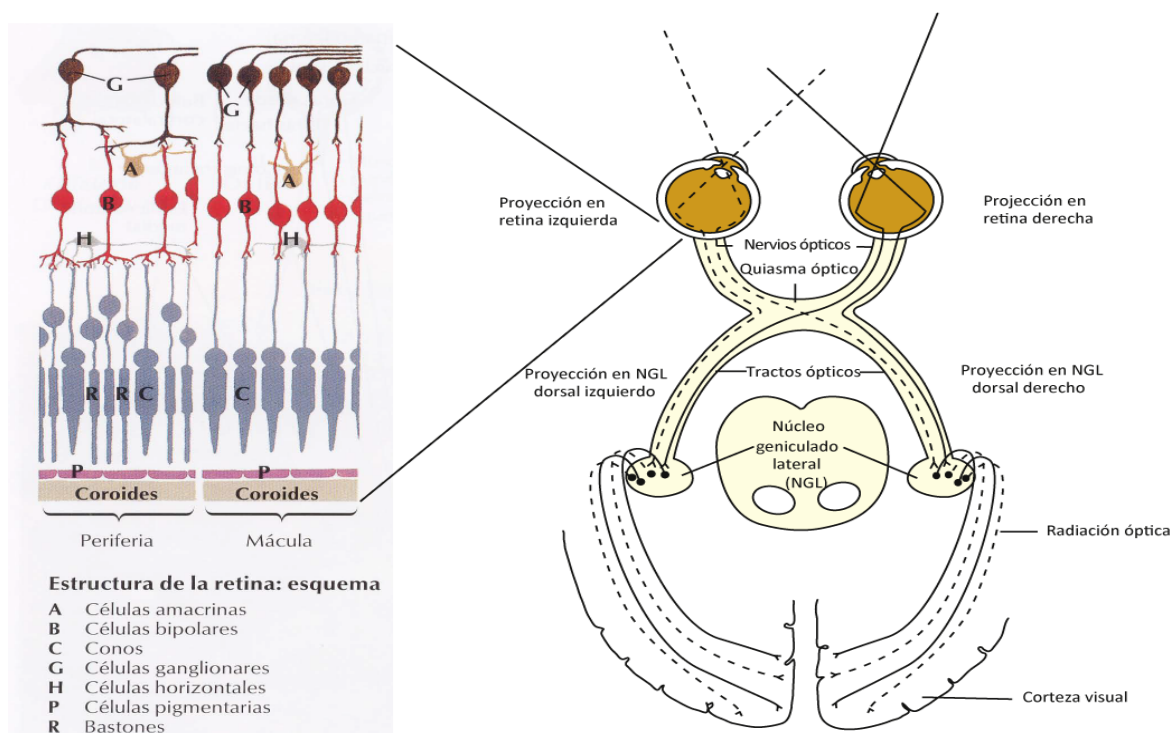


Figura I1. Esquema del proceso de la visión. Figura adaptada de Netter, 2012 y de Remington, 2012. Se detallan las estructuras implicadas en la visión.

1.2. Anatomía del globo ocular humano

El ojo está compuesto por el globo ocular y sus órganos anexos: músculos oculares, párpados, conjuntiva, aparato lacrimal y cejas. Nos centramos en el globo ocular compuesto por tres capas concéntricas que desde el exterior al interior son:

- **Túnica fibrosa**: Compuesta por la esclerótica en la parte posterior, capa blanca cuya función es la de protección y mantenimiento de la estructura; y la córnea en la parte anterior, capa transparente dónde se refractan los rayos de luz.
- **Túnica vascular**: Constituida por el iris y el cuerpo ciliar en la parte anterior, y la coroides en la parte posterior. Su función es la de vascularización.
- **Túnica neurosensorial**: Formada por la retina, encargada de transformar las radiaciones electromagnéticas en señales nerviosas (fototransducción) que se envían hacia el cerebro a través de las neuronas.

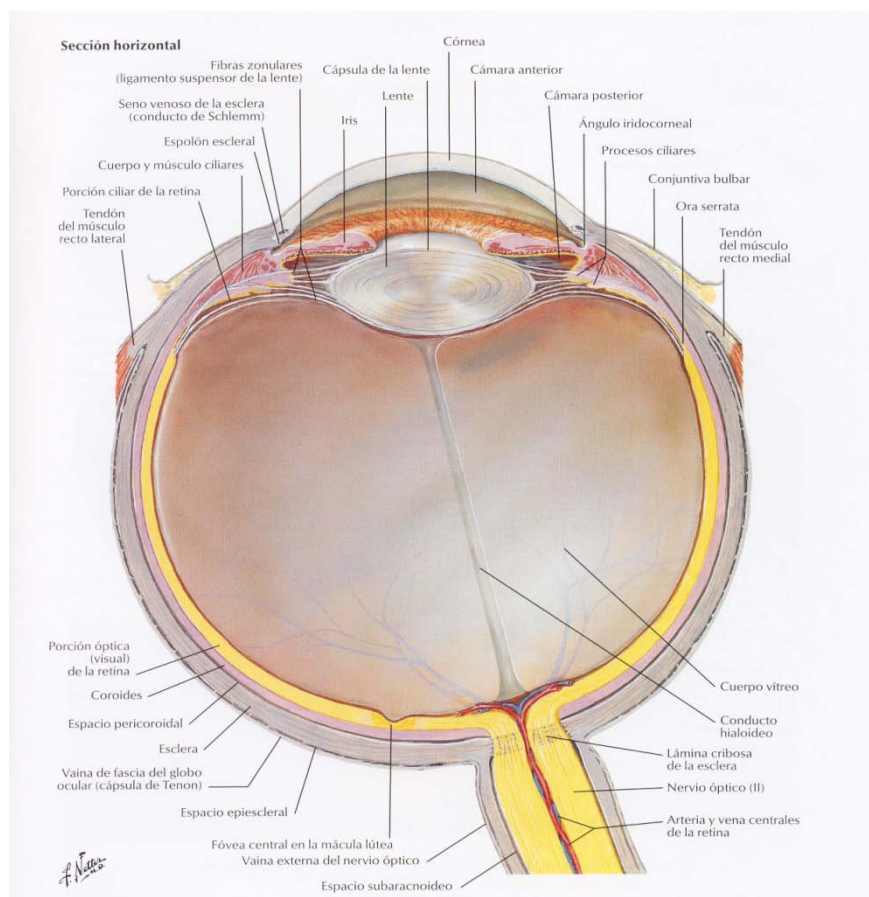


Figura 12. Anatomía del ojo humano. Figura tomada de Netter, 2012. Se señalan las principales estructuras del globo ocular humano.

El globo ocular está relleno de tres cámaras con fluidos:

- La cámara anterior: situada entre la córnea y el iris que contiene humor acuoso.
- La cámara posterior: espacio relleno de humor acuoso entre el iris y el cristalino.
- La cámara vítrea: espacio relleno de humor vítreo entre el cristalino y la retina.

Las principales estructuras oculares del **segmento anterior** son:

- Córnea: superficie externa transparente del ojo, es la primera lente del sistema ocular. Está compuesta por epitelio estratificado escamoso no queratinizado, un estroma de tejido conectivo y una monocapa de células endoteliales de no más de 1 mm de grosor.
- Iris: se trata del músculo circular coloreado con una apertura en el centro (pupila) que regula la cantidad de luz que penetra en el ojo. Se encuentra anterior al cuerpo ciliar y juega un papel importante en la dinámica acuosa y en la configuración del ángulo iridocorneal, ambos afectan a la presión intraocular y al riesgo de glaucoma. El epitelio pigmentario del iris es de origen neuroectodérmico. El iris tiene dos regiones: la región periférica constituida por el músculo dilatador y la región interna con el músculo esfínter, que se encarga de la contracción. Ambos músculos permiten la acomodación del iris para controlar cantidad de luz que entra hacia el interior del ojo. Los dos son ejemplos raros de músculos derivados del ectodermo. En cuanto a la membrana pupilar y al estroma del iris, ambos son de origen mesenquimal. Proceden de la tercera ola de células migrantes de la cresta neural (Williams and Bohnsack, 2015). Durante el segundo mes de gestación, entre la separación del cristalino y la aparición del epitelio del iris, se forma una fina capa mesodérmica anterior al cristalino (Mann, 1925). La parte central formará la membrana pupilar, la cual se elimina tras la gestación. Las células mesenquimales migran a lo largo del borde anterior del epitelio del iris junto a los melanocitos del estroma, que también derivan de la cresta neural (Sturn and Larsson, 2009). Estos melanocitos son los responsables del color del iris, concretamente se debe a la concentración del pigmento y al tipo de melanina que contienen (ratio eumelanina:feomelanina).
- Cristalino: se localiza detrás del iris y es la segunda lente que atraviesa la luz, como tal se encarga de que la imagen se proyecte definida en los fotorreceptores. Para

que esto suceda la forma del cristalino cambia debido a los ligamentos que se contraen y relajan buscando la acomodación del mismo.

Las principales estructuras oculares del **segmento posterior** son:

- **Retina:** se trata de la capa más interna del ojo, en ella se diferencian dos regiones: la ciega y la sensible. La primera denominada papila, es la zona por donde se proyectan los axones de las células ganglionares hacia el nervio óptico. La segunda es la mácula, la región de mayor nitidez, de aproximadamente 5 mm de diámetro. Esta última recibe e integra la información luminosa. La retina está compuesta por distintas capas o estratos, que desde el exterior al interior son las siguientes: la capa nuclear externa, compuesta por los receptores (conos y bastones); la capa plexiforme externa, zona de sinapsis con células de la siguiente capa: la capa nuclear interna, con células bipolares. El axón de estas neuronas se dirige hacia la capa plexiforme, donde sinapta con las células ganglionares, situadas en la capa con su mismo nombre (Figura 11).
- **Fóvea:** anatómicamente se trata de la depresión que se encuentra en el centro de la mácula. Su formación se debe a que las neuronas de la retina se apartan dejando solo fotorreceptores en el centro. En ella se encuentra la mayor concentración de conos de la retina, entre 199.000 y 300.000 conos por mm^2 , mide aproximadamente 1,5 mm de diámetro y en su 0,5 mm interior está libre de capilares, lo cual hace que la luz llegue sin obstrucciones. Se trata de la región de la retina especializada en la discriminación de los detalles y en la visión en color.
- **Nervio óptico:** Los axones de las células ganglionares de la retina convergen formando este nervio, que se encarga de enviar la señal nerviosa hasta el cerebro.

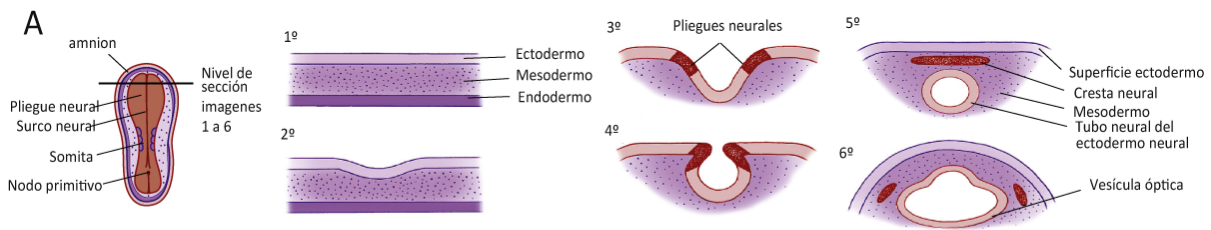
1.3. Embriología ocular

En la tercera semana del desarrollo embrionario tiene lugar la formación de las tres capas germinales primarias: ectodermo, mesodermo y endodermo (Figura I3.A), de las cuales las dos primeras formarán parte de las estructuras oculares. El ectodermo de la superficie dorsal del embrión experimenta un engrosamiento que dará lugar a la placa neural, la cual dará lugar a su vez al sistema nervioso central, incluyendo las estructuras visuales. Aproximadamente a los 18 días de gestación se forma un surco en el centro de la placa, sus bordes crecen hacia el centro envolviendo la ranura y creando así los pliegues neurales (Figura I3.A). La ranura se expande y los pliegues crecen el uno hacia el otro hasta fusionarse creando el tubo neural. Justo antes de esta fusión, un área de células de cada pliegue neural se separa del ectodermo formando las células de la cresta neural. Éstas forman agregados de células dentro del mesodermo que se encuentra rodeando al tubo neural (Figura I3.A). El tejido del tubo neural se denomina en este estadio ectodermo neural y su superficie, superficie neural. Ambas difieren en localización anatómica y en potenciales de diferenciación, tal como se ilustra en la Figura I3.B.

El neuroectodermo se diferencia en la retina, las capas posteriores del iris y el nervio óptico. El ectodermo de la superficie de la cabeza forma el cristalino y el epitelio corneal. El mesodermo, situado entre las dos capas anteriores, da lugar a las capas fibrosa y vascular del ojo. Las células de la cresta neural se diferencian en el endotelio corneal, la esclerótica y la coroides (Wright, 1997) (Figura I3B). En todo este proceso del desarrollo ocular en vertebrados tienen una función importante *PAX6* y los genes que regulan su transcripción (Mathers, 1997).

En la cuarta semana de desarrollo aparecen los surcos ópticos en los pliegues neurales, que se evaginan dando lugar a las vesículas ópticas proyectadas hacia el mesénquima vecino (Figura I3.C). Estas vesículas ópticas contactan con la superficie del ectodermo y envían una señal para que el ectodermo superficial adyacente a las vesículas se engrose formando las placodas del cristalino (Figura I3.C). La formación de la placoda del cristalino está seguida por la invaginación recíproca de la misma y de la vesícula óptica para establecer el primordio óptico, el cual se compone de la vesícula del cristalino, la copa óptica, la superficie del

ectodermo y el mesénquima vascular, a partir del cual se desarrollan los vasos sanguíneos hialoideos (Figura 13.C).



B

Superficie ectodérmica	Ectodermo neural	Cresta neural
Cristalino Epitelio corneal Epitelio conjuntivo Párpados (epitelio y glándulas) Epitelio ciliar Revestimiento sistema nasolagrimal	Epitelio pigmentario de la retina Retina neural Fibras nervio óptico Neuroglía Epitelio del cuerpo ciliar Epitelio del iris Músculos esfínter y dilatador del iris	Estroma corneal Endotelio corneal Esclera Estructura trabecular Células pigmentarias uveales Tejido conectivo uveal Músculo ciliar

C

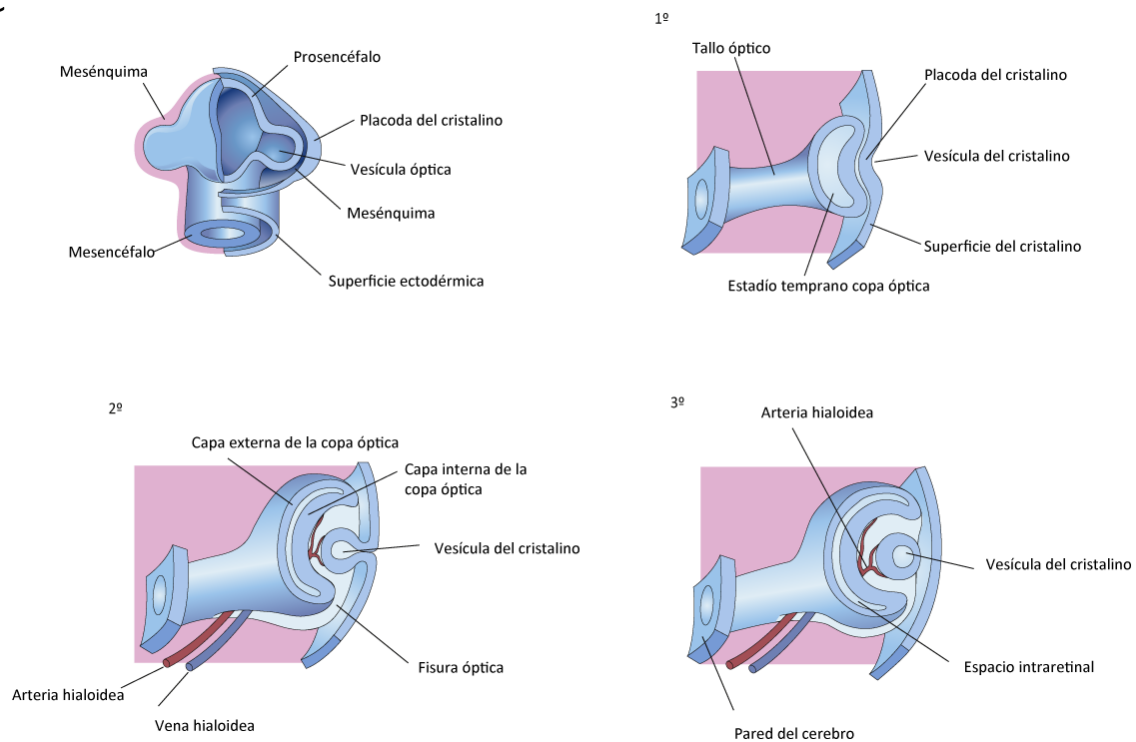


Figura 13. Desarrollo embrionario ocular. Figura adaptada de Remington, 2012. **A)** Formación de la superficie ectodérmica, el ectodermo neural y la cresta neural. **B)** Potenciales de diferenciación de la superficie ectodérmica, del ectodermo neural y de la cresta neural. **C)** Estadios tempranos de la formación de la copa óptica.

1.4. Mecanismos moleculares asociados al desarrollo ocular

La morfogénesis ocular es un proceso complejo que está dirigido por un programa genético altamente coordinado que media la diferenciación secuencial de los diferentes compartimentos oculares (Beccari, 2013) (Figura I4). En procesos iniciales de diferenciación, el factor transcripcional PAX6 juega un papel crucial como “master control” regulando la formación espacio-temporal del polo anterior y posterior (Shaham, 2012), en colaboración con otros reguladores transcripcionales (RAX, OTX2, SOX2, VAX2, PAX2) y señales morfogénicas BMP, FGF, ácido retinoico, TGF- β y Wnt (Furuta and Hogan, 1998; Wawersik, 1999; Mic, 2004; Smith, 2005; Morcillo, 2006; Rajagopal, 2009; Machon, 2010; Garcia, 2011; Grocott, 2011; Huang, 2015; Jidigam, 2015). Muchas de estas rutas están reguladas por PAX6 de forma directa (Figura I5), sin embargo, la red de señalización entre todas las células implicadas permanece sin resolver.

PAX6 es un factor de transcripción altamente conservado, que juega un rol clave durante la etapa embrionaria en la morfogénesis y el mantenimiento ocular y neural (Prosser and van Heyningen, 1998). Esta lista de genes diana de PAX6 llega hasta unos 200 (Coutinho, 2011; Shaham, 2012).

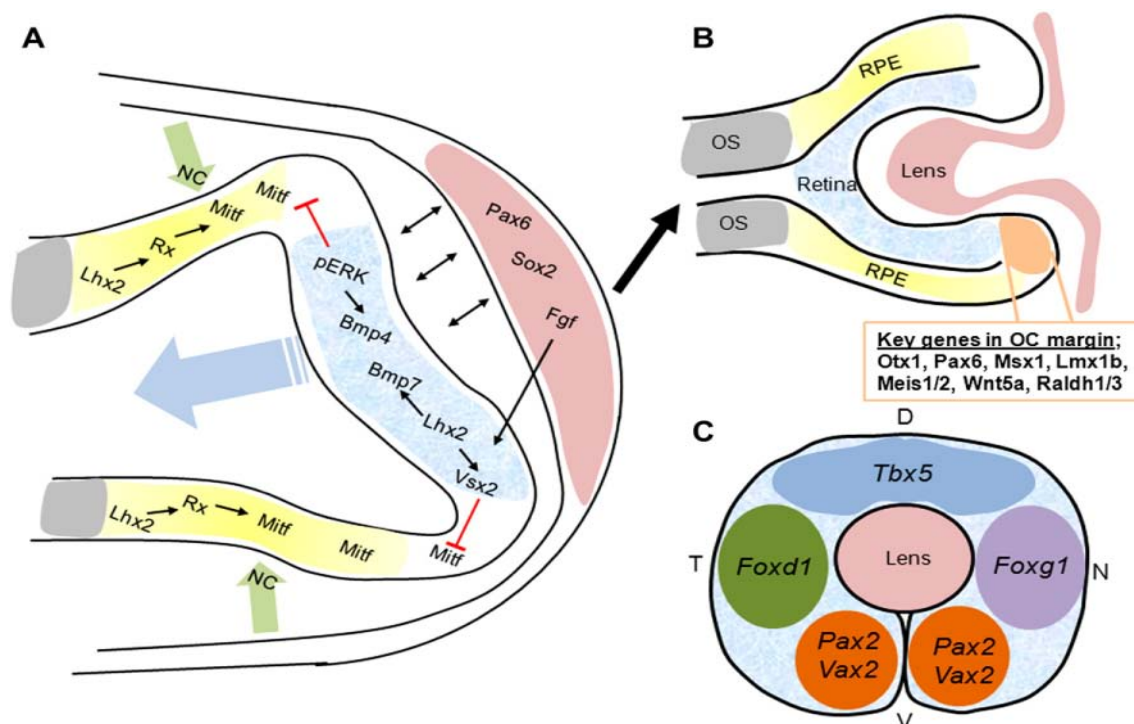


Figura I4. Procesos moleculares involucrados en el desarrollo ocular. Vease *Pax6* en el cristalino en A y como uno de los genes clave en el margen de la copa óptica en B. Figura tomada de Gregory-Evans, 2013.

Existe la hipótesis de que Pax6 regula al menos 3 capas:

1) Controla la expresión de otros factores de transcripción, que de forma conjunta regulan la expresión de genes diana en un mecanismo *feed-forward*, característico para la diferenciación terminal. Así, por ejemplo, Pax6, c-Maf, Prox1 y crystallinas regularían el desarrollo del cristalino; y Pax6, Mitf, Tyrp1, Tyr y Mlana participarían en el desarrollo del epitelio pigmentario de la retina (RPE) (Xie, 2011; Raviv, 2014; Fujimura, 2015; Sun, 2015).

2) Vías de transducción *upstream* a Pax6: señales FGF/MAPK/Erk, TGF- β y Hedgehog; y *downstream*: señales Notch, Wnt, TGF- β y ácido retinoico. Existen estudios en moscas donde se determinan redes donde se ponen de manifiesto estas interacciones (Desplan, 1997; Firth, 2009).

3) Función dual de Pax6 como activador y represor, en gran parte a través del reclutamiento de enzimas remodeladoras de cromatina. La especificidad de este reclutamiento está determinada por modificadores post-transcripcionales de Pax6 y factores de transcripción que se unen directamente a Pax6 en el ADN.

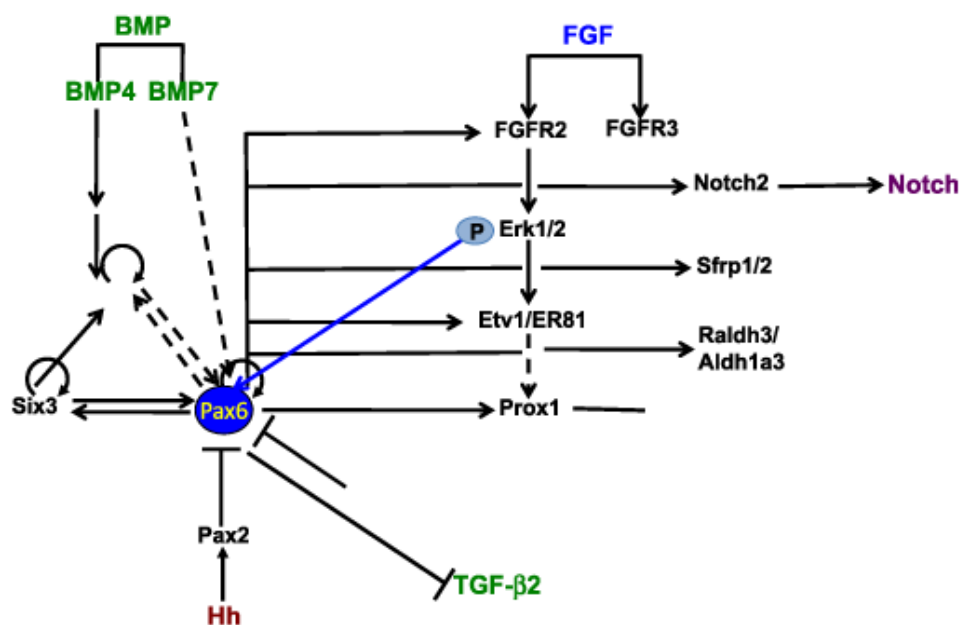


Figura 15. Señalización mediada por PAX6 detectada en organismos vertebrados. Figura tomada de Cvek, 2017.

1.5. Características del gen *PAX6*

1.5.1. Estructura génica y proteica

El gen *PAX6* (*paired box gene 6*) está localizado en el brazo corto del cromosoma 11, concretamente en la región 11p13. Este gen se extiende 22 Kb y consta de 14 exones, incluyendo al 5 alternativo (5a). *PAX6* codifica para una proteína de igual nombre que contiene dos dominios de unión al ADN, el dominio pareado (PD) (NH₂-terminal) y el homeodominio (HD), unidos por una región rica en glicina. Además, en el extremo COOH terminal de la proteína, hay un dominio rico en prolina-serina-treonina (PST) con función transactivadora. Éste es necesario para iniciar la transcripción y modular la unión al ADN por el HD (Ton CC, 1991; Glaser T, 1992; Zhang, 2011), que también regula la transcripción de los genes a los que se une dicha proteína. Por ello, *PAX6* se localiza generalmente en el núcleo de las células, donde ha sido detectado en varias ocasiones (Ploski, 2004; Tabata, 2018). En humanos, *PAX6* usa al menos 3 promotores (Azuma, 1999; Singh, 2002): P0 (exón 0), P1 (exón 1) and P α (intrón 4), lo que da lugar al menos a 3 isoformas del gen. La isoforma canónica (P0) de 422 amino ácidos (NM_000280), que contiene todos los dominios; la isoforma secundaria (P1) de 436 amino ácidos con la inclusión del exón 5a (NM_001604), que interrumpe el PD haciendo que reconozca una secuencia de ADN diferente; y finalmente la isoforma (P α), donde se pierde el PD.

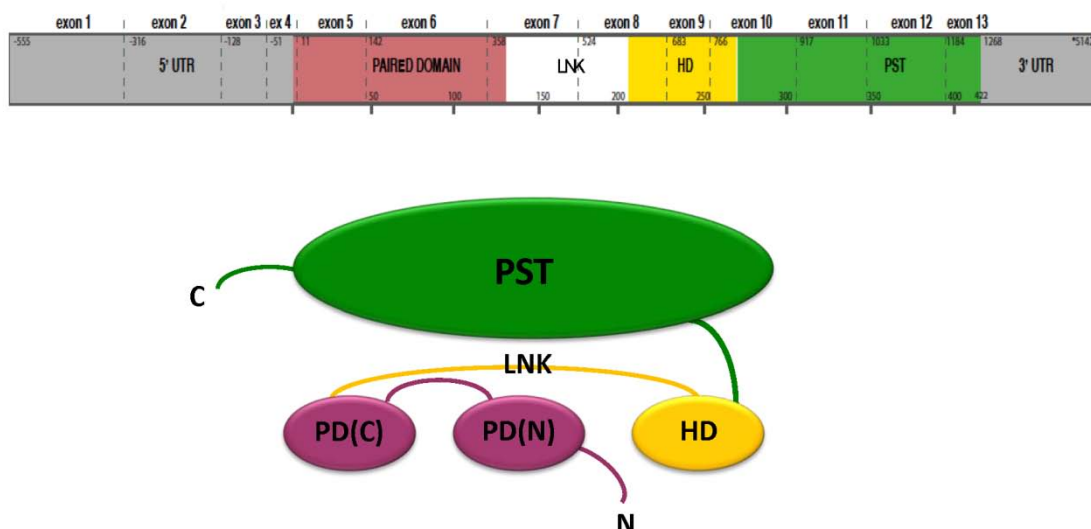


Figura 16. Representación del gen y la proteína *PAX6*. En la parte superior, se representa la estructura exónica del gen *PAX6*, incluyendo regiones UTRs no codificantes en gris y las diferentes regiones codificantes. Los números superiores del rectángulo indican la posición del ADN y los inferiores la posición proteica. En la parte inferior de la imagen, se representa la proteína y sus diferentes dominios funcionales. PD, dominio pareado; HD, homeodominio; PST, región rica en prolina/serina/treonina. PD(C): subdominio C-terminal del dominio pareado; PD(N): subdominio N-terminal del dominio pareado. Figura adaptada de Tzoulaki, 2005.

1.5.2. Regulación transcripcional de *PAX6*

La activación transcripcional de *PAX6* está controlada por la actividad coordinada de sus elementos *cis*-reguladores (CREs, del inglés *cis-regulatory elements*), que controlan su cantidad y localización espacio-temporal de forma específica, muy importante durante el desarrollo del embrión. Estos CREs son promotores, *enhancers*, *insulators* y elementos de unión (*boundary elements*) que se localizan desde en los intrones de *PAX6* hasta a cientos de kilobases aguas arriba y abajo del mismo, muchos de ellos en las regiones intrónicas de los genes vecinos *RCN1* y *ELP4* (Bhatia, 2013; Kammandel, 1999; Kleinjan, 2001; Kleinjan, 2004; Kleinjan, 2006; Mc Bride, 2011; Navratilova, 2009; Ravi, 2013). Hoy en día existen más de una treintena de *enhancers* de *PAX6* descritos, muchos de ellos identificados mediante el estudio de los puntos de ruptura de las translocaciones e inversiones en afectos de aniridia en los que *PAX6* permanece intacto (Fantes, 1995). Estos elementos están altamente conservados en vertebrados desde mamíferos a pez cebra (Kleinjan, 2004; Lakowski, 2007), llegando a tener algunos de ellos 450 millones de años. Además, se han identificado secuencias en regiones altamente conservadas como “agCNE” en antiguos gnathostomos (Bhatia, 2014).

En el año 2000 Lauderdale y cols. demostraron que los elementos 3' reguladores de *PAX6* son esenciales para su expresión génica utilizando células híbridas de humano y ratón tras la identificación de dos deleciones *de novo* aguas abajo a dicho gen (Lauderdale, 2000). El CRE más conocido es SIMO (Lauderdale, 2000), incluido en varias microdeleciones que excluyen a *PAX6* pero que se asocian al desarrollo de aniridia (Schwarz, 2000; Kleinjan, 2004; Bhatia, 2013). Recientemente se ha descrito una variante puntual en SIMO como causante de la aniridia, que afectaría al mecanismo autoregulatorio de *PAX6* (Bhatia, 2013). Todas estas evidencias demuestran la importancia de preservar la integridad de los elementos reguladores. Por lo tanto, las alteraciones en estos elementos pueden afectar a la dosis de *PAX6* y a su capacidad de respuesta ante señales externas.

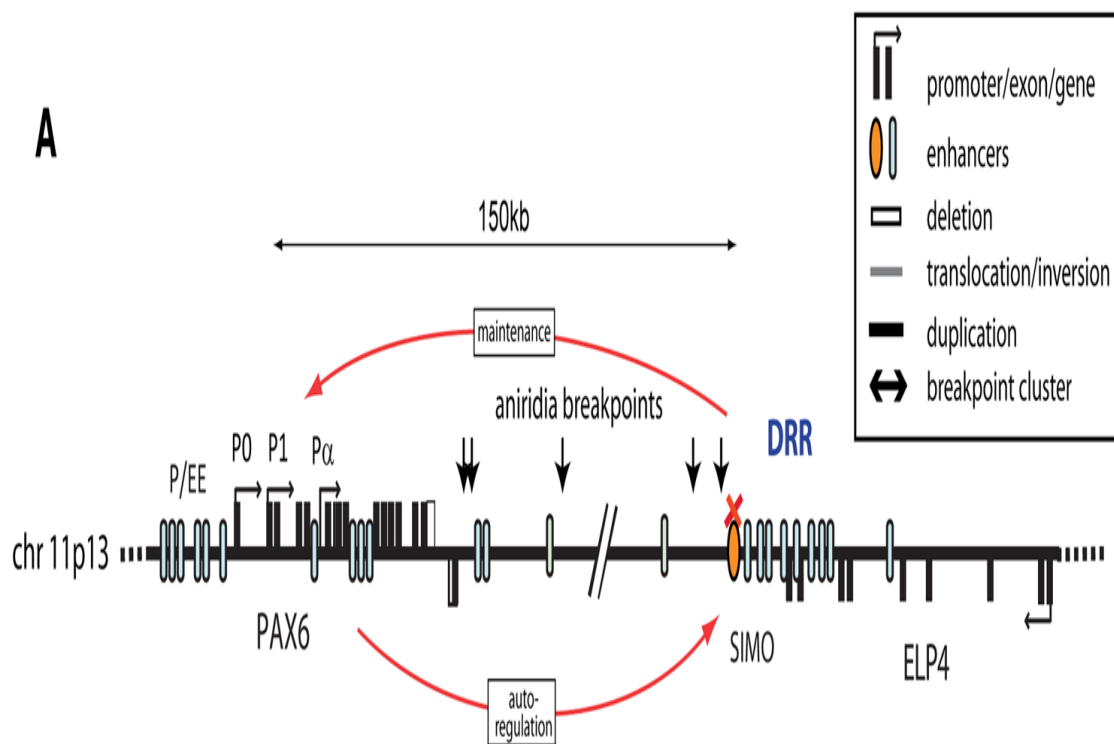


Figura 17. Representación del locus de PAX6 y sus elementos reguladores. Figura adaptada de Bathia, 2014b.

1.5.3. Expresión de PAX6 en diferentes tejidos

PAX6 no solo juega un rol clave durante la etapa embrionaria en la morfogénesis ocular y neural, sino también en etapa postnatal para el mantenimiento de estos sistemas (Prosser and van Heyningen, 1998). Por ello, PAX6 se expresa en retina, cristalino, córnea, iris, múltiples regiones del cerebro, también en médula espinal, epitelio olfatorio, cerebelo, además participa en la diferenciación de los islotes pancreáticos. Durante el resto de la vida se expresa en cerebelo, cerebro, en ojos participando en el mantenimiento del epitelio corneal, en el sistema olfatorio, en el sistema nervioso central, en intestino y en el páncreas en donde regula la función de los islotes α -pancreáticos (Simpson TI, 2002; Sivak, 2003; Hart, 2013; Yogarajah, 2016; Huettl, 2016; Maurya, 2017). PAX6 tiene también un papel importante en el desarrollo y la función de la glándula pituitaria, tanto en humano como en ratón (Terzic & Saraga-Babic, 1999; Kioussi, 1999; Bentley, 1999), así como de las glándulas pineales PAX6. Una representación de la expresión de PAX6 se detalla en la Figura 18.

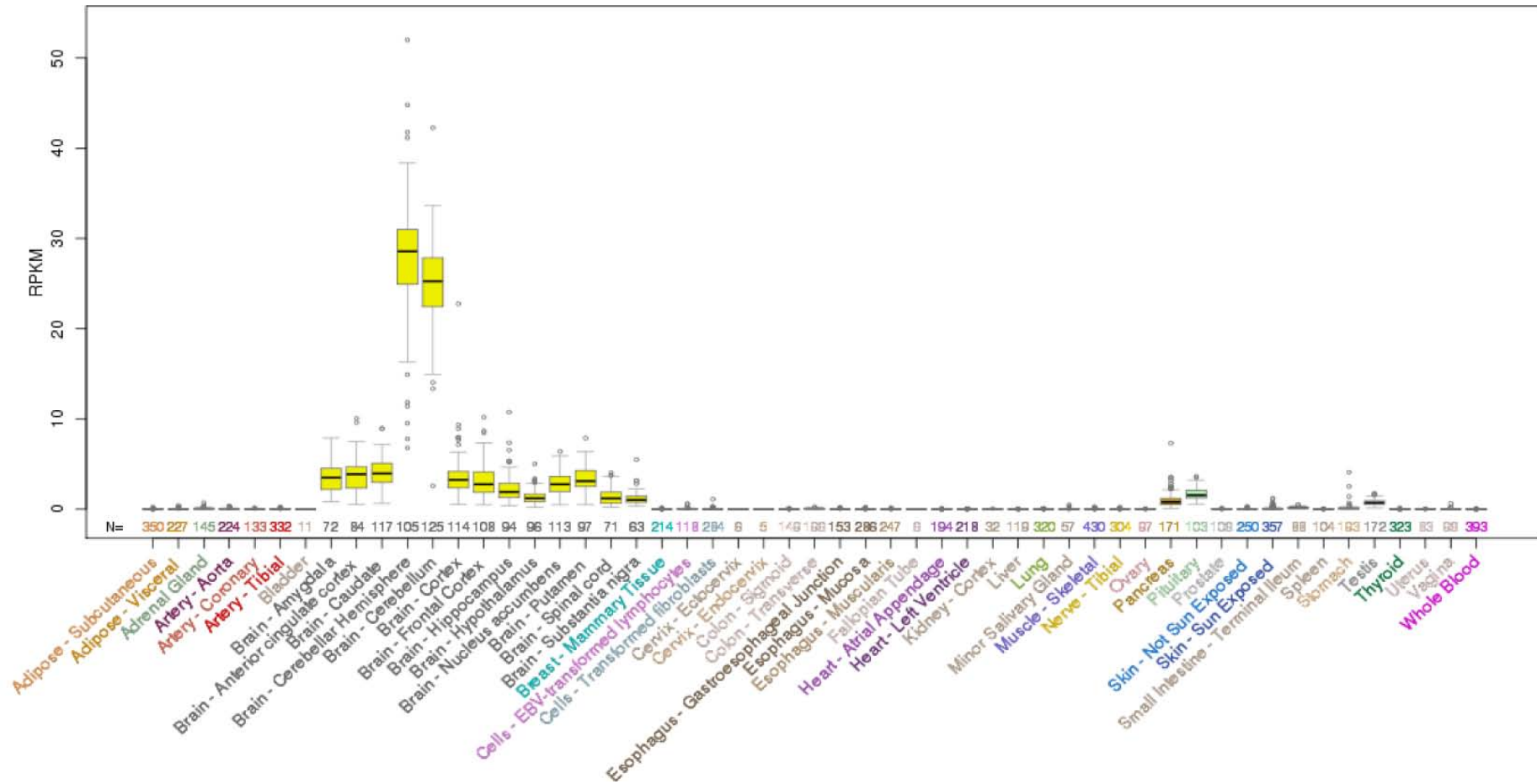


Figura 18. Representación de la expresión de la isoforma NM_000280.4 de PAX6 en diferentes tejidos humanos. PAX6 se expresa principalmente en cerebro y también en páncreas, pituitaria y testículos. Figura tomada de <https://commonfund.nih.gov/GTEX/>.

1.6. Defectos en PAX6 asociados a malformaciones oculares

Se ha descrito que una bajada en la dosis génica de PAX6 durante el desarrollo embrionario ocasiona una gran variedad de anomalías oculares en distintos modelos animales y en humanos (Figura 19) (van Raamsdonk, 2000; Nakayama, 2005; Davis-Silberman, 2005). No solo la haploinsuficiencia de este gen conduce a un desarrollo anómalo, existen ensayos en ratón que demuestran que su sobreexpresión también causa anomalías oculares como microftalmia, cataratas, microcórnea, defectos en la retina, en el iris y en los cuerpos ciliares (Schedl, 1996; Dora, 2008; Davis, 2009). Por lo tanto, se requiere de una dosis exacta del gen para cada tejido y proceso del desarrollo. En cuanto al requerimiento de la dosis exacta necesaria en determinados momentos del desarrollo, se ha descrito que la progresión de algunas de las anomalías asociadas a haploinsuficiencia de PAX6 se puede frenar o incluso revertir en modelos de ratón portadores de una mutación *nonsense* gracias al tratamiento con ataluren por vía tópica (Gregory-Evans, 2014) en estadios perinatales cuando aún no ha concluido el desarrollo ocular.

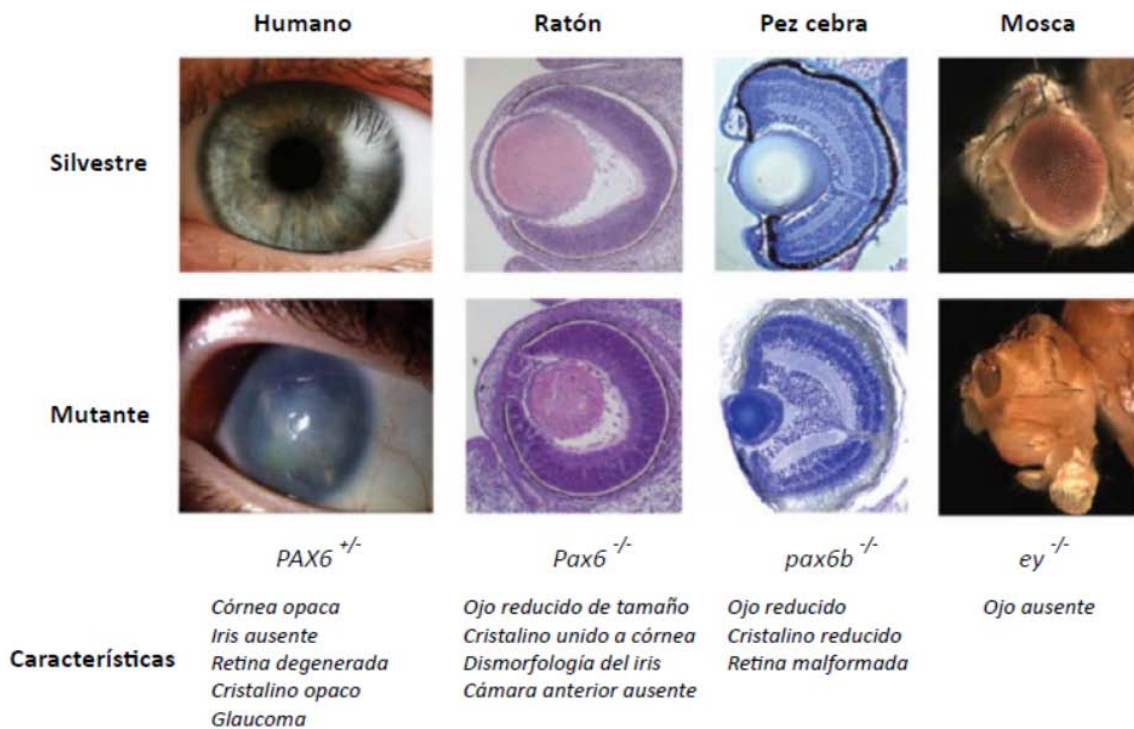


Figura 19. Variedad de anomalías oculares en distintos modelos animales y en humanos. Figura adaptada de Washington y cols, 2009.

En humanos, los defectos moleculares en *PAX6* han sido asociados también a distintos tipos de malformaciones oculares, siendo la más frecuente la aniridia congénita. En función del tipo molecular, se han descrito pacientes que presentan otros fenotipos incluyendo coloboma ocular, microftalmia, microcórnea, hipoplasia foveal aislada, anomalías de nervio óptico aisladas, queratopatía, cataratas congénitas, anomalía de Peters y anomalía de Morning glory (Mirzayans, 1995; van Heyningen and Williamson, 2002; Azuma, 2003; Tzoulaki, 2005; Neethirajan, 2006; Wang, 2012). Generalmente estos últimos casos de aniridia-like se deben a variantes *missense* en *PAX6*, mientras que la aniridia clásica está predominantemente asociada con variantes que truncan la proteína.

1.7. Aniridia congénita

La aniridia (OMIM 106210) es una enfermedad panocular congénita caracterizada principalmente por el desarrollo anómalo del iris y la fovea, que suele aparecer de manera bilateral, llevando a complicaciones visuales severas que pueden conducir a la ceguera. La aniridia fue descrita por primera vez en 1818 por Barratta, en 1926 se publicaron 14 casos (Rush, 1926) y en 1984 se describieron las características clínicas de esta enfermedad y las posibles hipótesis de su aparición (Nelson, 1984).

Se denomina aniridia tanto al signo clínico como a la condición en sí. La incidencia de este desorden en la población mundial es de 1 en 40.000-100.000 nacimientos (Hingorani, 2012), dependiendo de la población estudiada.

1.7.1. Manifestaciones oculares de la aniridia

El fenotipo clásico de aniridia se caracteriza principalmente por una hipoplasia del iris y la fovea, lo que lleva a una baja agudeza visual, generalmente de entre 20/100 y 20/200, aunque puede ser mejor si no se sufre nistagmo (20/60) (Hingorani, 2012). Con el progreso de la enfermedad aparecen cataratas, glaucoma y queratopatía asociadas a la aniridia. La hipoplasia de iris es la anomalía más característica de la enfermedad y puede ir desde la ausencia total a un defecto parcial (coloboma), pasando por pupilas excéntricas o deformes (*extropion uveae*) (Figura 110). El cristalino pierde transparencia, lo que conduce a la aparición de cataratas en el 50-85% de los casos, también puede existir en él vascularización remanente y/o una membrana pupilar persistente, sin embargo, la subluxación del mismo es menos frecuente. El glaucoma es bastante común (30-50% de los casos) y se suele

desarrollar en la adolescencia o en la etapa adulta. La queratopatía se debe principalmente a un fallo en las células troncales limbales, por lo que la córnea se vuelve opaca, se queratiniza y termina vascularizándose. Los pacientes suelen sufrir también malformaciones del polo posterior, principalmente una hipoplasia de la fovea o incluso su vascularización causada por un desarrollo anómalo de la misma. Además se ha observado hipopigmentación en la mácula e hipoplasia del nervio óptico aproximadamente el 10% de los enfermos.

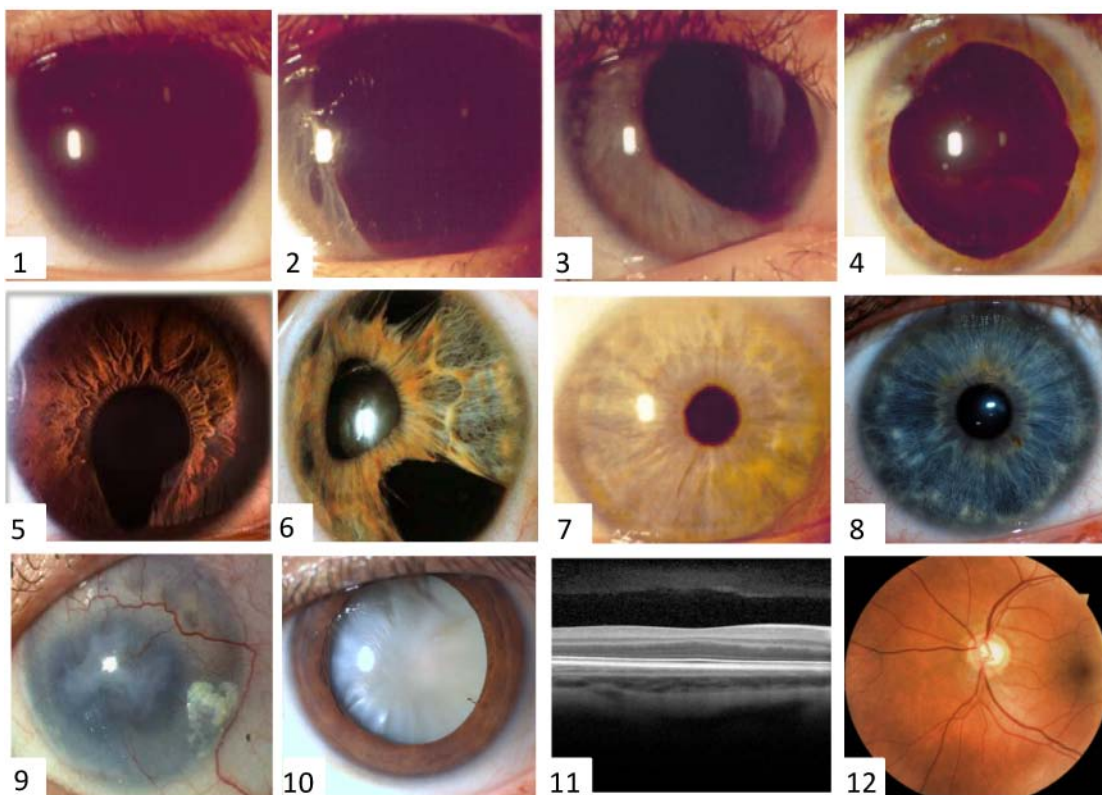


Figura 110. Fenotipos asociados a defectos en PAX6. Se muestran distintos ejemplos de anomalías del desarrollo ocular asociadas a variantes en PAX6, incluyendo: 1) aniridia completa, 2 y 3) hipoplasias o aniridias parciales, 4) atrofia iridiana, 5) coloboma de iris, 6) policoria, 7) corectopia, 8) iris normal, 9) queratopatía, 10) cataratas, 11) hipoplasia foveal e 12) hipoplasia de nervio óptico. Figura adaptada de Gronskov, 1999.

1.7.2. Fenotipos extraoculares asociados a aniridia

Algunos pacientes con aniridia aislada sufren síntomas extraoculares, el más común es la reducción del sentido del olfato (Sisodiya, 2001; Hingorani, 2012), menos usual es la dificultad auditiva debida a anomalías interhemisféricas que afectan al procesamiento auditivo central (Bamiou, 2007). También se han descrito anomalías neurológicas como la discapacidad intelectual y el autismo (Davis, 2008; Hu, 2015). Dichos trastornos se pueden explicar tras estudios con MRI que mostraron irregularidades en la comisura anterior, el córtex cingulado anterior, el cerebelo, los lóbulos temporal y occipital, el cuerpo calloso, la glándula pineal y el bulbo olfatorio (Sisodiya, 2001; Mitchel, 2003; Bamiou, 2007). Sin embargo, la función cognitiva suele ser normal (Thompson, 2004) y son raras las anomalías del comportamiento y el retraso en el desarrollo (Malandrini, 2001; Heyman, 1999). Más allá del sistema nervioso central, se ha detectado relación entre aniridia e intolerancia a la glucosa. Así, varios estudios indican un aumento en la prevalencia de diabetes en pacientes con aniridia en comparación con en el resto de la población (7% frente a 4,5%) (Netland, 2011; Menkle, 2015).

Además, la aniridia puede presentarse de forma aislada o como parte de un síndrome en 1/3 de los pacientes (Fischbach, 2005). El más conocido es el síndrome de WAGR (OMIM 194072), acrónimo de las siglas inglesas de las alteraciones que lo definen (*Wilms tumor, Aniridia, Genital abnormalities and Retardation*). Por lo tanto, este síndrome se caracteriza por el desarrollo de tumor de Wilms, aniridia, distintas anomalías genitourinarias y discapacidad intelectual. Estos afectos tienen un riesgo del 50-70% de desarrollar el tumor de Wilms, que es el cáncer renal infantil más común. Menos común es el síndrome de WAGRO, una de sus variantes en la que además está presente la obesidad.

1.8. Genética de la aniridia clásica

La aniridia clásica sigue un patrón de herencia autosómico dominante en 2/3 de los casos, el resto se tratan de casos esporádicos producidos por mutaciones *de novo* (Zhang, 2011; Hingorani, 2009). Por lo tanto, un afecto tiene un 50% de probabilidad de transmisión de la enfermedad a su descendencia. En 1991, Ton y cols, identificaron la región cromosómica asociada con el fenotipo de la aniridia gracias a estudios de ligamiento y clonaje posicional, mapeando clones de ADNc de pacientes con síndrome de WAGR portadores de deleciones en la banda 11p13 (Ton, 1991). Poco después, en 1992 se logró clonar el gen *PAX6* y se identificaron mutaciones en pacientes con aniridia (Glaser, 1992). Desde entonces, se considera que *PAX6* es el principal gen causante de aniridia (Figura 16).

La aniridia congénita se debe mayoritariamente a mutaciones puntuales en *PAX6*, pero hasta un 30% de los casos se deben a microdeleciones en la región 11p13 (Hall, 2019). Estas variaciones en el número de copias (CNVs) pueden afectar a dicho gen por completo o de manera parcial, incluso pueden incluir a genes contiguos. Las CNVs más frecuentes están asociadas al síndrome de WAGR e incluyen al menos a los genes *PAX6* y *WT1*. Este último sería el responsable del aumento en el riesgo de desarrollo del tumor de Wilms. En el caso del síndrome de WARGO, las deleciones se extienden hasta el gen *BDNF*, que se asocia a la aparición de la obesidad (Hingorani, 2012).

Se han descrito también casos de aniridia en los que las CNVs no afectan a *PAX6*, pero sí a sus regiones reguladoras. Adicionalmente, se han detectado algunos pacientes con translocaciones o inversiones que disrumpan la unidad de transcripción de este gen o sus elementos reguladores (Fantès, 1995; Gronskov, 2001; Crolla, 2002; Robinson, 2008; Redeker, 2008). Menos frecuentes, aunque no leves, serían las duplicaciones del gen. En el 2013 se reportó un caso de una duplicación de 877 Kb que incluía a *PAX6* y cursaba con severas consecuencias intra- y extraoculares (Schilter, 2013).

1.8.1. Espectro mutacional de PAX6

A día de hoy se han publicado cerca de 500 mutaciones en la literatura o recogidas en una base de datos mutacional de PAX6, *Human PAX6 Mutation Database* (LOVD) (Brown, 1998). De ellas, más de 300 son de tipos SNVs. Más del 90% de estas últimas llevan a la introducción de un codón de parada prematuro (PTC) en el marco abierto de lectura, a extensiones del dominio C-terminal, o a la sustitución de uno de sus aminoácidos (Hingorani, 2012).

Dentro de las mutaciones de codón de parada prematuro, se han descrito variantes *nonsense*, inserciones o deleciones *frameshift*, y variantes en los sitios canónicos de *splicing* de PAX6 (Gronskov, 1999; Hingorani, 2009; Robinson, 2008; Vincent, 2003; Tzoulaki, 2005). La mayoría se han descrito en los exones que codifican para el dominio pareado de unión al ADN (Figura I11). Estas variantes dan lugar a un ARNm aberrante conteniendo un PTC que es degradado mediante el sistema *nonsense mediated decay* (NMD), para prevenir la generación de proteínas truncadas. Este tipo de variantes dan lugar a alelos de pérdida de funcional (LOF) que conducen a la haploinsuficiencia de PAX6. Se han descrito 4 mutaciones *nonsense* que son *hotspot* (R203*, R240*, R261* y R317*) (Figura I11). Se trata de dinucleótidos CpG localizados en los exones 8, 9, 10 y 11, que representan el 21% de las mutaciones totales en el gen y el 59% de las variantes *nonsense* (Tzoulaki, 2005).

Las mutaciones C-terminales son aquellas que amplían el marco de lectura de manera que se alarga la proteína resultante, lo cual probablemente afecte a su función (Tzoulaki, 2005). Estos cambios no son detectados por el sistema NMD, que no actúa a partir de los 50 últimos pb del penúltimo exón (Hentze, 1999; Byers, 2002), y por tanto, no se espera que conduzcan a haploinsuficiencia.

Las mutaciones *missense* que provocan la sustitución de un aminoácido de la proteína PAX6 alterando su polaridad/carga y su estructura estérica, por lo que pueden afectar al plegamiento y la funcionalidad de la misma. En la mayoría de las ocasiones, suelen ser mutaciones de pérdida de función, pero también se ha descrito que algunas variantes *missense* pueden aumentar su afinidad de unión con el ADN con respecto al silvestre, actuando como dominantes negativos cuando se coexpresan

(Singh, 1998). Por tanto, no sería casualidad que cuatro quintas partes de las variantes *missense* se localicen en el dominio pareado de *PAX6*, modificando su unión a los genes diana (Hanson, 1999; van Heyningen, 2002; Azuma, 2003). Algunas de estas *missense* podrían dar lugar a formas hipomorfos (Tang; 1997; Favor, 2001; Vincent, 2003), las cuales mantienen algunas de las funciones pero posiblemente no regulen correctamente a los genes *downstream*.

Por otra parte, destacar que solo se han identificado un pequeño número de mutaciones en la región 5'UTR de *PAX6* (Glaser, 1992; Gronskov, 1999; Ticho, 2006; Robinson, 2008; Redeker, 2008; Skeens, 2011; Vasilyeva, 2017). Hasta el momento de esta Tesis, tan solo había 9 variantes descritas en la literatura y/o en LOVD, tal como se representa en la Figura I12.

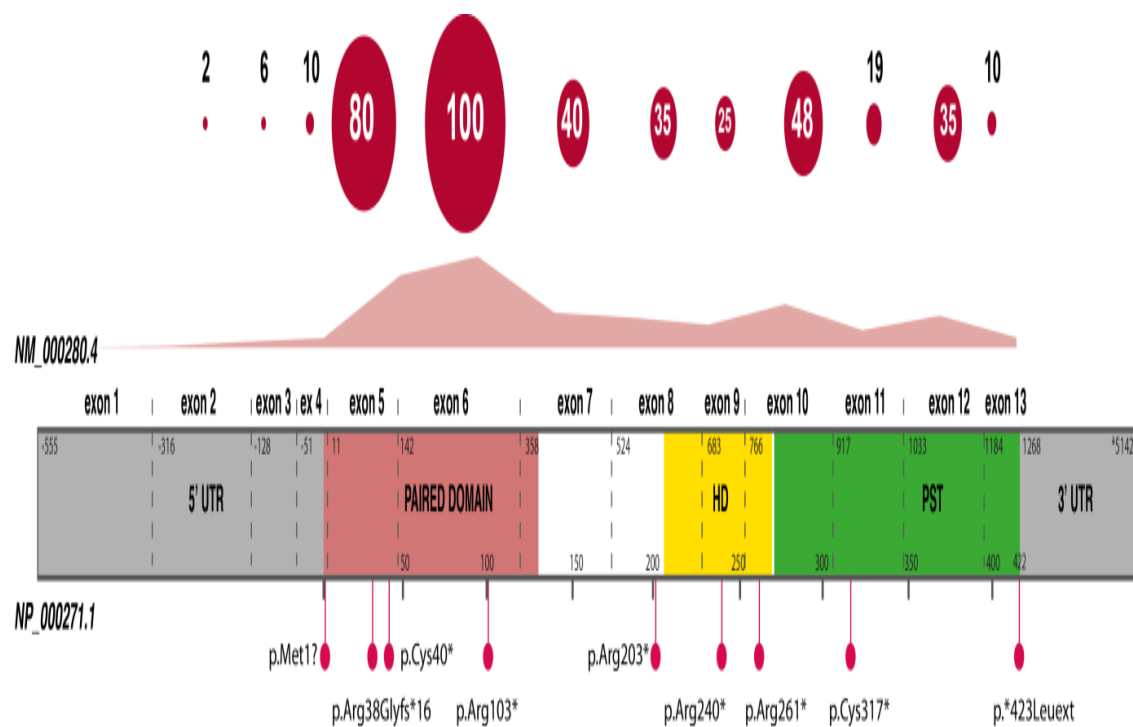


Figura I11. Representación del número de variantes publicadas para cada exón de *PAX6*. Se representa en gris las regiones 5' y 3' UTR; en ros, amarillo y verde, los exones que codifican los principales dominios funcionales de *PAX6*, pareado, homeodominio y el dominio rico en prolina-serina-treonina, respectivamente. En la parte superior de la imagen se representa el recuento de las variantes descritas en *PAX6*, por exones, en la base de datos LOVD. En la parte inferior se representan las variantes identificadas con más frecuencia.

Las mutaciones bialélicas en *PAX6* son excepcionales ya que la pérdida del gen en homocigosidad es habitualmente incompatible con la vida, provocando la muerte perinatal. (Glaser, 1994; Schmidt-Sidor, 2009). Sin embargo, existen varios casos publicados sobre embarazos que hayan llegado a término bajo estas condiciones. Dos neonatos con mutaciones bialélicas en *PAX6* murieron a los pocos días de nacer como consecuencia de las graves anomalías cerebrales, ambos con anoftalmia (Glaser, 1994; Schmidt-Sidor, 2009). Otro individuo heterocigoto compuesto sobrevivió hasta al menos los 4 años también con fenotipo ocular severo de microftalmia, microcefalia y además con trisomía 21, y diabetes *mellitus* (Solomon, 2009). Sin embargo, se han publicado otros dos casos con aniridia o coloboma sin anomalías sistémicas causados presuntamente por la combinación de un alelo hipomorfo moderado y otro *missense* (Gronskov, 1999; Chao, 2003).

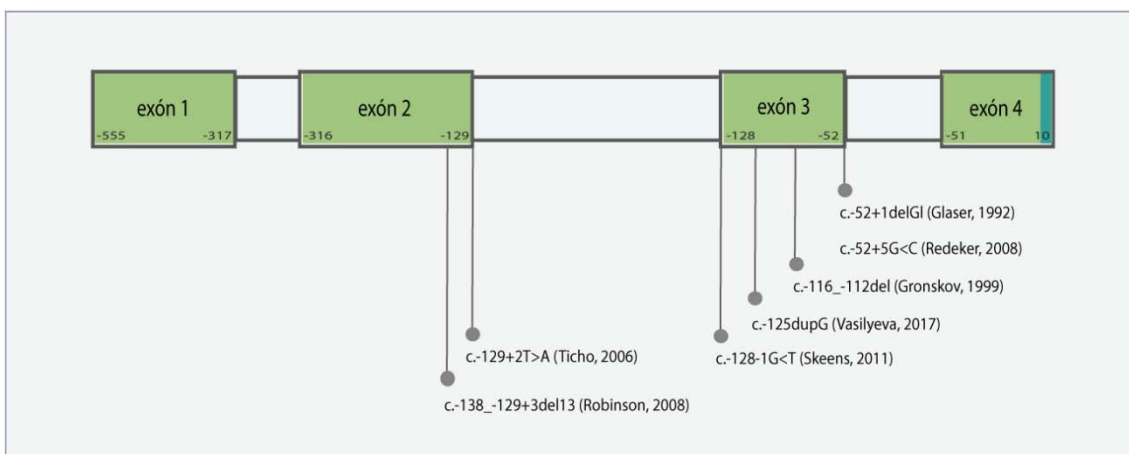


Figura I12. Representación de la localización de las variantes descritas en 5'UTR de *PAX6*. Datos obtenidos de la base de datos LOVD. Figura adaptada de Plaisancie & Tarilonte, 2018.

1.8.2. Correlación genotipo-fenotipo

Pese a que hace 28 años que se descubrió el gen causante de la aniridia, la correlación genotipo-fenotipo permanece aún sin resolver. Las variantes de *PAX6* conducen a un espectro ocular que depende de su naturaleza y de su localización en el gen (Lee, 2013). En un estudio realizado en el 2005, Tzoulaki describió que el 80% de los casos de aniridia están causados por mutaciones que crean un PTC y solo el 10% son debidas a mutaciones *missense*.

Generalmente se suele aceptar que las mutaciones que llevan a la introducción de un PTC están asociadas con fenotipos de aniridia clásica con hipoplasia de iris y de fovea. El fenotipo asociado a las variantes de tipo *run-on* suele ser aniridia, por tanto, se deduce que también generan alelos de pérdida de función. Sin embargo, los fenotipos atípicos, que en esta Tesis clasificaremos como *aniridia-like*, suelen estar asociados a variantes *missense* pero también a variantes no codificantes del gen (Tzoulaki, 2005; Hingorani, 2009). Estos fenotipos atípicos pueden ser más leves, como formas aisladas del coloboma de iris o de hipoplasia foveal, pero también han sido asociadas a formas más severas como la anomalía de Peters o microftalmia (Hanson, 1999; Azuma, 2003; Nallathambi, 2006; Thomas, 2014).

Se hipotetizó que las mutaciones que crean un PTC tras el HD dan lugar a proteínas truncadas cuyo mecanismo de acción es el de dominante negativo, ya que se cree que al permanecer intactos los dominios de unión al ADN la proteína se une a éste, pero no activa a los genes *downstream*, por tanto, interfiere con la función normal de la proteína (Singh, 1998; Duncan, 2000; Yasuda, 2002). De este tipo de variantes en principio se esperaba que diesen lugar a fenotipos más severos o marcadamente diferentes a los PTC anteriores al dominio PST. Sin embargo, más del 98% de los PTC, tanto anteriores como internos al dominio PST, se relacionaron con la aniridia en un estudio realizado en el año 2005 (Tzoulaki, 2005). De este trabajo se deduce que las mutaciones truncantes se asocian con fenotipos de aniridia independientemente de su localización en el gen, esto explica la falta de correlación fenotipo-localización en las mutaciones que provocan PTC, ya que todas llevarían a un mecanismo de haploinsuficiencia.

Por el contrario, sí podríamos esperar fenotipos más severos en mutaciones localizadas en regiones críticas, como los dominios de unión al ADN, en el caso de mutaciones dominantes negativas o de ganancia de función (Lee, 2013). Incluso se ha llegado a especular que las *missense* localizadas en el subdominio N-terminal del PD afectan a la cámara anterior y las del C-terminal afectan a la cámara posterior, sin embargo esta regla no siempre se cumple, se trata de un mecanismo más complejo (Azuma, 1996). En general las *missense* reducen la unión de *PAX6* al ADN, pero las hay que impiden dicha unión como la R125C (Axton, 1997). Otras *missense* aumentan la función de la proteína, este es el caso de la V54D, localizada en el exón 5a, que neutraliza el efecto negativo de este exón alternativo en la unión proteína-ADN.

En general, existe una gran variabilidad fenotípica asociada a la aniridia a nivel inter- e intrafamiliar, por lo que una misma mutación puede causar diferentes presentaciones y severidades en el desarrollo de la patología. Una de las explicaciones a la expresividad variable en la aniridia asociada a variantes en *PAX6* no solo sería el tipo de mutación sino la influencia genética de alelos en otros genes modificadores de *PAX6*, localizados tanto *upstream* como *downstream* (Lee, 2013). Otra hipótesis que podrían explicar este hecho es el mosaicismo en algunos casos con manifestaciones más leves o la existencia de un fenómeno de fenocopia de la aniridia (Lee, 2013).

1.9. Diagnóstico diferencial

Se han descrito patologías que presentan fenotipos solapantes con la aniridia, principalmente en las manifestaciones del segmento anterior. Así, algunos genes asociados a disgenesias del segmento anterior pueden causar hipoplasias de iris, similares a las presentes en algunos pacientes con aniridia y mutaciones en *PAX6*. Entre ellos destaca los genes *FOXC1* y *PITX2*, asociados a un espectro de anomalías del segmento anterior, desde formas aisladas de iridodisgenesia, anomalía de Rieger, anomalía de Peter, y a formas sindrómicas. El síndrome de Axenfeld-Rieger (OMIM 180500 y 601499), con herencia autosómica dominante, también presenta malformaciones del segmento anterior ocular como la ceguera por glaucoma en el 50% de los casos, y se puede presentar con alteraciones sistémicas, incluyendo hipoplasia maxilar y dental y piel umbilical redundante en las formas asociadas a mutaciones en

PITX2 o malformaciones cardíacas en las formas asociadas al gen *FOXC1* (Alkemade, 1966). Algunos pacientes con aniridia pueden presentar otras malformaciones características de la anomalía de Peters (OMIM 604229) que asocia anomalías en la cámara anterior, principalmente opacidad corneal (Peters, 1906). Esta patología puede deberse a mutaciones en los genes *PAX6*, *PITX2*, *CYP1B1* y *FOXC1*. El síndrome de Peters-Plus, que asocia retraso de crecimiento, baja talla, braquidactilia y retraso del crecimiento, además de otras malformaciones cardiovasculares y genitourinarias, tiene un patrón de herencia mayoritariamente autosómico recesivo por mutaciones en el gen *B3GLCT*.

En pacientes con aplasia iridiana, se debe realizar un diagnóstico diferencial con el síndrome de Gillespie (OMIM 206700). Este síndrome, caracterizado además por la presencia de ataxia cerebelar no progresiva y discapacidad intelectual, se caracteriza por una típica presentación iridiana de aplasia circumpupilar (Hingorani, 2012). Esta enfermedad no cursa con queratopatía, cataratas e hipoplasia foveal, características de la aniridia clásica. Este síndrome fue reportado por primera vez en 1965 por Frederick Gillespie y explica el 2% de los casos de aniridia-*like*. Este síndrome ha sido recientemente caracterizado genéticamente y representa una entidad distinta a la aniridia clásica causada por variantes dominantes y recesivas en el gen *ITPR1* (Gerber, 2016; McEntagart, 2016).

En otros casos, algunos recién nacidos con presencia de alteraciones iridianas, baja agudeza visual, fotofobia y nistagmo, en donde resulta difícil realizar una valoración oftalmológica, debería realizarse un buen diagnóstico diferencial para no confundirlo con otras anomalías como coloboma de iris, albinismo ocular, hipopsia foveal, cataratas congénitas, hipoplasia de nervio óptico aislada, etc. Para ello suele ser necesario además de una revisión oftalmológica en profundidad, la realización de estudios genéticos que permitan valorar los genes asociados a estas formas solapantes de la aniridia.

1.10. Estudios genéticos en la aniridia

Es importante realizar un estudio genético en los casos con sospecha clínica de aniridia. En base a esto se puede realizar un consejo genético al individuo en cuestión y para sus familiares afectos. También es importante el hecho de poder dar un adecuado consejo reproductivo, ya que en el caso de la aniridia se puede aplicar la técnica del diagnóstico genético preimplantacional (DGP) para poder seleccionar individuos sanos e implantarlos en el útero materno. Este estudio genético permite hacer un diagnóstico diferencial para descartar los fenotipos solapantes con la aniridia.

Los estudios de la aniridia tradicionalmente priorizan la búsqueda de deleciones que impliquen a *PAX6* y *WT1* con el motivo de descartar el síndrome de WAGR en casos de riesgo, es decir, en pacientes menores de 10 años sin antecedentes familiares (Hingorani, 2012; Blanco-Kelly, 2012; Richardson, 2016) (Figura I14). Para ello, se puede utilizar la técnica de MLPA (*Multiplex ligation-dependent probe amplification*) utilizando un ensayo específico para analizar estos genes y otros adyacentes en las regiones genómicas 11p13-14 (Redeker, 2018). Recientemente, se está extendiendo el uso de *arrays* de hibridación genómica comparada (aCGH), tanto en diseños comerciales como en diseños personalizados para analizar CNVs en el *locus* de WAGR (Blanco-Kelly, 2017; Franzoni, 2017), que permiten acotar los puntos de rotura con mayor precisión. En caso de identificar grandes deleciones, estas se pueden validar con la técnica FISH (*fluorescent in situ hybridization*) o con cariotipo de alta resolución. La técnica FISH también se suele utilizar para descartar la presencia de reordenamientos equilibrados que disrumpen *PAX6* sin producir variaciones en el número de copias (Crolla, 2002; Robinson, 2008).

En los casos negativos o en individuos con antecedentes familiares se realiza un estudio de mutaciones puntuales intragénicas en los exones codificantes de *PAX6* y en sus regiones de *splicing*. Tradicionalmente este estudio se ha realizado mediante secuenciación Sanger. Gracias al desarrollo de las nuevas técnicas de secuenciación masiva, actualmente se está implementando el uso de paneles de genes para el estudio de la aniridia (Plaisancie, 2018) o, como ocurre en nuestro laboratorio, el estudio de las regiones codificantes mediante secuenciación del exoma clínico. El uso

de NGS permite estudiar cientos de genes a la vez, lo cual está acelerando el proceso de caracterización genética en los casos de aniridia sin mutaciones en *PAX6*.

Con este algoritmo tradicional basado en MLPA y secuenciación Sanger de las regiones codificantes se logran caracterizar entre el 50 y el 80% de los individuos sospechosos de aniridia, en función de la población en estudio y la tecnología utilizada (Gronskov, 2001; Robinson, 2008; Hingorani, 2009; Lim, 2012; Dubey, 2015). Dada la implicación de los elementos *cis*-reguladores de *PAX6* en la aniridia (D'Elia, 2007; Bathia, 2013; Cvekl, 2017), el cribado adicional de los mismos y un correcto fenotipado permite aumentar la tasa de diagnóstico hasta el 90% en algunos estudios más recientes (Bobilev, 2016; Vasilyeva, 2017; Sannan, 2017; Souzeau, 2018). Por lo tanto, se estima que el 90% de los casos de aniridia podrían estar causados por alteraciones genéticas en el gen *PAX6*, ya sea en su región codificante, en las no codificantes o en sus *enhancers*.

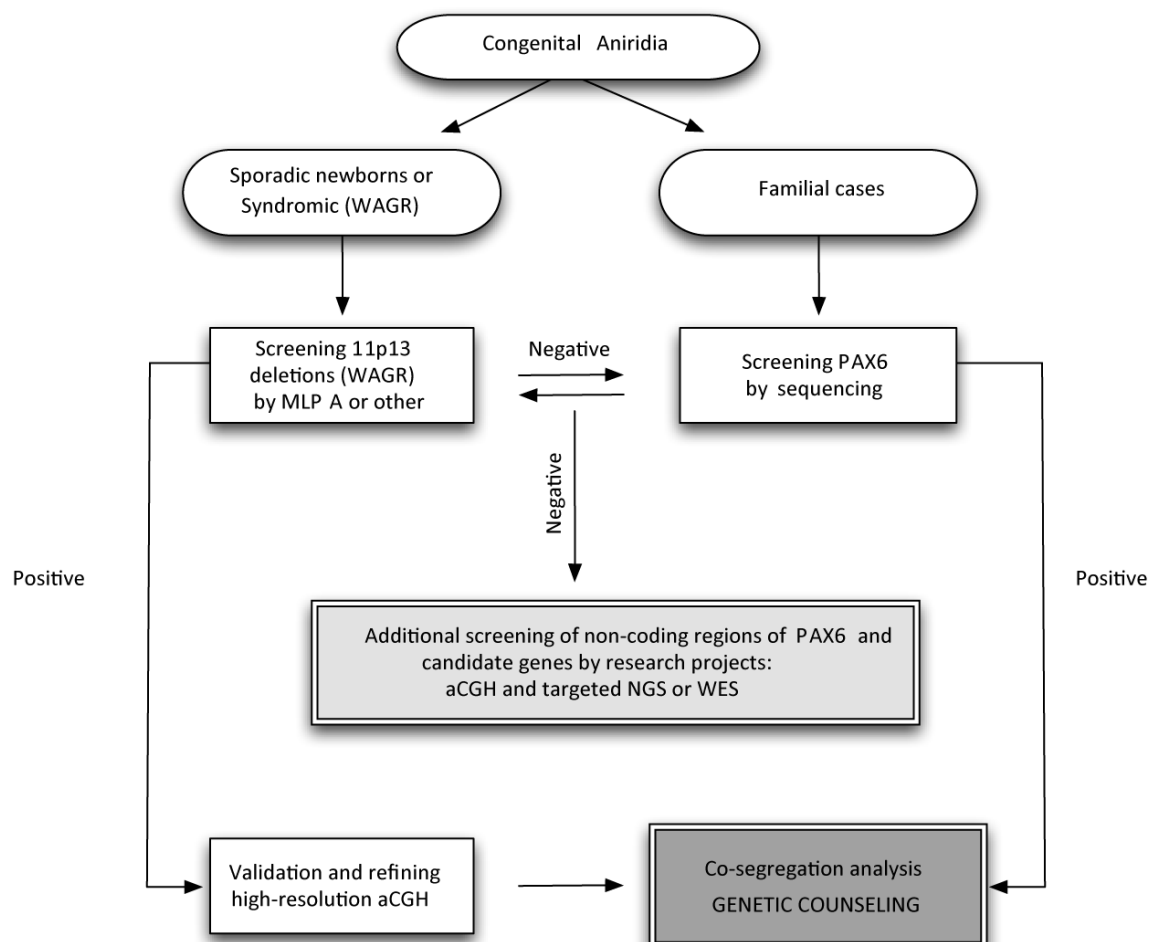


Figura 113. Algoritmo de diagnóstico genético en aniridia.

1.11. Casos de aniridia *PAX6* negativos

Dado que el 10% de los pacientes con aniridia no presentan mutaciones en *PAX6* tras su estudio exhaustivo, deben existir otros genes o mecanismos genéticos aún no identificados. En la identificación de las causas genéticas de los casos de aniridia no caracterizados genéticamente, el desarrollo de las nuevas tecnologías de secuenciación juega un papel fundamental. El abaratamiento de la NGS está generalizando el uso de WES como la alternativa más adecuada y efectiva para esclarecer nuevas causas genéticas de las enfermedades raras con etiología desconocida. En esta aproximación se utiliza una estrategia libre de hipótesis pues se analiza el exoma, es decir, todas las regiones codificantes del genoma. Por otro lado, la secuenciación del genoma completo (WGS, *whole-genome sequencing*) presenta la ventaja de cubrir también regiones no codificantes, como los intrones y UTRs, intrónicas, así como las regiones reguladoras; sin embargo su aún elevado coste hace que su uso no se haya extendido todavía a día de hoy.

Desde la identificación de *PAX6* en 1991 (Ton, 1991), existen pocos estudios que se hayan centrado en la identificación de nuevos *loci* para la aniridia. En 2015, se identificó mediante WGS, al gen *TRIM44*, como la causa de la aniridia en una familia china (Zhang, 2015). Este estudio determinó además que la mutación *missense* identificada inhibe la expresión de *PAX6*, lo que indica que podría ser regulador negativo de la transcripción de *PAX6*, y por tanto, lo hace candidato a ser un nuevo gen asociado a aniridia.

Dado el gran solapamiento fenotípico entre la aniridia y otras disgenesias del segmento anterior, algunos de los genes asociados a estas patologías son potenciales candidatos a causar aniridia. Sin embargo, existen pocos trabajos que hayan abordado el cribado de genes implicados en desarrollo ocular en cohortes de pacientes sin mutaciones en *PAX6* (Lim, 2012; Ansari, 2016). Del estudio de genes como *PITX3*, *FOXE3*, *FOXD3* y *NR2E1*, se obtuvieron resultados negativos (Corso-Diaz, 2012; Kloss, 2012). En otros casos sin mutaciones en *PAX6*, se han identificado mutaciones en otros genes asociados a disgenesias del segmento anterior, principalmente en *FOXC1* y *PITX2* (Khan, 2008; Ito, 2009; Karthikeyan, 2014; Ansari, 2016). Dado que muchas de estas

variantes se han identificado en pacientes con fenotipos aniridia-like, esto nos lleva a subrayar la importancia de un buen diagnóstico diferencial.

Dadas las limitaciones en el abordaje genético tradicional para secuenciar las regiones intrónicas y reguladoras, resulta imprescindible abordar el estudio de pacientes “*PAX6-negativos*” mediante estrategias que permitan la identificación de alteraciones en éstos o en nuevas regiones reguladoras. Otro aspecto a valorar en el futuro es la posible presencia de mosaicismo somático, que podría explicar algunos de los casos sin mutaciones en *PAX6*.

1.12. Manejo clínico y perspectivas terapéuticas

En la actualidad no existe un tratamiento de la aniridia como tal, pero debido a que es una enfermedad que afecta a muchas estructuras oculares sí que existen soluciones quirúrgicas o tratamientos paliativos para algunas de las complicaciones. Los errores de refracción se corrigen mediante lentes o gafas, estas últimas recomendadas en casos de problemas corneales, que si además están tintadas reducen la fotofobia. La terapia de oclusión funciona para corregir el estrabismo o la ambliopía anisométrica. Existen ayudas para el estudio como son las lupas para la lectura de texto. Los niños con síndrome de WAGR necesitan seguimiento renal siendo necesaria la cirugía en caso de tumor renal. El glaucoma se trata con medicación tópica y si no funciona se puede recurrir a la cirugía. La extracción de cataratas se realiza en casos de cristalino extremadamente opaco (Lee, 2008), ya que puede no mejorar del todo la agudeza visual debido a la existencia de otras estructuras oculares dañadas. En los casos de síndrome de fibrosis anirídica se necesita de cirugía para eliminar la membrana.

Los problemas corneales, asociados al 90% de los casos de aniridia, son la principal causa de la progresiva pérdida de visión en la aniridia. El tratamiento es paliativo y se basa en el uso de lubricantes, mucolíticos y en la oclusión lagrimal puntual en primera instancia, ya que la cirugía corneal no suele tener buen pronóstico, siendo realizada en casos en los que la opacidad corneal reduce significativamente la visión. Una de las opciones novedosas para tratar la queratopatía asociada a la aniridia es el trasplante *ex vivo* de células troncales limbales (Lee, 2008; Hall, 2019). Otra opción para la queratopatía es el trasplante autólogo de epitelio de la mucosa oral, técnica

prometedora pero con resultados variables (Shortt, 2011). En caso de fallo del trasplante la última opción sería la queratoprótesis, que consiste en la implantación de una córnea artificial.

Aunque *a priori* la reparación de anomalías congénitas asociadas a la aniridia parece utópica, parte de la discapacidad visual que presentan los pacientes en edad adulta se debe a complicaciones secundarias, por lo que sería interesante modular o incluso evitar su aparición. En la actualidad, el desarrollo de nuevos protocolos basados en medicina regenerativa con el uso de células madres son muy prometedores en el tratamiento de la queratopatía asociada a la aniridia. Recientemente, se ha publicado un modelo *in vitro* de células troncales del limbo procedentes de un donante sano, modificadas mediante edición genómica usando el sistema CRISPR-Cas9 para introducir la mutación *nonsense* (p.E109*), en donde se ha ensayado una terapia sustitutoria con proteína recombinante de PAX6 soluble, con resultados muy prometedores (Roux, 2018).

Por otro lado, se está desarrollando en Canadá un ensayo clínico en fase 2 (NCT02647359) para probar la eficacia del Ataluren (PTC-124, Translarna), suministrado por vía tópica en solución START (0,9% cloruro sódico, 1% *tween* 80, 1% *ataluren*, 1% *carboximetilcelulosa*) (Wang, 2015). El Ataluren es una molécula *readthrough* que restaura la traducción en presencia de codones PTCs ya probado en otras patologías como la fibrosis quística (Kerem, 2014) y la distrofia muscular de Duchenne (McDonald, 2017). Aunque su eficacia ha sido cuestionada a lo largo de los años, el Ataluren está aprobado como medicamento huérfano en pacientes con mutaciones *nonsense* en el gen *DMD*. Existen evidencias de su eficacia en otras patologías oculares *in vitro* e *in vivo*, lo que abre un abanico de posibilidades para su administración oftálmica.

El uso de esta molécula demostró resultados prometedores en un estudio preclínico en el modelo de ratón *Pax6^{Sev+/-}*, portador de una variante *nonsense* (Gregory-Evans, 2014; Wang, 2017). La manipulación perinatal de la dosis de Pax6 mediante la supresión del PTC provocado por mutaciones *nonsense* no solo ha frenado, sino que ha revertido las malformaciones en córnea, cristalino y retina, llegando a mejorar la

respuesta del ERG (electrorretinograma) en la retina de estos animales tratados (Gregory-Evans, 2014). Estos hallazgos ponen de manifiesto que las estructuras oculares presentan una inesperada plasticidad postnatal que puede aprovecharse para revertir defectos congénitos. Dado que las variantes *nonsense* representan una importante causa genética de aniridia, y entre ellas se encuentran variantes *hotspot* identificadas en decenas de pacientes en todo el mundo (Hingorani, 2012), esta estrategia puede ser muy prometedora para el tratamiento de la aniridia. A pesar de estos resultados tan esperanzadores, hay que ser cautos con las posibilidades reales para la aplicabilidad clínica del Ataluren en pacientes, pues no se han publicado estudios en células derivadas de pacientes, por lo que existen muchos aspectos por resolver antes de que esta terapia, si se demuestra realmente su eficacia, pueda extenderse. Este ensayo clínico con el uso de Ataluren en niños con aniridia, estaba previsto que concluyera en 2018, pero a fecha de la escritura de esta Tesis Doctoral, los primeros resultados aún no han sido aún publicados.

La teoría sobre la reversión postnatal abre nuevas expectativas para el desarrollo de tratamientos que permitan la recuperación funcional en aniridia y otras MOC (Sahel, 2014), bien con moduladores de PTCs o con otras estrategias terapéuticas dirigidas a otros tipos de mutaciones, como podría ser el uso de oligos antisentido dirigidos a modular variantes de *splicing*.

Objetivos

El **objetivo principal** de esta Tesis Doctoral ha sido el estudio de nuevos mecanismos genéticos causantes de la aniridia en una cohorte española de pacientes con resultados no concluyentes durante el estudio del gen *PAX6*.

Los **objetivos concretos** desarrollados en este trabajo fueron:

1. Estudiar las **regiones no codificantes** de *PAX6* y otros **genes con fenotipo solapante** con el fin de identificar **nuevas variantes causantes de aniridia**.
2. **Caracterizar funcionalmente** las variantes de significado clínico incierto en *PAX6* con potenciales efectos sobre el *splicing* con motivo de evaluar su posible patogenicidad.
3. Estudiar la implicación del **mosaicismo** en aniridia para explicar casos de recurrencia y discordancia fenotípica intrafamiliar.
4. **Obtener el espectro mutacional** en la cohorte española de pacientes con aniridia.
5. Evaluar la **eficacia del algoritmo de diagnóstico genético** tras la introducción de nuevas técnicas de secuenciación masiva.

Materiales y Métodos

3.1. Pacientes y muestras biológicas

3.1.1. Descripción de la cohorte del estudio

Esta Tesis Doctoral se centra principalmente en el estudio de una cohorte de 98 familias, mayoritariamente de origen español, que fueron remitidas al Servicio de Genética del Hospital Universitario Fundación Jiménez Díaz (FJD, Madrid) entre 1994 y 2019 para la realización de estudios genéticos de aniridia congénita. Antes de la inclusión en este trabajo de tesis, se revisó la historia clínica del paciente, incluyendo anamnesis de síntomas oculares y sistémicos, además de los antecedentes familiares en al menos tres generaciones. Se consideraron casos familiares aquellos en los que se determinó un patrón de herencia autosómico dominante en al menos dos generaciones consecutivas de individuos afectados. Se consideraron casos esporádicos aquellos en los que no existían antecedentes familiares de la enfermedad.

El principal criterio clínico de inclusión en este estudio fue la presencia de hipoplasia completa o parcial del iris y/u otras anomalías iridianas. Los pacientes con **aniridia clásica** presentan una manifestación panocular caracterizada por hipoplasia del iris y de la fovea, acompañadas de otros hallazgos, incluyendo hipoplasia del nervio óptico, subluxación del cristalino, queratopatía por disfunción limbar, cataratas precoces y glaucoma (Hingorani, 2012; Casas-Llera, 2018; Voskresenskaya, 2017; Yahalom, 2018; Skeen, 2011). Los pacientes con **formas atípicas o `aniridia-like`** incluidos en este trabajo presentan formas leves de hipoplasia de iris y otras anomalías iridianas (coloboma, atrofia, corectopia, policoria, etc), acompañados o no de otras presentaciones, como anomalía de Peters, microftalmia, hipoplasia foveal, glaucoma, etc. Algunos pacientes de esta cohorte presentaban otros síntomas sugestivos del **síndrome de WAGR**, como tumor de Wilms, anomalías genito-urinarias y/o retraso intelectual.

Todos los pacientes incluidos en este trabajo tenían realizado un estudio genético previo siguiendo las recomendaciones y el algoritmo diagnóstico descrito en 2012 por Blanco-Kelly y cols. en la "Guía para el estudio genético de la Aniridia" en 2012 (Figura I13). Este consistió en el cribado de microdeleciones en la región 11p13 y/o mutaciones puntuales en los exones codificantes de *PAX6* en el caso índice de cada

familia utilizando las técnicas clásicas de MLPA (*Multiplex ligation-dependent probe amplification*) y secuenciación Sanger, respectivamente.

Algunos de los objetivos desarrollados en esta Tesis Doctoral fueron llevados a cabo dentro de trabajos colaborativos y por ello, se incluyeron familias procedentes de otras dos cohortes: i) una cohorte francesa de 27 familias con aniridia procedentes del Hospital Universitario Pûrpan de Toulouse (Toulouse, Francia) y ii) una cohorte de 36 familias españolas procedentes del Hospital Universitario Ramón y Cajal (Madrid, España).

Todos los individuos fueron incluidos en este estudio tras la obtención de un consentimiento informado firmado directamente por ellos o por sus tutores legales. El Comité de Ética de la Investigación de cada institución aprobó el estudio, en el que se siguieron los criterios éticos de los principios de la Declaración de Helsinki.

3.1.2. Muestras de pacientes y controles

Las muestras utilizadas en este estudio fueron específicamente obtenidas de forma prospectiva o proceden de la colección de muestras del Biobanco del Hospital Universitario FJD. Todas ellas fueron anonimizadas tras su recogida. Se utilizaron muestras de ADN obtenidas principalmente de sangre periférica, además de saliva, orina y/o semen en el análisis de mosaicismo.

En los estudios de expresión de *PAX6* se realizaron con **células linfoblastoides** obtenidas mediante centrifugación en gradiente de densidad utilizando *Lymphocyte Separation Medium* (Lonza, Bornem, Belgium) a partir de 2 mL de sangre periférica fresca (< 24 horas) recogida por venopunción en un tubo utilizando EDTA o heparinilitio como coagulantes. A partir de estas células, se establecieron cultivos primarios, o bien se generaron líneas celulares estables mediante un protocolo de inmortalización utilizando virus de Epstein-Barr (EBV), que fue realizado en colaboración con el Biobanco del CIBERER (CIBERER-Biobank, Valencia).

Los cultivos primarios y establecidos de linfocitos procedentes de pacientes y controles fueron cultivados en botellas de 25 ó 75 cm² a temperatura de 37°C en una atmósfera de humedad relativa del 95% y 5% de CO₂, en medio RPMI1640 con L-

glutamina (2mM) (Lonza) suplementado con suero fetal bovino (FBS) (Gibco, Thermo Fisher Scientific) al 10% (v/v), además de penicilina y estreptomina (Gibco) al 1% (v/v).

3.1.3. Líneas celulares establecidas

En este trabajo se utilizaron líneas celulares comerciales, incluyendo una línea humana de epitelio pigmentario adulto ARPE19 (*Adult Retinal Pigment Epithelial*) y una línea establecida de células humanas embrionarias renales HEK293T (*Human Embryonic Kidney*). Estas líneas fueron cultivadas en placas estandarizadas de 6 pocillos en condiciones similares a las descritas en el apartado anterior. Las células HEK293T fueron cultivadas con medio DMEM con 0,584 g/L de L-glutamina y 4,5 g/l de glucosa (Lonza). Las células ARPE19 fueron cultivadas en medio DMEM-F12 (Ham) (1:1) con 2,5 mM de L-glutamina y 15 mM de HEPES (Lonza). Estas células adherentes fueron recogidas con una solución de tripsina 0,25% (v/v) y EDTA 0,02% (v/v) antes de ser sedimentadas mediante centrifugación.

3.2. Aislamiento de ácidos nucleicos

En la obtención de muestras de **ADN genómico** (ADNg) a partir de sangre periférica y saliva se utilizaron los sistemas de extracción automatizados *BioRobot EZ1* (QIAGEN, Germany) y/o *MagNA Pure Compact System* (Roche Life Science, Germany). Para la extracción de ADNg procedente de orina se utilizó un método manual utilizando el kit comercial *NucleoSpin Tissue* (Macherey-Nagel). Este mismo kit fue utilizado en la extracción de ADNg a partir de muestra seminal, utilizando las siguientes modificaciones: i) congelación del semen a -80°C durante al menos 24 horas y ii) digestión con 600 µg proteinasa K, añadiendo 25 µL Tritón X100 (0,1%) y 0,1M DTT en el tampón de lisis.

El **ARN total** fue extraído de forma manual a partir las distintas líneas celulares anteriormente descritas usando el reactivo comercial *TRIzol reagent* (Ambion, Thermo Fisher) siguiendo las instrucciones del fabricante. La integridad de este ARN se evaluó mediante electroforesis en gel de agarosa (1% en TBE), seguido de una cuantificación mediante espectrometría utilizando el *NanoDrop 1000 Spectrophotometer* (Thermo

Fisher Scientific). A partir del ARN total, se llevó a cabo una reacción de transcripción reversa utilizando el kit *SuperScript IV Reverse Transcriptase* (Invitrogen, Thermo Fisher Scientific) siguiendo las instrucciones de la casa comercial.

3.3. Análisis de CNVs mediante arrays de CGH y MLPA

Los análisis de CNVs realizados en esta Tesis Doctoral fueron realizados mediante distintas aproximaciones, incluyendo *arrays* de hibridación genómica comparada (aCGH), MLPA y secuenciación Sanger.

3.3.1. aCGH

Esta técnica fue realizada en 33 casos índice sin variantes puntuales ni CNVs afectando a *PAX6* con el fin de realizar un cribado de microdeleciones en las regiones *cis*-reguladoras, así como en otros genes asociados a MOC. Para ello, se utilizaron dos diseños personalizados realizados con la tecnología *Agilent SurePrint* (Agilent Technologies, California, USA):

i) *WAGR-array*: desarrollado previamente para estudiar específicamente el *locus* asociado al síndrome de WAGR en la región 11p13-p15 (Blanco-Kelly, 2017). El *array* con formato 8x15K está formado por 15.744 sondas para estudiar 5 Mb de esta región (hg19, chr11:29.750.000-34.749.999).

ii) *MOC-array*: con un formato 4x180K, fue diseñado previamente para el estudio de 150 genes relacionados con el desarrollo ocular, incluyendo el *locus* WAGR (Ceroni, 2018). Este diseño contiene 130.000 sondas para cubrir las regiones génicas y la región 11p13, con un espaciado medio de 330 pb, además de 38.000 sondas *backbone*, distribuidas homogéneamente a lo largo de todos los cromosomas con un espaciado medio de 69 Kb.

Los *arrays* fueron realizados siguiendo el protocolo recomendado por la casa comercial. Se digirieron 500 ng de ADN tanto de los individuos en estudio como de las referencias del mismo sexo pareadas (Promega, Madison, USA), con las enzimas *RsaI* y *AluI* durante 2h a 37°C. Tras la inactivación de las enzimas a 65°C durante 20 min, las muestras de ADN de pacientes y de referencia se marcaron con los fluoroforos Cy5-dUTP y Cy3-dUTP, respectivamente utilizando el kit *Genomic DNA Enzymatic Labelling*

(Agilent). Tras una purificación utilizando las columnas *Microcon Ym-30* (Merck Millipore Corporation, Germany), las muestras marcadas fueron hibridadas en los cristales a 65°C durante 24h. Finalmente los cristales fueron lavados y escaneados en el escáner de fluorescencia *SureScan Microarray Scanner* de Agilent. Se utilizó el programa *Agilent CytoGenomics* para el análisis de las imágenes y el cálculo del *log2-ratio* de las intensidades de fluorescencia relativas para cada sonda en el paciente (Cy5 rojo) *versus* referencia (Cy3-verde). Una ratio de -1 indica una delección heterocigota (pérdida de 1 copia) de una sonda, 0 presenta un patrón normal (2 copias) y una ratio de +0.5 indica una amplificación heterocigota (3 copias). La presencia de una CNV vino determinada por el hallazgo de al menos 3 sondas consecutivas delecionadas o amplificadas.

3.3.2. MLPA

Se utilizó la técnica de MLPA para validar y posteriormente segregar en las familias las CNVs identificadas mediante *arrays* de CGH. Para ello, se utilizaron dos kits comerciales (MRC-Holland, Netherlands):

i) kit *SALSA MLPA probemix P219-B3*, con sondas principalmente para el gen *PAX6*, también para los genes *WT1*, *RCN1*, *BDNF*, *FSHB*, *DCDC1*, y *ELP4*, que flanquean a *PAX6*.

ii) kit *SALSA MLPA probemix P054-C2*, con sondas para los genes *FOXC1* y *PITX2* entre otros.

El procedimiento se realizó siguiendo el protocolo recomendado por el fabricante. Primero se llevó a cabo una amplificación por PCR y los fragmentos resultantes se separaron por electroforesis capilar utilizando el equipo *ABI 3130xl Genetic Analyzer* (Applied Biosystems, USA). El análisis se realizó con el programa *Coffalyser* (MRC-Holland) para calcular la ratio de la señal de cada sonda entre paciente *vs* muestra control. Una ratio de 1 indica la presencia de dos copias (individuo diploide normal), ratios inferiores a 0,6 o superiores a 1,4 indican delecciones (1 copia) y duplicaciones (3 copias) en heterocigosis de la secuencia diana, respectivamente.

3.4. Técnicas de secuenciación masiva

En este trabajo se realizó el cribado de 33 casos índices sin variantes patogénicas ni CNVs en las regiones codificantes de *PAX6* utilizándose distintas estrategias basadas en secuenciación masiva: i) panel personalizado para el estudio de *PAX6* y otros genes; ii) la secuenciación del exoma completo (WES).

3.4.1. Panel personalizado de genes asociados a MOC

Se utilizó una estrategia de captura basada en la tecnología Haloplex, diseñada de forma previa a esta Tesis Doctoral, que permite secuenciar:

i) el *locus* completo de *PAX6*, incluyendo UTRs, exones e intrones de todas las isoformas descritas en *RefSeq*.

ii) 20 regiones *cis*-reguladoras, descritas en el momento del diseño (Williams, 1998; Kammandel, 1999; Kleinjan, 2001; Griffin, 2002; Kleinjan, 2004; McBride, 2011; Bathia, 2013; Ravi, 2013; Bhatia, 2014) recogidas en el Anexo 1

iii) los exones codificantes y no codificantes, UTRs y 20 pb de las regiones intrónicas de *splicing* de todas las isoformas *Refseq* de otros 120 genes asociados a distintas formas de MOC.

Este diseño está formado por 19.705 amplicones específicamente diseñados para amplificar 1.168 regiones diana. El diseño final presenta un tamaño de 450 Kb y cubre teóricamente el 99% de las regiones diana. La preparación de librerías se realizó a partir de 450 ng de ADN utilizando el sistema de enriquecimiento *HaloPlex* (Agilent Technologies) siguiendo el protocolo de la casa comercial, tal como se describe en la Figura M1. Brevemente, esta consistió en una digestión enzimática de la muestra, seguida de una circularización de los fragmentos digeridos mediante sondas biotiniladas que específicamente se hibridan a ambos extremos de las secuencias diana. Las moléculas circulares obtenidas fueron purificadas con bolas magnéticas *Dynabeads* (Invitrogen, Thermo Fisher Scientific) unidas a estreptavidina (Invitrogen) y más tarde fueron amplificadas mediante PCR. Las librerías de ADN finalmente obtenidas contienen las secuencias diana con códigos de barra específicos de cada muestra y las secuencias adaptadoras para ser secuenciadas utilizando la tecnología

sequencing by synthesis (SBS) de *Illumina* (San Diego, USA). La secuenciación, llevada a cabo en equipos *MiSeq* o *NextSeq500*, se realizó de forma *paired-end* (2x150pb) a una profundidad mínima de 400X.

3.4.2. Secuenciación del exoma completo

Dos de las familias incluidas en este trabajo fueron previamente analizadas mediante WES, incluyendo un total de 4 individuos afectados y 3 sanos. Las librerías fueron preparadas usando el kit *Agilent Sure Select Human All Exon v5* (Agilent Technologies,) y se secuenciaron de forma *paired-end* (2x100pb) en un secuenciador *HiSeq2500* (Illumina) a profundidad de 60-100X, siguiendo los protocolos de las casas comerciales.

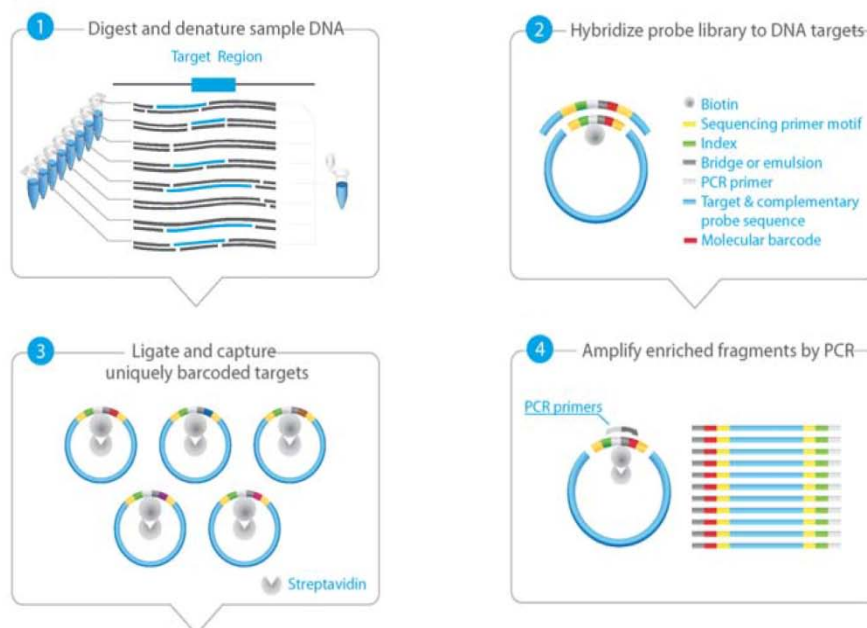


Figura M1. Preparación de librerías de ADN siguiendo el método de captura de HaloPlex. Imagen adaptada del protocolo de la guía de Agilent Technologies Figura 2, versión B0 June, 2015, pgn 27.

3.4.3. Análisis bioinformático

Los análisis bioinformáticos de los diferentes datos de NGS fueron realizados utilizando procedimientos estandarizados con una *pipeline* específicamente implementada en nuestro laboratorio (Ceroni, 2018) tal como se describe en la Figura

M2, que utiliza *BWA* (Li & Durbin, 2009) para el alineamiento de las secuencias al genoma de referencia hg19, *GATK* (McKenna, 2010) para el *calling* de las variantes tipo SNVs e indels y la herramienta de anotación *ANNOVAR* (Wang, 2010). Las variantes en *PAX6* fueron nombradas en base a la isoforma canónica *RefSeq PAX6 NM_000280.4*.

Se realizó un análisis de CNVs a partir de datos de NGS utilizando una estrategia de profundidad de lectura (*read depth*). Se utilizó la herramienta *BEDTools* (Quinlan and Hall, 2010) para obtener la profundidad de lectura media de cada región diana. Posteriormente, se utilizó el método de detección de CNVs de la herramienta *CoNVaDING* (Johansson, 2016) utilizando parámetros por defecto del programa. Primero, se calculó la ratio de profundidad de lectura de cada región capturada entre cada muestra con un control formado por todas las muestras procesadas en la misma carrera de secuenciación. Las regiones con ratios inferiores a 0,65 o superiores a 1,4 fueron consideradas candidatas a presentar deleciones o duplicaciones, respectivamente. Solo estas regiones candidatas fueron consideradas para ser validadas por mostrar diferencias significativas (Ceroni, 2018).

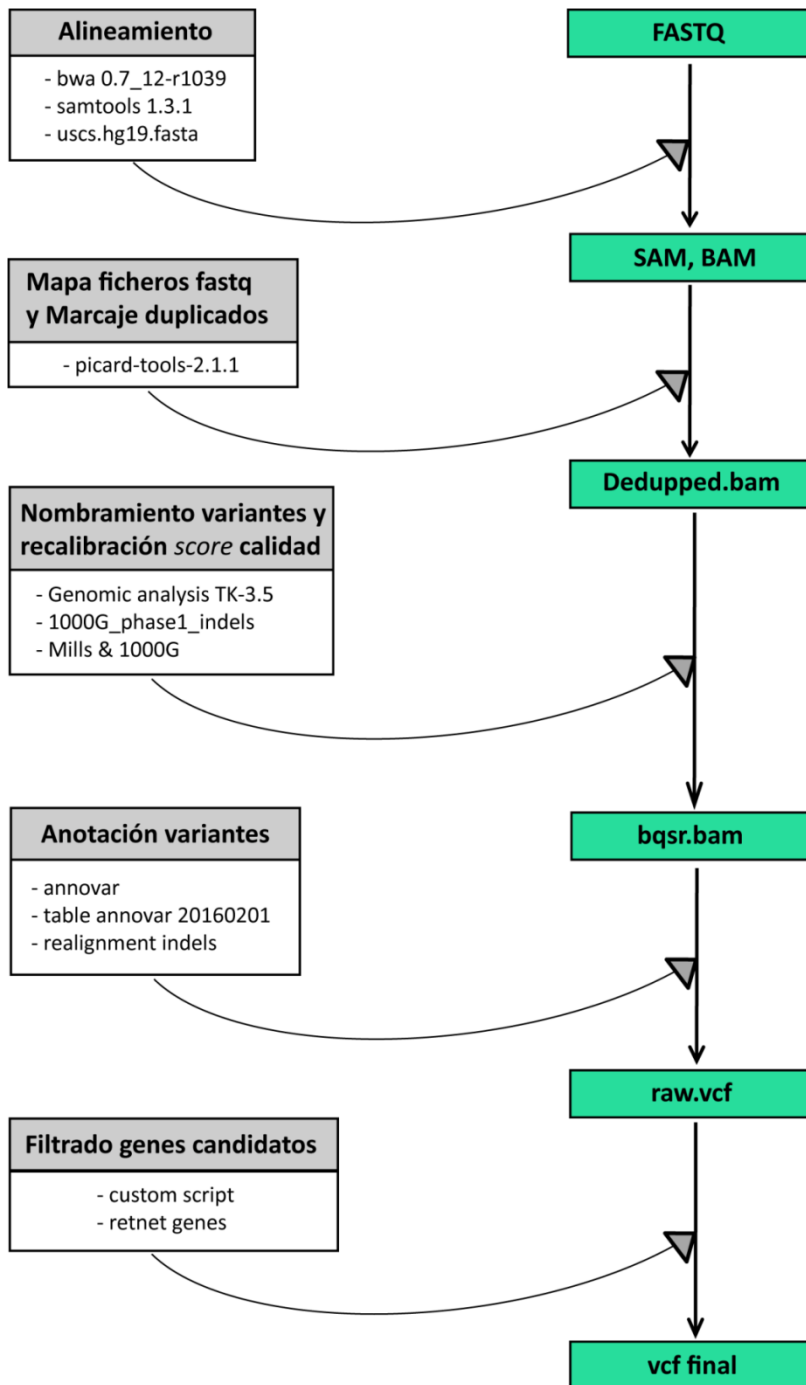


Figura M2. Representación de la pipeline bioinformatica implementada en nuestro laboratorio. Se describe el proceso de transformación de los datos crudos del secuenciador a un archivo vcf.

3.5. Priorización y clasificación de las variantes

Se priorizaron variantes raras en heterocigosis en el gen *PAX6* y, en aquellos casos en los que no se identificaron variantes candidatas, se consideraron posibles variantes asociadas con otros fenotipos, principalmente en genes de disgenesia del segmento anterior. La priorización de variantes potencialmente patogénicas fue realizada teniendo en cuenta los siguientes criterios:

i) variantes previamente descritas en distintas bases de datos de mutaciones (*HGMD*, *ClinVar* y *PAX6 LOVD Mutation Database*).

ii) variantes nuevas o raras en el *locus PAX6* y regiones reguladoras con una frecuencia alélica mínima (MAF) de <0,002% en las bases de datos poblacionales *gnomAD* y *CIBERER Spanish Variant Server*.

iii) en muestras sin variantes en *PAX6*, se priorizaron variantes nuevas o raras en exones y regiones canónicas de *splicing* con una frecuencia inferior al 0,01% o 1% en genes con herencia dominante o recesiva, respectivamente.

Se realizó una valoración de la patogenicidad de las variantes identificadas utilizando distintos predictores *in silico*, como CADD (Rentzsch, 2019), DANN (Quang, 2015) y otros específicos según las consecuencias esperadas de cada variante:

i. en la evaluación de los posibles efectos sobre sitios canónicos de *splicing*, se utilizaron *Human Splicing Finder (HSF)*, *Splice Site Finder-like (SSF)*, *MAxEntScan*, *NNSPLICE* y *GeneSplicer*, implementados en *Alamut*.

ii. en la evaluación de los posibles efectos sobre los elementos reguladores de *splicing* (SREs) se utilizaron *ESEfinder* y *Ex-skip*, implementados en *Alamut*.

iii. la conservación de las posiciones nucleotídicas y/o proteicas afectadas por las variantes fue evaluada mediante *GERP* y *PhyloP100way*, utilizando la herramienta web *Varsome* (Kopanos, 2018).

iv. para evaluar los efectos del cambio de aminoácidos en variantes tipo *missense*, se utilizaron *SIFT*, *Polyphen*, *Mutation Taster* y *Align GVGD*, implementados en *Alamut Visual* v.2.12 (Interactive Biosoftware, Sophia Genetics, Switzerland). Además, a las

variantes *missense* localizadas en el exón 6 se analizó *in silico* del efecto del cambio aminoacídico sobre la estructura 3D de la proteína cristalizada mediante el programa *PyMol* (DeLano, 2002) usando el modelo cristalográfico “PAX6 PBD ID 6PAX” previamente publicado por Xu y cols, en 1999.

Todas las variantes fueron clasificadas de acuerdo a las recomendaciones del *American College of Medical Genetics and Genomics (ACMG)* (Richards, 2015) como: patogénicas (clase 5), probablemente patogénicas (clase 4), variantes de significado incierto (clase 3), probablemente benignas (clase 2) o benignas (clase 1). Todas las variantes clasificadas como clase 3, 4 y 5 fueron confirmadas mediante secuenciación Sanger y segregadas en sus respectivas familias utilizando muestras disponibles de familiares afectados y sanos. Se priorizaron variantes *de novo* en casos esporádicos y variantes con herencia dominante co-segregando en todos los individuos afectados en casos familiares.

La **clasificación de la causalidad** de las variantes detectadas se estableció de la siguiente manera:

i) variantes probablemente no causales: variantes raras en las que la segregación en la familia las excluye como responsables de la patología.

ii) variantes de significado incierto (*vus*): se incluyen aquellas con efectos posiblemente patogénicos según predictores *in silico* cuya segregación en la familia no está disponible o no es informativa, o variantes con efectos no conocidos (variantes nuevas *in-frame*, UTRs, sinónimas, intrónicas y regiones reguladoras).

iii) variantes probablemente causales: variantes nuevas con predicción *in silico* y/o *in vitro* patogénica y que segregan con la patología en familias informativas.

iv) variantes causales: aquellas asociadas previamente a aniridia en la literatura y/o en bases de datos.

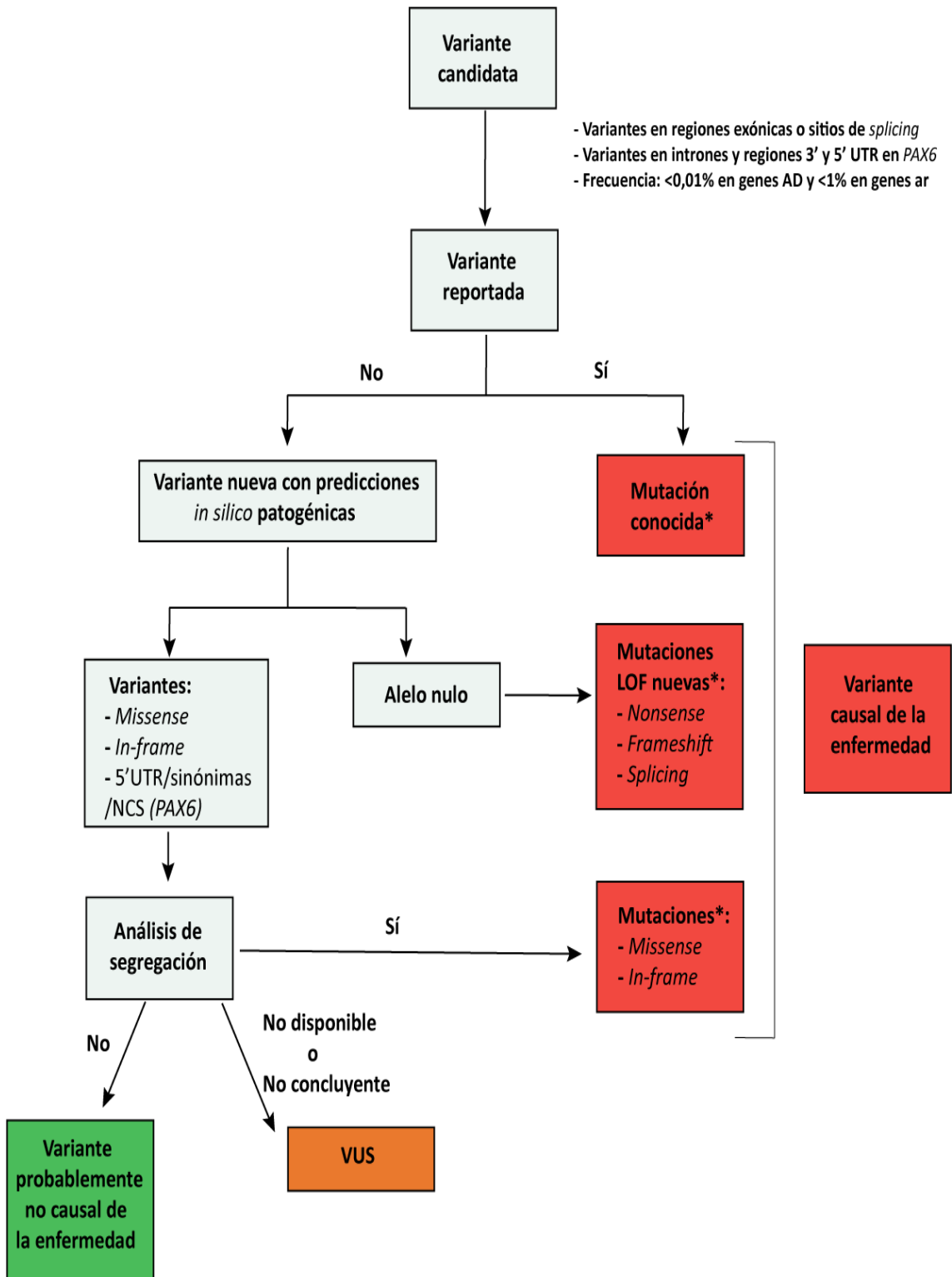


Figura M3. Algoritmo llevado a cabo para la clasificación de las variantes. * Análisis de segregación cuando fue posible. AD: autosómico dominante; ar: autosómico recesivo. Imagen adaptada de Fernandez-San Jose, 2015.

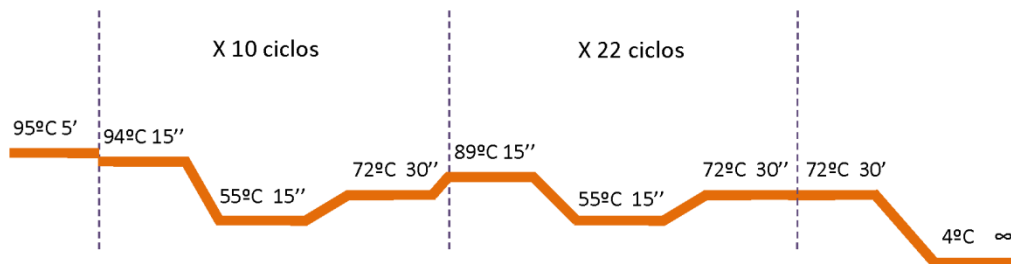
3.6. Secuenciación Sanger

El ADN_g fue amplificado mediante PCR utilizando el kit comercial *FastStart Taq DNA Polymerase* (Roche, Life Science) en termocicladores *GeneAmp PCR System 2700* (Applied Biosystems, CA, USA) usando cebadores específicos para amplificar las regiones en estudio (Anexos 2 y 3). Los productos obtenidos se visualizaron mediante electroforesis en gel de agarosa (2% en TBE 1X) y seguidamente fueron purificados enzimáticamente con *ExoSAP-IT* USB (Affymetrix, California, USA). Después, se realizó una reacción de secuenciación empleando el kit *Big Dye Terminator v3.1 Cycle Sequencing* (Applied Biosystems), que fue posteriormente purificada en placas de filtración utilizando el kit comercial *Montage SEQ96 Sequencing Reaction Cleanup* (Merck Millipore). La secuenciación fue realizada en un secuenciador automático *ABI 3130xl Genetic Analyzer* (Applied Biosystems) y las secuencias obtenidas fueron finalmente analizadas mediante el programa *Staden Package* (Staden, 2003).

3.7. Análisis de haplotipos

Se llevó a cabo un análisis de haplotipos para analizar la posible existencia de un efecto fundador en tres familias portadoras de la misma variante en *PAX6*. Para ello, se utilizaron 6 marcadores tipo microsatélite, STR (*short tandem repeat*), polimórficos y con elevada heterocigosidad, seleccionados en la base de datos *GeneLoc* (Anexo 4). Estos STRs flanquean a *PAX6* a una distancia máxima de hasta 34,5 Mb. Para el genotipado de los microsatélites se realizó mediante PCR multiplex donde los fragmentos fueron amplificados por separado utilizando parejas de cebadores con el *forward* marcado con fluorescencia. Seguidamente los productos de la PCR fueron separados mediante electroforesis capilar en el equipo *ABI 3130xl Genetic Analyzer* (Applied Biosystems) y posteriormente analizados mediante el programa *GeneMapper v3.5* (Life Technologies). En la construcción de los haplotipos se utilizó el programa *Cyrillic v2.1* (Cyrillic Software, Wallingford, UK).

H ₂ O	buffer	dNTPs	Cebador F	Cebador R	DNA	FastStart	Vf
8,1 µL	1,5 µL	3 µL	0,25 µL	0,25 µL	1,5 µL	0,4 µL	15 µL



3.8. Construcción de minigenes

Se utilizó una estrategia de *exon-trapping* para realizar un análisis *in vitro* de variantes que potencialmente alteran el *splicing* de *PAX6*. Para ello, se utilizó el vector pSPL3 (Figura M4), que contiene el promotor SV40, regiones de *splicing* donadora (SD6) y aceptora (SA2) derivadas de HIV/Tat humano flanqueantes a una región *polylinker*, en donde se incorpora mediante clonaje el exón o los exones de interés del gen a analizar. En la obtención de los minigenes, se siguió el protocolo de Desviat y cols, 2012 con pequeñas modificaciones.

Las secuencias genómicas a estudiar se obtuvieron mediante amplificación de PCR a partir de ADNg de un individuo control utilizando el kit *FastStart Taq polymerase* (Roche, Basel, Switzerland) y cebadores específicos (Anexos 2 y 3) para amplificar los exones en estudio y 200 nucleótidos intrónicos adyacentes. Tras su purificación con el kit *QIAquick Gel Extraction* (Qiagen, Hilden, Germany), los fragmentos fueron clonados en vectores TOPO-TA (Invitrogen), o de forma alternativa, directamente en el vector pSPL3 utilizando una estrategia de digestión/ligación con el kit *Rapid DNA Ligation* (Roche) siguiendo las recomendaciones de la casa comercial. Durante este trabajo se obtuvieron las siguientes construcciones, detalladas en la Figura M4:

i) **midigen para la región 5'UTR de PAX6** (*pSPL3-PAX6-5UTR-WT*), que incluye los exones 1 al 4 de *PAX6*. Debido al tamaño del intrón 2 de 3902 pb, esta construcción fue obtenida mediante el ensamblado de dos construcciones intermedias: un minigen para los exones 1 y 2 y un segundo minigen para los exones 3 y 4. Estas dos

construcciones fueron obtenidas mediante clonaje en un vector TOPO-TA a partir de los amplicones de interés en bacterias *E. coli* TOPO-10. El midigen 5'UTR se obtuvo posteriormente mediante subclonaje en pSPL3 de estos dos minigenes intermedios utilizando una estrategia de digestión / ligación con combinaciones de 2 enzimas de restricción: NotI - MreI y MreI - BamHI. El midigen finalmente obtenido no presenta la región intrónica completa del exón 2, tal como se representa en las Figuras M4 y M5.

ii) **minigen del exón 6 de PAX6** (*pSPL3-PAX6-ex6-WT*), que incluye además de este exón, ± 200 nucleótidos adyacentes a ambos lados, tal como se describe en las Figuras M4 y M6. Este minigen fue obtenido mediante clonaje directo del amplicón en el vector pSPL3 mediante digestión / ligación utilizando NotI - BamHI.

Las variantes c.-129+1G>A del exón 2, c.-128-2delA, c.-52+1G>A, c.-52+5delG y c.-52+3_-52+6delins del exón 3 y c.174C>T, c.178T>C, c.249T>G, c.326_331dup, c.333C>A, c.356G>T y c.357+6T>G del exón 6 fueron introducidas en las construcciones *wild-type* mediante mutagénesis dirigida utilizando el sistema *QuikChange Lighting Site-Directed Mutagenesis* (Agilent) con cebadores específicos (Anexos 2 y 3), siguiendo las recomendaciones del fabricante. Las posiciones de las variantes estudiadas están indicadas en las Figuras M5 y M6.

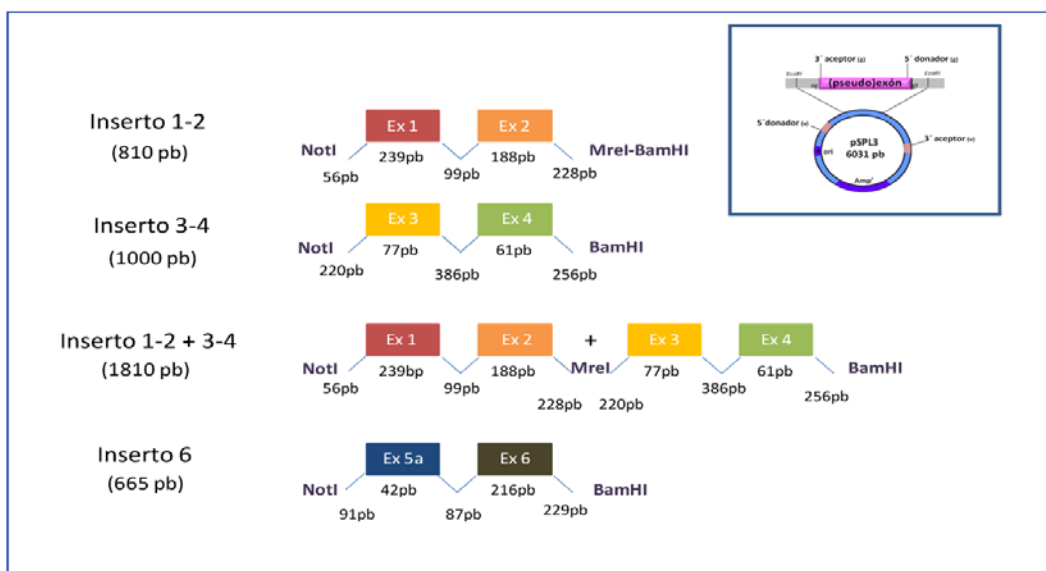


Figura M4. Insertos de PAX6 introducidos en el vector pSPL3 para la transfección de células HEK293T y ARPE19. Se indica el tamaño exacto de cada exón e intrón introducido y las enzimas de restricción utilizadas durante el proceso de clonaje.

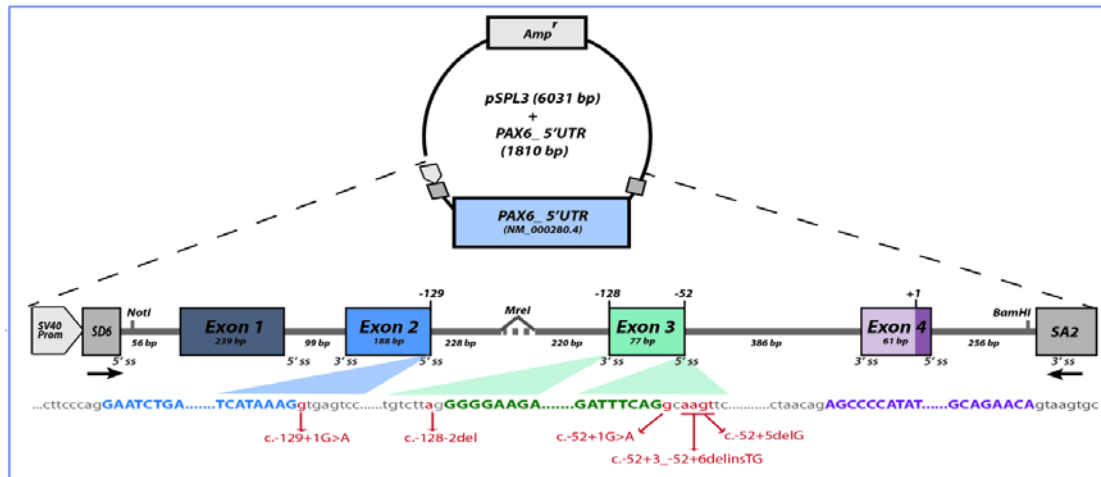


Figura M5. Estructura del constructo silvestre “pSPL3-PAX6-5UTR-WT” utilizado en ensayos *in vitro* de *splicing*. Se señalan las posiciones de las variantes introducidas mediante mutagénesis dirigida.

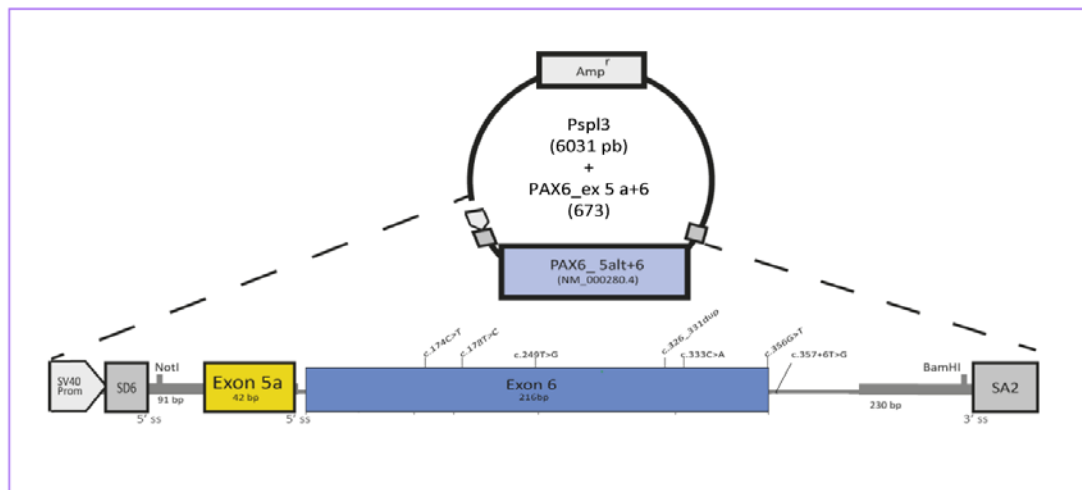


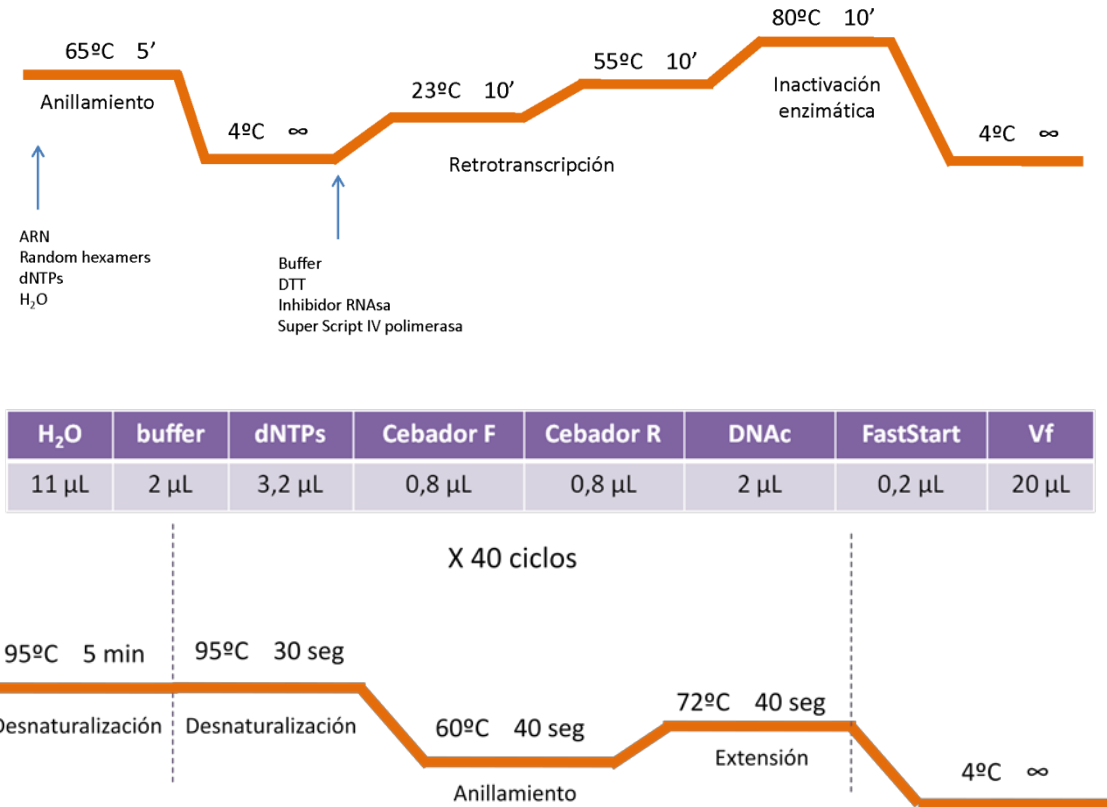
Figura M6. Estructura del constructo silvestre “pSPL3-PAX6-ex6-WT” utilizado en ensayos con *minigenes*. Se señalan las posiciones de las variantes introducidas mediante mutagénesis dirigida.

3.9. Análisis de expresión

3.9.1. Análisis de expresión de minigenes

Los ensayos de minigenes se llevaron a cabo por duplicado para cada constructo con 4×10^5 células HEK293T y ARPE19 en placas P6. Las células fueron transfectadas con construcciones silvestre, mutantes así como el vector pSLP3 vacío como control negativo. Las transfecciones se llevaron a cabo en medio completo utilizando *Jet-PEI in vitro DNA Transfection Reagent* (Polyplus Transfection, NY, USA) siguiendo el protocolo del fabricante, añadiendo 2 o 3 μg de la construcción plasmídica para las células HEK293T y ARPE19, respectivamente. Las células fueron recogidas a las 24 horas de la

transfección. El análisis de *splicing* se llevó a cabo mediante RT-PCR a partir de 2,5 µg de ARN total utilizando cebadores específicos del vector pSPL3 o del inserto (Anexos 2 y 3), utilizando las siguientes condiciones:



La expresión de *GAPDH* fue analizada como control interno. Los amplicones fueron separados por un gel de agarosa (2% en TAE), purificados con el kit *NucleoSpin Gel and PCR Clean-up* (Macherey-Nagel, Duren, Germany) y posteriormente secuenciados utilizando cebadores específicos del vector o del constructo (Anexos 2 y 3). En los casos con patrones complejos de *splicing*, los amplicones fueron clonados con el sistema TOPO-TA, de forma previa a su secuenciación.

Paralelamente, se realizó una semi-cuantificación de los transcritos obtenidos mediante RT-PCR utilizando un cebador SD6 (*forward*) marcado con 6-FAM, utilizando las mismas condiciones que las descritas. Los amplicones (1 µL) fueron analizados por triplicado mediante electroforesis capilar en el equipo *ABI 3130xl Genetic Analyzer* (Applied Biosystems). En función del tamaño de los fragmentos resultantes, previamente visualizados en la electroforesis, se utilizaron los marcadores de tamaño *ROX-1000* (Asuragen) o *LIZ-500* (Applied Biosystems) para los productos de los

constructos “*pSPL3-PAX6-5UTR-WT*” y “*pSPL3-PAX6-ex6-WT*”, respectivamente. Se utilizó el programa *GeneMapper v3.5* para el análisis del tamaño de los fragmentos y la cuantificación de la abundancia relativa de cada transcrito mediante el área de los picos.

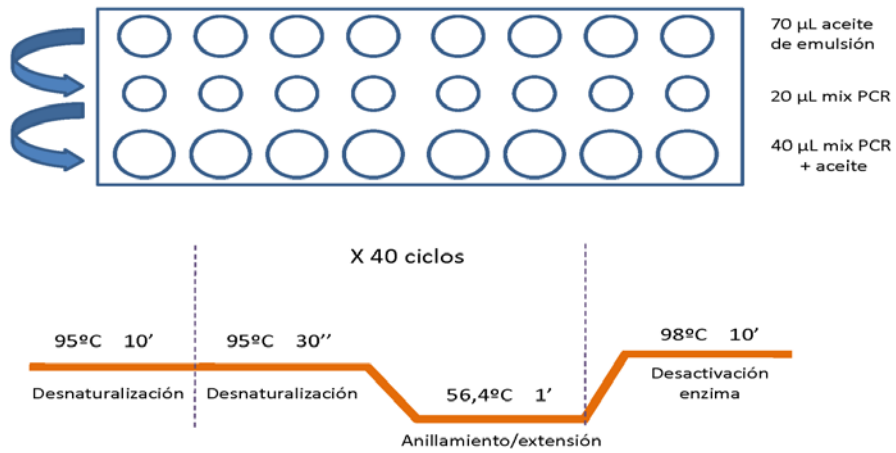
3.9.2. Análisis de *splicing* en muestras de pacientes

Se cultivaron $7,5 \times 10^5$ células linfoblastoides inmortalizadas por EBV en placas P6 tratándolas durante 6 horas con distintos inhibidores del sistema NMD, incluyendo: i) 0,5 y 0,75 mg/mL cicloheximida (Sigma-Aldrich, USA); iii) 100 y 300 µg/mL emetina (Sigma); v) 0,5 y 50 µM NMN Inhibitor NMDI14 (Calbiochem, Merk Millipore); vii) 5 y 10 µM wortmanina (Sigma). El análisis de expresión se realizó mediante RT-PCR a partir de entre 2,5 y 5 µg de ARN total utilizando cebadores diseñados para amplificar dos tipos de fregmento, exón 1 al 5 y exón 1 al 8 de *PAX6* (Anexo 2).

3.10. Análisis de mosaicismo

El análisis del mosaicismo se realizó utilizando ADNg procedente de muestras de sangre, saliva, orina y/o semen. Se utilizaron ensayos *TaqMan SNP* (Thermo Fischer Scientific) específicos para genotipar el alelo silvestre y el mutado. Los ensayos para las variantes c.120C>A, c.174C>T y c.178T>C en el gen *PAX6* fueron diseñadas específicamente con sondas para marcar con VIC el alelo silvestre y con FAM el alelo mutado. En el caso de la variante c.771G>A se utilizó una sonda comercial prediseñada (rs121907929, ID:C_152371166_10), donde el marcaje del alelo silvestre y WT está diseñado a la inversa de lo descrito para las anteriores variantes.

El análisis por PCR digital (*Droplet Digital PCR, ddPCR*) se realizó en el sistema *QX200* (Bio-Rad). Primero, se llevó a cabo la generación de 20.000 microgotas utilizando el aceite de emulsión *Oil for probes* (Bio-Rad) en el *QX200 Droplet Generator* (Bio-Rad). Después, se realizó la reacción de PCR en emulsión en un volumen de 20 µL conteniendo: 10 µL de *2x ddPCR Supermix* (Bio-Rad Laboratories, Hercules, USA), 900 nM de los cebadores, 250 nM de sonda marcada con VIC (para el alelo silvestre), 250 nM de la sonda marcada con FAM. La PCR se llevó a cabo en un termociclador *C1000 Touch Termal Cycler* (Bio-Rad) utilizando las siguientes condiciones:



Posteriormente, se procedió a la lectura de la fluorescencia en el lector *QX200 Droplet Reader* (Bio-Rad). Finalmente se analizó la fracción del alelo mutado con el programa *Quantasoft v1.7* (Bio-Rad), con el método de análisis de detección de eventos raros (*Rare Event Detection*). Todas las muestras fueron analizadas por cuadruplicado.

3.11. Bases de datos y herramientas bioinformáticas utilizadas

Bases de datos
ClinVar: https://www.ncbi.nlm.nih.gov/clinvar/
dbSNP: https://www.ncbi.nlm.nih.gov/projects/SNP/
Ensembl: http://www.ensembl.org
GeneLoc: https://genecards.weizmann.ac.il/geneloc/
gnomAD: http://gnomad.broadinstitute.org/
HGMD: http://www.hgmd.cf.ac.uk
RetNet: https://sph.uth.edu/retnet
UCSC: https://genome.ucsc.edu/
Predictores <i>in silico</i>
Variantes de splicing
ESEFinder: https://rulai.cshl.edu/
GeneSplicer: https://www.cbc.umd.edu/software/GeneSplicer/gene_spl.shtml
Human Splicing Finder: https://www.umd.be/HSF3/
MaxEntScan: https://genes.mit.edu/burgelab/maxent/Xmaxentscan_scoreseq.html
NNSPLICE: https://omictools.com/nnsplICE-tool
Variantes missense
Align GVGD: https://www.mutationtaster.org/
CADD: https://cadd.gs.washington.edu/
MutationTaster: https://www.mutationtaster.org/
M-CAP: https://bejerano.stanford.edu/mcap/
Polyphen: https://genetics.bwh.harvard.edu/pph2/
PyMOL: https://pymol.org
SIFT: https://sift.jcvi.org/
Herramientas para el análisis bioinformático de NGS
ANNOVAR: https://annovar.openbioinformatics.org/
BWA (<i>Burrows-Wheeler Aligner</i>)-MEM: https://bio-bwa.sourceforge.net/
CoNVaDING: https://github.com/molgenis/CoNVaDING
DNANexus: https://www.dnanexus.com/
ENSEMBL Variant Effect Predictor: https://www.ensembl.org/vep
GATK (<i>Genome Analysis Tool Kit</i>): https://software.broadinstitute.org/gatk/

Tabla M1. Bases de datos y herramientas bioinformáticas utilizadas en este trabajo.

Resultados

En esta Tesis Doctoral, se ha estudiado una cohorte de 98 familias con sospecha de aniridia aislada o sindrómica, todas ellas remitidas a lo largo de los últimos 25 años a la consulta de genética del Hospital Universitario Fundación Jiménez Díaz para estudio molecular de *PAX6* utilizando el algoritmo de diagnóstico descrito previamente en el apartado de Introducción (Figura I13).

Con este abordaje en el que sólo se estudiaban las regiones codificantes de *PAX6*, además de las CNVs en la región 11p13, se detectaron variantes patogénicas en el 63% de las familias (Figura R1). La mayoría son variantes de pérdida de función en la región codificante de *PAX6* (49%). El 14,3% presentan microdeleciones que afectan únicamente a dicho gen o incluyen además a genes contiguos asociados al síndrome de WAGR, principalmente al gen *WT1*. En un 10% de las familias se identificaron variantes de significado incierto en *PAX6*. Por lo tanto, en 1 de cada 4 familias no se identificaron potenciales mutaciones en *PAX6*.

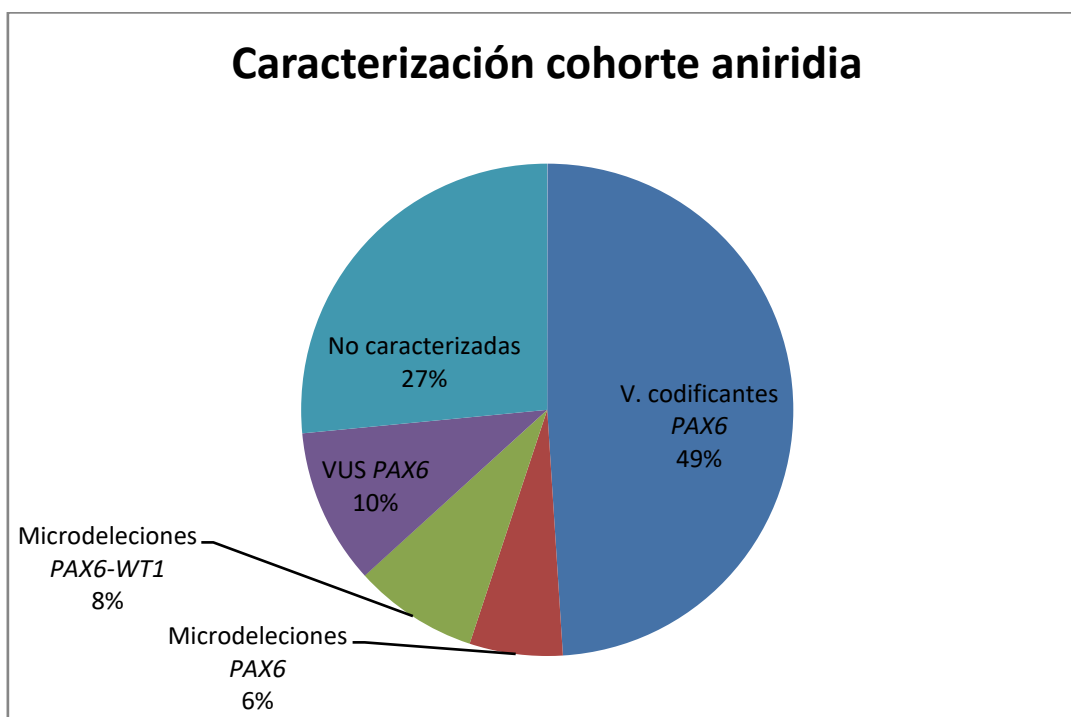


Figura R1. Representación del espectro mutacional en la cohorte de familias españolas con sospecha de aniridia aplicando el algoritmo clásico de diagnóstico. Los datos, representados en forma porcentual, fueron obtenidos mediante el estudio del diagnóstico genético realizado de forma previa a este trabajo de tesis doctoral. Los tipos de variantes identificadas son: variantes codificantes en *PAX6*, microdeleciones totales o parciales de *PAX6*, microdeleciones de genes contiguos en la región 11p13 (incluyendo entre otros a *PAX6* y *WT1*), variantes de significado clínico incierto y familias no caracterizadas, en las que no se detectaron variantes en los exones codificantes de *PAX6*, siendo consideradas "PAX6 negativas".

En este trabajo de Tesis Doctoral, se ha procedido al estudio de 33 familias con aniridia en las que no se había obtenido un resultado concluyente, bien por la ausencia de mutaciones en los exones codificantes de *PAX6* o por presentar VUS en los mismos. Para ello, hemos realizado estudios funcionales para caracterizar las variantes tipo VUS identificadas en *PAX6* con el fin de mejorar su clasificación patogénica. Además, se ha llevado a cabo un cribado de las regiones no codificantes y reguladoras de *PAX6*, así como de otros genes asociados a patologías relacionadas, para identificar nuevas variantes potencialmente patogénicas.

Este planteamiento se ha realizado gracias al diseño de nuevas herramientas de diagnóstico personalizadas para el estudio de la aniridia basadas en técnicas genómicas, como *arrays de CGH* y paneles de secuenciación masiva para el análisis del *locus* de *WAGR* y hasta otros 120 genes asociados a MOC.

Por otro lado, hemos querido profundizar en la implicación del mosaicismo en patologías asociadas a *PAX6* para explicar casos de recurrencia en hermanos afectos procedentes de padres asintomáticos, así como en casos de variabilidad fenotípica intrafamiliar. Para ello, se ha empleado una nueva técnica basada en PCR digital que realiza una cuantificación más precisa de alelos de baja frecuencia en casos con sospecha de mosaicismo en *PAX6*.

Aparte de estudiar nuestra cohorte de familias, en este trabajo hemos podido ampliar el estudio a otros pacientes de dos cohortes procedentes del Hôpital Purpan de Toulouse y del Hospital Universitario Ramón y Cajal de Madrid.

Durante este trabajo, se han llevado a cabo las siguientes actividades:

- i) Revisión de la patogenicidad de las variantes en *PAX6* identificadas en nuestra cohorte.
- ii) Estudios de casos sin mutaciones en los exones codificantes de *PAX6* mediante el cribado de las regiones no codificantes de *PAX6* y de otros genes asociados a MOC.
- iii) Caracterización funcional de VUS localizadas en el exón 6 de *PAX6*.

iv) Identificación y validación de posibles casos de mosaicismo en familias con variantes en *PAX6*.

v) Integración de los resultados para obtener una visión global de la enfermedad.

En base a los datos obtenidos, el apartado de Resultados de esta Tesis Doctoral ha sido dividido en cuatro apartados independientes:

4.1. Implicación de variantes en regiones no codificantes en *PAX6* y/o variantes en otros genes en los casos de aniridia “*PAX6* negativos”.

4.2. Caracterización funcional de VUS que afectan al *splicing* del exón 6 de *PAX6*.

4.3. Estudio de la implicación del mosaicismo parental en familias con variantes en *PAX6*.

4.4. Actualización del espectro mutacional en nuestra cohorte de aniridia

4.1. Implicación de variantes en regiones no codificantes en *PAX6* y/o variantes en otros genes en los casos de aniridia “*PAX6* negativos”

En el primer apartado de Resultados de la presente Tesis Doctoral, nos hemos centrado en intentar caracterizar 33 familias sin variantes patogénicas en *PAX6*. Todas las familias estudiadas en este apartado tienen aniridia clásica o un fenotipo de sospecha de aniridia-like, y resultaron negativas o con resultado no concluyente tras un cribado de *PAX6* siguiendo el algoritmo clásico de diagnóstico.

Para ello, se han realizado distintos abordajes para estudiar otros genes relacionados con estos fenotipos así como las regiones no codificantes y reguladoras de *PAX6*. Inicialmente se analizaron mediante secuenciación Sanger los 5 exones codificantes de *TRIM44*, nuevo gen que fue identificado en 2015 como causante de la aniridia en una familia de origen chino (Zhang, 2015), sin embargo, no se identificaron mutaciones en nuestra cohorte.

Además, esta cohorte de casos negativos en *PAX6* fue analizada utilizando paneles de *arrays* de CGH y/o NGS, específicamente diseñados para analizar el *locus* completo de *PAX6*, incluyendo UTRs, regiones *deep intronic* y elementos *cis*-reguladores, así como otros 120 genes asociados a patologías del desarrollo ocular.

Durante este trabajo, se identificaron tanto deleciones afectando a *enhancers* localizados en 3' del gen *PAX6*, así como variantes no codificantes en *PAX6*, principalmente localizadas en los exones que forman el 5'UTR. En estas últimas, quisimos profundizar en los mecanismos patogénicos asociados utilizando estudios de expresión *ex-vivo* o *in vitro* mediante minigenes. En este primer apartado, también se incluyeron las variantes identificadas en un cribado en 27 familias francesas que fue realizado en el laboratorio del Dr. Calvas y Chassaing en el Hôpital Purpan de Toulouse. Todas ellas con un fenotipo de aniridia aislada y sin variantes en la región codificante de *PAX6*.

Finalmente, en nuestra cohorte se han identificado variantes patogénicas en otros genes asociados a disgenesias del segmento anterior.

4.1.1. Identificación de microdeleciones afectando elementos *cis*-reguladores

En este trabajo, gracias al análisis de CNVs realizado mediante paneles personalizados de aCGH y/o NGS, se identificaron microdeleciones que solo afectan a las regiones reguladoras de *PAX6* en 5 casos índice con aniridia aislada en nuestra cohorte (Tabla R1). Estas CNVs fueron posteriormente validadas y segregadas en las familias utilizando una segunda técnica, MLPA y/o aCGH, en función de la aproximación utilizada para su identificación. Estas deleciones pueden aparecer como eventos *de novo*, tal como se observa en 2 casos, o pueden haber sido heredadas de un progenitor. En todos los casos descritos en este trabajo se dispone de la acotación mínima delecionada mediante aCGH.

La Figura R2 describe la identificación y el refinamiento de los puntos de corte de estas 5 microdeleciones identificadas en nuestra cohorte. Además, en 1 de ellas, los puntos exactos de corte se secuenciaron utilizando cebadores diseñados específicamente en 5' y 3' para cada CNV, tal como se refleja en esta figura. Las características de estas microdeleciones, su localización y sus dimensiones concretas se especifican en la Tabla R1. En general, se han identificado microdeleciones de entre 502 y 680 Kb situadas a una distancia entre 90 y 120 Kb respecto al último exón de *PAX6*, por tanto, son teloméricas al gen, tal como se representa en el esquema de la Figura R3. Todas estas deleciones afectan a elementos *cis*-reguladores situados en las regiones intrónicas de *ELP4*, no estando implicados ni *PAX6* ni *WT1*, tal como se observa en la Figura R3. En esta misma figura, se representan no solo las 5 microdeleciones identificadas en nuestra cohorte, sino también otras 5 identificadas en la cohorte francesa de nuestros colaboradores. Además, en la parte inferior, se describe la localización de otras 16 microdeleciones de las regiones 3' de *PAX6* identificadas y acotadas por aCGH publicadas en la literatura (Figura R3). La comparación en estas 26 CNVs de los puntos de corte definidos por el tamaño mínimo de la deleción obtenido por aCGH, nos permitió redefinir la anteriormente descrita región crítica para la regulación transcripcional de *PAX6* de 244 Kb (Ansari, 2016). Esta nueva región crítica de tan solo 18 Kb (chr11:31.648.248-31.666.340) únicamente contiene el elemento *cis*-regulador E180, lo que pone en evidencia su posible implicación en la patogénesis de la aniridia.

Paciente	Sexo	Edad (años)	Posición genómica deleción (hg19)	Tamaño (Kb)	Distancia a PAX6 (Kb)	Segregación
ANI-067	F	10	11:31.083.877-31.704.548	621	±102	AD
ANI-008	F	61	11:31.147.306-31.714.853	568	±91	AD
ANI-021	F	15	11:31.186.493-31.698.208	512	±108	<i>De novo</i>
ANI-105	F	10	11:31.184.352-31.686.948*	502	±119	<i>De novo</i>
ANI-118	F	52	11:31.000.501-31.680.548	680	±125	N.D.

Tabla R1. Resumen de las microdeleciones en la región cromosómica 11p13 que afectan a elementos 3' reguladores de PAX6 identificadas en nuestra cohorte. Se detalla la distancia de la deleción al último exón de PAX6. *: posición exacta de puntos de ruptura acotados mediante LR-PCR y posterior secuenciación de los puntos de corte de las deleciones. N.D.: No hay datos.

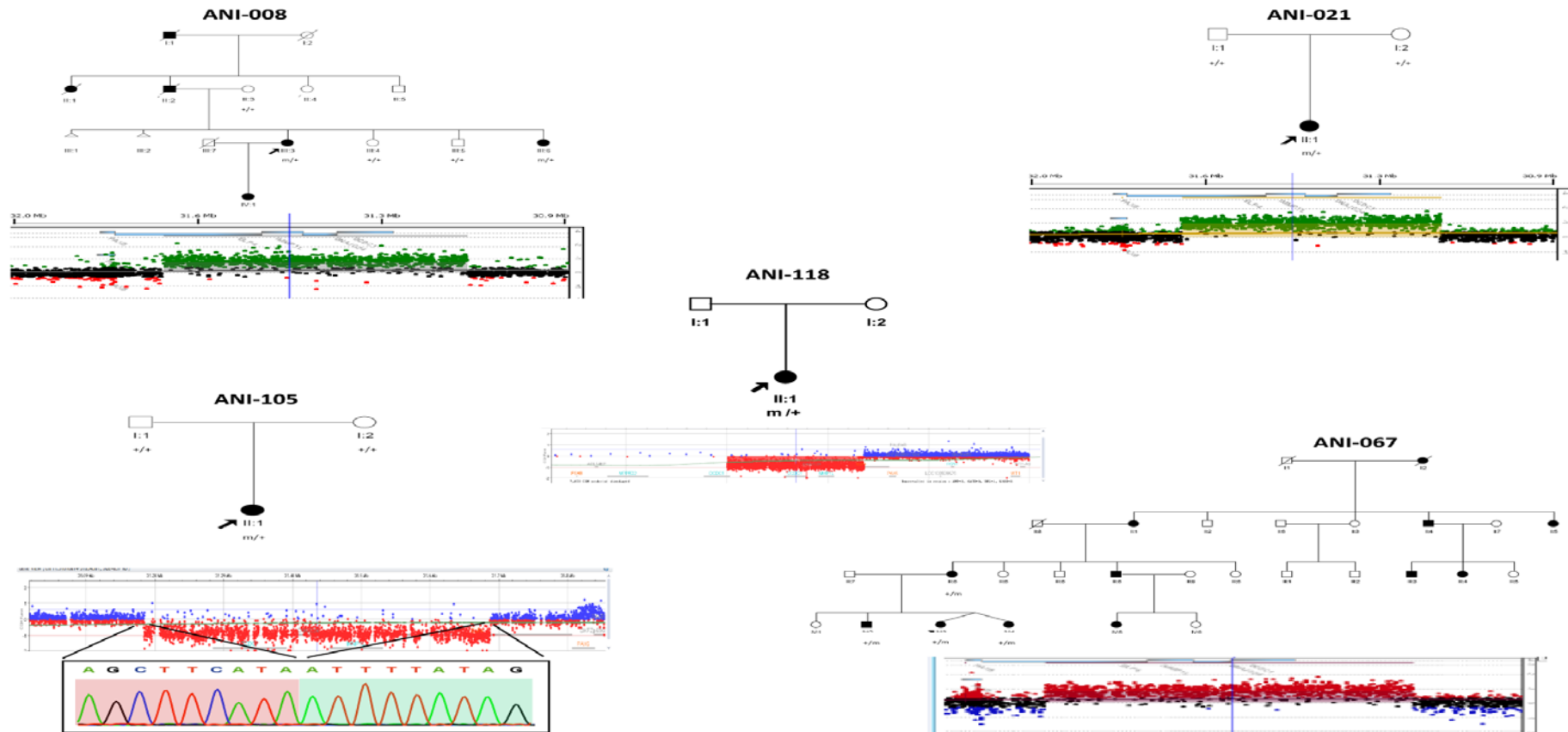


Figura R2. Segregación de las familias con deleciones 3' a PAX6. Se muestran los resultados de los aCGH y en la familia ANI-105 también la acotación de los puntos de ruptura exactos mediante PCR.

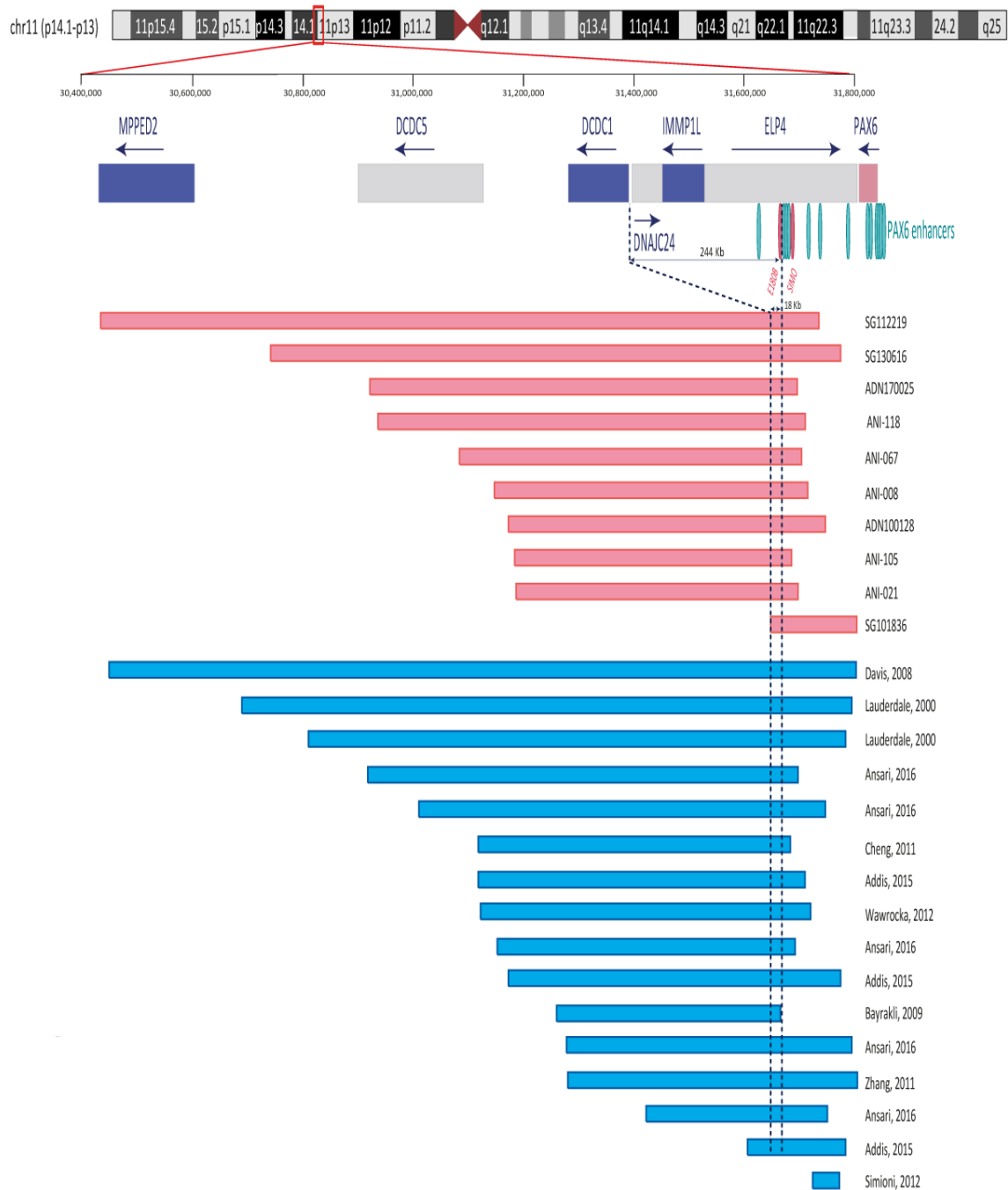


Figura R3. Identificación de deleciones en regiones reguladoras 3' a PAX6 en pacientes con aniridia aislada. Localización de PAX6 y sus enhancers descritos en el cromosoma 11. Las barras rojas representan la posición genómica de las deleciones en nuestras familias. Las barras azules muestran las coordenadas genómicas de 16 familias adicionales recogidas de la literatura (Lauderdale, 2000; Davis, 2008; Bayaki, 2009; Chen, 2011; Zhang, 2011; Simioni, 2012; Wawrocka, 2012; Ansari, 2016). Las barras verdes indican grupos de regiones cis-reguladoras localizadas en posiciones intrónicas de ELP4 y PAX6. Los elementos SIMO y E180 se representan en rojo. Las líneas verticales representan la región mínima de 18 Kb (chr11:31648248-31666340).

4.1.2. Cribado de variantes puntuales en las regiones no codificantes de *PAX6*

En los 28 pacientes restantes, el análisis de las regiones UTR, intrones y *enhancers* del gen *PAX6* mediante un panel de NGS, permitió identificar 3 variantes potencialmente patogénicas en el 5'UTR en 5 casos índice (Tabla R2), lo que explicaría el 18% (5/28) de los pacientes de este estudio.

En este cribado estaban incluidos 10 pacientes con estudio diagnóstico no concluyente debido a la identificación de una variante de significado clínico incierto en *PAX6*. En ninguno de estos pacientes se identificó ninguna otra variante en el *locus* de *PAX6* de las previamente identificadas mediante cribado clásico.

En otros 13 pacientes, este estudio no identificó ninguna variante en *PAX6*, incluyendo intrones, UTRs y regiones reguladoras, que pudiera ser potencialmente candidata a explicar el fenotipo de aniridia. Por lo tanto, el 39% (13/33) de los pacientes de este trabajo, son realmente *PAX6* negativos.

En los siguientes apartados, profundizaremos en las características genéticas y clínicas asociadas a las variantes identificadas en el 5'UTR de *PAX6* y además, describiremos los resultados de su caracterización funcional.

Familia	Sexo	Edad (años)	Fenotipo ocular	Herencia	SNV	Localización	Segregación	Referencias
ANI-006	M	33	Aniridia	Familiar	c.-129+1G>A	Intrón 2	En madre afecta	Nueva
ANI-069	M	34	Aniridia	Familiar	c.-129+1G>A	Intrón 2	En madre afecta	Nueva
ANI-088	M	6	Aniridia	Familiar	c.-129+1G>A	Intrón 2	En padre afecto	Nueva
ANI-057	M	51	Aniridia incompleta con anomalía de Peters	Esporádico	c.-128-2delA	Intrón 2	<i>De novo</i>	Axton, 1997 Robinson, 2008 Vasilyeva, 2017
ANI-012	M	46	Aniridia	Familiar	c.-52+5delG	Intrón 3	En padre afecto	Nueva

Tabla R2. Resumen de variantes no codificantes encontradas en nuestra cohorte de aniridia y en fenotipos relacionados. No existen datos de la frecuencia alélica de estas variantes en la base de datos gnomAD.

4.1.3. Identificación de variantes probablemente patológicas en el 5'UTR de *PAX6*.

Durante el cribado de las regiones no codificantes de *PAX6*, se identificaron 3 variantes distintas en la región 5'UTR, localizadas en 2 de los 3 exones no codificantes del gen (Tabla R2 y Figura R4). Todas eran variantes nuevas con la excepción de una deleción en la región aceptora del exón 3, c.-128-2delA, descrita previamente en varios trabajos anteriores (Axton, 1997; Gronskov, 1999; Redeker, 2008; Robinson, 2008).

Estas variantes fueron detectadas en 5 familias segregando de forma autosómica dominante en varios afectados con aniridia clásica (Figura R5), excepto en un caso esporádico (ANI-57) portador de una variante *de novo* que presentaba un diagnóstico de aniridia-like con anomalía de Peters (Tabla R2). A la vista de las segregaciones y dada su ausencia en bases de datos poblacionales, fueron clasificadas como probablemente patológicas siguiendo los criterios de patogenicidad del ACMG.

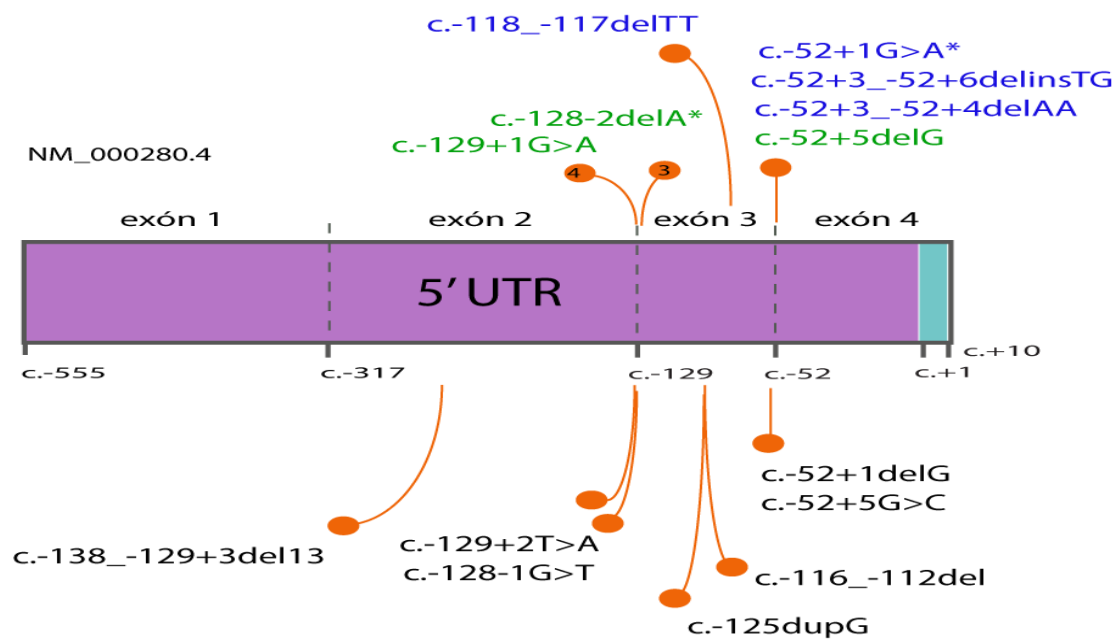


Figura R4. Representación esquemática de variantes en 5'UTR en el gen *PAX6*. En la sección superior, se muestra la posición de las variantes detectadas en este estudio. *: variantes previamente descritas en la literatura. En verde las variantes descritas en nuestra cohorte de la FJD; en azul las variantes descritas en la cohorte francesa; la variante c.-52+5delG fue descrita en ambas cohortes. En la sección inferior, se representan las variantes previamente descritas en la literatura y/o en la base Human *PAX6* mutation database.

Destacar que una de las variantes aparece de forma recurrente en nuestra cohorte de pacientes. La variante c.-129+1G>A se identificó en 3 familias no emparentadas en las que se descartó un posible efecto fundador mediante un estudio de haplotipos. El estudio de 6 marcadores STR situados en el cromosoma 11 cercanos al *locus* de *PAX6*. Además, el primer individuo afecto con aniridia en todas las familias aparece como descendiente de progenitores sanos, lo que parece apoyar que estas variantes aparecieran como eventos *de novo* independientes (Figura R5).

Resultados similares fueron obtenidos por el grupo del Dr. Calvas en una cohorte de 27 familias sin mutaciones codificantes en *PAX6*. Utilizando una estrategia de NGS basada en la amplificación de este gen completo y regiones reguladoras (Plaisancié & Tarilonte, 2018), se identificaron 6 variantes diferentes en el 25% de estos pacientes. Estas incluyen las variantes recurrentes c.-129+1G>A y la c.-128-2delA, también identificadas en nuestra cohorte. En la Figura R4 se representan las 7 variantes distintas identificadas en estas dos cohortes de pacientes.

La mayoría de las variantes identificadas en la región 5'UTR de *PAX6*, no solo en este trabajo sino también las descritas previamente en la literatura, se localizan en los sitios canónicos de *splicing* de los exones no codificantes 2 y 3 (Figura R4). Por esta razón, evaluamos *in silico* el posible efecto a nivel del *splicing* de las 7 variantes identificadas en ambas cohortes de pacientes. Excepto la delección c.-118_-117delTT identificada en una familia francesa, el resto de las variantes parecen susceptibles de provocar cambios en el procesamiento del pre-ARNm mediante la alteración directa de los sitios canónicos de estos 2 exones, tal como indican distintos predictores de *splicing* (Tabla R3).

Con objeto de validar las predicciones obtenidas *in silico* para estas variantes, durante este trabajo se llevaron a cabo estudios de expresión mediante análisis de RT-PCR. Para ello, se utilizó ARN procedente de linfocitos inmortalizados de pacientes portadores y/o de dos tipos de líneas celulares, HEK293T y ARPE19, tras su transfección con midigenes para la región 5'UTR de *PAX6* de tipo silvestre y para cada una de las variantes en estudio.

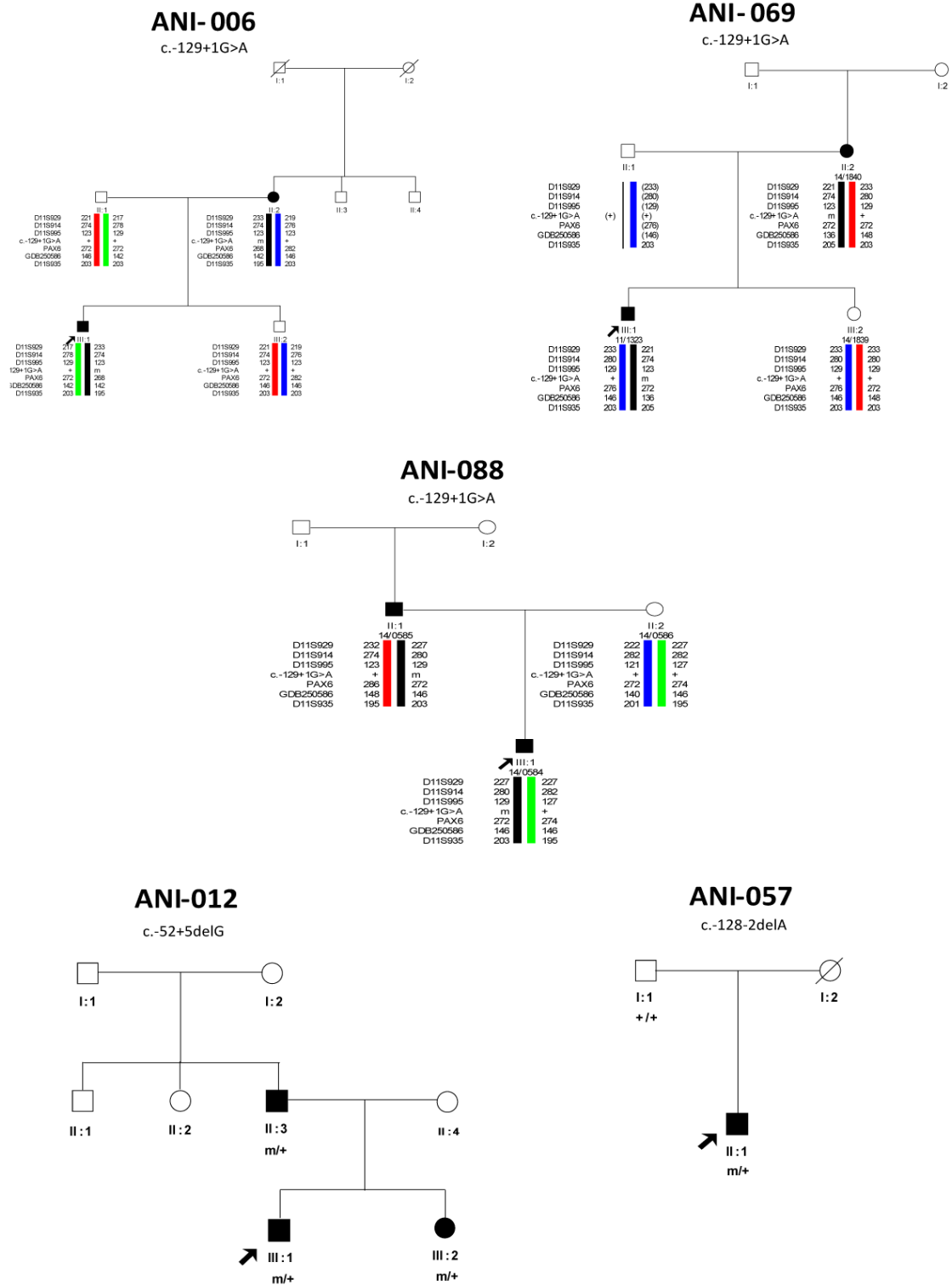


Figura R5. Segregaciones de las tres variantes identificadas en el 5'UTR de PAX6 en cinco familias españolas con aniridia. En las tres familias portadoras de la variante c.-129+1G>A, un análisis de haplotipos realizado con 6 marcadores STRs descartó un posible efecto fundador.

Variante	Nº pacientes / Nº familias	Localización	HSF	MaxEntScan	SSF	NNSPLICE	Resumen predicción <i>in silico</i>	Resultados minigenes isoformas (%)*
c.-129+1G>A	7/4	5'SS canónico exón 2	97,9 ⇒ — (-100%)	10,5 ⇒ — (-100%)	95,3 ⇒ — (-100%)	1 ⇒ — (-100%)	Disrupción del 5'SS y un sitio ESE	ΔE2 (9%), ΔE2-E3 (12%), ΔE2-E4 (32%), (Ins_I2)_254 (19%), (Ins_I2)_406 (6%) y CT (4%)
c.-128-2delA	9/6	3'SS canónico exón 3		10,7 ⇒ — (-100%)	86,7 ⇒ — (-100%)	1 ⇒ — (-100%)	Disrupción del 3'SS y dos ESEs	ΔE3 (96%) y CT (4%)
c.-118_-117delTT	2/1	Exón 3	---	---	---	---	Sin alteración	N.D.
c.-52+1G>A	6/4	5'SS canónico exón 3	---	---	100 ⇒ — (-100%)	---	Disrupción del 5'SS y creación de un nuevo sitio ESE	ΔE3 (100%)
c.-52+3_-52+4delAA	2/1	5'SS canónico exón 3	---	---	100 ⇒ — (-100%)	---	Disrupción del 5'SS y creación de un nuevo sitio ESE	N.D.
c.-52+3_-52+6delinsTG	2/1	5'SS canónico exón 3	---	---	100 ⇒ — (-100%)	---	Disrupción del 5'SS y creación de un nuevo sitio ESE	ΔE3 (100%)
c.-52+5delG	2/1	5'SS canónico exón 3	---	---	100 ⇒ 87,2 (-12,8%)	---	Alteración del 5'SS y disrupción del sitio ESE	ΔE3 (100%)

Tabla R3. Predicciones de splicing y resultados *in vitro* de variantes identificadas en el 5'UTR de PAX6. Se indican los scores de splicing predichos con 4 programa diferentes, incluyendo Human Splicing Finder (HSF); MaxEntScan; Splice Site Finder (SSF) y NNSPLICE, así como el valor. 3'SS = aceptor (3' acceptor splice site); 5'SS= donador (5' donor splice site); ESE= Exonic Splicing Enhancer; ESS= Exonic Splicing Silencer. *: indica porcentaje de cada isoforma detectada en un ensayo de PCR semi-cuantitativa. ΔE2, ΔE3, ΔE2-4: delección de los exones 2, 3 y del 2 al 4, respectivamente. CT: transcrito canónico. (Ins_I2)_254 y (Ins_I2)_406: inserción del intrón 2, 254 pb y 406 pb, respectivamente. N.D.: ensayo de minigen no disponible.

Tal como se describe en la Figura M5 de Material y Métodos, esta construcción de tipo midigen fue realizada con el vector de tipo *exon-trapping* pSPL3, en donde se incluyeron los tres primeros exones no codificantes de *PAX6* y el exón 4, primer exón codificante del gen. Por lo tanto, estos midigenes incluyen no solo los exones cuyo *splicing* era susceptible de verse modificado por la acción de las variantes en estudio, sino que contiene además las regiones intrónicas y exónicas contiguas. Sin embargo, el intrón 2 no pudo ser incluido entero en el vector, debido a su gran tamaño de unas 4 Kb, solo están incluidas 228 y 220 pares de bases adyacentes a las regiones donadoras y aceptoras, respectivamente. En estos estudios *in vitro*, el análisis de patrón de *splicing* para la construcción silvestre en ambos tipos celulares mostró una única isoforma de ARNm con el tamaño esperado, la cual se corresponde con el transcrito canónico (CT) de 589 pb (Figura R6). Por lo tanto, el midigen desarrollado para la región 5'UTR parece reproducir correctamente el *splicing* de *PAX6* descrito en condiciones fisiológicas.

A continuación, vamos a describir de manera individual los resultados obtenidos para la caracterización funcional de las variantes localizadas en el exón 2 y exón 3.

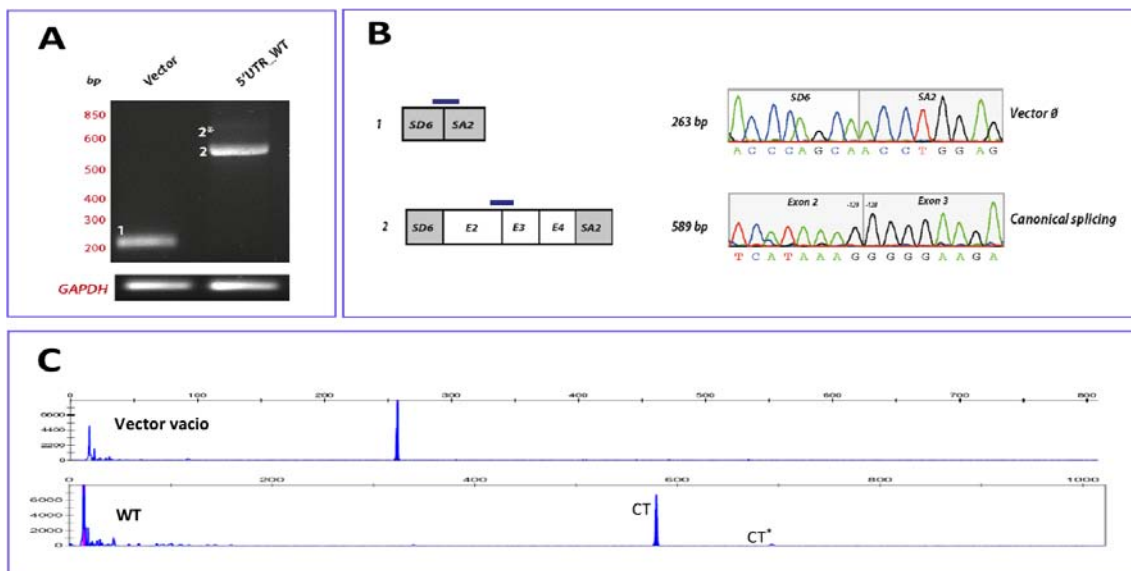


Figura R6. Análisis de *splicing* para la construcción silvestre mediante un ensayo *in vitro* con un midigen para el 5'UTR de *PAX6*. Se muestran los resultados obtenidos de la transfección de ARPE19 (similares a los de HEK293T) con el vector vacío utilizado como control. A) Análisis mediante electroforesis en gel de agarosa para *PAX6* (*GAPDH* como control interno). B) Representación esquemática de los eventos de *splicing*. Se muestran los electroferogramas obtenidos a partir de la secuenciación de las isoformas detectadas. C) Análisis mediante PCR- semicuantitativa. *: indica un artefacto producido por la inserción de 114 pb del vector en el inserto de *PAX6*.

4.1.4. Caracterización funcional de la variantec.-129+1G>A en el exón 2

Tal como se comentó anteriormente, la variante c.-129+1G>A fue identificada por primera vez en este trabajo en 4 familias con aniridia, tres de ellas de nuestra cohorte y una adicional en la cohorte francesa. Este cambio conduciría a la total interrupción del donador canónico del exón 2 no codificante, produciendo un *skipping* aberrante de dicho exón, tal como predicen 4 herramientas *in silico* (Tabla R3). El estudio de expresión de esta variante se pudo realizar en paralelo, tanto *in vitro* mediante el midigen anteriormente descrito, como en ARN procedente de LCL derivadas de pacientes portadores de esta variante.

En el estudio *in vitro*, la variante c.-129+1G>A dio lugar a un complejo patrón de productos de *splicing* formado por al menos 6 isoformas aberrantes diferentes, tal como se muestra en la Figura R6. Todas ellas fueron analizadas mediante un método semicuantitativo de electroforesis capilar y posteriormente secuenciadas, lo que permitió confirmar la naturaleza de los diferentes mecanismos aberrantes de *splicing* originados por la interrupción del donador canónico del exón 2 (Tabla R3). Se han caracterizado 5 isoformas aberrantes originadas por deleciones de regiones exónicas o retenciones intrónicas (Figura R7). Estas incluyen 3 isoformas delecionadas:

i) isoforma $\Delta E2$ de 401 pb causada por el *skipping* del exon 2, tal como se predecía *in silico*.

ii) isoforma $\Delta E2+E3$ de 324 pb, producida por el *skipping* del exón 2 y 3

iii) isoforma $\Delta E2+E3+E4$ de 263 pb, como consecuencia del *skipping* completo de los exones 2 a 4 (dos exones no codificantes y el primer exón codificante). Esta isoforma tiene el mismo tamaño que el *splicing* originado por el vector vacío.

Se han identificado otras dos isoformas aberrantes de mayor tamaño que el CT como consecuencia de retenciones parciales del intrón 2 (Ins_I2):

iv) isoforma Ins_I2(254) de 843 pb debido a la retención de 254 pb.

v) isoforma Ins_I2(406) de 995 pb, por la retención de 406 pb.

Estas isoformas podrían haberse originado aparente por el uso alternativo de dos donadores críticos intrónicos localizados en estas posiciones. Por otro lado, se detectó residualmente la isoforma canónica en los experimentos realizados con la línea ARPE19, aunque representando menos del 5% de los transcritos (Tabla R3, Figuras R7 y R8).

En líneas generales se detectaron las mismas isoformas en las dos líneas transfectadas, sin embargo se observó un desbalance en su proporción (Tabla R3, Figuras R7 y R8). La diferencia más llamativa se observa en la expresión de la isoforma $\Delta E2+E3+E4$, que ocurre de forma residual en la línea HEK293, sin embargo, parece constituir el transcrito mayoritario (32%) en la línea retiniana.

Por otro lado, se establecieron líneas celulares inmortalizando los linfocitos de los casos índice de las familias ANI-006 y ANI-069, portadores de la variante c.-129+1G>A en heterocigosis. A pesar las predicciones *in silico* y los resultados obtenidos en el estudio *in vitro*, no se logró detectar alteración alguna en el *splicing* mediante análisis de RT-PCR para analizar distintos amplicones del ARNm incluyendo los exones 1 al 5 o del 1 al 7. La ausencia de detección de transcritos anómalos, incluyendo el predicho del *skipping* del exón 2, podría deberse a varias causas, incluyendo su degradación mediante el sistema NMD (*nonsense mediated-decay*) por la presencia de un codón de parada prematuro en el transcrito resultante. Por ello, se estudió la expresión de *PAX6* en LCL tratadas con varios tipos de inhibidores directos e indirectos del sistema NMD por separado y usando distintas concentraciones durante un periodo de 6 horas, sin que de nuevo detectásemos la presencia de isoformas aberrantes. En la Figura R9, se muestra un ejemplo de algunos de los experimentos realizados en estas líneas celulares durante este trabajo.

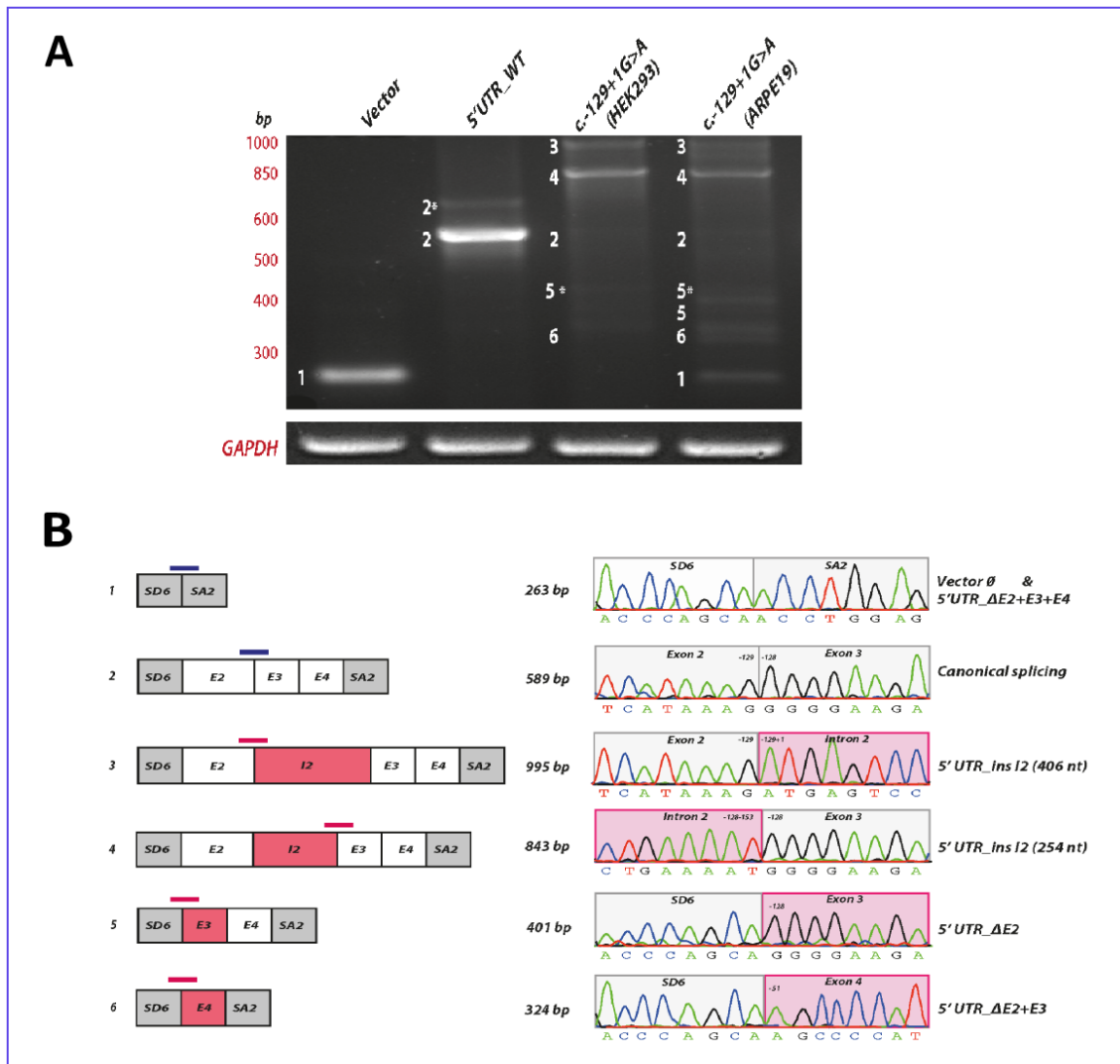


Figura R7. Análisis de splicing para la variante c.-129+1G>A mediante un ensayo in vitro con un midigene para el 5'UTR de PAX6. Se muestran los resultados obtenidos de la transfección de dos tipos celulares (HEK293T y ARPE19) con construcciones silvestres y mutadas, así como el vector vacío utilizado como control. La expresión fue valorada mediante RT-PCR. **A)** Análisis mediante electroforesis en gel de agarosa del patrón de splicing obtenido para PAX6 y la expresión de GAPDH como control interno. **B)** Representación esquemática de los eventos de splicing identificados para cada minigene. Los exones de PAX6 y los sitios de splicing del vector pSLP3 están representados con rectángulos blancos y grises, respectivamente. En rojo se representan los exones involucrados en el splicing aberrante. Se muestran los electroferogramas obtenidos a partir de la secuenciación de las isoformas detectadas a nivel del sitio de splicing canónico (gris) o aberrante (rojo). $\Delta E2$, $\Delta E2+3$, $\Delta E2+E4$: deleción exón 2, 2 y 3 y del 2 al 4, respectivamente. *Ins_I2(254)* e *Ins_I2(406)*: inserción de 254 y 406 pb del intrón 2, respectivamente. *: Isoformas artefactuales obtenidas debido al uso alternativo de un sitio aceptor críptico en el vector pSLP3 que da lugar a la introducción en cada isoforma detectada de 114 pb de secuencias adicionales procedentes del vector a continuación de la secuencia de PAX6.

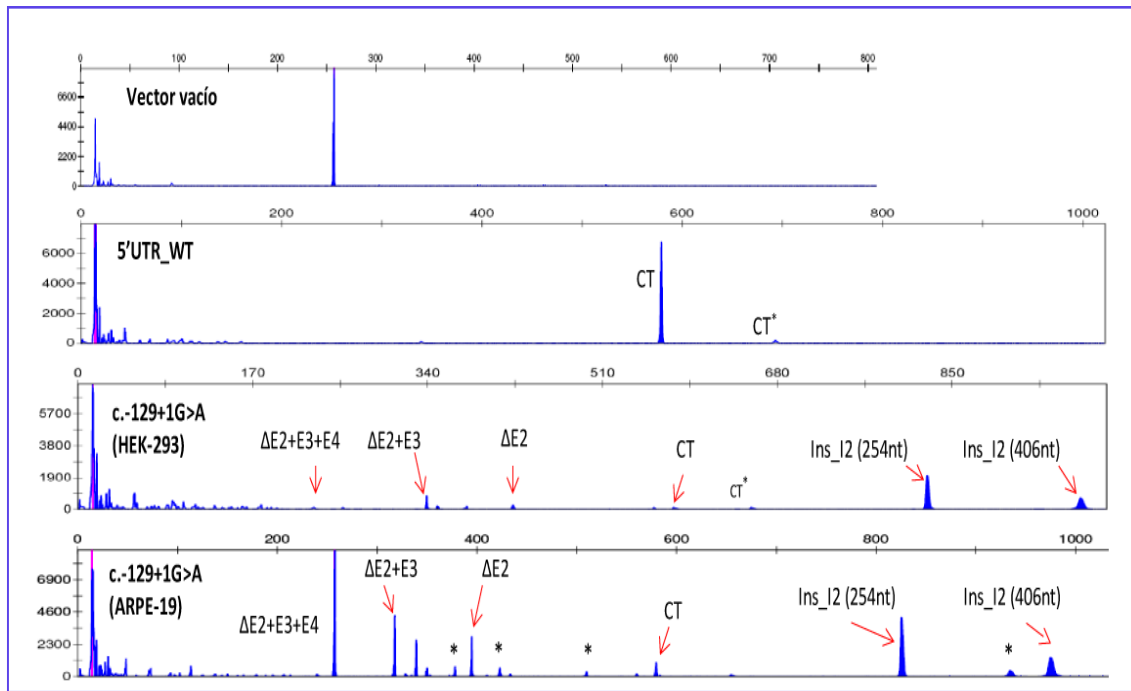


Figura R8. Diferentes isoformas resultantes del análisis de splicing de la variante c.-129+1G>A mediante PCR-semicuantitativa. Se muestran los resultados obtenidos de la transfección de dos tipos celulares (HEK293T y ARPE19) con construcciones silvestres y mutadas, así como el vector vacío utilizado como control. La expresión fue valorada mediante RT-PCR y un posterior análisis semicuantitativo mediante electroforesis capilar.

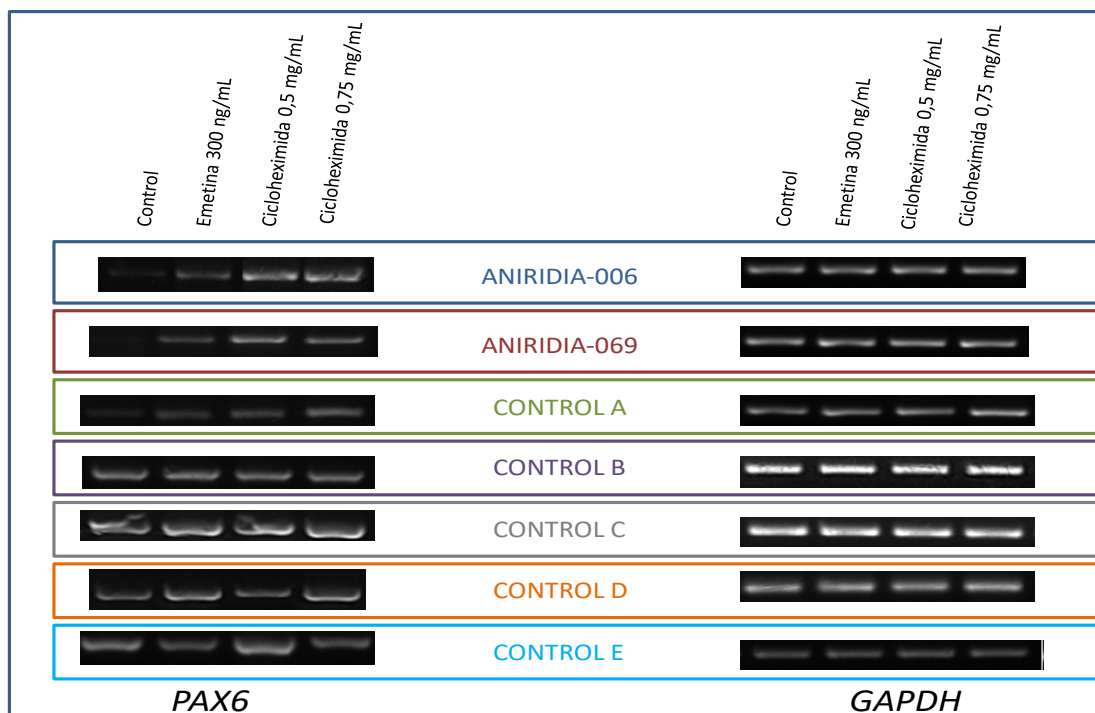


Figura R9. Análisis de expresión de isoformas de PAX6 tras inhibición del sistema NMD en portadores de la mutación c.-129+1G>A. Se analizó la expresión de PAX6 en dos afectados y 5 individuos control mediante RT-PCR con cebadores para los exones 1 y 5 en linfocitos inmortalizados tras distintos tratamientos con inhibidores indirectos del sistema NMD durante 6 horas. En todos los casos se detecta un fragmento de 383 pb correspondiente a la isoforma NM_000280 de PAX6.

4.1.5. Caracterización funcional de variantes localizadas en el exón 3

Como se ha descrito anteriormente, en este trabajo se han identificado en nuestra cohorte 2 variantes en el exón 3, además de otras 4 variantes en pacientes franceses. Para todas ellas, excepto la variante c.-118_-117delTT, se predice una alteración del *splicing* de este exón con solo uno de los predictores utilizados (Tabla R3). En el caso de la variante c.-128-2delA, se pronostica la disrupción de sitio aceptor canónico del exón 3. En otras 3 variantes localizadas en el sitio donador c.-52+1G>A, c.-52+3_-52+4del, c.-52+3_-52+6delinsTG, se predice además la creación de un nuevo sitio de unión a un elemento regulador del *splicing* de tipo ESE (*Exonic Splicing Enhancers*). En el caso de la variante c.-52+5delG, solo se predice un debilitamiento del 12% del donador canónico.

En este trabajo, se realizó un estudio funcional de 4 de estas 6 variantes, la variante recurrente c.-128-2delA y 3 de las variantes localizadas en el donador. No se pudo analizar la variante c.-118_-117delTT debido a dificultades con la mutagénesis dirigida durante la obtención de una construcción específica, ni la variante c.-52+3_-52+4del, dado que fue identificada *a posteriori* por nuestros colaboradores franceses, una vez que este estudio ya había concluido. El estudio de expresión de las 4 variantes analizadas se realizó mediante estudios *in vitro* (Figura R10) utilizando el midigen anteriormente descrito (Figura M5). En el caso ANI-057 portador de la variante c.-128-2delA, se analizó además ARN procedente de LCL. De forma similar a lo descrito para la variante c.-129+1G>T, este análisis de expresión no fue concluyente, pues no se pudo detectar alteración en el patrón de *splicing* con las condiciones utilizadas, en presencia o ausencia de inhibidores del sistema NMD.

En las 4 variantes analizadas *in vitro*, se confirmaron errores en el procesamiento del *splicing* en los ensayos realizados con ambas líneas celulares, detectándose una isoforma prácticamente mayoritaria de 512 pb causada por el *skipping* completo del exón 3 (Figura R10B). En las variantes c.-128-2delA y c.-52+5del además se detectó residualmente el transcrito canónico (<5% del total de las isoformas) (Figura R11).

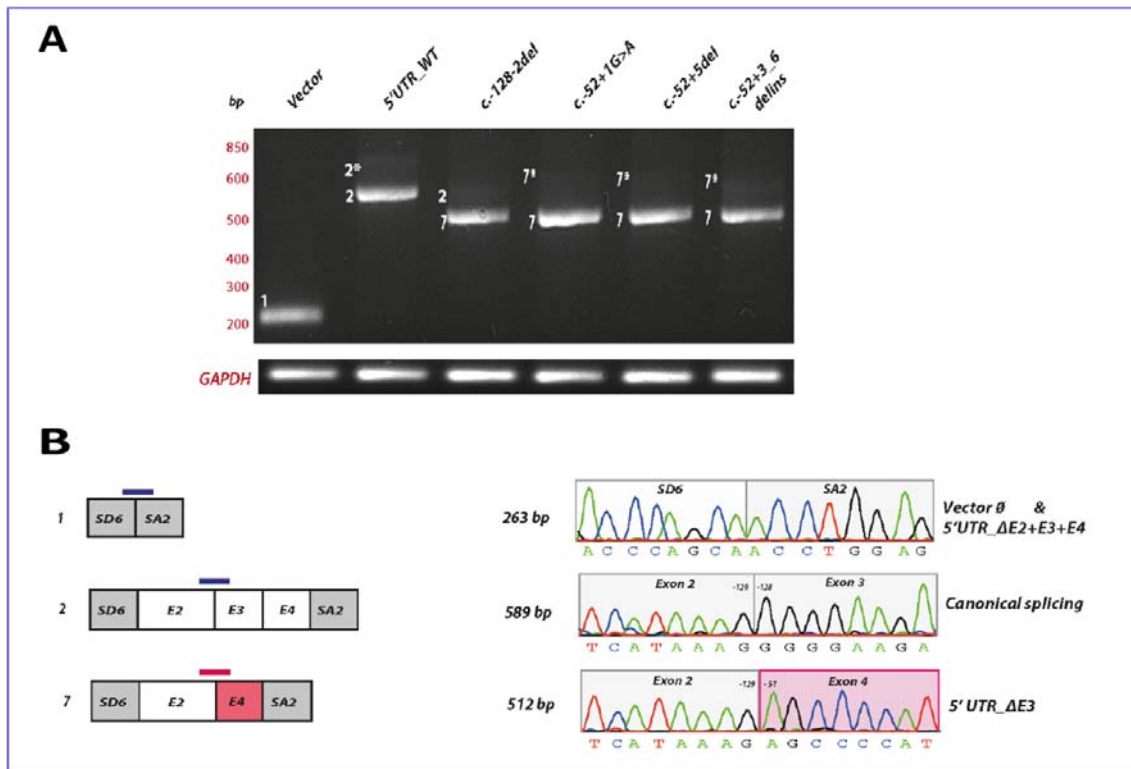


Figura R10. Análisis de splicing para las variantes del exón 3 c.-128-2delA, c.-52+1G>A, c.-52+3_52+6delinsTG y c.-52+5delG mediante un ensayo in vitro con un midigen para el 5'UTR de PAX6. Se muestran los resultados obtenidos de la transfección de ARPE19 con construcciones silvestres y mutadas, así como el vector vacío utilizado como control. La expresión fue valorada mediante RT-PCR. A) Análisis mediante electroforesis en gel de agarosa del patrón de splicing obtenido para PAX6 y la expresión de GAPDH como control interno. B) Representación esquemática de los eventos de splicing identificados para cada minigen. Los exones de PAX6 y los sitios de splicing del vector pSLP3 están representados con rectángulos blancos y grises, respectivamente. En rojo se representan los exones involucrados en el splicing aberrante. Se muestran los electroferogramas obtenidos a partir de la secuenciación de las isoformas detectadas a nivel del sitio de splicing canónico (gris) o aberrante (rojo). ΔE2: deleción exón 2; Ins_12: inserción intrón 2.

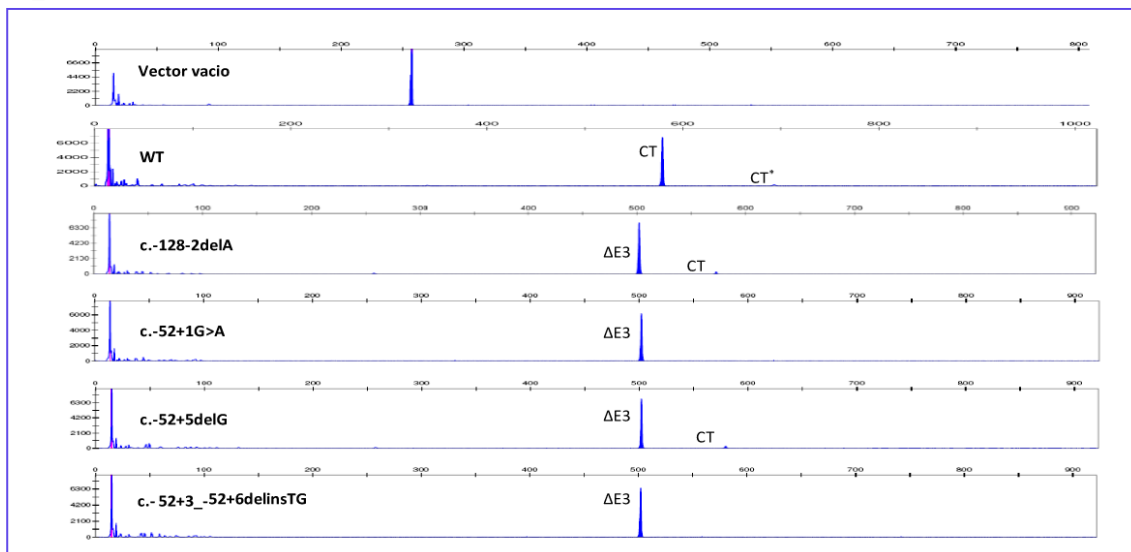


Figura R11. Diferentes isoformas resultantes del análisis de splicing de las variantes del exón 3 mediante PCR-semicuantitativa. Transcritos obtenidos mediante RT-PCR y una posterior amplificación mediante PCR realizada con los cebadores SD6-FAM y SA2 para su posterior análisis semicuantitativo mediante electroforesis capilar.

En resumen, nuestro trabajo ha permitido identificar y caracterizar funcionalmente 5 variantes en 5'UTR de *PAX6*. Por un lado, hemos caracterizado una variante que da lugar a múltiples transcritos aberrantes. Por otro lado, hemos comprobado que distintas variantes situadas en el donador o en el aceptor del exón 3 dan lugar al mismo transcrito (Figura R12).

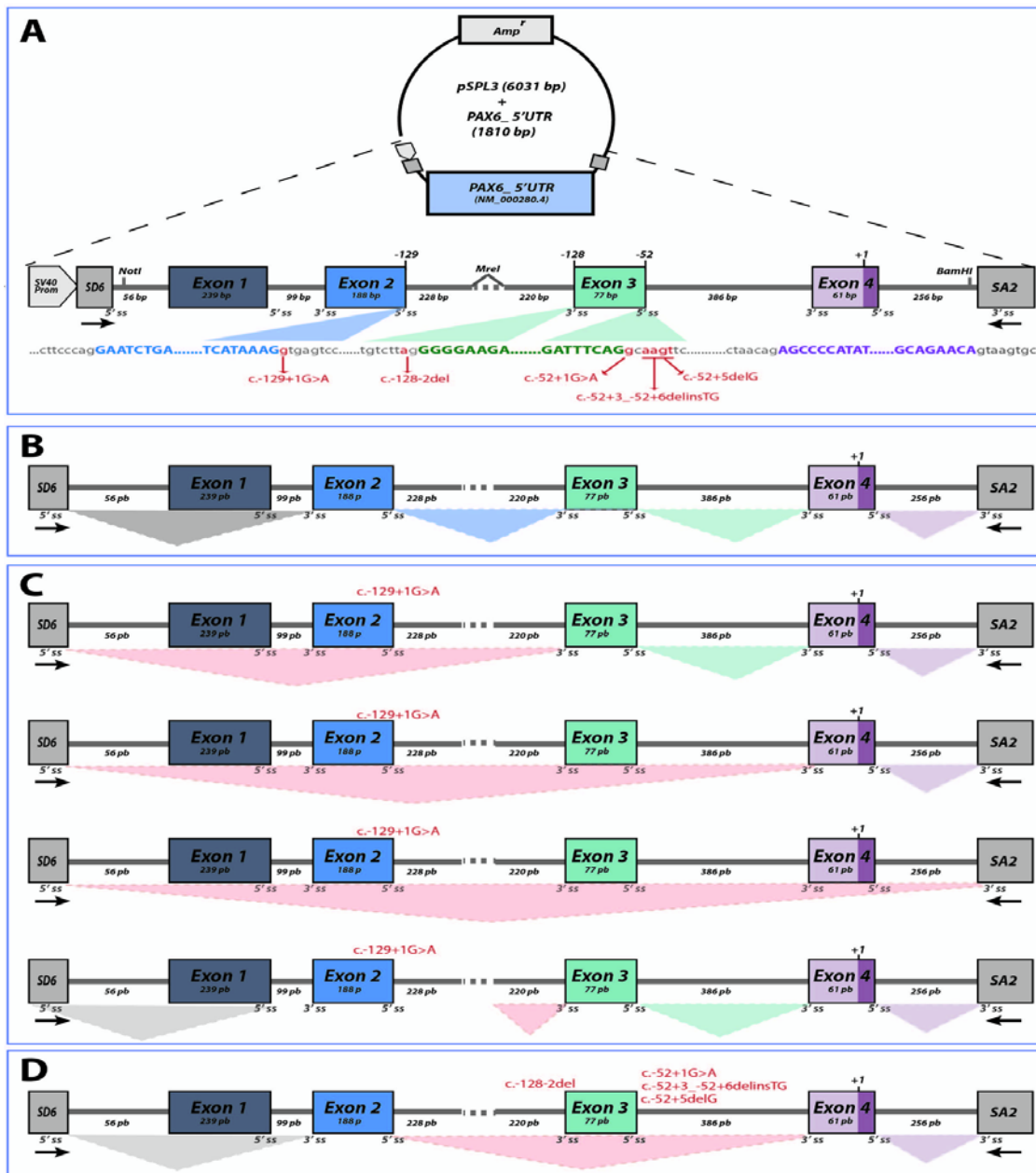


Figura R12. Representación esquemática del patrón de splicing obtenido para los exones del 5'UTR para cada isoforma. A) Características de la construcción y localización de las variantes estudiadas. B) Splicing silvestre. C) Splicing aberrante producido por la variante del exón 2. D) Splicing aberrante producido por las variantes del exón 3. Las flechas: indican la localización de los cebadores específicos del vector (SD6 y SA2) utilizados durante la RT-PCR. Los triángulos rojos invertidos indican splicing aberrante. El resto de triángulos invertidos indican un splicing canónico y su color se debe al exón anterior.

4.1.6. Identificación de variantes patogénicas en otros genes de disgenesia del segmento anterior

Las herramientas de análisis de aCGH y/o NGS utilizadas en el cribado de los 14 casos *PAX6* negativos permitieron el análisis de hasta 120 genes adicionales que están implicados en el desarrollo de otros tipos de anomalías oculares. En 4 de las 14 familias en las que confirmamos la ausencia de variantes en *PAX6*, se pudo identificar alteraciones patogénicas en genes asociados a formas solapantes de disgenesia del segmento anterior. Entre ellas, se ha identificado una mutación en heterocigosis en el gen *PITX2* y 3 variantes heterocigotas en el gen *FOXC1*.

Paciente	Sexo	Edad (años)	Genes involucrados	Variante	Segregación	Fenotipo
ANI-023	F	10	<i>FOXC1</i>	NM_001453.2:c.609delC; p.(Ala204Argfs*111)	AD (materna)	Aniridia-like, madre con anomalía de Axenfeld-Rieger
ANI-059	F	61	9 genes <i>DUSP22, IRF4, EXOC2, HUS1B, LOC285768, FOXQ1, FOXF2, FOXC1, GMDS</i>	Deleción de 1,4 Mb chr6:339.743-1.773.695	<i>De novo</i>	Aniridia glaucoma congénito
MD-245	F	15	<i>DUSP22, IRF4, EXOC2, HUS1B, LOC285768, FOXQ1, FOXF2, FOXC1, GMDS etc.</i>	Anillo del cromosoma 6 Deleción de 1,6 Mb (chr6:339.743-1.900.682) y deleción de 9 Mb (chr6:162.054.374-170.921.089)	<i>De novo</i>	Aniridia sindrómica
ANI-114	F	10	<i>PITX2</i>	NM_153427:c.625delT; p.(Ser209Profs*30)	AD (paterna)	Aniridia-like, hipoplasia maxilar, piel umbilical redundante

Tabla R4. Variantes patogénicas identificadas en los genes *FOXC1* y *PITX2*. Se detalla la información clínica y familiar de los cuatro pacientes portadores, los cuales presentan un fenotipo solapante al de *PAX6*.

4.1.6.1. Variantes en el gen *FOXC1*.

En este trabajo, hemos identificado 3 casos índice portadores de variantes patogénicas en *FOXC1*, gen asociado a un amplio espectro de fenotipos de disgenesia del segmento anterior. Entre ellas, se incluye una variante puntual y la delección completa del gen en 2 pacientes. Las CNVs fueron validadas utilizando un aCGH personalizado, lo que permitió definir su tamaño, los puntos de rotura y los genes involucrados en las mismas.

En el caso índice de la familia ANI-059 se identificó una delección de más de 1,4 Mb en la región cromosómica 6p25.3 que afecta a 9 genes entre ellos *FOXC1* (Figura R13). La posterior segregación familiar de esta CNV mediante MLPA reveló que se trataba de un evento con aparición *de novo*. Este paciente es un caso de aniridia-like que presentaba una aniridia incompleta con glaucoma congénito, lo que podría concordar con presentaciones no sindrómicas asociadas a la haploinsuficiencia de *FOXC1*. No se detectaron anomalías extra-oculares asociadas al síndrome de Axenfeld-Rieger.

Se identificó una segunda CNV afectando a *FOXC1* en un caso esporádico de aniridia de la familia MD-245. Un análisis de aCGH identificó dos deleciones de 1,6 Mb y 9 Mb afectando a las regiones teloméricas 6p25 y 6q26-27, respectivamente del cromosoma 6. Tal como se observa en el cariotipo mostrado en la Figura R13, se trata de un anillo en el cromosoma 6, que apareció de manera espontánea en el paciente.

Por último, se ha identificado una tercera variante *frameshift* p.(Ala204Argfs*111) (Kelberman, 2011) en dos miembros de una familia con sospecha de aniridia-like, ANI-023 (Figura R13). El caso índice y su madre tienen una marcada variabilidad fenotípica, la primera está diagnosticada de aniridia y se sospecha que su madre pueda tener síndrome de Axenfeld-Rieger. Dada esta sospecha clínica en una de las afectas, se cribó el gen *FOXC1* mediante secuenciación Sanger, sin que se identificaran mutaciones. De forma similar, un análisis de WES realizado como primera aproximación de NGS tampoco había permitido identificar la variante causal. La razón de ambos resultados negativos es que esta variante se encuentra en una región altamente repetitiva, difícil de secuenciar, que no estaba bien representada en el análisis de WES. El uso de un

panel de captura a mayor profundidad de lectura permitió que se cubriera mejor esta región y con ello, se reveló la existencia de esta variante (Figura R13).

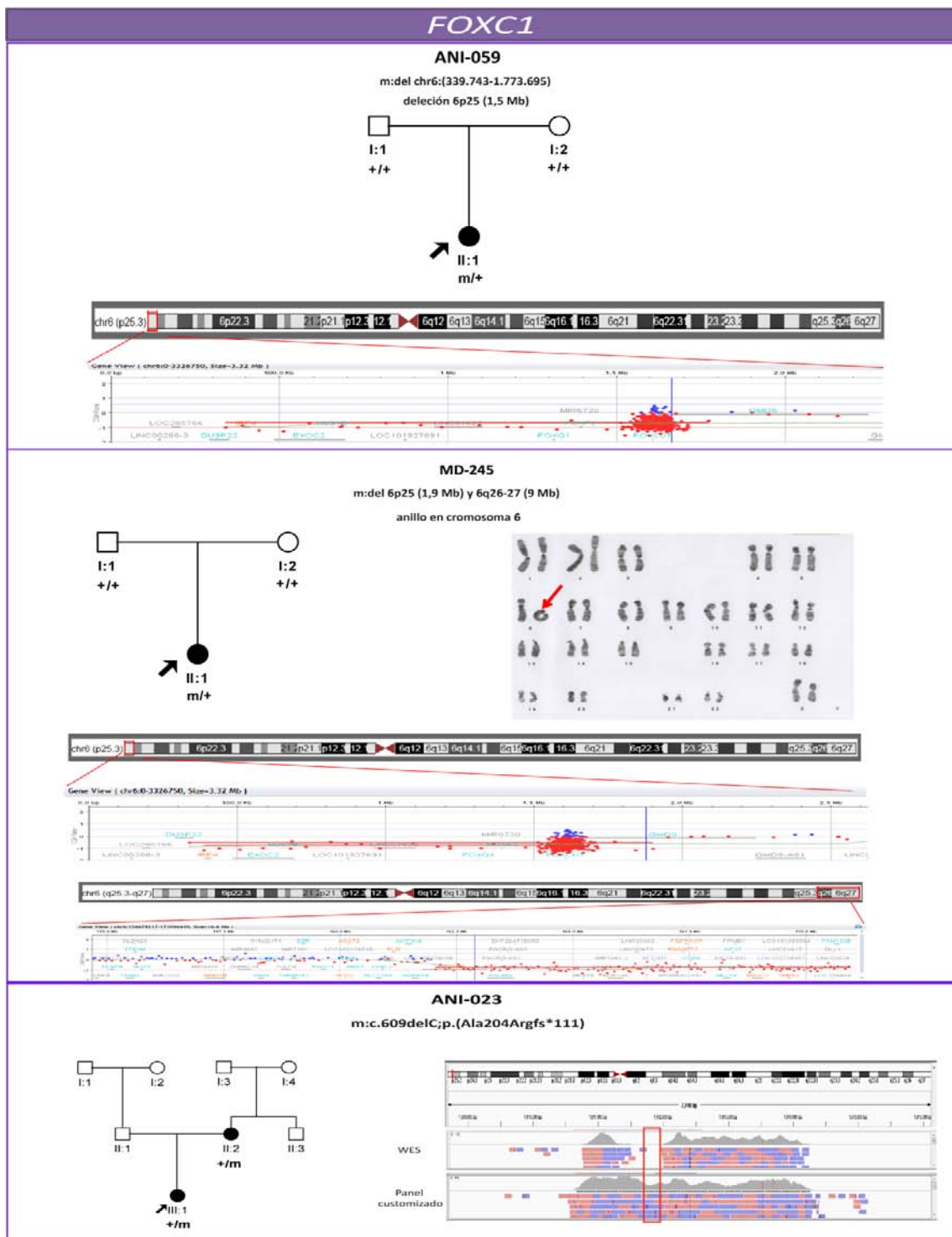


Figura R13. Identificación y segregación de variantes en FOXC1 en tres pacientes con aniridia. Representación de dos CNVs identificadas mediante aCGH en los pacientes ANI-059 y MD-245. En este caso, se confirmó la presencia de un anillo cromosómico mediante cariotipo. Se describe la identificación de una SNV por nuestro panel personalizado en una región que no estaba cubierta en un análisis previo mediante WES, tal como se muestra en el recuadro rojo en la ANI-023.

4.1.6.2. Variantes identificadas en el gen *PITX2*

En este trabajo detectamos una variante nueva de tipo *frameshift* p.(Ser209Profs*30) localizada en el exón 5 del gen *PITX2* en el caso índice de la familia ANI-114, con sospecha de aniridia congénita por la presencia de hipoplasia de iris, queratopatía, disfunción limbar y glaucoma congénito. La segregación familiar identificó que el padre, aparentemente asintomático, presentaba la misma variante que su hija (Figura R14). Una posterior revisión clínica indicó que el padre e hija presentaban otras características asociadas al síndrome de Axenfeld-Rieger, como piel redundante periumbilical, hipoplasia maxilar inferior, agenesia maxilar superior y dientes hipoplásicos. Adicionalmente, el padre presentaba presión intraocular elevada y había sido diagnosticado de glaucoma en la 4ª década de la vida, síntoma que no había sido relacionado con la patología congénita de su hija. Todos estos síntomas son concordantes con el fenotipo asociado a las variantes en *PITX2*. Este estudio confirmó por un lado la expresividad variable de las variantes en este gen, y por otro lado que síntomas aparentemente aislados en el *probandus* y su padre realmente tienen su origen en una patología genética con herencia autosómica dominante, lo que ha tenido gran implicación en el consejo genético de esta familia.

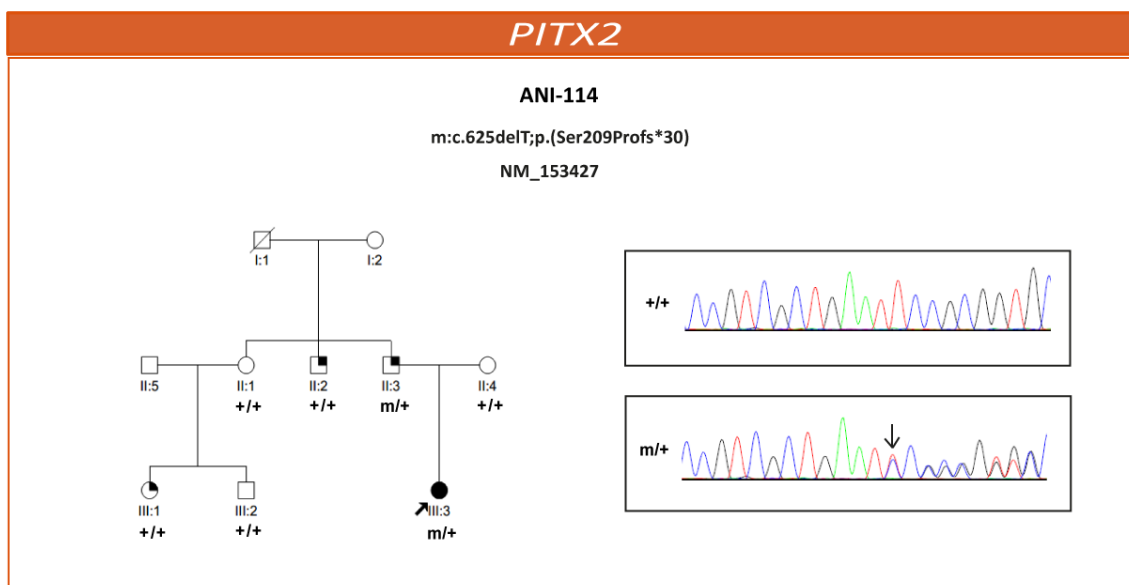


Figura R14. Identificación y segregación de variantes en el gen *PITX2*. Se muestra la segregación familiar y el electroferograma Sanger de la variante c.625delT;p.(Ser209Profs*30) (NM_153427) en la familia ANI-114.

4.2. Caracterización funcional de VUS que afectan al *splicing* del exón 6 de *PAX6*

En este apartado, nos hemos centrado en el análisis de las variantes de significado incierto identificadas en *PAX6* en alrededor del 10% de los casos de nuestra cohorte de pacientes. Tal como se observa en la Figura R15, dos tercios de ellas se localizan en una de las regiones *hotspot* de *PAX6*, el exón 6, que codifica gran parte del dominio pareado de la proteína.

Las VUS encontradas en el exón 6 de *PAX6* en nuestra cohorte de pacientes son:

i) dos variantes sinónimas c.174C>T;p.(Gly58=) y c.249T>G;p.(Val83=).

ii) dos variantes *missense* c.178T>C;p.(Tyr60His) y c.356G>T;p.(Ser119Ile).

iii) una duplicación *in-frame* c.326_331dup;p.(Glu109_Gly110dup).

iv) una variante de *splicing* c.357+6T>G que está situada fuera de los nucleótidos canónicos del sitio donador.

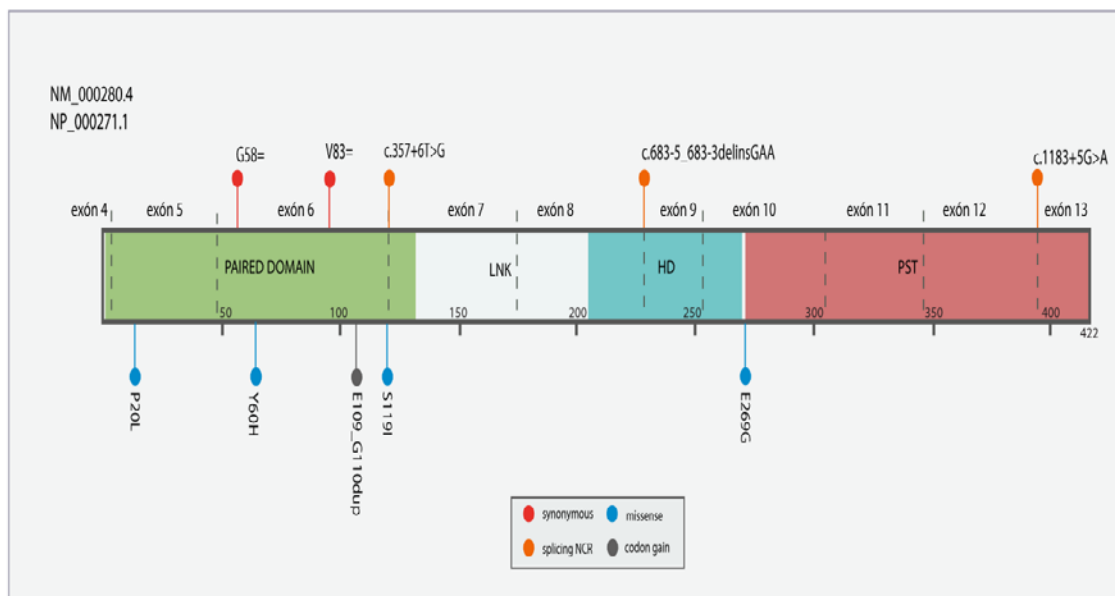


Figura R15. Representación de la localización de las 10 VUS identificadas en *PAX6*. En la parte superior de la imagen aparecen las variantes no codificantes: dos variantes sinónimas y tres variantes de *splicing* en nucleótidos no canónicos. En la parte inferior de la figura aparecen las variantes codificantes que no dan a alelos nulos: cuatro variantes *missense* y una variante de ganancia de dos codones. Se representan los exones codificantes de *PAX6* y los dominios de la proteína a la que da lugar. HD: homeodominio, PST: dominio prolina-serina-treonina, NCR: non-coding *splicing* región.

Tal como se describe en la Tabla R5, ninguna de ellas estaba descrita en el momento de su identificación en bases de datos mutacionales (*HGMD* y *PAX6 LOVD Database*) ni está actualmente en bases de datos poblacionales (gnomAD y Spanish Variant Server). Además, todas ellas afectan a nucleótidos conservados evolutivamente y fueron predichas como potencialmente deletéreas en programas *in silico* de predicción de patogenicidad. En el caso de las dos variantes *missense*, afectan a residuos altamente conservados, cuya sustitución se predice como patogénica en 3 programas diferentes. Según criterios ACMG, fueron clasificadas como clase 3. Estudios *in silico* realizados con 4 predictores de *splicing*, resumidos en la Tabla R6, revelaron que todos estos cambios a nivel nucleotídico tienen capacidad de crear o potenciar 5 sitios donadores crípticos diferentes en el exón 6, o alterar el sitio donador canónico. Por ello, se predice que podrían modificar el proceso de *splicing* de dicho exón. En la Figura R16, se esquematizan tanto la localización de las VUS como los potenciales donadores crípticos situados en el exón 6.

Durante este trabajo, se valoraron otras posibles variantes en el exón 6 que pudieran estar alterando el *splicing* a través de estos potenciales sitios crípticos, tanto identificadas en nuestra cohorte como descritas en bases de datos de mutaciones de *PAX6*. Encontramos en la base de datos *PAX6 LOVD* una tercera variante sinónima, c.333C>A;p.(Val111=), que Murton et als, 2000 identificaron en un paciente con aniridia. De forma similar a las otras variantes en estudio de nuestra cohorte, esta variante tiene potencial para activar un donador críptico del exón 6 (Tablas R5 y R6).

Con el fin de validar las predicciones obtenidas *in silico* para estas variantes, durante este trabajo se llevaron a cabo estudios de expresión para valorar posibles isoformas de ARNm aberrantes por el uso alternativo de los potenciales donadores crípticos. Estos estudios fueron realizados mediante el análisis de RT-PCR en ARN procedente de linfocitos inmortalizados de pacientes o en dos tipos de líneas celulares, HEK293T y ARPE19, tras su transfección con minigenes para el exón 6 de *PAX6* utilizando construcciones de tipo silvestre y para cada una de las variantes en estudio, tal como se describe en la Figura M6 de Material y Métodos.

Cambio nucleótido	Cambio residuo	Familia	Herencia	Clasificación ACMG	DANN	CADD	GERP	PhyloP100	SIFT	Mutation Taster	Align GV-GD	Referencias
c.174C>T	p.(G58=)	ANI-062	AD	3	0,63	12,7	5,35	----	----	----	----	Filatova, 2019
c.178T>C	p.(Y60H)	ANF-060	AD*	3	0,99	27,7	5,35	6,023	D	D	Class C65	Tarilonte, 2018
c.249T>G	p.(V83=)	ANI-048	AD*	3	0,663	6,33	5,35	----	----	----	----	ND
c.326_331dup	p.(E109_G110 dup)	ANI-053	<i>de novo</i>	3 o 4	----	----	5,35	----	----	----	----	ND
c.333C>A [#]	p.(V111=)	N.D.	N.D.	2 o 3	0,73	13,2	5,35	----	----	----	----	Murton, 2000
c.356G>T	p.(S119I)	ANF-037	Sospecha AD	3 o 4	0,99	34	5,35	6,023	D	D	Class C65	ND
c.357+6T>G	p.(?)	ANI-001	AD	3	0,95	20,1	5,35	----	----	----	----	ND

Tabla R5. Información de patogenicidad de las VUS en el exón 6 de PAX6. Se muestran datos de segregación, clasificación de patogenicidad según los criterios ACMG y predicciones in silico utilizando las siguientes herramientas patogenicidad: DANN, CADD; herramientas de conservación: GERP y PhyloP100 y herramientas para variantes missense: SIFT, Mutation Taster y Aling GV-GD; [#]: Variante recuperada de la base de datos LOVD de mutaciones de PAX6. *: sospecha de mosaicismo. D: deletéreo. ND: no datos.

Resultados

Cambio nucleótido	HSF	MaxEntScan	NNSPLICE	SSF	Posición predicción	Predicción <i>splicing in silico</i>
c.174C>T	— ⇒ 84,8 (+100%)	— ⇒ 6,9 (+100%)	— ⇒ 0,98 (+100%)	74,1 ⇒ 76,8 (+3,7%)	c.173	Creación de un nuevo 5'SS críptico en c.173
c.178T>C	76,5 ⇒ — (-100%)	7,3 ⇒ — (-100%)	0,74 ⇒ — (-100%)	72,1 ⇒ — (-100%)	c.177	Pérdida de un 5'SS críptico en c.177. Alteración de sitio ESE SF2.
c.249T>G	80,4 ⇒ 91,3 (+13,5%)	4,4 ⇒ 8,1 (+86,7%)	0,51 ⇒ 1,00 (+94,6%)	70,1 ⇒ 82,7 (+18%)	c.250	Fortalecimiento del 5'SS críptico en c.250
c.326_331dup	79,5 ⇒ — (-100%)	0,78 ⇒ — (-100%)	-----	-----	c.331	Disrupción de un 5'SS críptico en c.331 Creación de un nuevo sitio ESS (hnRNPA1)
c.333C>A [#]	79,5 ⇒ 84,6 (+6,3%)	0,78 ⇒ 7,6 (+100%)	— ⇒ 0,98 (+100%)	— ⇒ 77,8 (+100%)	c.331	Creación o fortalecimiento de un 5'SS críptico en c.331
c.356G>T	83,2 ⇒ 83,3 (+0,1%)	10,1 ⇒ 10,4 (+3,7%)	0,99 ⇒ 0,98 (-1,5%)	77,4 ⇒ 77,9 (+0,6%)	c.357+1	No alteración 5'SS canónico. Alteración sitio ESE SRp40/SRp55.
c.357+6T>G	_ ⇒ 81,7 (+100%)	— ⇒ 4,3 (+100%)	-----	— ⇒ 77,4 (+100%)	c.357+6	Creación nuevo 5'SS en c.357+6. Debilitamiento secundario del donador natural c.357+1

Tabla R6. Predicciones in silico del potencial espliceogénico de variantes en el exón 6 de PAX6. Scores de las predicciones de *splicing* para el donador canónico (c.357+1) o 5'SS crípticos obtenidos a partir de 4 herramientas de *splicing in silico*: Human Splicing Finder (HSF); MaxEntScan; Splice Site Finder (SSF) y NNSPLICE; 5'SS: donador (del inglés *splicing site*); ESE: *exonic splicing enhancer* (potenciador del *splicing*), ESS: *exonic splicing silencer* (silenciador del *splicing*), D: *deletéreo*; [#]: Variante recuperada de la base de datos de LOVD de mutaciones de PAX6.

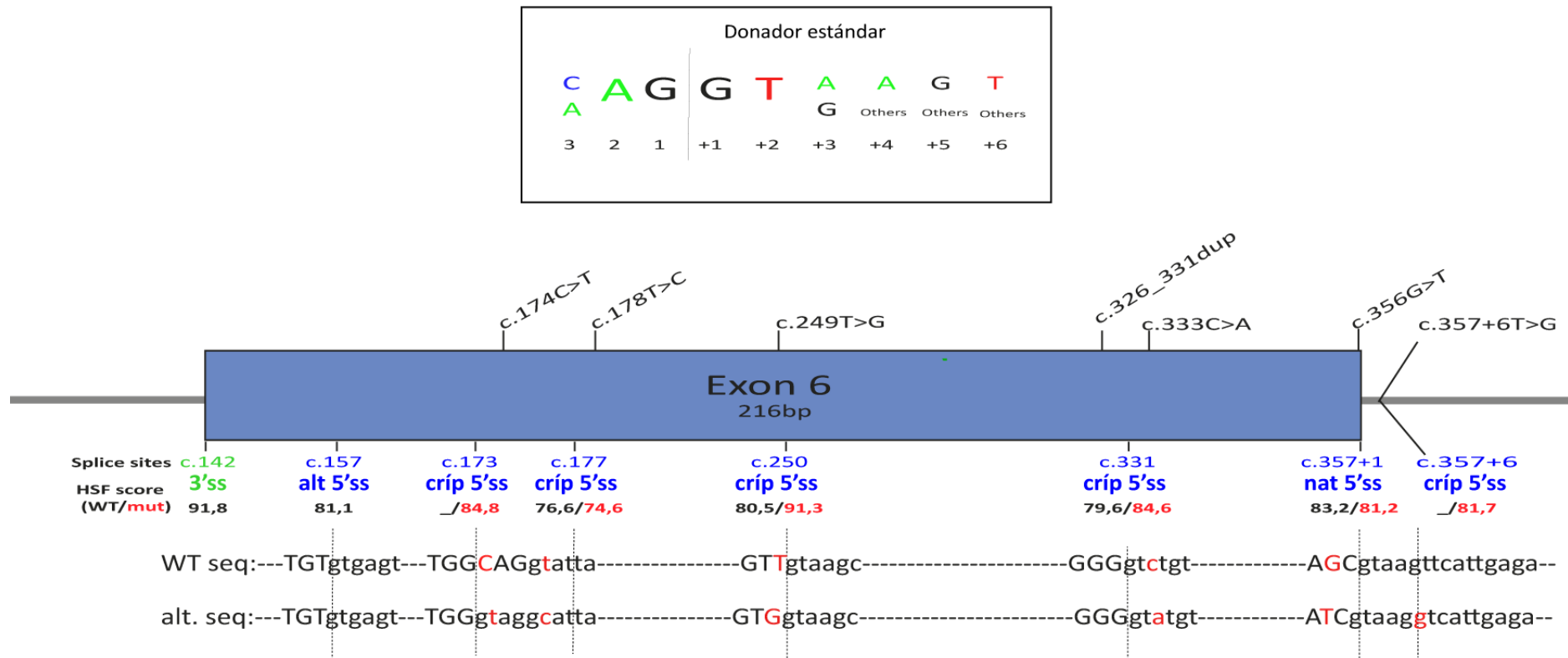


Figura R16. Representación esquemática del potencial espliceogénico para el donador canónico y los donadores alternativos identificados en el exón 6 de PAX6. En la parte superior se muestran las distintas variantes en estudio. En la parte inferior se describen las posiciones de los sitios de splicing: el aceptor canónico (3'ss) está representado en verde y los donadores (5'ss) predichos representados en azul, incluyendo el sitio canónico (c.357+1), un sitio alternativo en posición c.157 (alt 5'ss) y los 5 crípticos (críp 5'ss). Se indican además las posiciones del dinucleótido GT (+1) y el score obtenido en la herramienta in silico Human Splicing Finding (>75). WT: secuencia silvestre; mut: secuencia mutada.

El resultado del estudio funcional en la construcción silvestre fue el uso de forma mayoritaria del donador canónico, localizado en la posición c.357+1, tal como se detecta tras un análisis semicuantitativo la isoforma canónica (CT, del inglés *canonical transcript*) en el 86% del total de transcritos obtenidos. Además, se detectaron otras dos isoformas minoritarias $\Delta 6_{181}$ (12%) y $\Delta 6_{27}$ (2%) (Figura R17 y Anexo 6), que se deben al uso alternativo de dos donadores crípticos situados en posiciones c.177 y c.331, respectivamente. Estos resultados, que fueron similares en las dos líneas celulares transfectadas, podrían confirmar la debilidad del donador canónico del exón 6, c.357+1, tal como se predice en los análisis *in silico* previamente realizados.

A continuación, vamos a describir de manera individual los resultados obtenidos para cada una de las variantes estudiadas en el exón 6. Un resumen de todos los hallazgos se encuentran en los Anexos 5 y 6.

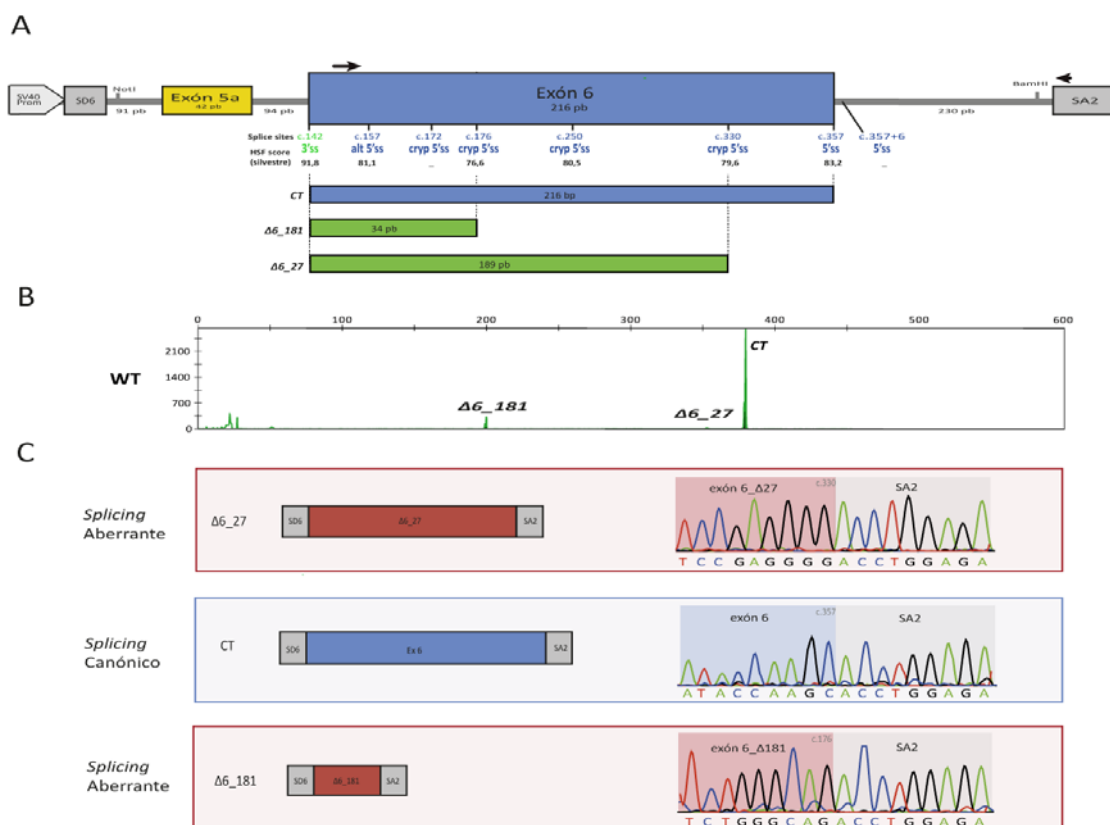


Figura R17. Resultados del estudio funcional en la construcción silvestre mediante el minigen "PAX6_ex6". A) Representación esquemática del splicing del exón 6 de PAX6 predicho para la secuencia silvestre, localización de los potenciales sitios de splicing y el resultado del análisis *in silico* así como de las isoformas de splicing derivadas. B) Análisis de los transcritos obtenidos mediante RT-PCR semicuantitativa. C) Secuenciación Sanger de las isoformas detectadas. $\Delta 6_{27}$ y $\Delta 6_{181}$: deleción de 27 y 181 pb, respectivamente; CT: transcrito canónico.

4.2.1 Identificación de variantes exónicas silentes que crean donadores crípticos exónicos

Previamente a este trabajo, se habían identificaron mediante secuenciación Sanger dos variantes sinónimas c.174C>T;p.(Gly58=) y c.249T>G;p.(Val83=), en heterocigosis en los casos índice de las familias ANI-062 y ANI-048, respectivamente. En un estudio posterior mediante paneles de NGS y/o aCGH, en estos pacientes no se identificaron otras posibles variantes patogénicas en *PAX6* ni en ninguno de los 120 genes relacionados con MOC estudiados.

La variante sinónima **c.174C>T;p.(Gly58=)** segrega en los 3 individuos afectados con aniridia congénita de una familia de dos generaciones (ANI-062, Figura R18.A), en la que existe además sospecha de mosaicismo germinal en los padres sanos de dos hermanos afectados. Los predictores *in silico* indican que el cambio C>T en c.174 conlleva la creación de un nuevo sitio donador alternativo (GT) en la posición c.173 de una fuerza ligeramente mayor al obtenido para el sitio canónico c.357+1 (*scores* HSF: 84.8 y 83.2, respectivamente) (Figura R18.B). Estos datos revelan que esta variante sinónima es candidata a modificar el *splicing*, favoreciendo la utilización de este nuevo donador frente al canónico, lo que podría conducir a que se produjese un *skipping* parcial de este exón.

Esta hipótesis fue confirmada mediante estudios de expresión mediante RT-PCR en una línea LCL derivada del individuo heterocigoto III:3 (Figura R18.A), así como en un análisis *in vitro* mediante minigenes. En ambos casos, el análisis funcional de la variante c.174C>T revela la aparición de una única isoforma aberrante, $\Delta 6_{185}$. Su posterior secuenciación confirmó la delección de las últimas 185 pb del exón 6 como consecuencia de la creación del donador predicho en posición c.173 (Figura R18.C). El resultado del minigen indica además que el sitio canónico estaría totalmente silenciado por el nuevo donador, debido a que no se pudo detectar la isoforma silvestre.

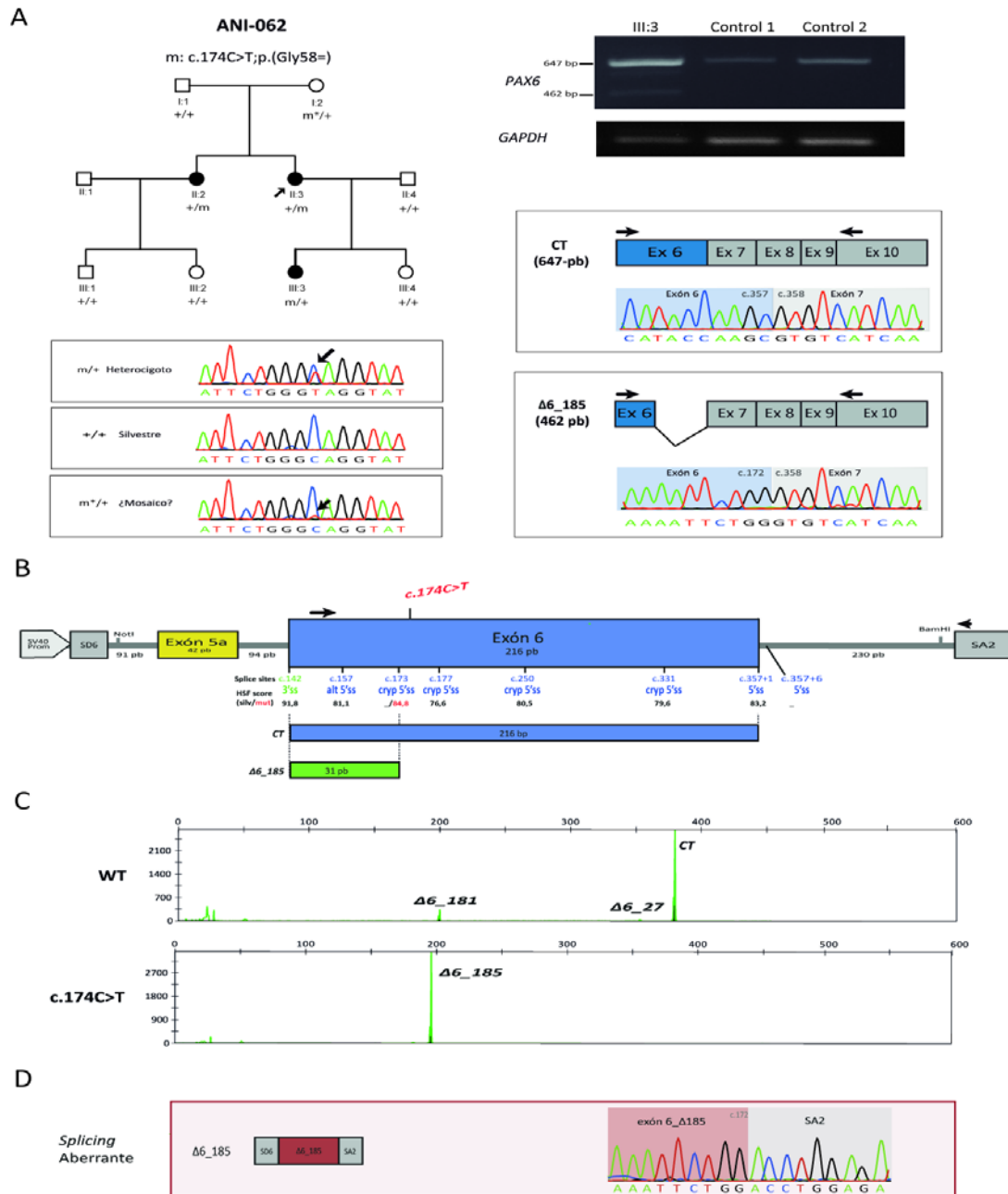


Figura R18. Estudio funcional de la variante c.174C>T. A) Análisis de segregación en la familia ANI-062, se muestran los electroferogramas para un individuo silvestre (+/+), uno heterocigoto (m/+) y un posible mosaico (m*/+) para la variante. En el panel de la izquierda se muestran los resultados del análisis de expresión a partir de LCLs procedentes del individuo heterocigoto III:3 y de dos controles sanos. El estudio se realizó mediante RT-PCR utilizando cebadores para los exones 6 y 10 de PAX6 y GAPDH como control interno. En el análisis mediante electroforesis en agarosa se detectaron dos isoformas en el individuo III:3: el fragmento esperado de 647 pb, correspondiente a la isoforma CT derivada del splicing correcto y un fragmento adicional de 462 pb (isoforma Δ6_185). La posterior secuenciación Sanger de esta isoforma confirmó el skipping parcial de 185 pb debido al uso del donador c.173. **B)** Representación esquemática del splicing del exón 6 predicho para la variante, de la localización de los sitios de splicing canónico y crípticos, del resultado del análisis in silico así como de las isoformas de splicing derivadas. **C)** Análisis semi-cuantitativo de RT-PCR mediante electroforesis capilar de los transcritos obtenidos en los experimentos de minigenes para las construcciones silvestre y de la variante, utilizando los cebadores c.148-HEX y SA2. **D)** Secuenciación Sanger de la isoforma Δ6_185 obtenida en el análisis de minigenes. Δ6_27, Δ6_181 y Δ6_185: delección de 27, 181 y 185 pb, respectivamente; CT: transcrito canónico.

En conclusión, estos análisis funcionales demuestran que esta variante silente desequilibra el *splicing* produciendo una delección parcial del exón 6, cuya consecuencia proteica sería la creación de un codón de parada prematuro 12 residuos aguas abajo de la variante (p.(Arg59Valfs*12)). Sin embargo, esta proteína aberrante no llegaría a formarse ya que la isoforma $\Delta 6_{185}$ sería presumiblemente degradada por el sistema NMD (*nonsense mediated decay*).

La variante **c.249T>G;p.(Val83=)** es la segunda sinónima identificada en nuestra cohorte, concretamente en una familia de dos generaciones de afectados de aniridia congénita (ANI-048, Figura R19.A). Aunque esta variante segrega en un padre y su hija, no se pudo completar el análisis de segregación en la familia debido a la imposibilidad de obtener muestras de ADN en los otros 3 miembros afectados. Esta familia, al igual que la anterior, podría ser un caso de mosaicismo germinal debido a la presencia de la patología en dos hermanos afectados descendientes de una pareja asintomática. Los predictores *in silico* de *splicing* indican que esta variante aumenta el potencial de un donador alternativo localizado en la posición c.250 (*score* HSF: 80,5 a 91,3), superando en más de un 8% al donador canónico c.357+1 (Figura R19.B). De forma similar a la anterior variante, esta variante silente altera el *splicing* del exón 6 produciendo un *skipping* parcial de este exón.

El análisis funcional del *splicing* del cambio c.249T>G mediante minigenes confirmó las predicciones. En el análisis de expresión se observó un transcrito mayoritario $\Delta 6_{108}$, de menor tamaño que el CT por la delección de 108 pb aguas abajo de la variante. Esta isoforma se origina como consecuencia de la potenciación del donador c.250 (Figura R19.C-D). Esta isoforma defectuosa, que representa el 91% del total de transcritos, conduce a una delección *in-frame* de 36 aa (p.(Lys86_Ser121del)) en el PD de la proteína (Anexo 6). Además, se observaron otras dos isoformas minoritarias, una isoforma elongada I6_5 y la isoforma canónica en un 6% y un 3%, respectivamente (Figura R19.C). La elongación de 5 pb del exón 6 podría ser explicada por la activación secundaria de un donador críptico que se predice en los análisis *in silico* en posición intrónica c.357+6 (Figura R16). Estos hallazgos confirman que esta variante potencia un sitio críptico exónico, lo que a su vez silencia indirectamente el sitio donador canónico casi por completo.

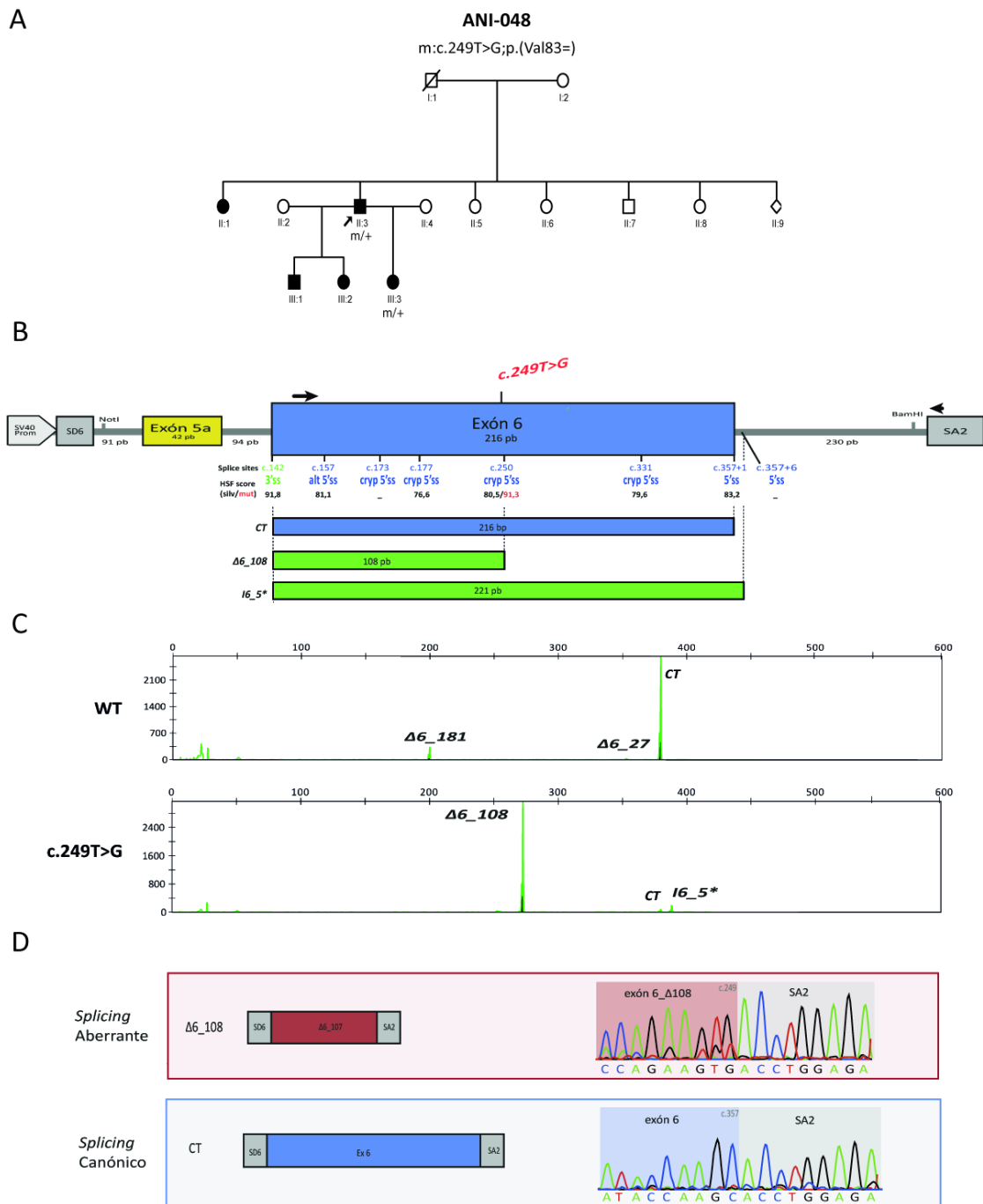


Figura R19. Estudio funcional de la variante c.249T>G. A) Análisis de segregación en la familia ANI-048. **B)** Representación esquemática de la predicción de splicing del exón 6 para la variante, así como la representación de las isoformas obtenidas in vitro. **C)** Resultados de un análisis de expresión mediante RT-PCR a partir de una construcción con la variante. **D)** Secuenciación Sanger de las isoformas obtenidas Δ6_27, Δ6_108 y Δ6_181: delección de 27, 108 y 181 pb, respectivamente (ordenadas según frecuencia); CT: transcrito canónico; I6_5: inserción de 5 pb intrónicas; *: isoforma no secuenciada.

Por último, se realizó un estudio funcional mediante minigenes para la variante sinónima **c.333C>A;p.(Val111=)**, localizada cerca de un donador críptico predicho en la posición c.331. Esta variante fue detectada por primera vez en un paciente con aniridia del cual desconocemos su genealogía (Murton, 2000). Los predictores de *splicing* calculan para esta variante un aumento del potencial espliceogénico del donador críptico c.331 en un 5% (score HSF de 79,6 a 84,6), lo que hace que pueda ser potencialmente más fuerte que el 5'SS canónico (Figura R20.A).

El ensayo *in vitro* con la variante c.333C>A reveló, también en esta ocasión, un patrón de *splicing* diferente al de la construcción silvestre. En esencia se trata de las mismas isoformas identificadas en el constructo silvestre, pero con un notable desbalance, siendo la isoforma mayoritaria la delección de 27 pb ($\Delta 6_{27}$) representando el 84% de los transcritos, seguida por un 13% de la isoforma canónica y un 3% de la delección de 181 pb ($\Delta 6_{181}$). Con lo cual, las isoformas canónica y $\Delta 6_{181}$ descendieron un 83% y un 9%, respectivamente con respecto a la construcción silvestre, en cambio la Δ_{27} aumentó un 82% (Figura R20 y Anexo 6). Por ello, afirmamos que esta sinónima potencia el donador alternativo en posición c.331, lo cual daría lugar a una proteína aberrante con una delección *in-frame* de 9 residuos p.(Cys112_Val120del) (Anexo 6).

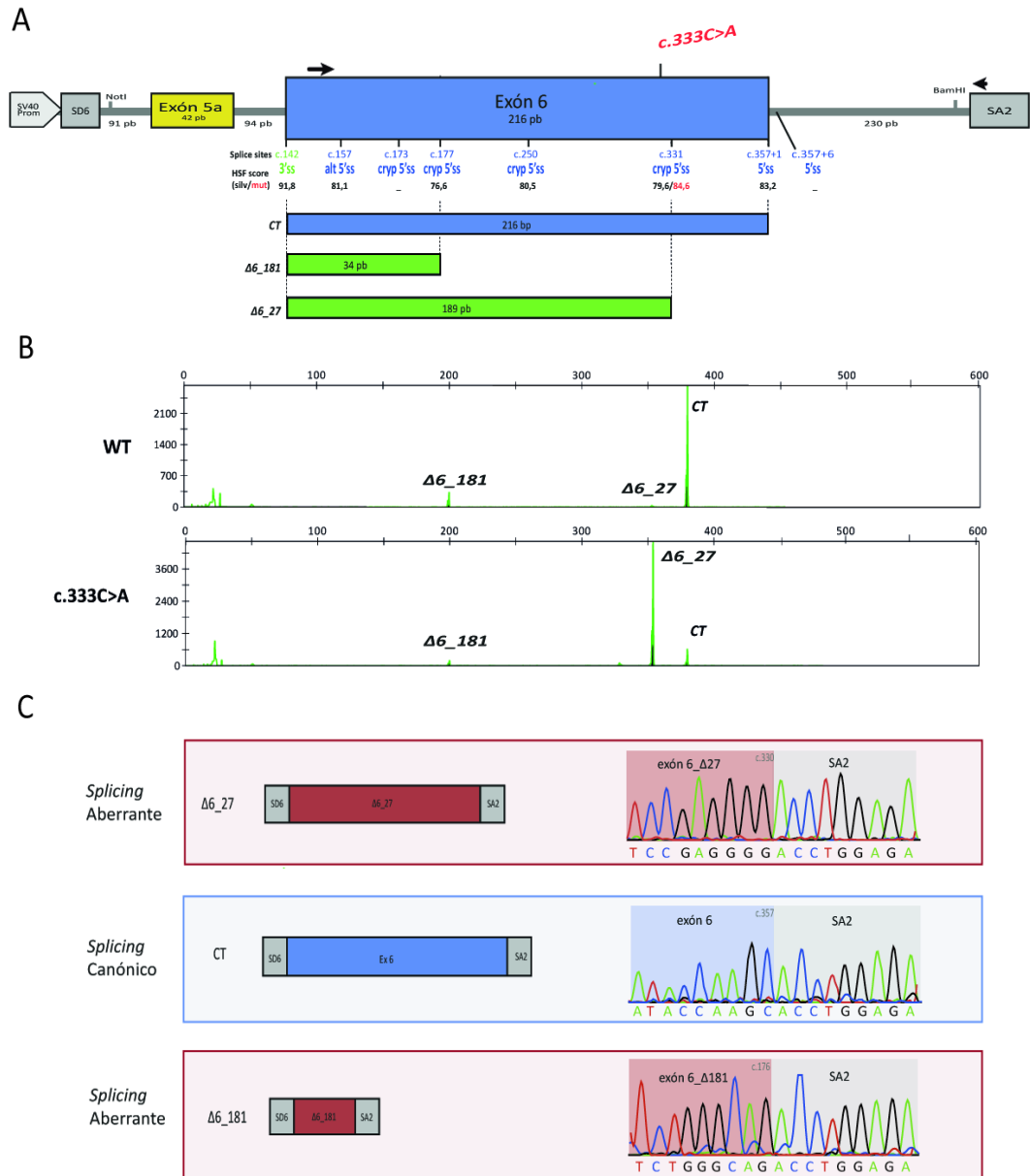


Figura R20. Estudio funcional de la variante c.333C>A. A) Representación esquemática de la predicción de splicing del exón 6 para la variante, así como de las isoformas obtenidas in vitro. **B)** Resultados del análisis semi-cuantitativo de RT-PCR mediante electroforesis capilar realizado con los cebadores c.148 y SA2 a partir de una construcción con la variante. **C)** Secuenciación Sanger de las isoformas obtenidas (ordenadas según frecuencia). Δ6_27, y Δ6_181: delección de 27, 181 pb, respectivamente; CT: transcrito canónico.

4.2.2. Creación de un nuevo donador críptico intrónico por una variante situada en sitio no canónico de *splicing*

La variante **c.357+6T>G** fue identificada en la familia ANI-001, donde aparece *de novo* en el individuo II:1 y fue transmitida a su único hijo afecto III:2 (Figura R21.A). Entre ambos individuos afectados existe una importante variabilidad fenotípica. El padre fue diagnosticado de anomalía de Peters debido a la presencia bilateral de leucomas centrales congénitos, mientras que su hijo presenta una aniridia clásica, además de malformaciones congénitas en sus extremidades inferiores. En este último, se identificó una segunda variante en heterocigosis de origen materno c.1159G>A;p.(Glu387Lys) en el gen *CYP1B1* (NM_000104.3), asociada probablemente al glaucoma congénito presente en su madre.

La variante c.357+6T>G se encuentra fuera de los nucleótidos canónicos del donador, sin embargo, los predictores señalan una posible modificación del *splicing* debido a un doble efecto. Esta variante debilitaría ligeramente al sitio donador canónico c.357+1 (*score* HSF de 83,2 a 81,2). Además, el cambio T>G conduce a la creación de un nuevo GT en posición c.357+6 (Figura R21.B), con un potencial espliceogénico similar al del donador.

A pesar de estas predicciones, en el estudio funcional mediante minigenes se identificó un patrón complejo de *splicing* formado por más de 5 isoformas diferentes, observándose un claro predominio de dos de ellas (Figura R21.C). La isoforma $\Delta 6_{108}$ representa la mitad de los transcritos totales y la $\Delta 6_{181}$ un tercio de los mismos. Se observa además el silenciamiento casi total del donador canónico, confirmado por la presencia de la isoforma CT en tan solo un 3%. Por tanto, se deduce que esta variante conduce indirectamente a la activación en paralelo de dos donadores exónicos distantes, situados en las posiciones c.250 y c.177, respectivamente.

Este ensayo *in vitro* también confirma la creación de un nuevo donador intrónico, c.357+6, gracias a la detección de una isoforma elongada I6_5, que representa el 10% del total de transcritos.

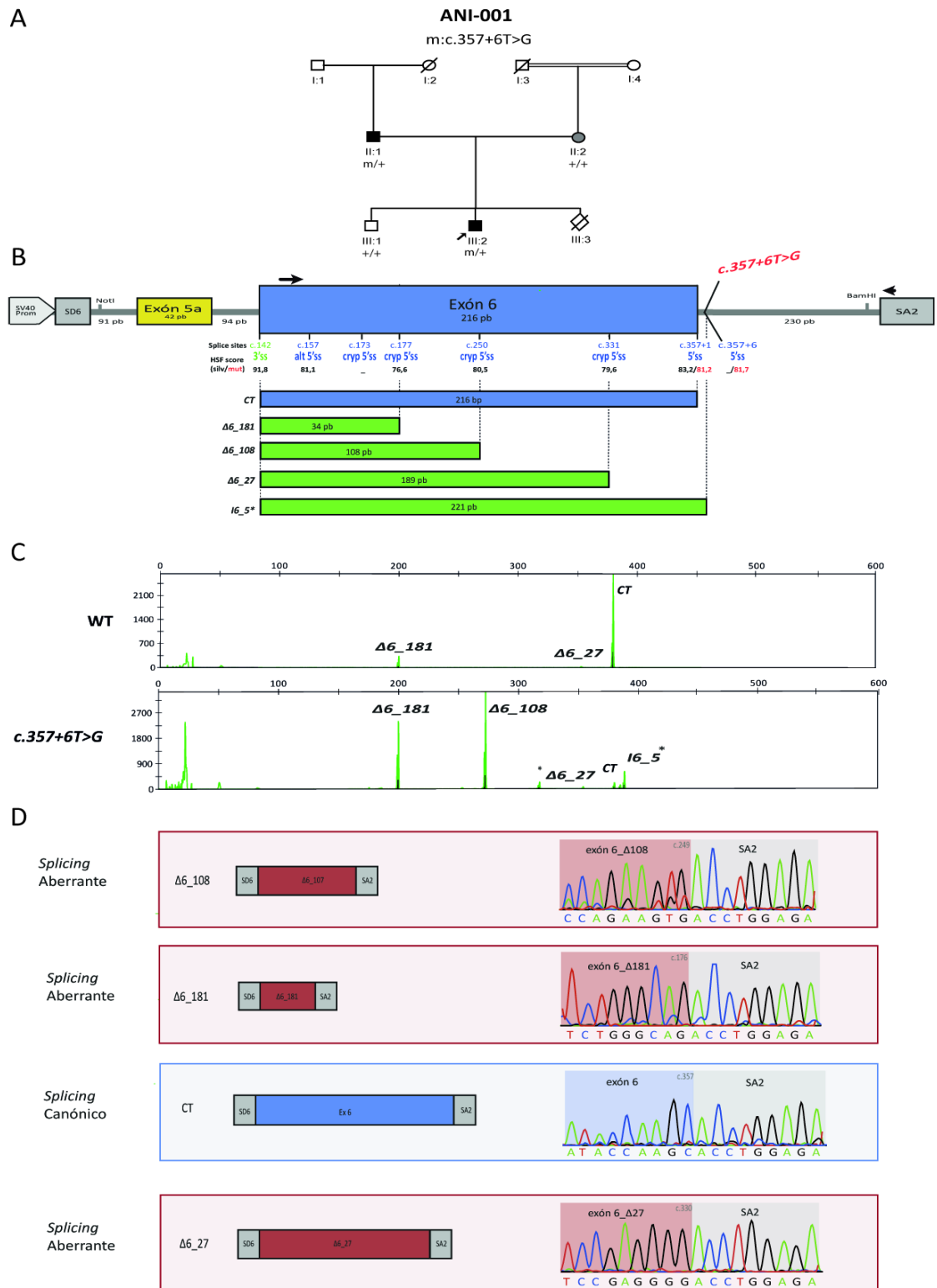


Figura R21. Estudio funcional de la variante c.357+6T>G. A) Análisis de segregación en la familia ANI-001. El individuo silvestre II:2, en gris, presenta glaucoma congénito. **B)** Representación esquemática de la predicción de splicing del exón 6 para la variante y de las isoformas obtenidas. **C)** Resultados del análisis semi-cuantitativo mediante RT-PCR mediante electroforesis capilar realizado con los cebadores c.148 y SA2 a partir de una construcción con la variante. **D)** Secuenciación Sanger de las isoformas obtenidas ordenadas según frecuencia. Δ6_27, Δ6_108 y Δ6_181: delección de 27, 108 y 181 pb, respectivamente; CT: transcrito canónico; I6_5: inserción de 5 pb intrónicas; *: isoformas no secuenciadas.

4.2.3. Una pequeña duplicación conduce a un defecto de *splicing* mediado por la activación de un ESS

La duplicación *in-frame* c.326_331dupAGGGGG se identificó en un caso esporádico de aniridia congénita. Su posterior segregación familiar confirmó que se trata de una variante de aparición *de novo* (familia ANI-053, Figura R22.A). Los predictores de *splicing* no refieren que altere el potencial espliceogénico del donador críptico anteriormente identificado en posición c.331 (Figura R22.B), aunque sí se detecta una potencial alteración a nivel de elementos reguladores del *splicing* (SRE) al crearse un nuevo sitio de unión a hnRNPA1, un elemento silenciador del *splicing* (ESS) (Tabla R6).

En el ensayo funcional mediante minigenes, la duplicación c.326_331dup mostró un resultado muy similar al producido por la variante c.357+6T>G. Se observó un marcado efecto deletéreo al suprimirse por completo la expresión de la isoforma canónica (Figura R22.C). Ligadas a esta variante se identificaron varias isoformas alternativas, incluyendo de forma mayoritaria la $\Delta 6_{108}$, que 2/3 del total de transcritos, la retención de 5 pb del intrón (I6_5) en un 18%, además de otras formas más minoritarias. La isoforma $\Delta 6_{181}$ se observó en la misma proporción que la obtenida en el constructo silvestre (Anexo 6 y Figuras R22 y R17); por tanto, la aparición de esta isoforma no se considera que sea una consecuencia de la duplicación.

También debemos destacar que el donador en c.331, cuya secuencia está inmersa en la duplicación, no fue utilizado de forma alternativa en este ensayo de minigenes. Al contrario de lo que se observó para la mayoría de las variantes estudiadas, incluida la secuencia silvestre, en este caso no detectamos el transcrito $\Delta 6_{27}$ derivado del uso aberrante de este donador críptico. Esto podría indicar que este donador se disrumpe con la modificación de las secuencias adyacentes.

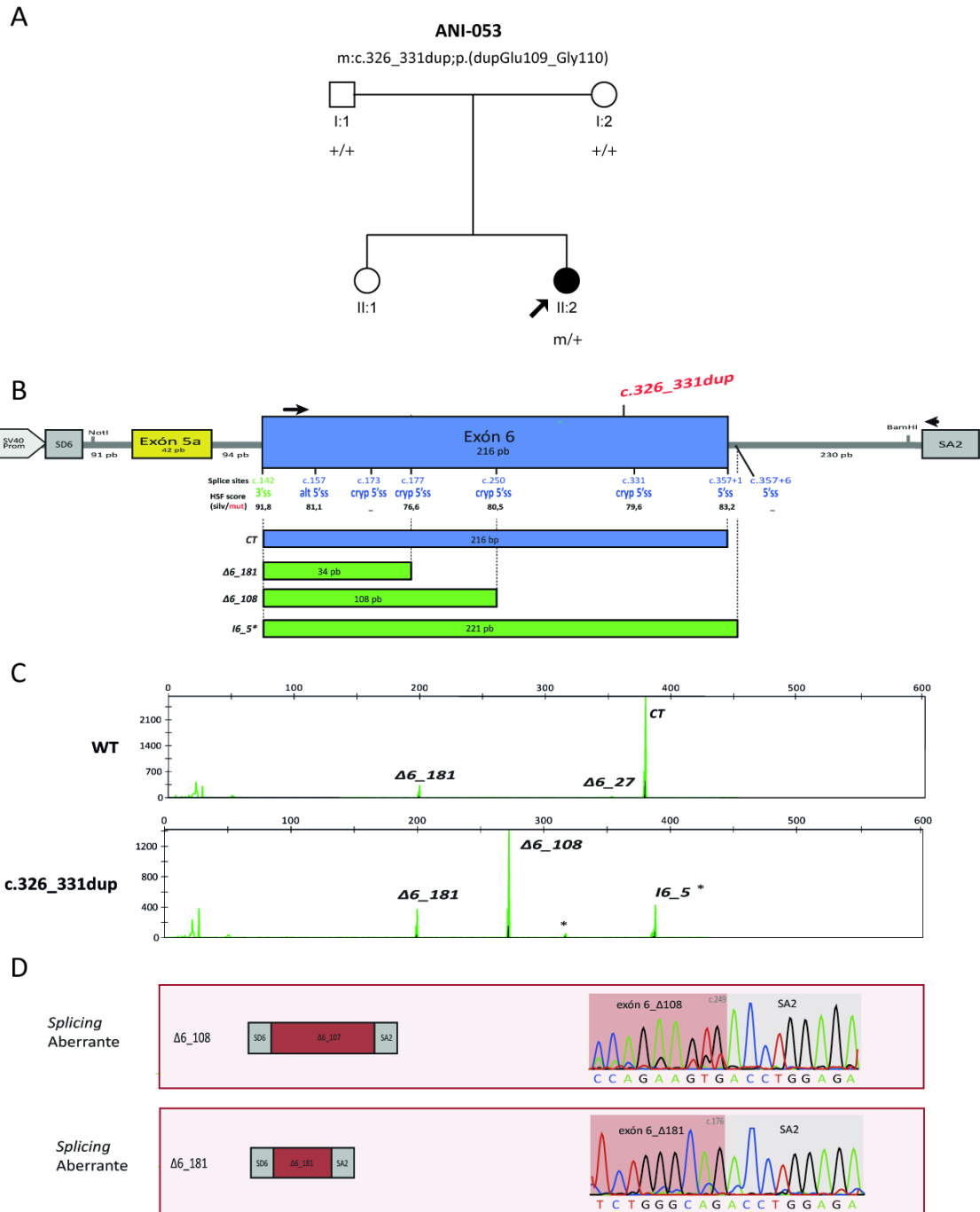


Figura R22. Estudio funcional de la variante c.326_331dup. **A)** Análisis de segregación en la familia ANI-053. **B)** Representación esquemática de la predicción de splicing del exón 6 para la variante, así como de las isoformas obtenidas in vitro. **C)** Resultados del análisis semi-cuantitativo de RT-PCR mediante electroforesis capilar realizado con los cebadores c.148 y SA2 a partir de una construcción con la variante. **D)** Secuenciación Sanger de las isoformas obtenidas ordenadas según frecuencia. Δ6_108 y Δ6_181: delección de 108 y 181 pb, respectivamente; CT: transcrito canónico; *: isoformas no secuenciadas.

4.2.4. La posible alteración proteica producida por dos variantes *missense*

En este estudio sobre el *splicing* del exón 6 de *PAX6*, se han estudiado también dos variantes de tipo *missense*, c.178T>C y c.356G>T, ambas localizadas en dos sitios de *splicing*, el donador críptico c.177 y el donador canónico c.357+1, respectivamente (Figura R16).

La variante *missense* **c.178T>C;p.(Tyr60His)**, fue identificada en un caso índice con corectopia y microftalmia, gracias al cribado del panel de 121 genes asociados a distintas formas de MOC. Este cambio está altamente conservado en *PAX6*, no está descrito previamente en bases de datos poblacionales y segrega en su hermano y en su hijo, ambos afectados (ANF-060; Figura R23.A). Tal como refleja el árbol genealógico, esta familia representa un tercer caso de recurrencia de hermanos afectados descendientes de padres asintomáticos en los que sospechamos de mosaicismo germinal para esta variante en *PAX6*. Esta variante se localiza contigua a dos de los donadores alternativos identificados durante este trabajo, c.173 y c.177 (Figura R23), por lo tanto afectaría a una región muy enriquecida en sitios de unión a elementos SREs. Aunque los predictores *in silico* indican que este cambio anularía al donador alternativo más próximo, c.177, se predice la alteración de un sitio de unión a un elemento potenciador del *splicing* tipo SF2 (Tabla R6).

La segunda variante *missense* estudiada, **c.356G>T;p.(Ser119Ile)**, se identificó en el caso índice de una familia con un fenotipo complejo de microftalmia con múltiples malformaciones del segmento anterior y posterior (Tabla R7). A pesar de existir una aparente transmisión autosómica dominante (ANF-037, Figura R23.A), esta herencia no pudo ser confirmada por falta de ADN. Esta variante, localizada en un nucleótido no canónico del donador, no modifica su potencial espliceogénico. Sin embargo, altera sitios de unión a potenciadores del *splicing* de tipo SRp40/SRp55 (Tabla R6).

Aunque en principio no se predice una alteración del sitio donador, ambas variantes tienen potencial de modificar sitios de unión a elementos reguladores de *splicing*. Por ello, se valoró su posible implicación en la alteración del *splicing* mediante el análisis de minigenes. En ambas variantes, el patrón de *splicing* del minigen no se ve

modificado con respecto al constructo silvestre, obteniéndose un perfil de splicing similar a los observados en la construcción silvestre (Figura R23).

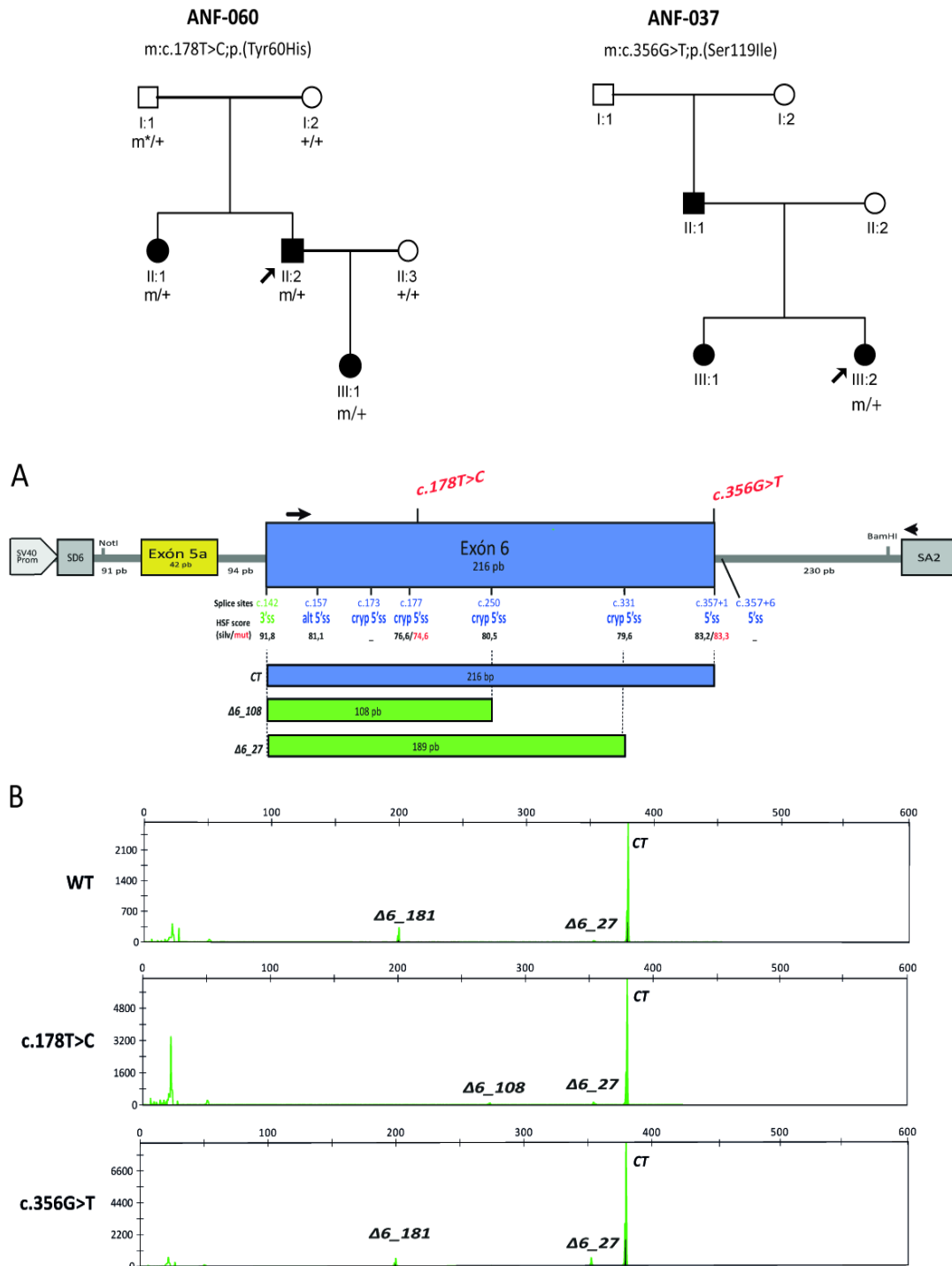


Figura R23. Estudio funcional de las variantes c.178T>C y c.356G>T. A) Análisis de segregación de las variantes c.178T>C y c.356G>T en las familias ANF-060 y ANF-037, respectivamente. **B)** Representación esquemática de la predicción de splicing del exón 6 para la variante, así como de las isoformas obtenidas *in vitro*. **C)** Resultados del análisis semi-cuantitativo de RT-PCR mediante electroforesis capilar con los cebadores c.148 y SA2 a partir de una construcción con las variantes. **D)** Secuenciación Sanger de las isoformas obtenidas ordenadas según frecuencia. Δ6_27 y Δ6_108: delección de 27 y 108 pb, respectivamente; CT: transcrito canónico.

A pesar de que no hemos podido confirmar la hipótesis de una posible alteración del *splicing* del exón 6, ambos cambios *missense* se predicen como deletéreos a nivel proteico en varios predictores (Tabla R5). Estas variantes conducen a la sustitución de residuos altamente conservados en el dominio pareado (PD) de la proteína PAX6, lo cual podría alterar su estructura 3D, tal como se observa en la Figura R24 donde se muestra la modelización de la consecuencia de ambos cambios aminoacídicos.

La variante c.178T>C provoca el reemplazo de un residuo polar Tirosina en la posición 60 por un aminoácido básico y menos voluminoso de Histidina, lo cual cambiaría las propiedades y la superficie de la proteína (Figura R24.D-E). En esta representación 3D, se aprecia también que el residuo 60 silvestre interacciona con un conjunto de residuos de gran importancia funcional (Figura R24.D). Entre ellos se produce la interacción con el residuo Gly64, cuya alteración se ha descrito como uno de los cambios más deletéreos en la proteína PAX6 (Udhaya, 2015).

La variante c.356G>T conduce a la sustitución de un aminoácido polar, Serina, en posición 119 por Isoleucina, residuo más voluminoso e hidrofóbico (Figura R24.B-C). La representación de la sustitución *in silico* muestra que esta posición se encuentra en contacto directo con el ADN y que el cambio disrumpe 4 de las 7 interacciones que el residuo silvestre establece, dos con la molécula de ADN y otras dos con el residuo Serina122.

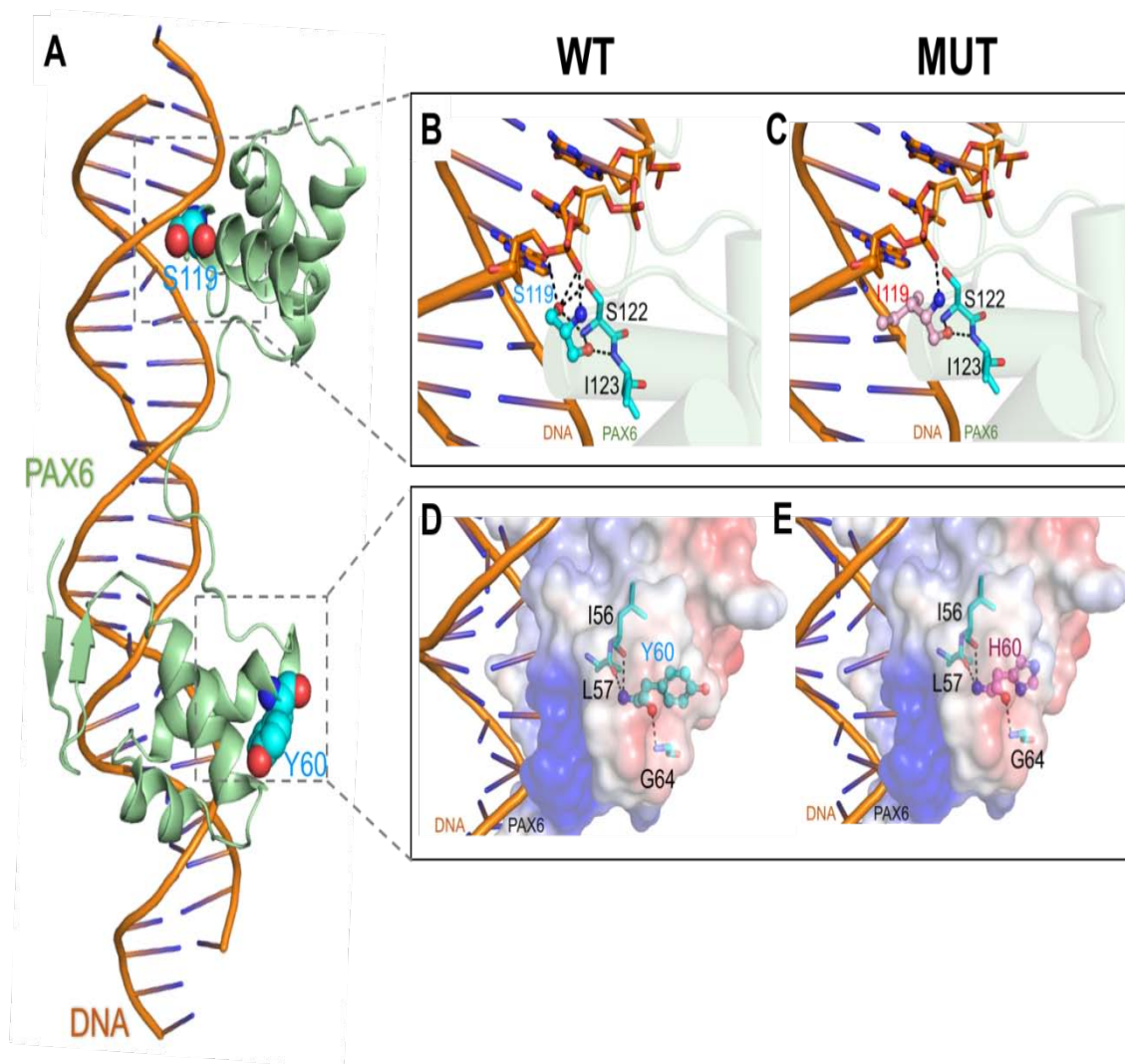


Figura R24. Modelización de las variantes missense Tyr60His y Ser119Ile en la estructura 3D de la proteína PAX6 cristalizada unida a una molécula de ADN. A) Visión del conjunto del dominio pareado (DP) de PAX6 unido al ADN, mostrando en verde PAX6 y en naranja el ADN. Los recuadros B-E muestran una ampliación de las regiones donde se localizan ambas variantes. B-C) Se muestra en posición 119 el residuo Serina silvestre (azul cian) e Isoleucina mutado (rosa claro), respectivamente. Se observa la pérdida de 4 de las 7 interacciones del residuo 119 por la presencia de Isoleucina, dos con el ADN y dos con el residuo S122. La mutación colocaría un grupo hidrofóbico próximo a la molécula de ADN y crearía choques estéricos debido a que el residuo Isoleucina es más voluminoso que la Serina. D-E) Se muestra en posición 60 el residuo Tirosina silvestre (azul cian) y el Histidiana mutado (rosa claro), respectivamente. El cambio hace que la polaridad de la superficie sea más positiva. El residuo mutado presenta un anillo más pequeño, lo cual deforma la proteína. Se observa la interacción entre el residuo 60 y el G64. Se muestra la superficie de la proteína con la carga electrostática (azul, positiva; roja, negativa). Los residuos de interés se muestran con barras azul cian. Los potenciales puentes de hidrógeno potenciales se representan con líneas discontinuas. Los átomos de oxígeno y de nitrógeno se muestran en rojo y azul oscuro, respectivamente.

4.2.5. Clasificación patogénica de variantes en función al defecto de *splicing* y posibles correlaciones fenotípicas

A la vista de los resultados obtenidos en los estudios de valoración del *splicing*, se ha establecido una clasificación de severidad de las variantes atendiendo a la consecuencia proteica de las distintas isoformas aberrantes obtenidas (Anexo 6). Las 5 variantes en las que se pudo confirmar su implicación en el *splicing* presentan un rango variable de transcritos de *PAX6* con una ratio variable de isoformas canónicas *versus* aberrantes.

Se han consideraron como las variantes más severas aquellas que pierden por completo la expresión de la isoforma canónica, lo que incluye las variantes c.174C>T y c.326_331dup. La variante c.174C>T, pese a ser un cambio sinónimo, ha sido clasificada como la más severa de todas las variantes en estudio, debido a que solo se obtiene un transcrito aberrante que causaría una proteína truncada p.(Arg59Valfs*12). Se predice que esta isoforma truncada no llegaría a formarse pues el pre-ARNm sería degradado por el sistema NMD, lo cual conduciría a la haploinsuficiencia de *PAX6*.

Las variantes c.326_331dup y c.357+6T>G producen una mezcla de transcritos aberrantes que darían lugar tanto a delecciones *frameshift* como *in-frame*, siendo consideradas como moderadamente severas. Por último, las variantes c.249T>G y c.333C>A fueron predichas como menos severas dado que podrían producir niveles residuales de la isoforma correcta, además de que conducirían principalmente a isoformas con delecciones *in-frame* del exón 6 y por tanto, a la delección parcial del dominio PD.

A pesar de este intento de clasificación en función de las posibles consecuencias moleculares, no parece existir ninguna correlación genotipo-fenotipo evidente. Todos los portadores heterocigotos de estas 5 variantes sufren nistagmo, baja agudeza visual, cataratas de aparición precoz y opacidad corneal como consecuencia de la deficiencia de las células limbares (Tabla R7). El fenotipo del iris varía ampliamente entre los individuos estudiados, presentando anomalías de distinta severidad, incluyendo aniridia completa, hipoplasia, coloboma o atrofia iridiana. Esta variabilidad no parece correlacionar con el tipo de consecuencia asociada al defecto en *PAX6*, pues no solo ocurre entre portadores de distintas variantes sino también entre individuos de la misma familia e incluso entre ambos ojos de un mismo individuo.

En la Figura R25.A-J, se muestra el ejemplo de una portadora en heterocigosis de la variante presuntamente más severa c.174C>T. La paciente presenta una marcada asimetría entre ambos iris con una hipoplasia prácticamente total en el ojo derecho, y un iris rudimentario con policoria en el izquierdo. Una presentación similar de aniridia parcial y coloboma de iris se observa en el portador III:3 de la variante c.249T>G de la familia ANI-048. Hay que destacar que estas diferencias en la expresividad de esta variante parecen aún más marcadas en su padre (II:2), el cual presenta un iris completo con únicamente pequeños defectos estructurales y corectopia (Tabla R7). Ambos presentan el resto de características asociadas a la aniridia, incluyendo queratopatía y baja visión. Ésta última en el caso de la hija se debe a una hipoplasia foveal, cataratas y glaucoma.

Esta expresividad fenotípica intra-familiar también está presente en dos individuos heterocigotos para la variante c.357+6T>G de la familia ANI-001, que manifiestan diferencias en defectos iridianos y corneales (Figura R25.I-L). Por un lado, el caso índice presenta aniridia clásica acompañada de otras malformaciones esqueléticas congénitas (Tabla R7 y Figura R25.I-J), mientras que su padre presenta atrofia y coloboma iridiano (Tabla R7 y Figura R25.K-L). Este último paciente fue diagnosticado clínicamente de anomalía bilateral de Peters debido a la presencia congénita de leucoma central en ambas córneas.

Resultados

Familia Paciente	Variante	Última revisión (años)	Defecto iris	Hipoplasia foveal	Cataratas precoces	Queratopatía	Otras características oculares	Características extraoculares
ANI-001 II:1	c.357+6T>G	38	Atrofia iridiana y sinequias iridocorneales (B), coloboma inferior (OI)	ND	(+) B	(+) B Leucoma central congénito		
ANI-001 III:2	c.357+6T>G	14	Aniridia completa	ND	(+) B	(+) B	Miopía magna y estrabismo	Anomalías congénitas esqueléticas de la extremidad inferior derecha y epilepsia.
ANI-048 II:2	c.249T>G	68	Iris completo con estructura anormal y corectopia (B)	ND	(+) B	(+) B	Endotropía (OD) y glaucoma	Diabetes (38 años)
ANI-048 III:3	c.249T>G	28	Coloboma nasal (OD) Aniridia parcial (OI)	(+) B	(+) OI	(+) B	Miopía magna y endotropía	Obesidad
ANI-053 II:2	c.326_331dup	1	Aniridia completa	(+) B	(-)	(-)		
ANI-62 III:3	c.174C>T	23	Aniridia parcial (B; más severa en OD)	(+) B	(+) B	(+) B	Hipoplasia nervio óptico (B) y miopía magna	
ANF-037 III:2	c.356G>T p.(Ser119Ile)	10	Corectopia	(+) B	(+)	(-)	Microftalmia (B), embriotoxon posterior, membrana pupilar, corectopia, papila pálida y miopía magna.	Ligero retraso mental, convulsiones, sobrepeso e hiperinsulinemia.
ANF-060 II:1	c.178T>C p.(Tyr60His)	43	Corectopia*	(-)	(+)	(-)	Microftalmia (B) e hipermetropía elevada	

Table R7. Características clínicas de los individuos portadores de las variantes estudiadas en el exón 6. B: bilateral; OI: ojo izquierdo; OD: ojo derecho; ND: no disponible.*: corectopia en los 3 afectados de la familia ANF-060, supuestamente secundaria a iridectomía. Todos los pacientes presentan nistagmo.

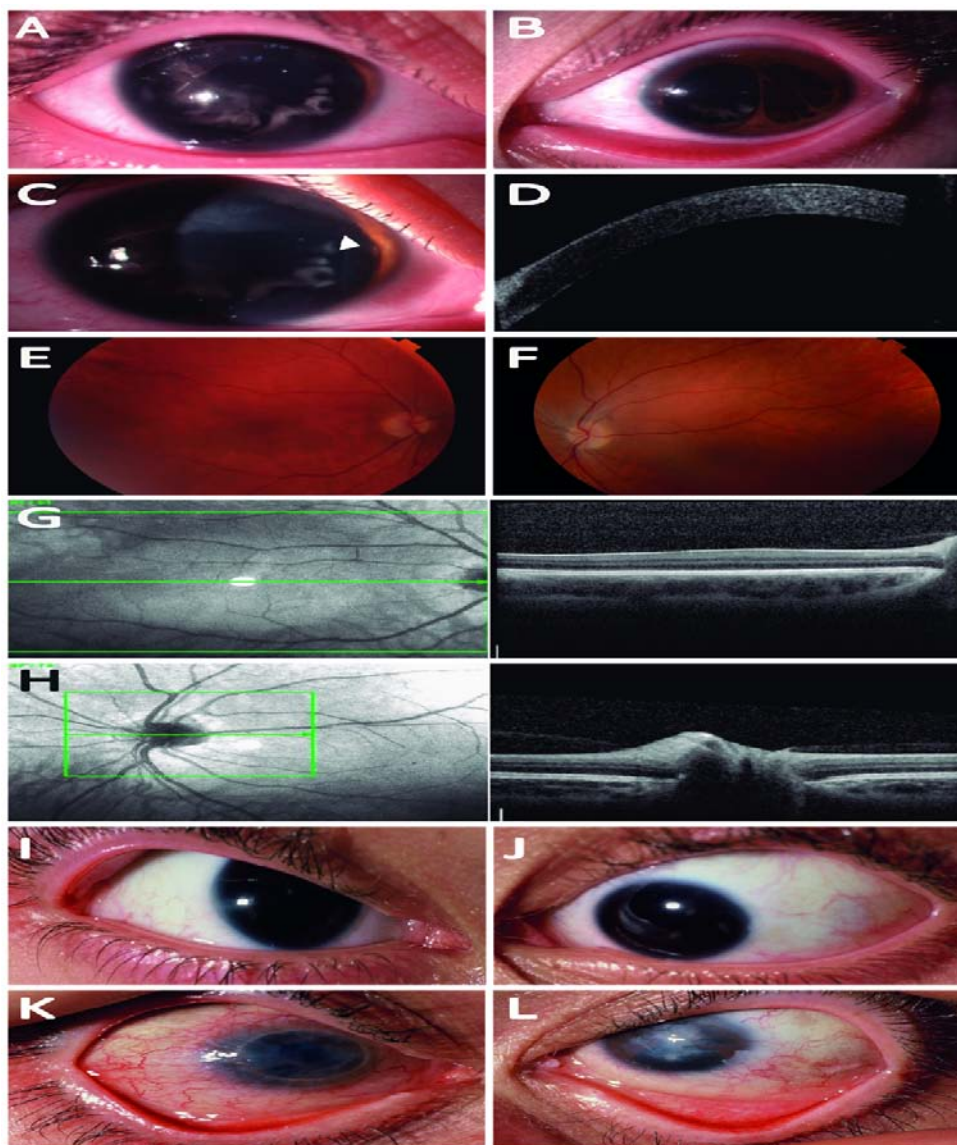


Figura R25. Hallazgos oftalmológicos en individuos afectados de las familias ANI-062 y ANI-001. A-H) Datos clínicos del individuo III:3 de la familia ANI-0062 portador de la variante c.174C>T a los 23 años. **A y C)** Fotografías del ojo derecho muestran un iris nasal remanente, catarata cortical y deficiencia de células limbares con fibrosis subepitelial y pannus. **B)** Fotografía del ojo izquierdo muestra un iris rudimentario, catarata polo anterior y deficiencia de células limbares con ligera opacidad corneal periférica. **D)** Imagen de tomografía de coherencia óptica (OCT) del segmento anterior del ojo derecho muestra hiperreflectividad epitelial, correspondiéndose con un área de fibrosis corneal. **E-F)** Fotografías del fondo de ojo muestran discos ópticos de pequeño tamaño, fundus teselado y reflejo foveal anormal en ambos ojos. **G)** Imagen de una sección horizontal de OCT de la macula derecha muestra una hipoplasia foveal de grado IV, caracterizada por ausencia de depresión foveal, extrusión de las capas internas de la retina, alargamiento del segmento exterior y ensanchamiento de la capa nuclear externa. **H)** OCT del nervio óptico del ojo izquierdo muestra un disco óptico de pequeño tamaño. **I-L)** Datos clínicos de los pacientes II:1 y III:2 de la familia ANI-001, portadores de la variante c.357+6 T>G. **I-J)** Fotografías de ambos ojos del individuo III:2 a los 14 años muestra ausencia de iris, deficiencia leve de células limbares con ligera opacidad corneal periférica. **K)** Fotografía del ojo derecho del individuo II:1 a los 38 años muestra un injerto corneal como resultado de una queratoplastia penetrante realizada a los 36 años. **L)** Fotografía del ojo izquierdo del paciente II:1 muestra un leucoma central muy vascularizado, coloboma iridiano inferior, polioria, opacidad corneal central y periférica.

4.3. Estudio de la implicación del mosaicismo parental en *PAX6*

A raíz de las sospechas de mosaicismo parental en tres de las familias portadoras de VUS en el exón 6 de *PAX6* (ANI-062, ANI-048 y ANF-060), decidimos profundizar en las implicaciones de posibles variantes post-cigóticas en la etiopatogenia de la aniridia o de patologías relacionadas.

Primero, realizamos una búsqueda de otras familias portadoras de mutaciones en *PAX6* en nuestra cohorte en las que hubiera una posible sospecha de mosaicismo. Los criterios de inclusión en este estudio fueron: i) familias formadas por una pareja asintomática con más de un hijo afecto y ii) familias en donde existiera una notable variabilidad fenotípica entre un primer caso esporádico y sus descendientes afectados, teniendo una presentación más severa en la segunda generación.

Según estos criterios seleccionamos cuatro familias portadoras de variantes en *PAX6*, las tres anteriormente descritas con recurrencia de aniridia en hijos de padres asintomáticos y una adicional (ANI-100) con variabilidad fenotípica. Desafortunadamente, en la familia ANI-048 no pudimos realizar el estudio de mosaicismo debido al fallecimiento del padre y a la falta de colaboración de la madre con problemas de salud asociados a su elevada edad. En las tres familias en las que disponíamos de muestras de los progenitores con sospecha de mosaicismo, se comprobó la paternidad mediante análisis de haplotipos utilizando microsatélites de tipo STRs. Por lo tanto, descartamos la existencia de una falsa paternidad en estas familias.

Gracias a una colaboración con la Dra. Villamar del Hospital Universitario Ramón y Cajal, en este trabajo se incluyó una familia adicional con padre e hija diagnosticados de aniridia clásica, en la que un estudio preliminar reveló un posible alelo en mosaicismo en el progenitor para una variante patogénica en *PAX6* identificada en su hija.

En estas cuatro familias, la segregación de la variante en *PAX6* reveló en los electroferogramas la presencia de trazas del alelo mutante en el ADN de sangre periférica en uno de los padres, lo que podría indicar la presencia de un mosaico (Figuras R26-29). En todos los casos, se observó un pico para el nucleótido mutado de un tamaño inferior al observado en el hijo heterocigoto. Este desbalance en los picos

electroforéticos fue también detectado posteriormente en un ensayo de SNaPshot realizado por la Dra. Villamar en 3 de los 4 individuos potencialmente portadores de mosaicismo. Este análisis de SNaPshot no fue realizado en la familia ANI-062.

Con la finalidad de confirmar el mosaicismo, se utilizó una estrategia de genotipado mediante sondas Taqman utilizando la técnica de PCR digital con el fin de cuantificar de forma precisa la frecuencia alélica del alelo mutado vs el silvestre. Esta aproximación nos permitió confirmar el desbalance alélico observado previamente mediante técnicas cualitativas (Anexo 7 y Tabla R8). Aunque el estudio se realizó utilizando ADN de sangre, en algunos casos se pudo obtener otras muestras somáticas con distinto origen embrionario, incluyendo orina y saliva, así como muestra de semen del progenitor de la familia ANI-100.

En todas estas familias se pudo confirmar la presencia en uno de los progenitores de un alelo somático de baja frecuencia, entre 11% y 29%. La presencia de este alelo mosaico en tejidos somáticos así como su transmisión vertical a la descendencia en estas familias nos permiten concluir que se trata de casos de mosaicismo gonosomal.

A continuación, se procede a describir con más detalle los hallazgos clínico-genéticos de estas cuatro familias con confirmación de mosaicismo gonosomal, lo que explica tanto el origen como la transmisión de la enfermedad. Los datos clínicos de estas 4 familias se resumen en la Tabla R8.

Familia	Variante causante	Paciente	Edad	F.A.	Anomalia iris	Nistagmo	Cataratas precoces	Glaucoma	Queratopatía	Otros signos clínicos
ANI-062	c.174C>T p.(Gly58=)	I:1 (madre)	90	Mosaico 11% sg 17% sv	No	No	No	No	No	Fotofobia.
		II:3	51	HET	Aniridia parcial	N.D.	N.D.	N.D.	N.D.	
		III:3	24	HET	Aniridia parcial (B, > OD)	+ (B)	+ (B)	+ (B)	Insuficiencia limbal (B, > OD)	Hipoplasia del nervio óptico y alta miopía.
ANI-100	c.120C>A p.(Cys40*)	I:1 (padre)	59	Mosaico 13% sg/sv/o 29% sm	Coloboma (B)		+ (B)			
		II:2	25	HET	Aniridia completa (B)	+ (B)	+ (B)	+ (B)	Opacidad corneal con insuficiencia limbal (B)	Transplante limbal y glaucoma. Implante de válvula Ahmed (OD).
ANF-060	c.178T>C p.(Tyr60His)	I:1 (padre)	80	Mosaico 18% sg 29% sv						Asintomático, sin anomalías oculares.
		II:1	43	HET	Corectopia secundaria*	+ (B)	+ (B)			Microftalmia (B) e hipermetropía elevada.
		II:2	38	HET	Corectopia secundaria *	+ (B)	+ (B)			Microftalmia (B) e hipermetropía elevada.
		III:1	10	HET	Corectopia secundaria *	+ (B)	+ (B)			Microftalmia (B) e hipermetropía elevada.
ANI-RyC	c.771G>A p.(Trp257*)	I:1 (padre)	50	Mosaico (18% sg)	Aniridia completa (B)	+ (B)	+ (B)	+ (B)		Cirugía de cataratas.
		II:1	29	HET	Aniridia completa (B)	+ (B)	+ (B)	+ (B)	Insuficiencia limbal (B)	Retraso crecimiento; cirugía de cataratas y Amaurosis (OD).

Tabla R8. Características clínicas y genéticas de varios individuos pertenecientes a familias con sospecha de mosaicismo en variantes en PAX6. F.A.: frecuencia alélica; B: bilateral; OD: ojo derecho; sg: sangre; sv: saliva; sm: semen; or: orina; HET: heterocigoto. *: corectopia secundaria a iridectomía.

4.3.1. Descripción clínico-genética de las familias con mosaicismo en *PAX6*

4.3.1.1. Familia ANI-062

La familia ANI-062, tal como se ha descrito anteriormente, está formada por 3 generaciones, dos de ellas con al menos un individuo afecto. La aniridia congénita aparece por primera vez en la segunda generación con recurrencia en dos hermanas afectas descendientes de una pareja sin aparente afectación ocular. La transmisión de la aniridia por parte del caso índice (II:4) a una de sus hijas indica un posible patrón de herencia autosómico dominante (Figura R26.A), que fue confirmado gracias al hallazgo de la variante sinónima c.174C>T;p.(Gly58=) en el exón 6 de *PAX6*. Esta variante fue clasificada como probablemente patogénica gracias al estudio funcional descrito en el anterior apartado de esta Tesis Doctoral (Figura R18).

Durante el estudio de segregación de esta variante en la familia mediante secuenciación Sanger, se identificó una traza del alelo mutante en la madre, asintomática (I:2, Figura R26.B). Con el fin de confirmar la sospecha de mosaicismo y descartar un posible artefacto o contaminación, se procedió a realizar una cuantificación de la fracción alélica mediante ddPCR. Se pudo confirmar que la presencia del alelo mutado en saliva fue alrededor de un 6% superior que en sangre, 17% y 11%, respectivamente (Figura R26.C).

Este individuo mosaico, no ha presentado a lo largo de su vida ninguna de las complicaciones asociadas a la aniridia, como queratopatía, cataratas o glaucoma, que sí están presentes en sus familiares heterocigotos (Tabla R8). Esto fue confirmado en una reciente evaluación oftalmológica, realizada ya en su novena década de la vida, en donde solo refirió una leve fotofobia sin que se observasen ninguno de los signos malformativos característicos de la aniridia, como la hipoplasia iridiana en el segmento anterior, y la hipoplasia del nervio óptico y de la fovea en el segmento posterior, que están presentes en sus familiares afectados, tal como hemos descrito en el anterior apartado en su nieta (Figura R25).

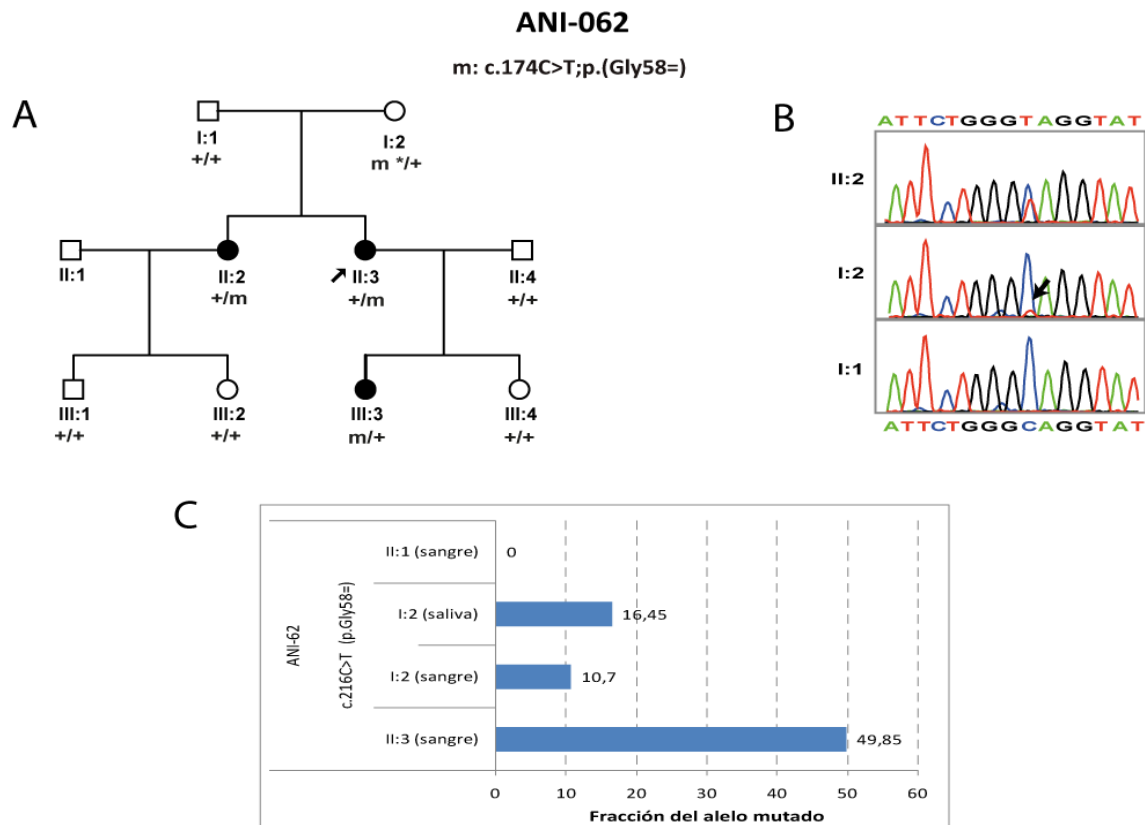


Figura R26. Identificación de mosaicismo en la familia ANI-062. A) Segregación de la variante c.174C>T de PAX6 en la familia. +: alelo silvestre; m: alelo mutado; m*: alelo mosaico. **B)** Electroferograma del resultado de la secuenciación Sanger como indicio de la sospecha de mosaicismo. La flecha descendente: indica la traza del alelo mutado en el individuo sospechoso de mosaicismo I:2. **C)** Cuantificación de la fracción del alelo mutado obtenida mediante ddPCR en muestras de sangre y saliva de varios miembros de la familia.

4.3.1.2. Familia ANI-100

En la familia ANI-100, existe una gran variabilidad fenotípica entre el *probandus* y su padre (Figura R27.A). El individuo II:2 tiene una hipoplasia completa de iris acompañada de opacidad corneal debida a una insuficiencia limbar, cataratas de debut temprano y glaucoma. A consecuencia de esto, el *probandus* sufre baja visión y nistagmo. Sin embargo, su padre (I:1), asintomático durante sus primeras décadas de vida, fue diagnosticado de coloboma de iris y cataratas tras el nacimiento de su hijo con aniridia congénita (Tabla R8).

El caso índice se estudió con un panel de NGS de genes relacionados con MOC, identificándose la variante c.120C>A;p.(Cys40*) en heterocigosis localizada en el exón 5 de PAX6, previamente identificada como mutación *hotspot* (Vincent, 2003). La posterior segregación mediante la secuenciación Sanger mostró trazas del alelo

mutante en el padre con desbalance en el pico electroforético con respecto al alelo silvestre (Figura R27.B). Este hallazgo en el padre junto con el fenotipo mucho más leve que el de su hijo, heterocigoto confirmado, hizo que nos planteásemos la posibilidad de que se tratase de un individuo mosaico.

El mosaicismo parental fue confirmado mediante ddPCR, observándose diferencias en la fracción alélica de la mutación en tejidos somáticos con respecto a la línea germinal. Esta pudo ser estudiada gracias a la obtención de muestra de ADN a partir de semen. Tanto en sangre como en saliva y en orina, el grado de mosaicismo era de tan solo el 13%, sin embargo, la presencia de la mutación se duplica en semen (Figura R27.C). Por tanto, en esta familia se pudo establecer de manera fehaciente que el riesgo de transmisión de la enfermedad a la descendencia es de 1/3.

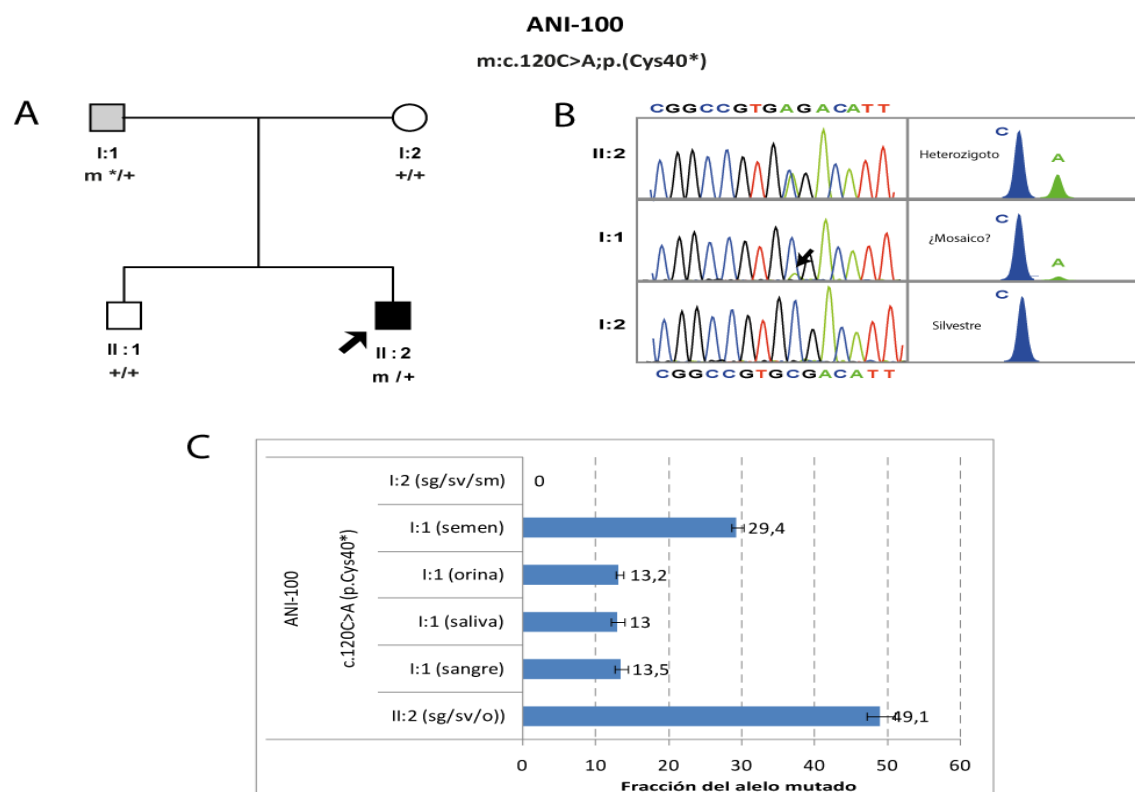


Figura R27. Identificación de mosaicismo en la familia ANI-100. A) Segregación de la variante c.120C>A de PAX6. +: alelo silvestre; m: alelo mutado; m*: alelo mosaico. **B)** Electroferograma del resultado de la secuenciación Sanger y del SNaPshot como indicios de la sospecha de mosaicismo. La flecha descendente: indica la traza del alelo mutado en el individuo sospechoso de mosaicismo I:1). **C)** Cuantificación de la fracción del alelo mutado obtenida mediante ddPCR sobre distintos tejidos de varios miembros de la familia.

4.3.1.3. Familia ANF-060

Se trata de una familia de tres generaciones, con una patología ocular que sigue un patrón de herencia autosómico dominante, en donde existe una recurrencia de dos hermanos afectados procedentes de una pareja asintomática (Figura R28.A). Todos los individuos afectados son fenotípicamente homogéneos, presentan microftalmia con cataratas congénitas y una hipermetropía elevada en ambos ojos. Como consecuencia de la baja agudeza visual, todos cursan con nistagmo. Los tres afectados sufren una corectopia, que según los informes clínicos disponibles anteriores a este estudio, no sería congénita sino secundaria a una iridectomía (Tabla R8).

En el caso índice (II:2) de esta familia, se identificó en heterocigosis una nueva variante *missense* c.178T>C;p.(Tyr60His) en el exón 6 de *PAX6*, gracias al estudio de un panel de 121 genes asociados a MOC. Tal como se ha descrito en el primer apartado de esta Tesis, esta variante se considera probablemente patogénica debido a su segregación y a la predicción del cambio aminoacídico como deletéreo (Tabla R5) y altamente dañino para la estructura 3D del dominio pareado de la proteína *PAX6*, tal como se describió en el anterior apartado (Figura R24).

Durante la segregación del resto de individuos de la familia, se identificaron trazas del alelo mutado en el padre sano (I:1), mostrando un tamaño de pico en la electroforesis que aparentemente se correspondería a un tercio del silvestre (Figura R28.B). Un posterior genotipado mediante ddPCR permitió confirmar el mosaico con una fracción del alelo mutado 1,5 veces superior en saliva con respecto a sangre, 29% frente a 18%, respectivamente (Figura R28.C).

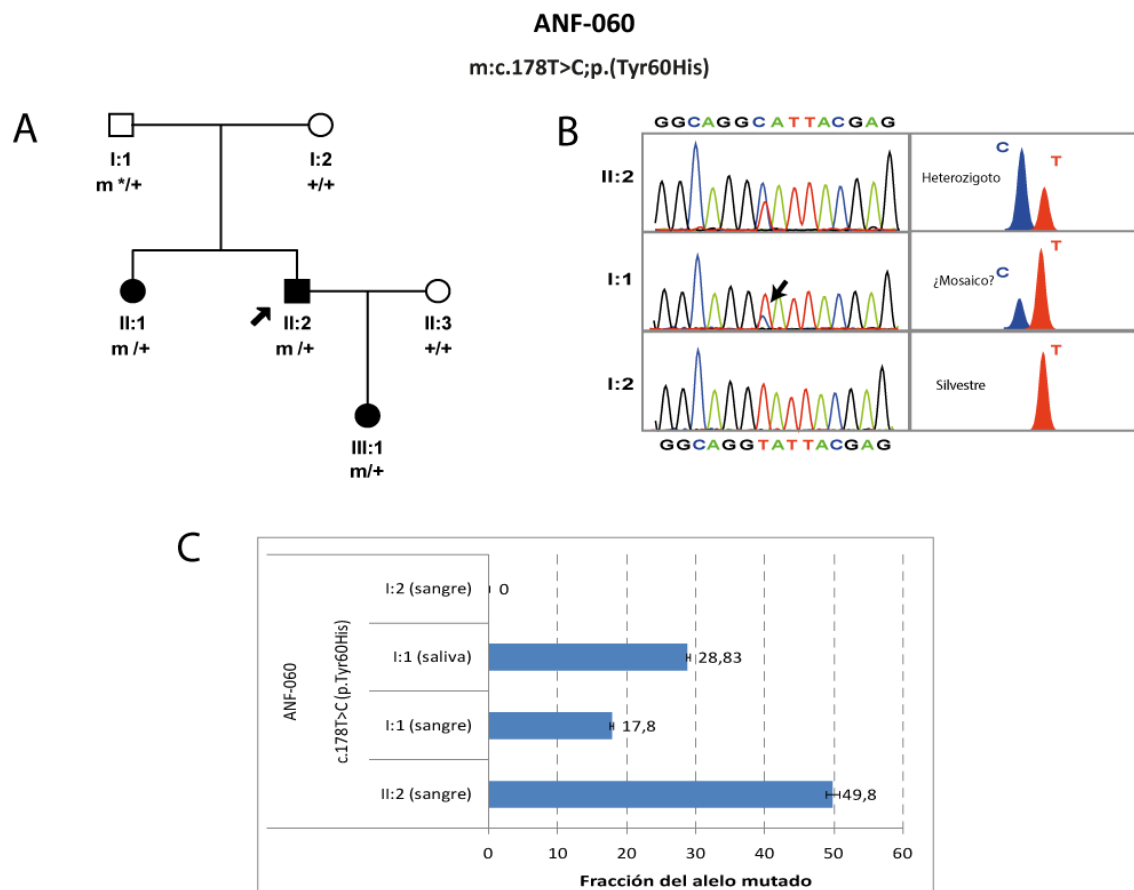


Figura R28. Identificación de mosaicismo en la familia ANF-060. A) Segregación de la variante c.178T>C de PAX6. +: alelo silvestre; m: alelo mutado; m*: alelo mosaico. **B)** Electroferograma del resultado de la secuenciación Sanger y del SNaPshot como indicios de la sospecha de mosaicismo. La flecha descendente indica la traza del alelo mutado en el individuo sospechoso de mosaicismo I:1. **C)** Cuantificación de la fracción del alelo mutado obtenida mediante ddPCR sobre distintos tejidos de varios miembros de la familia.

4.3.1.4. ANI-RyC

Esta familia, formada por un padre y una hija afectados de aniridia congénita (Figura R29.A), procede de una cohorte de pacientes con disgenesia de segmento anterior reclutados y estudiados por la Dra. Villamar en el Hospital Ramón y Cajal. En el caso índice (II:1), se había identificado una variante nula p.(Trp257*) en el exón 11 de PAX6 mediante secuenciación de NGS. Durante la segregación familiar, el padre (I:1) no mostró la variante en heterocigosis tal como se hubiese esperado de una transmisión autosómica dominante, sino que se identificó una traza del alelo mutado con pico electroforético de un tamaño similar al ruido de fondo (Figura R29.B). Un posterior ensayo de SNaPshot confirmó esta diferencia cualitativa entre el pico mutado y el silvestre en el padre en comparación con su hija heterocigota (Figura R29.B). Posteriormente, se pudo cuantificar mediante ensayo de genotipado mediante ddPCR

la presencia de este alelo mutado en casi un 18% en muestra de sangre del individuo I:1, lo cual ratificó el evento de mosaicismos detectado mediante Sanger.

Al contrario que en las tres familias anteriores, el individuo con mosaicismos presenta una sintomatología similar a la de su hija afectada de aniridia congénita. Padre e hija padecen nistagmo como consecuencia de una baja agudeza visual, además de cataratas y glaucoma. Si bien es cierto que el caso índice sufre amaurosis en su ojo derecho y tuvo un retraso leve del crecimiento en su niñez (Tabla R8). Su padre, pese a presentar tan solo un 18% del alelo mutado en sangre (Figura R29.C), sufre aniridia congénita con hipoplasia completa de iris, cataratas y glaucoma, aunque aparentemente con un desarrollo más leve que su hija. A sus 50 años de edad, no parece haber desarrollado signos de la aniridia clásica, como la queratopatía asociada a la aniridia, al contrario de lo observado en su hija en su segunda década de vida.

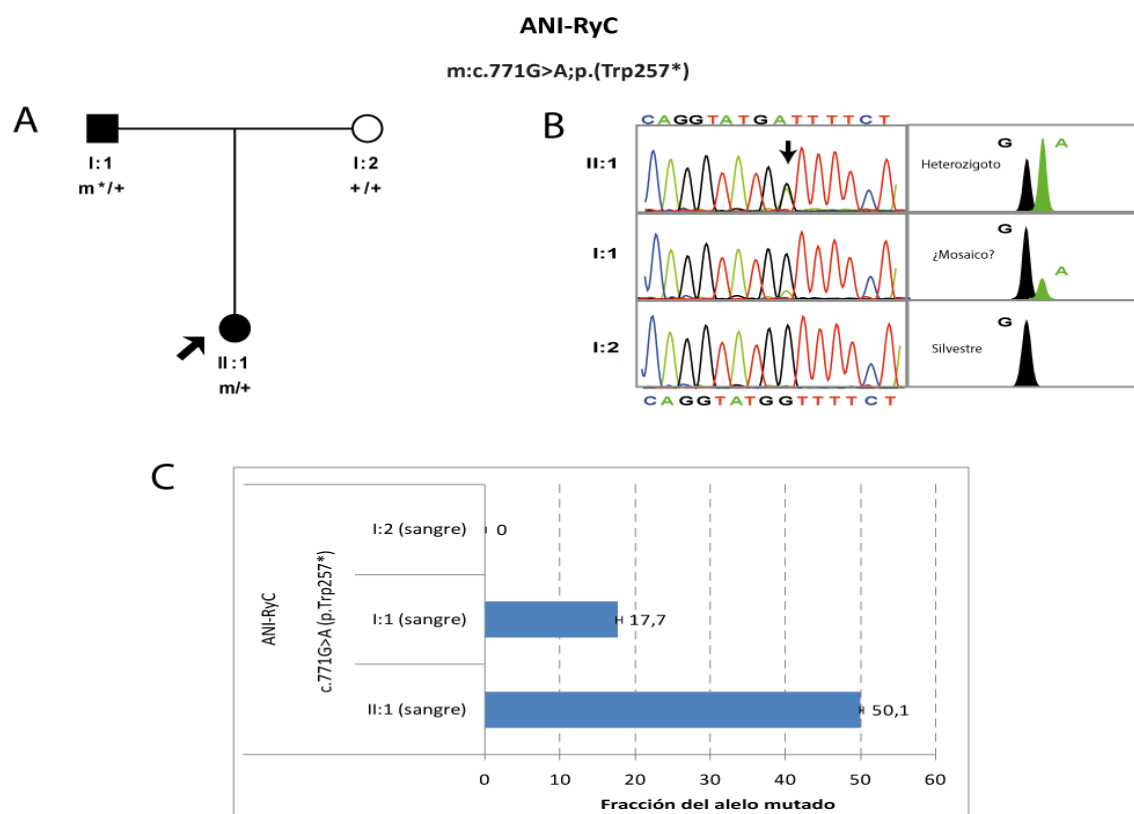


Figura R29. Identificación de mosaicismos en la familia ANI-RyC. **A)** Segregación de la variante c.771G>A de PAX6. +: alelo silvestre; m: alelo mutado; m*: alelo mosaico. **B)** Electroferograma del resultado de la secuenciación Sanger y del SNaPshot como indicios de la sospecha de mosaicismos. La flecha descendente: indica la traza del alelo mutado en el individuo sospechoso de mosaicismos I:1. **C)** Cuantificación de la fracción del alelo mutado obtenida mediante ddPCR en sangre de varios miembros de la familia.

4.4. Actualización del espectro mutacional en nuestra cohorte de aniridia

Para finalizar, queremos exponer una visión general de los resultados genéticos en nuestra cohorte de 98 pacientes con fenotipos de aniridia y aniridia-like tras los estudios realizados en esta Tesis doctoral en 33 de los 36 pacientes sin diagnóstico genético concluyente.

El cribado tanto de las regiones no codificantes de *PAX6* como de otros genes solapantes, permitió el diagnóstico de 14 de los 36 pacientes previamente considerados como “*PAX6* negativos”, gracias a la identificación de variantes patogénicas en el 5'UTR o en los *enhancers* de *PAX6*, así como en otros genes.

Otros 10 pacientes presentaban variantes de significado incierto en *PAX6*, sin que se haya identificado ninguna otra causa genética. Tras realizar un estudio funcional de la posible implicación en la alteración del *splicing* o alteraciones en la estructura 3D de la proteína hemos podido reclasificar como patogénicas 6 de las 10 VUS previamente identificadas. Esto hace que el porcentaje de pacientes de nuestra cohorte de portadores de VUS haya disminuido del 10 al 4%.

Por lo tanto, combinando los resultados obtenidos previamente en un estudio diagnóstico con los derivados de este estudio de investigación, se ha identificado variantes potencialmente patogénicas o VUS en el 88% (86/98) de las familias, siendo *PAX6* el gen claramente mayoritario (Figura R30). En nuestra cohorte, el 80% (82/98) de los casos presentan alteraciones claramente causales en *PAX6* y el 4% variantes en *FOXC1* o *PITX2*. Esto implica un aumento en el rendimiento diagnóstico del 20%, desde un 63% al 84% obtenido en este trabajo. Por lo tanto, un 12% de los pacientes de esta cohorte permanecen sin diagnóstico genético concluyente, entre ellos 3 pacientes que no pudieron ser incluidos en el cribado de las regiones no codificantes.

De manera general, la aniridia clásica se debe a alteraciones del gen *PAX6*, el fenotipo WAGR a deleciones que incluyen a *PAX6* y *WT1* y los casos con fenotipo aniridia-like, tanto a variantes atípicas en *PAX6*, incluyendo *missense*, variantes en el 5'UTR y de *splicing*, pero también presentan variantes en *FOXC1* y *PITX2*.

Atendiendo al espectro mutacional identificado en nuestra cohorte, descrito en la Figura R30, el 55% de los pacientes presentan variantes en regiones codificantes, el 5%

presentan variantes en la región 5'UTR y el 19% presentan CNVs, de las cuales el 6% afectan solamente al gen *PAX6*, el 8% afectan también a sus genes contiguos (asociados al síndrome de WAGR) y el 5% restante afectan únicamente a sus regiones reguladoras (Figura R30.A).

Atendiendo a la herencia observada, 1/3 presentan una herencia autosómica dominante y 2/3 son casos esporádicos (Figura R30.B). Tal como se ha podido establecer en este trabajo, el 3% presentan mosaicismo parental confirmado molecularmente y en el restante 3% no tenemos datos familiares que nos permitan determinar su herencia.

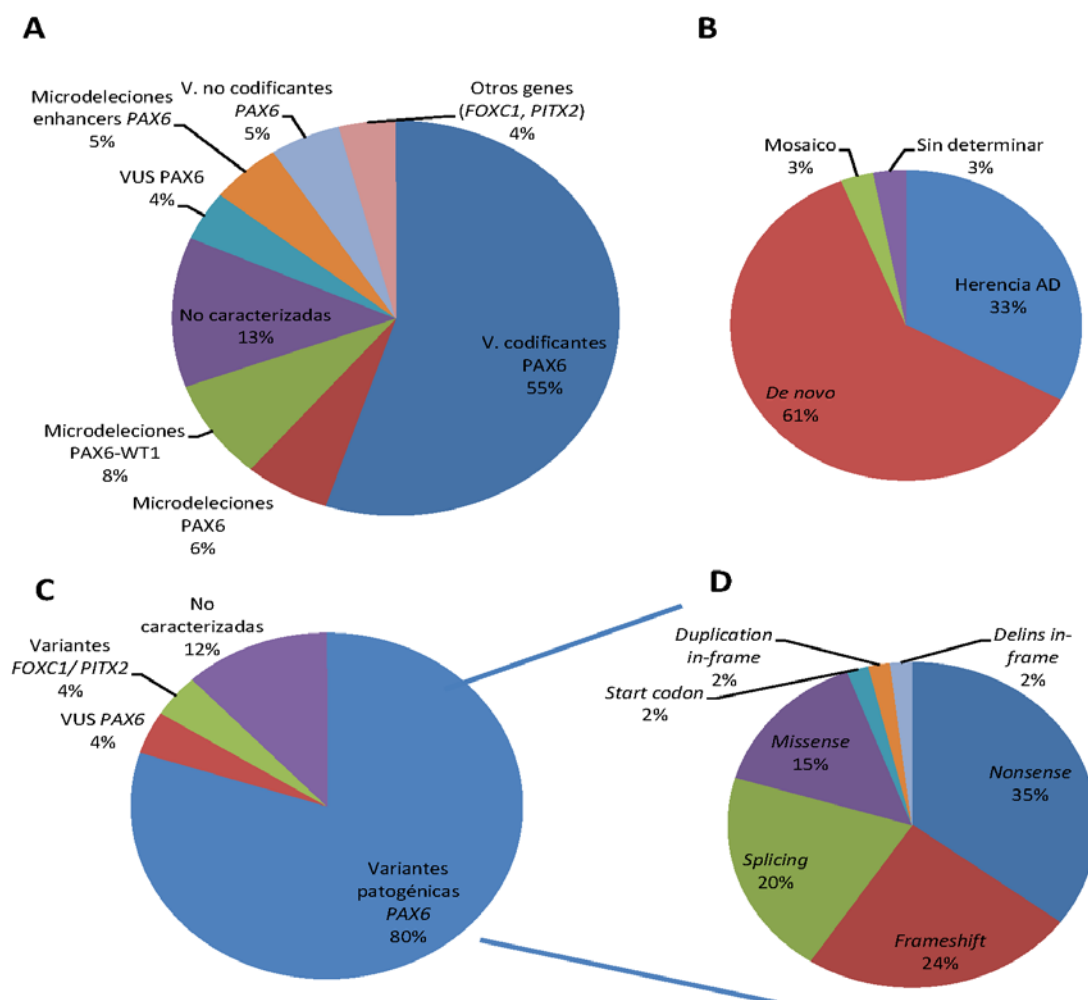


Figura R30. Resumen de las características genéticas de la cohorte de 98 familias con aniridia y aniridia-like estudiadas en este trabajo. A) Tipos de variantes detectadas en nuestra cohorte. B) Patrón de herencia. C) Genes implicados. D) Espectro mutacional en *PAX6*.

Discusión

El objetivo principal de esta Tesis ha sido profundizar en el conocimiento sobre las bases genéticas y moleculares de la aniridia congénita mediante la identificación y la caracterización de nuevos mecanismos genéticos causantes de la aniridia en una cohorte de pacientes que representa la mayor descrita hasta la fecha en nuestro país.

Para ello, primero hemos evaluado el rendimiento diagnóstico del algoritmo genético utilizado de forma rutinaria en nuestro laboratorio siguiendo las recomendaciones de Blanco-Kelly (Blanco-Kelly, 2012). Este protocolo analiza de forma secuencial SNVs en el gen *PAX6* y microdelecciones de la región 11p13, priorizando un análisis u otro en función de las características familiares y la edad del paciente. En la época pre-NGS, este análisis se realizaba mediante la secuenciación Sanger de los exones codificantes y un diseño específico de MLPA para analizar CNVs en *PAX6*, y otros genes contiguos implicados en el síndrome de WAGR. Tras la implementación de la técnica de NGS en nuestro laboratorio, este estudio se realiza actualmente utilizando la aproximación del exoma clínico que permite analizar de forma simultánea SNVs y CNVs en *PAX6*. En el 73% de los pacientes se identifican variantes que afectan a la secuencia codificante de *PAX6*, cifras similares a las descritas en otros trabajos donde el rendimiento varía entre 50-90% en función de la técnica de cribado utilizada (Gronskov, 2001; Robinson, 2008; Hingorani, 2009; Lim, 2012; Dubey, 2015; Bobilev, 2016; Vasilyeva, 2017; Sannan, 2018; Souzeau, 2018).

Esta estrategia diagnóstica no detecta alteraciones patogénicas en el 25% de los pacientes de nuestra cohorte. A la vista de estos resultados, en esta Tesis nos hemos centrado en aportar luz a la caracterización genética de estos pacientes considerados "*PAX6* negativos", ampliando el estudio genético analizando otras regiones que normalmente no suelen considerar en los estudios rutinarios, como son las regiones no codificantes (intrones y UTRs) y las regiones reguladoras. En ambos casos, estas regiones juegan un papel clave en la regulación de la expresión de *PAX6*. De hecho, trabajos previos habían identificado algunos pacientes con aniridia portadores de variantes en estas regiones. Así, se han identificado delecciones de los elementos reguladores situados en la región 3' de *PAX6*, sin llegar a afectar a la región codificante en numerosos pacientes (Lauderdale, 2000; Davis, 2008; Bayrakli, 2009; Cheng, 2011; Zhang, 2011; Simoni, 2012; Wawrocka, 2012; Addis, 2015; Ansari, 2016). Además, se

ha descrito una variante puntual en la secuencia del *enhancer* SIMO en un caso esporádico (Bathia, 2013). Por otro lado, previamente a nuestro estudio, se habían detectado 9 variantes en el 5'UTR de *PAX6* en distintos trabajos (Glaser, 1992; Axton, 1997; Gronskov, 1999; Ticho, 2006; Redeker, 2008; Robinson, 2008; Skeens, 2011; Vasilyeva, 2017). Sin embargo, es muy probable que este tipo de variantes estén infradetectadas en aniridia y podrían explicar algunos de los casos *PAX6* negativos de nuestra cohorte. Por otro lado, el estudio de estos pacientes puede darnos la clave para descubrir nuevos genes o mecanismos implicados en la patofisiología de esta enfermedad. Este fue el caso de una familia de origen chino en la que se identificó una variante *missense* en *TRIM44*, gen asociado por primera vez a la aniridia (Zhang, 2015). Ensayos *in vitro* realizados por los autores de este trabajo, concluyeron que este nuevo gen, situado a 4 Mb de *PAX6*, sería un regulador negativo de la expresión de *PAX6* y la variante p.G155R identificada conduciría a que se potencie dicha inhibición. En este mismo sentido, no se ha descrito ninguna otra variante asociada a aniridia en este gen hasta la fecha, por lo que en un futuro será necesario profundizar en la implicación real de este gen en el desarrollo de esta patología.

Por otro lado, dada la gran variabilidad fenotípica de la aniridia, principalmente en su signo más característico, la hipoplasia de iris, su estudio genético debería incluir un diagnóstico diferencial con otras patologías del segmento anterior que pueden presentar anomalías iridianas solapantes. Un ejemplo de ello es la hipoplasia de iris presente en el Síndrome de Gillespie, que recientemente ha sido asociado a variantes en el gen *ITPR1* (Gerber, 2016; McEntagart, 2016; Carvalho, 2017; Dentici, 2017; Zambonin, 2017). Durante esta Tesis, hemos identificado una variante de significado clínico incierto en este gen en uno de nuestros pacientes *PAX6* negativo. En el futuro, se deberán realizar estudios de segregación y análisis funcionales con el fin de conocer si dicha variante podría ser la causa de la clínica del paciente.

El estudio realizado durante esta Tesis Doctoral en los casos sin mutaciones en *PAX6* de nuestra cohorte nos ha permitido identificar variantes probablemente patogénicas, tanto en regiones no codificantes de *PAX6*, como en genes asociados al síndrome de Axenfeld-Rieger (*FOXC1* y *PITX2*).

En un porcentaje considerable del 10% de nuestra cohorte, las variantes identificadas en regiones codificantes de *PAX6* fueron consideradas como variantes de significado clínico incierto tras una revisión de la patogenicidad siguiendo las recientes recomendaciones de la ACMG. Por lo tanto, el resultado del estudio genético no terminó de ser concluyente. Esto afecta principalmente a variantes *missense*, *in-frame*, sinónimas y otras localizadas en sitios no canónicos del *splicing*, para las que no se puede inferir directamente que causen haploinsuficiencia en *PAX6*. Esto ocurre de forma similar con las variantes identificadas en el 5'UTR de *PAX6* durante esta Tesis Doctoral. La mayoría de ellas afectarían al *splicing* de un exón no codificante, sin que se pueda concluir claramente cuáles son las consecuencias sobre la expresión del transcrito canónico de *PAX6*. Por ello, en este trabajo hemos querido también profundizar en la caracterización de los mecanismos asociados a estas variantes de significado incierto, principalmente aquellas en las que se predecía que podrían alterar el *splicing* de *PAX6*, lo que nos ha permitido reclasificar 10 de las 12 variantes estudiadas en base a los resultados de alteración del *splicing* natural de *PAX6*. Las únicas variantes que no provocan alteraciones en el *splicing* son dos *missense*, que sin embargo fueron incluidas en este estudio por situarse cerca de potenciales donadores del exón 6. Ante los resultados inconcluyentes obtenidos en los estudios de *splicing* para estas variantes, su modelización *in silico* usando un modelo cristalográfico indicó que estos cambios no sinónimos conducen a alteraciones en plegamiento en la estructura 3D de la proteína, por lo que también se pudieron reclasificar como probablemente patogénicas.

Por último, en esta Tesis hemos podido confirmar por primera vez la existencia de 4 individuos mosaicos en familias con afectos de aniridia y microftalmia, lo que nos ha permitido descubrir la causa de un patrón de herencia inusual en unos casos y/o de la variabilidad fenotípica intra-familiar en otros.

5.1. Caracterización genética de casos de aniridia “*PAX6* negativos”

En esta Tesis Doctoral hemos recopilado 33 casos índice sin mutaciones en la región codificante de *PAX6*, que han sido estudiados con nuevas herramientas de aCGH y NGS específicamente diseñadas para cribar de manera exhaustiva no solo dicho gen por

completo y sus *enhancers*, cuya secuencia, previamente identificada por distintos autores (Williams, 1998; Kammandel, 1999; Kleinjan, 2001; Griffin C, 2002; Kleinjan DA, 2004; McBride DJ, 2011; Bathia, 2013; Ravi V, 2013; Bhatia S, 2014), fue también incluida. Este diseño, además incluye otros 120 genes relacionados con MOC. Este cribado, nos ha permitido identificar variantes probablemente patogénicas, tanto en regiones no codificantes de *PAX6*, como en genes asociados al síndrome de Axenfeld-Rieger en el 10% y el 4% de los pacientes, respectivamente.

Por un lado, hemos podido identificar CNVs en 5 de los pacientes, todas ellas localizadas regiones 3' reguladoras de *PAX6* y con un tamaño entre 500 y 700 Kb. Resulta interesante destacar que ninguna de las deleciones identificadas incluyera a *PAX6*, localizado a 90-125 Kb de distancia, aunque sí al elemento SIMO, un *enhancer* altamente conservados a lo largo de la evolución que ha sido implicado en la regulación transcripcional de *PAX6* de forma dependiente de tejido (Bathia, 2014; Bathia, 2015). Estos autores demostraron que la disrupción de este elemento, no solo mediada por su deleción completa sino también por un cambio puntual, G>T en posición chr11:31,685,945, provoca aniridia mediante la desregulación directa de la expresión de *PAX6* (Bathia, 2013). Por ello, se ha postulado que la disrupción de este elemento podría ser la principal responsable de la aniridia en pacientes portadores de microdeleciones de las regiones 3' reguladoras de *PAX6* (Ansari, 2016, Bathia, 2013; Franzoni, 2017). SIMO se encuentra localizado dentro de una región de 244 Kb que parece ser crítica para la regulación transcripcional de *PAX6* y que fue definida por Ansari en 2016 tras el análisis de los puntos de corte de las distintas microdeleciones identificadas hasta la fecha en pacientes con aniridia. Esta región crítica estaba acotada en 3' y 5' por deleciones identificadas por Ansari y cols en 2016 y Bayrakli y cols en 2009, respectivamente.

La comparación de los puntos de corte acotados por aCGH de 26 microdeleciones identificadas en esta región, no solo en nuestra cohorte y en pacientes franceses, sino también las descritas previamente en la literatura, nos ha permitido acotar aún más esta región crítica. Excepto un paciente descrito por Simoni y cols., 2012, el resto de los pacientes comparten una región delecionada mínima en común de tan solo 18 Kb en la posición chr11:31648248-31666340. Tal como se observa en la Figura R3, esta

nueva región mínima se ha acotado en el extremo 3' gracias a la identificación de una microdelección de 155 Kb en una familia francesa y en 5' por la delección descrita previamente por Bayrakli y cols., 2009. También es importante destacar que esta nueva región crítica no incluye a SIMO. Hasta el día de hoy, solo se han descrito 3 casos previamente publicados portadores de microdeleciones de las regiones 3' de *PAX6* que no incluyeran a este elemento regulador (Bayrakli, 2009; Cheng, 2011; Simioni, 2012). Nuestro trabajo evidencia que podrían existir otros *enhancers* importantes, aparte de SIMO, implicados en el desarrollo de aniridia.

El único elemento *cis*-regulador conocido de *PAX6* incluido en esta nueva región crítica es el *enhancer* E180, un elemento ultraconservado evolutivamente en los últimos 450 millones de años (Bathia, 2014). En un modelo de transgénesis en pez cebra, este elemento parece jugar un importante papel en la regulación transcripcional de *PAX6*, a pesar de que está situado a 130 Kb del mismo en la región intrónica de *ELP4*, tal como describieron Bathia y cols. Este regulador media una expresión espacio-temporal de *PAX6* de forma dependiente de tejido, conduciendo a su activación en los ganglios trigéminos y en las neuronas de la espina dorsal, además de en tejido ocular durante la embriogénesis (Bathia, 2014). Al contrario de lo descrito para SIMO, hasta el momento no se ha identificado ningún paciente en donde la aniridia se origine como consecuencia de únicamente la disrupción de E180. Durante el cribado de los *enhancers* de *PAX6* realizado en este trabajo, no se identificaron variantes puntuales en este elemento en la cohorte de pacientes franceses (Plaisancie, 2018). En nuestra cohorte, este elemento no estaba incluido en panel de NGS utilizado, ya que en el momento del diseño este regulador aún no había sido caracterizado. En un futuro, será necesario profundizar en la posible implicación de este elemento en nuestra cohorte de 12% de pacientes sin mutaciones en *PAX6*.

De forma similar a otros trabajos, debemos subrayar el hecho de que en ninguna de las dos cohortes estudiadas se detectaron CNVs ni variantes puntuales que afectaran a la región genómica entre *RCN1* y *PAX6*, lo que indicaría que los elementos 5' reguladores no parecen presentar un papel importante en el desarrollo de aniridia. Por otro lado, nuestros colaboradores franceses identificaron 11 variantes raras en algunos elementos 3' reguladores de *PAX6* en 8 pacientes de su cohorte (Plaisancié & Tarilonte,

2018). Sin embargo, fueron consideradas como VUS debido a que no se pudo hacer una interpretación de patogenicidad más precisa por la falta de ADN parental para el análisis de segregación. Además, su implicación también se ve limitada debido a que las variantes identificadas en realidad están descritas en bases de datos poblacionales con frecuencias relativamente altas, entre 0,003% y 0,9%, comparado con la frecuencia de aparición del fenotipo de aniridia (entre 0,002% y 0,001%). Por lo tanto, excepto la variante anteriormente mencionada en SIMO (Bathia, 2013), hasta el momento no se ha identificado ninguna otra variante puntual potencialmente patogénica en los elementos reguladores de *PAX6*.

En este trabajo, tampoco pudimos detectar otras variantes en las regiones 3'UTR ni intrónicas en nuestra cohorte. Sin embargo, estas regiones podrían tener una implicación en el desarrollo de aniridia. Así, en la cohorte francesa, se detectaron tres variantes *deep-intronic* en *PAX6* para las que se predice la creación de nuevos sitios de *splicing* (Plaisancié & Tarilonte, 2018). Sin embargo, una vez más la falta de segregación hizo que fueran clasificadas como VUS. Con el fin de dilucidar la patogenicidad de estas variantes, en estos momentos se están llevando a cabo estudios *in vitro* para dos variantes del intrón 6, similares a los realizados en esta Tesis Doctoral, los cuales esperamos que nos ayuden a establecer un posible diagnóstico concluyente en estas familias.

Por el contrario, nuestro trabajo ha revelado la existencia de variantes en el 5'UTR de *PAX6* en 5 familias no emparentadas de nuestra cohorte. Ese tipo de variantes en los exones no codificantes parece representar una causa frecuente de aniridia explicando alrededor del 14% de los casos *PAX6* negativos y el 5% del total de familias de nuestra cohorte, con una frecuencia similar a la observada en la cohorte francesa de nuestros colaboradores. En la base de mutaciones de *PAX6*, sólo el 1,3% de los pacientes con aniridia caracterizados (14/1067) portan este tipo de variantes. Por lo tanto, pensamos que este tipo de variantes parecen estar realmente infradetectadas en la mayoría de los estudios realizados anteriormente. Una de las razones más evidentes que podrían explicar este hecho es la ausencia de cribado de la región 5'UTR de *PAX6* en la mayoría de estudios diagnósticos realizados en la época pre-NGS en las

distintas cohortes estudiadas (Gupta, 1999; Vincent, 2003; Neethirajan, 2004; Zhang, 2011; Pérez-Solórzano, 2017).

El espectro mutacional en las dos cohortes para las variantes en el 5'UTR de *PAX6*, es también muy similar. En conjunto, se han identificado 12 familias portando 7 mutaciones diferentes en los exones 2 y 3 de *PAX6* (Plaisancie & Tarilonte, 2018). Excepto dos casos *de novo*, estas variantes están asociadas a formas autosómicas dominantes, donde las variantes segregan en los casos familiares. Hasta ahora, solo se habían descrito 9 variantes en estas regiones (Figura R4), muchas de ellas reportadas una única vez en familias aisladas (Glaser, 1992; Axton, 1997; Gronskov, 1999; Ticho, 2006; Redeker, 2008; Robinson, 2008; Skeens, 2011; Vasilyeva, 2017). La mayoría de las variantes identificadas se sitúan en sitios canónicos de *splicing* de los exones no codificantes 2 y 3, por lo que afectarían al procesamiento del pre-ARNm de *PAX6*, tal como se confirman en los estudios predictivos *in silico* realizados. La excepción son dos pequeñas deleciones (c.-118_-117del y c.-116_-112del), que aunque próximas al aceptor del exón 3, no parecen alterar el *splicing* en estos estudios.

De acuerdo con la guía del ACMG (Richards, 2015), las variantes c.-52+1G>A, c.-128-2delA y c.-129+1G>A son consideradas patogénicas debido a que: (1) se han encontrado en varias familias afectas de aniridia, (2) afectan a sitios canónicos de *splicing* y (3) se predicen como patogénicas mediante herramientas *in silico*. Además, en el caso de la variante c.-52+1G>A, existen evidencias experimentales que demuestran sus efectos patogénicos. Esta variante conduce al *skipping* de los exones 3 al 6 de *PAX6*, tal como demostraron Gronskov K. y cols. en estudios de expresión realizados en LCL de portadores heterocigotos (Gronskov, 1999).

Otras 3 variantes, c.-52+3_-52+4delAA, c.-52+3_-52+6delinsTG y c.-52+5delG, fueron identificadas en este trabajo por primera vez, y afectan a nucleótidos no canónicos del donador del exón 3. Aunque las herramientas de predicción *in silico* suelen detectar correctamente las interrupciones de las regiones consenso de *splicing*, resulta más difícil pronosticar las consecuencias de la interrupción de nucleótidos no canónicos. En general, no resulta sencillo de determinar ni la naturaleza específica de los transcritos resultantes en procesos aberrantes de *splicing*, ni la distribución de los

mismos, ni la posible implicación de otros sitios críticos de *splicing* o de otros elementos reguladores SREs (Spurdle, 2008). Por ello, es necesaria la realización de ensayos *in vivo* o *in vitro* para poder comprender el significado clínico de este tipo de variantes. En nuestro caso, esta interpretación se complica por la implicación de exones no codificantes, en los cuales no está claro cual es el papel que desempeñan en la regulación transcripcional y/o traduccional de *PAX6*.

Por todo ello, con el fin de explicar cuáles son los mecanismos patológicos asociados a estas variantes no codificantes del 5'UTR, se realizaron estudios funcionales para distintas variantes. Dada la dificultad en la obtención de muestras accesibles de nuestros pacientes en donde se pudiera valorar la expresión de *PAX6*, en este trabajo nos decantamos principalmente por realizar ensayos *in vitro* de *splicing* mediante minigenes en distintas líneas celulares. Estos estudios, basados en una estrategia *exon-trapping*, proporcionan una buena aproximación para comprobar lo que ocurre a nivel de pre-ARNm, tanto cuando se utilizan minigenes que incorporan el exón de interés como estrategias más actuales de midigenes que incluyen las regiones exónicas e intrónicas adyacentes. Esto mejora la reproducibilidad de modelización de los mecanismos de *splicing* al utilizarse condiciones más fisiológicas (Sanz, 2010; Acedo, 2015; van der Klift, 2015; de la Hoya, 2016; Fraile-Bethencourt, 2017; Nanan, 2017; Sangermano, 2017; Filatova, 2019).

En este trabajo, el uso de una estrategia de midigenes para la región 5'UTR de *PAX6* nos ha permitido medir el impacto de las variantes que afectan al *splicing* de los exones no codificantes de *PAX6*. Nuestra construcción no sólo contiene los 3 exones no codificantes, sino mantiene el contexto genómico al incluir las regiones intrónicas involucradas y el primer exón codificante. Esto nos ha permitido detectar varias isoformas aberrantes de *splicing* que afectan además a la región codificante de *PAX6*, como hemos observado para la nueva variante c.-129+1G>A localizada en el exón 2. El análisis con midigenes mostró un complejo patrón de eventos de *splicing* con diferencias aparentemente asociadas al origen de la línea utilizada. Se identificaron isoformas causadas no solo por el *skipping* predicho del exón 2, sino también por la delección de los exones contiguos, además de retenciones parciales de dicho intrón. La isoforma más abundante en la línea de retina ARPE19 fue la delección de los exones 2 al

4, seguido de la retención parcial del intrón 2. Esta última fue la isoforma predominante en HEK293T. Con estas evidencias experimentales, confirmamos la patogenicidad de esta variante, por lo que fue considerada la causa de la aniridia en 7 afectos procedentes de 4 familias no emparentadas en los que se había detectado esta mutación.

En esta misma línea, el análisis de midigenes realizado para la variante c.-52+1G>A mostró el *skipping* aberrante del exón 3. Nuestro experimento no reproduce fielmente las alteraciones en el *splicing* observadas por Gronskov K. y cols. pues esta construcción no incluye los exones 5 y 6, que también estarían involucrados en el *splicing* aberrante originado por esta variante. A pesar de estas limitaciones, nuestro midigen representa una buena estrategia para estudiar la implicación de otras variantes localizadas en el 5'UTR, cuando no está disponible una línea de LCL de pacientes en la que valorar la expresión de *PAX6*. En este sentido, un efecto similar fue observado *in vitro* para las otras 3 variantes del exón 3 estudiadas. Estas incluyen la anteriormente reportada variante c.-128-2delA (Axton, 1997; Robinson, 2008; Vasilyeva, 2017), encontrada en dos familias de este estudio, así como las variantes localizadas en nucleótidos no canónicos del donador, c.-52+3_-52+6delinsTG y c.-52+5delG. Dadas estas evidencias experimentales y su segregación en varios individuos afectos o su aparición *de novo*, en el caso de la variante c.-52+5delG, consideramos que serían causantes de aniridia.

Durante este trabajo, tuvimos la oportunidad de establecer LCL de tres individuos heterocigotos para las variantes c.-129+1G>A y c.-128-2delA, en donde se realizaron estudios de expresión de forma paralela al análisis *in vitro*. Desafortunadamente, a pesar de realizar distintas aproximaciones experimentales, sólo pudimos observar el ARNm producido por el alelo silvestre. La ausencia de detección de transcritos anómalos en estas células podría deberse a varias causas, incluyendo su degradación mediante el sistema NMD por la presencia de un codón de parada prematuro en el transcrito resultante, tal como se ha observado en otras variantes de *splicing* (Vincent, 2003). A raíz de esta hipótesis, decidimos estudiar la expresión de *PAX6* tras la inhibición directa e indirecta del mecanismo NMD. Sin embargo, no fuimos capaces de obtener resultados diferentes a los anteriores. Esta aproximación realizada sobre LCL

ha sido exitosa en la evaluación de otras variantes (Axton, 1997; Gronskov, 1999; Vincent, 2003), permitiendo detectar numerosas formas de *splicing* en *PAX6* (Vincent, 2003), tal como hemos observado también en esta Tesis Doctoral para la variante silente c.174G>A. Sin embargo, pueden existir otras limitaciones experimentales y técnicas que estén impidiendo detectar las isoformas aberrantes, entre ellas, el uso de técnicas de baja sensibilidad para la detección de transcritos poco abundantes (Whiley, 2014). El uso de técnicas de expresión cuantitativa con mayor sensibilidad, como la ddPCR o incluso RNA-seq, podrían ayudar a determinar la posible existencia de transcritos aberrantes en estas células que no estén siendo detectados por la técnica de RT-PCR utilizada. En este sentido, un análisis de RNA-seq podría además determinar el conjunto completo de isoformas presentes de forma no sesgada e independiente a los cebadores utilizados durante las técnicas basadas en PCR. Otra posibilidad es que realmente esta variante pueda tener diferencias en su expresividad de forma dependiente de tejido, tal como se ha evidenciado en nuestros experimentos de minigenes. Por otro lado, este modelo celular utilizado puede que no sea el más apropiado dado que *PAX6* se expresa en los linfocitos de forma ilegítima tras su transformación con virus EBV (Gronskov 1999; Gronskov, 2001; Vincent, 2003). Esto mismo fue verificado por nuestro grupo comparando la expresión de *PAX6* en cultivos primarios de linfocitos y en LCL de los mismos individuos, en donde solo se detectaron bajos niveles de expresión de *PAX6* tras la inmortalización. Una alternativa a estas líneas sería el uso de otro tipo de tejidos accesibles en donde valorar la expresión de *PAX6*. Los fibrocitos no serían una opción, ya que también hemos comprobado la ausencia de expresión de este gen. Sin embargo, se podrían valorar estudios de expresión con ARN procedente de células corneales y/o conjuntivales, obtenidas mediante técnicas de impresión corneal (Shortt, 2008), células limbares (Latta, 2018; Roux, 2018; Latta, 2019), lágrimas y/u en el futuro, otros modelos celulares basados en iPSCs derivados de pacientes (Bremond-Gignac, 2018), que permitieran realizar estudios de *splicing* en condicionales aparentemente más fisiológicas. Otra alternativa podría ser la obtención de modelos celulares utilizando la tecnología CRISPR (Burnight, 2017; Kitazawa, 2017; Roux, 2018; Shen, 2018; Shi, 2019).

Nuestro trabajo no nos permite valorar si los transcritos aberrantes detectados en los experimentos de minigenes serían deletéreos *in vivo*. En general, se piensa que los transcritos que afectan al inicio canónico del ORF serán deletéreos, como la isoforma identificada que implica el *skipping* del exón 4 de *PAX6*, al tener un impacto directo sobre la estructura de la proteína, o en la expresión a nivel del ARNm. Sin embargo, se podrían estar utilizando ORF alternativos, lo que dificulta determinar las consecuencias patológicas (Spurdle, 2008). En otras isoformas detectadas que implicarían solo a los exones no codificantes, desconocemos cual es el papel que representan estos exones en la expresión de *PAX6*. Se esperaría que diesen lugar a la proteína completa, aunque podría ocurrir un desequilibrio en los niveles de los transcritos por desregulación de la transcripción de *PAX6* o indirectamente por afectar a la estabilidad del pre-ARNm. En este sentido, se sabe que cambios que afectan a la estructura y a la longitud de los pre-ARNm podrían también modificar la estructura secundaria y la estabilidad del ARNm maduro (Mignone, 2002). Esta podría ser otra causa probable de la aparente falta de expresión *in vivo* de las isoformas aberrantes en LCL de los pacientes.

Nuestro trabajo evidencia que los errores de *splicing* de los exones no codificantes, que no suelen ser analizados de forma rutinaria, deben de tenerse en cuenta durante el diagnóstico genético de la aniridia. Estos podrían explicar hasta el 25% de los pacientes negativos tras un estudio rutinario, si tenemos en cuenta los datos conjuntos de las dos cohortes estudiadas (Plaisancie & Tarilonte, 2018). A pesar de que la mayoría son mutaciones particulares de cada familia, con excepción de alguna variante recurrente, todas ellas parecen compartir un mecanismo patogénico similar. Además, nuestro estudio revela que las variantes en otras regiones no codificantes son realmente minoritarias, aunque podrían explicar algún caso de la cohorte francesa siempre que se pudieran realizar estudios de segregación y funcionales que confirmaran su patogenicidad.

Este cribado además nos ha permitido descartar la existencia de alguna otra variante no codificante en *PAX6* en 23 pacientes de nuestra cohorte. Entre ellos se incluyen 10 pacientes portadores de variantes de significado clínico incierto en *PAX6*, anteriormente identificadas. Por lo tanto, nuestro cribado refuerza la hipótesis de la causalidad de estas variantes en estos pacientes. Por otro lado, podemos concluir que

el 48% (16/33) de los pacientes de este estudio específico, es decir, el 16% (16/98) del total de la cohorte, son *PAX6* negativos, pues no muestran ningún cambio que implique al *locus PAX6*. Sin embargo, la ampliación del estudio a otros genes asociados a formas solapantes de disgenesia del segmento anterior nos ha permitido caracterizar un 4% adicional de pacientes.

Finalmente, este cribado de pacientes sin diagnóstico ha revelado que solo el 12% de nuestros pacientes siguen a la espera de ser caracterizados genéticamente. Se asumen distintas hipótesis para explicar la ausencia de variantes patogénicas en estos pacientes. Primero, podrían existir otras variantes patogénicas que no hayan podido ser capturadas con las estrategias de análisis utilizadas, como las localizadas en: (1) regiones de baja cobertura, (2) en elementos reguladores de *PAX6* aún sin identificar, y por tanto, no incluidos en nuestro panel de NGS, (3) en las regiones no codificantes de otros genes asociados a MOC, dado que no están incluidas en el panel, (4) en nuevos genes candidatos para aniridia aún sin identificar o en otros genes asociados a MOC no incluidos en el panel. Además, estos pacientes podrían ser mosaicos somáticos con la variante mutante en cuestión localizada en tejidos oculares y/u otros de difícil acceso, que de no aparecer en sangre no podría ser detectada en un estudio rutinario. Por último, podrían presentar reordenamientos que no impliquen pérdida en el número de copia, como por ejemplo, translocaciones o inversiones equilibradas, que no pueden ser detectadas con las estrategias utilizadas en este trabajo. En algunos casos de nuestra cohorte, se pudo realizar un cariotipo excluyéndose la presencia de reordenamientos balanceados. En un futuro, el uso de la secuenciación completa del genoma (WGS), que permite el análisis simultáneo de variantes en nuevos genes, la captura de las regiones intrónicas y reguladoras e incluso el análisis de reordenamientos, podría ayudarnos a explicar algunos de los raros pacientes de nuestra cohorte sin caracterización genética.

5.2. Caracterización funcional y reclasificación clínica de VUS en *PAX6*

En los pacientes portadores de variantes de significado clínico incierto en *PAX6*, nuestro cribado anterior permitió descartar la presencia de otras causas genéticas, lo que podría apoyar que realmente estas VUS pudieran explicar la aniridia. Con el fin de

conseguir un resultado concluyente sobre su patogenicidad y ayudar en el diagnóstico y posterior consejo genético en las familias, una segunda línea de estudio de esta Tesis fue la realización de análisis funcionales en estas variantes tipo VUS.

Para ello, se revisaron de forma exhaustiva los datos que apoyaban la patogenicidad de las VUS identificadas en nuestra cohorte, utilizando para ello: nuevos criterios de clasificación basados en ACMG, datos de segregación, datos actuales de frecuencias poblacionales y un estudio exhaustivo con distintos predictores de patogenicidad. Con esta revisión, identificamos 10 variantes en la región codificante de *PAX6*, en las que no se obtuvo una clasificación de patogenicidad de tipo 4/5. Entre ellas, había variantes *missense*, *in-frame*, sinónimas y variantes localizadas en nucleótidos no canónicos de *splicing* (NCS). En conjunto, estas variantes representan el 10% de los casos de nuestra cohorte de pacientes. Destacar que la mayoría se localizan en el exón 6 de *PAX6*, *hotspot* en donde se han descrito más de 100 mutaciones, entre ellas numerosas variantes de *splicing*. Según los estudios *in silico* realizados, 6 de estas variantes pueden alterar potencialmente el *splicing* del exón 6 mediante la creación o potenciación de 5 sitios donadores crípticos o incluso a través de elementos reguladores del *splicing*.

Tal como se ha argumentado anteriormente en el estudio de las variantes del 5'UTR, existen dificultades para inferir los efectos sobre el *splicing* de las variantes localizadas en regiones NCS, es decir aquellas situadas fuera de las regiones ultraconservadas GT/AG. Además, existen pocos estudios que profundicen en los mecanismos de *splicing* de las variantes NCS en *PAX6*, bien mediante el uso de ARN de LCL derivadas de pacientes (Hanson, 1993; Axton, 1997; Gronskov, 1999; Vincent, 2003), o alternativamente por medio de aproximaciones *in vitro* con minigenes (Weisschuh, 2012; Miao, 2017; Zhang, 2017; Filatova, 2019), principalmente debido a las dificultades para confirmar su patogenicidad en un contexto clínico.

En el caso del exón 6, estudios anteriores utilizando LCL de pacientes revelaron que las variantes c.357+1G>A y c.357+2insT conducen a la delección parcial del mismo mediante la activación indirecta de un donador críptico localizado 108 pares de bases aguas arriba (Hanson, 1993; Axton, 1997). La consecuencia proteica es una delección *in-*

frame de 36 residuos del dominio pareado de unión al ADN. Estudios *in silico* realizados en la secuencia silvestre y en las mutadas indican que este donador canónico del exón 6 es débil y se predicen hasta 6 donadores alternativos, cinco de ellos exónicos, con una potencia ligeramente inferior (Figura R13). De hecho, uno de estos donadores crípticos en la posición c.157 da lugar a una isoforma alternativa con 201 nucleótidos del exón 6 delecionados que co-existe con la isoforma canónica en LCL de controles y en ojos de vacas y aves (Gronskov, 1999).

A la vista de estos antecedentes, en el presente trabajo nos planteamos la evaluación funcional de estas variantes VUS así como la profundización en los mecanismos de regulación del *splicing* de este exón 6. No solo se incluyeron estas 6 variantes de tipo VUS, sino que además valoramos si otras variantes en este exón podrían alterar asimismo el *splicing* a través de estos potenciales sitios crípticos. De hecho, incluimos una variante sinónima que Murton y cols, 2000 habían identificado en un paciente con aniridia, la cual también mostró potencial de alterar el *splicing* mediante la activación del donador críptico c.331.

Para profundizar en el funcionamiento del mecanismo de *splicing* de estas VUS, nuestra primera opción fue realizar un estudio de expresión en LCL de pacientes portadores, sin embargo, sólo pudimos realizarlo en una única familia. Como alternativa, desarrollamos un estudio *in vitro* basado en el uso de un minigen del exón 6 en células ARPE19 derivadas de retina humana. De esta manera, detectamos patrones anómalos de *splicing* en 5 de las 7 variantes estudiadas, la excepción fueron las 2 *missense*. El espectro de los transcritos aberrantes detectados incluye al menos 5 deleciones parciales del exón 6 y una retención de los primeros 5 nucleótidos intrónicos. Nuestro trabajo confirma que la causa principal de estas isoformas es la activación directa de un dinucleótido GT alternativo que podría utilizarse preferentemente o en paralelo al sitio donador natural.

Atendiendo a los resultados derivados de nuestro trabajo y de otros autores (Hanson, 1993; Axton, 1997; Filatova, 2019), podemos afirmar que existen 6 donadores alternativos en este exón (Figura R13). Estos incluyen el sitio en c.157 anteriormente mencionado (Gronskov, 1999), que hasta el momento no ha sido

implicado en el desarrollo de aniridia, así como otros 5 sitios crípticos que conducirían a mecanismos patogénicos. Nuestros resultados confirman que el donador alternativo en posición c.250 no solo se activaría por la disrupción del dinucleótido GT canónico (Hanson, 1993; Axton, 1997), sino también directamente por el cambio c.249T>G e indirectamente por la desregulación de elementos SRE. Utilizando este último mecanismo, una pequeña duplicación c.326_331 localizada 27 pb aguas arriba en una región enriquecida para sitios de unión a SREs, crea un nuevo sitio para unión a hnRNPA1, un potencial silenciador del *splicing*. La consecuencia es el silenciamiento del donador canónico. Este hallazgo subraya el importante papel de las regiones ricas en SRE en el sutil desbalance entre los donadores canónicos y alternativos (Brillen, 2017). Por otro lado, hemos podido confirmar el uso alternativo de un sexto donador críptico en posición c.173 en una línea de LCL de una paciente portadora de la variante sinónima c.174C>T, el cual ha sido recientemente reportado en ensayos *in vitro* (Filatova, 2019). Nuestro estudio ha permitido la identificación de otros 3 nuevos sitios crípticos c.177, c.331 y c.357+6, los cuales son activados como consecuencia de diferentes cambios nucleotídicos. Por lo tanto, nuestro trabajo subraya que la activación de donadores crípticos podría ser un mecanismo patogénico subestimado y recurrente en aniridia.

A pesar de la plétora de eventos de isoformas alternativas y de 5'SS crípticos aquí identificados, es probable que los errores de *splicing* detectados en este exón se deban básicamente a la aparente debilidad del 5'SS canónico. Esto vendría indicado por su baja potencia en las predicciones *in silico*, un uso sub-óptimo en el minigen control y principalmente por la presencia de transcritos alternativos constitutivos mediados por el donador críptico c.157. En este exón, cualquier evento que altere el frágil equilibrio entre el donador natural y los diferentes GT existentes, podría conducir a una alteración del *splicing*. En esta situación, nuestro estudio apunta que la modulación del *splicing* en este exón mediada por SREs podría jugar un papel importante en los mecanismos patogénicos. Un mecanismo similar fue descrito para una variante patogénica en el exón alternativo 5a, que altera la ratio de inclusión de dicho exón en la proteína, tras afectar a un sitio SRE (Epstein, 1994).

Por otro lado, no hemos detectado en nuestra cohorte de pacientes ninguna variante de *splicing* que pudiera alterar el sitio aceptor del exón 6. Aunque los sitios 3'SS han sido descritos como aparentemente menos dañinos en otras patologías (Epstein, 1994; Buratti, 2011; Sangermano, 2017), Filatova y cols. caracterizaron recientemente 2 variantes NCS que alteran el aceptor canónico mediante la activación de dinucleótidos A/T crípticos situados en posiciones intrónicas cercanas (Filatova, 2019). Además, en el intrón 6 se predicen otros aceptores y donadores crípticos, tal como ha puesto de manifiesto la identificación de dos variantes en el intrón 6, c.357+136G>A y c.357+334G>A en la cohorte de pacientes franceses (Plaisancié & Tarilonte, 2018). De forma similar, se ha descrito un ratón con microftalmia causada por una variante en este mismo intrón de *PAX6* que conduce a la inserción de un pseudoexón críptico (Puk, 2013). Estas variantes identificadas están localizadas fuera de las regiones intrónicas incluidas en el minigen utilizado en nuestro trabajo. Con el fin de profundizar en los mecanismos implicados en estas variantes, actualmente se está llevando a cabo su estudio *in vitro* utilizando un nuevo midigen para los exones 5 al 7, que recientemente hemos obtenido en nuestro laboratorio. Resultados preliminares indicarían que el escenario de la regulación del *splicing* de este exón podría ser aún más complejo de lo que inicialmente habíamos previsto.

La mayoría de las isoformas aberrantes aquí identificadas podrían conllevar a una delección *in-frame* del dominio PD de unión de ADN; por lo tanto, causarían una pérdida de la funcionalidad de *PAX6*, aún por determinar. La severidad de los fenotipos asociados a estas variantes podría correlacionarse con la ratio entre formas de *splicing* normales vs aberrantes, tal como se ha descrito en otras patologías oculares (Drivas, 2015). A pesar de realizar un intento de clasificación de la gravedad en base a las isoformas de *splicing*, realmente es difícil establecer correlaciones entre el genotipo y el fenotipo de los pacientes portadores de estas variantes. Con la excepción de las dos variantes *missense*, asociadas a microftalmia, en las que hemos excluido un papel en el *splicing*, la mayoría de los portadores del resto de variantes presentan manifestaciones clásicas de aniridia. Sin embargo, se observó un fenotipo del iris altamente variable entre distintos portadores de la misma variante, e incluso una marcada asimetría entre los ojos del mismo paciente (Figura R25).

Por último, este trabajo amplía el espectro de las variantes patogénicas que conducen a la aniridia, identificando nuevas variantes exónicas y NCS que alteran el *splicing* natural del exón 6 de *PAX6*. Por ello, cualquier tipo de cambio nucleótido y en cualquier posición podría alterar potencialmente el procesamiento del pre-mARN de dicho exón. Dada la sobrerrepresentación mutacional descrita en este exón, es altamente probable que otras variantes silentes o no codificantes hayan sido descartadas como patogénicas en otros estudios. Nuestro trabajo confirma también que las variantes sinónimas en *PAX6*, normalmente consideradas como benignas o VUS según los criterios estándar de clasificación, pueden no ser neutras y, por lo tanto, ante su identificación en un paciente con aniridia, se debería valorar la posibilidad de un error de *splicing*.

Nuestro trabajo también revela que los errores del *splicing* mediados por sitios donadores críticos podrían representar mecanismos patógenos recurrentes en la aniridia. Con lo cual, existe la posibilidad de que una proporción de los casos con aniridia sin caracterizar puedan explicarse por defectos similares. Por otro lado, nuestro estudio remarca la importancia de explorar el mecanismo de los sitios críticos de *splicing*, ya que es la base para plantear en un futuro nuevos tratamientos emergentes basados en oligonucleótidos antisentido que permitan corregir el *splicing* defectuoso (Albert, 2018; Cicediyan, 2019; Levin, 2019). En conclusión, el estudio funcional realizado nos ha permitido reclasificar 5 variantes como probablemente patogénicas. Todo ello pone de manifiesto la utilidad de la evaluación funcional de este tipo de variantes en su clasificación patogénica y, en la obtención de un diagnóstico y consejos genéticos más precisos.

5.3. Implicación del mosaicismo parental en la aniridia

Durante el desarrollo de esta Tesis Doctoral, hemos identificado posibles casos de mosaicismo parental en 5 familias portadoras de variantes en *PAX6* como causa de la variabilidad fenotípica y/o de la penetrancia incompleta encontradas en ellas. En los 4 casos en donde fue posible realizar un estudio molecular, el individuo mosaico, con un fenotipo parcial o incluso asintomático, presentaba el alelo "mutado" aparentemente desequilibrado en un análisis por secuenciación Sanger. Esta sospecha de mosaicismo

fue confirmada mediante el uso de una técnica de mayor sensibilidad para cuantificar alelos de baja frecuencia. La puesta a punto de la técnica de PCR digital en este proyecto nos permitió obtener una cuantificación del grado de mosaicismo de una forma mucho más precisa. En todos los casos, el alelo mutante en *PAX6* presentó una proporción inferior al 50%, valor esperado para un individuo heterocigoto. Los individuos mosaicos mostraron una relativa alta frecuencia del alelo mutado, entre el 11 y el 29,5%, en los tejidos estudiados procedentes de diferentes líneas somáticas, independientemente de la manifestación o no de la patología.

En los últimos años, el desarrollo de tecnologías de genotipado con mayor sensibilidad está permitiendo la detección de un número inesperadamente elevado de variantes en mosaicismo. Este tipo de variantes, que aparecen debido a errores post-cigóticos en la replicación del ADN en estadíos embrionarios tempranos, representan una nueva causa para la aparición de variantes *de novo* (Acuna-Hidalgo, 2015; Acuna-Hidalgo, 2016). Existen numerosos estudios que están poniendo de relieve el importante papel de estas variantes post-cigóticas en el desarrollo de enfermedades dominantes, explicando hasta el 10% de las enfermedades raras del desarrollo como epilepsia, autismo, o discapacidad intelectual (Acuna-Hidalgo, 2015; Stosser, 2017; Myers, 2018). De igual manera, estas variantes podrían estar involucradas en el desarrollo de enfermedades oculares congénitas (Ohuchi, 2018). Así, existen algunos trabajos en donde se han detectado variantes en mosaicismo asociadas al desarrollo de coloboma o microftalmia (Ceroni, 2018). Nuestro trabajo pone de manifiesto que este mecanismo explicaría al menos el 4% de las familias de nuestra cohorte. Sin embargo, pensamos que este fenómeno podría estar infravalorado dado que en este trabajo solo nos hemos centrado en familias en las que existía una sospecha previa de mosaicismo, bien por evidencias moleculares o por la existencia de recurrencia de aniridia en hijos descendientes de padres sanos. En un futuro, estudios en familias con mutaciones aparentemente *de novo* utilizando técnicas de alta sensibilidad, como la NGS a gran profundidad, estudios de célula única o ddPCR, podrían ayudarnos a identificar alelos mutados de baja o muy baja frecuencia que sean indetectables mediante secuenciación Sanger y, por lo tanto, hayan pasado desapercibidos en estudios rutinarios. Esto permitiría mejorar la estimación de la frecuencia del

mosaicismo y dilucidar la relevancia real de las variantes post-cigóticas en la patogénesis de las enfermedades asociadas a *PAX6*.

Destacar que dos tercios de los pacientes con aniridia son esporádicos (Netland, 2011; Lee, 2013), en los que se asume que son portadores de variantes *de novo*, normalmente no detectadas en muestras de sangre de los progenitores. Sin embargo, nuestro trabajo sugiere que la transmisión parental de variantes post-cigóticas podría ser un mecanismo patogénico frecuente asociado a mutaciones en *PAX6*, el cual estaría infradetectedo en los cribados que se realizan en la rutina diagnóstica. Por ello, en un futuro, se deberá plantear la posibilidad del mosaicismo parental en casos esporádicos con aparentes mutaciones *de novo*, aunque ambos progenitores sean en principio sanos.

Por otro lado, nuestro trabajo pone en evidencia la importancia del estudio del riesgo de transmisión en mosaicos para optimizar el consejo genético de la aniridia. En este sentido, el hallazgo de mosaicismo parental en una familia con aniridia conlleva grandes implicaciones para el consejo genético. Así, el análisis de muestra seminal de varones con mosaicismo gonosomal y/o germinal permitiría determinar el riesgo real de transmisión de la enfermedad. En una de las familias, el análisis del ADN procedente de semen del individuo mosaico permitió calcular que el riesgo real de transmisión de la variante post-cigótica a una futura descendencia era de 1/3, mucho mayor de lo que se esperaría para un caso esporádico, en el que se considera prácticamente despreciable (Acuna-Hidalgo, 2016). En las otras familias, no se pudo valorar el riesgo exacto de transmisión. Pese a ello, la transmisión del alelo mutado a los descendientes indica claramente la presencia de mosaicismo gonosomal, por lo que asumimos que estos individuos presentan el alelo mutado en una proporción significativa en su línea germinal.

Por otro lado, la transmisión de un alelo mosaico podría también explicar casos de recurrencia en familias con progenitores aparentemente sanos, como el caso de la familia descrita en este trabajo u otros casos publicados (Reed, 1955; Gronskov, 1999). El tema del mosaicismo germinal se postula hace años como la explicación más plausible ante casos de enfermedad recurrente en descendientes afectados procedentes

de progenitores sanos y sin antecedentes familiares. Distintos autores han identificado distintas familias con sospecha de mosaicismo germinal en aniridia (Reed, 1955; Gronskov, 1999) o en casos de microftalmia asociadas a mutaciones en *PAX6* (Deml, 2016; Riera, 2017). Sin embargo, estas sospechas no pudieron ser confirmadas molecularmente bien por una limitada disponibilidad de tejidos germinales y/o somáticos o bien por razones de sensibilidad analítica. En esta Tesis describimos 2 nuevas familias con recurrencia con fenotipo asociado a *PAX6* en hermanos descendientes de padres asintomáticos, en las cuales se pudo confirmar el mosaicismo gonosomal. Recientemente, Bai y cols, confirmaron un mosaicismo germinal en una muestra seminal, que no se daba en la línea somática del padre sano de dos afectos de aniridia (Bai, 2018)

Por otro lado, solo se habían descrito algunos casos excepcionales de mosaicismo somático para microdeleciones en la región 11p13 (Robinson, 2008; Erez, 2010; Huynh, 2017; Souzeau, 2018). Por primera vez, hemos identificado un caso de mosaicismo somático para una variante puntual en *PAX6* en un paciente con aniridia. Este hallazgo también conlleva una modificación en el consejo genético inicial, como ha ocurrido en una familia de este estudio. El riesgo de transmisión del alelo mutado debería ser menor del 50% esperado para un portador heterocigoto, tal como se postulaba antes del estudio.

Este trabajo describe por primera vez la existencia de variantes en mosaicismo como causa de la variabilidad intrafamiliar en patologías asociadas a *PAX6*, como aniridia, coloboma, y microftalmia. La variabilidad fenotípica es un fenómeno bien establecido en los defectos asociados a *PAX6* (Hanson, 1999; Hingorani 2012), que podría ser atribuible a modificadores genéticos y epigenéticos, pleiotropía y/o dosis génica (Schedl, 1996; Hanson, 1999; Vincent, 2003; Chou, 2015; Cvekl, 2017; Yasue, 2017). El mosaicismo somático podría explicar también parte de la variabilidad fenotípica en la aniridia, de manera similar a lo descrito en otras patologías oculares dominantes (Faivre, 2006; Beryozkin, 2016). Así, se ha descrito que algunos individuos mosaicos para microdeleciones de *PAX6* de bajo grado asocian fenotipos leves (Erez, 2010; Huynh, 2017). En nuestro trabajo, en todas las familias con confirmación de mosaicismo de nuestro trabajo se observan diferencias inter-generacionales en cuanto

a severidad del fenotipo ocular. En todas ellas, la causa más probable de esta variabilidad es la presencia del alelo mosaico en *PAX6* en un padre asintomático o con una presentación incompleta.

Los dos individuos mosaicos portadores de variantes *nonsense*, presentan distintas manifestaciones clínicas, pese a que ambos poseen el alelo mutante en torno al 15% en la línea somática. Uno de ellos presenta aniridia, aunque con sintomatología menos severa que la observada en su hija heterocigota. En cambio, el otro individuo presenta defectos iridianos prácticamente asintomáticos que no fueron detectados hasta la edad adulta tras el diagnóstico de su hijo. Por tanto, los niveles del alelo mutante en células sanguíneas de origen mesodérmico no son un indicador fiable del tipo y severidad de las manifestaciones oculares. Este último individuo presenta niveles similares de la mutación en otras células somáticas de distintos orígenes embrionarios, incluyendo saliva (ectodermo) y células vesicales (endodermo). Por ello, se podría pensar que la mutación debería haberse originado en los primeros estadios del desarrollo y estaría presente en una proporción similar en tejidos oculares. Esto indicaría la existencia de un efecto umbral para el desarrollo de las distintas manifestaciones fenotípicas de la aniridia.

Existen otros factores que podrían influir en la expresión fenotípica de los mosaicos como el efecto mutacional específico de cada variante o la función específica de *PAX6* en las líneas celulares con mosaicismo, así como el momento concreto del desarrollo donde ocurre la mutación (Shaham, 2012; Gregory-Evans, 2014; Cvekl, 2017). Siguiendo esta línea, el umbral de actividad transcripcional de *PAX6* requerido para la diferenciación de las distintas estructuras oculares es variable a lo largo de la embriogénesis ocular, tal como se ha confirmado en modelos animales transgénicos o modificados genéticamente (Davis-Silverman, 2005; Yasue, 2017). En este sentido, se ha descrito que el mosaicismo somático de *PAX6* causa distintos grados de malformaciones oculares en embriones murinos editados con CRISPR en función del momento de aparición de la mutación y su proporción (Yasue, 2017). Por lo tanto, la identificación de individuos mosaicos mostrando diferencias en la expresividad fenotípica, tanto en humanos como en modelos animales, apoya la idea de que la dosis de *PAX6* juega un papel importante en el desarrollo ocular y en la modulación

fenotípica. De acuerdo con nuestros resultados y con los modelos de CRISPR previamente publicados, deducimos que el mosaicismo somático podría ser la causa de algunos de los casos de aniridia sin mutaciones en *PAX6* o incluso explicar formas incompletas de aniridia o casos leves de colobomas iridianos aislados (Hanson, 1999). Sin embargo, dadas las limitaciones para obtener ciertos tejidos diana, existe una gran probabilidad de no poder detectar mosaicos de bajo grado o aquellos únicamente localizados en tejidos oculares (Ohuchi, 2018).

En resumen, nuestro trabajo confirma la hipótesis de que el mosaicismo germinal en aniridia puede explicar la recurrencia de hijos afectados procedentes de parejas asintomáticas. Además, demostramos por primera vez que algunos casos de variabilidad fenotípica intrafamiliar en aniridia pueden estar causados por la presencia de mosaicismo parental. En consecuencia nuestros hallazgos tienen importantes implicaciones clínicas para el consejo genético y el seguimiento oftalmológico en la aniridia congénita y otras malformaciones oculares relacionadas.

Finalmente, nuestro trabajo reafirma la necesidad de tener en consideración mosaicismo somático y/o gonosomal como factor genético para explicar: casos extraordinarios de recurrencia, casos de variabilidad fenotípica intrafamiliar, algunos casos con eventos de aparición *de novo* e incluso pacientes con aniridia sin aparente mutación en *PAX6*.

5.4. Actualización del espectro mutacional en nuestra cohorte de aniridia

En este trabajo, hemos realizado un exhaustivo análisis genético en las familias de nuestra cohorte ampliando el estudio diagnóstico previamente realizado. Esta ampliación consistió en el análisis de las regiones no codificantes de *PAX6* y en el estudio de otros genes asociados a patologías relacionadas, lo que nos ha permitido calcular el rendimiento actual del diagnóstico genético en la aniridia, utilizando distintos algoritmos de estudio.

Además, hemos evaluado la patogenicidad de las variantes de significado incierto mediante estudios funcionales, lo que nos ha permitido la reclasificación de la mayoría de ellas a variantes probablemente patogénicas.

Por todo ello, este trabajo proporciona datos actualizados del espectro mutacional en la aniridia en nuestra cohorte, la mayor recopilada en población española, así como la prevalencia de los genes asociados. Otro aspecto que ha sido puesto en evidencia es la identificación de casos de mosaicismo parental en nuestra cohorte.

Las CNVs que implican regiones aguas abajo de *PAX6* son una causa frecuente de aniridia, en nuestra cohorte representan un 14% de los casos sin alteración en *PAX6* (5/36). Este trabajo señala la necesidad de implementar el estudio de variantes en estas regiones reguladoras en el diagnóstico genético de la aniridia, lo cual se podría solventar fácilmente mediante el uso de *arrays* de CGH que incluyeran estas regiones. En nuestro laboratorio no sólo utilizamos *arrays*, tal como publicamos recientemente (Blanco-Kelly, 2017), también disponemos de paneles de NGS con gran profundidad de lectura para la captura de *enhancers* 3', que permiten analizar simultáneamente CNVs y SNVs con una única herramienta. Nuestra estrategia permite agilizar el diagnóstico molecular tanto de aniridia, como del síndrome de WAGR a un relativo bajo coste mediante aproximaciones robustas.

Esta Tesis Doctoral no ha profundizado en las posibles correlaciones genotipo-fenotipo que puedan existir en nuestra cohorte. Dada la existente variabilidad fenotípica de las alteraciones iridianas y al solapamiento con otras disgenesias del segmento anterior, la realización de este estudio requiere de una exhaustiva revisión de fenotipos, lo que ha quedado fuera del objeto de este trabajo. Sin embargo, en los resultados del análisis genético realizado en nuestra cohorte, se puede observar que las alteraciones en *PAX6*, tanto SNVs como CNVs, dan lugar principalmente a aniridia clásica. Este fenotipo puede asociarse a formas de síndrome de WAGR cuando se produce una delección de genes contiguos afectando a *PAX6* y *WT1*. Por otro lado, los casos caracterizados que estaban clasificados como aniridia-like presentan mutaciones en varios genes, principalmente variantes atípicas en *PAX6*, incluyendo *missense*, variantes en el 5'UTR y de *splicing*, pero también presentan variantes en *FOXC1* y *PITX2*.

A la vista de estos resultados, en esta Tesis Doctoral queremos proponer un nuevo algoritmo para el diagnóstico genético de la aniridia. Tal como se describe en la Figura

D1, este combina el cribado de las regiones codificantes de *PAX6* y de otros genes, que se realiza actualmente en nuestro laboratorio mediante exoma clínico, con el análisis de las regiones no codificantes de *PAX6* (Figura D1) en los casos negativos. El estudio genético de la aniridia debería incluir un cribado de SNVs y CNVs, bien de forma simultánea o por separado en función de las técnicas utilizadas. En el caso de utilizarse NGS como técnica de cribado, este análisis podría ser realizado simultáneamente mediante estrategias de exoma clínico, siempre que se incluya el gen *WT1* para el análisis del síndrome de WAGR y un análisis específico de CNVs. En cualquier caso, las CNVs identificadas deberían ser verificadas mediante una segunda técnica, recomendándose el aCGH para una mejor definición de los puntos de corte y la identificación de posibles deleciones de las regiones 3' reguladoras de *PAX6*. En el análisis de las regiones no codificantes, que no suelen estar incluidas en las estrategias de exoma clínico, deberían ser cribadas con diseños personalizados o en un futuro, mediante el análisis de WGS.

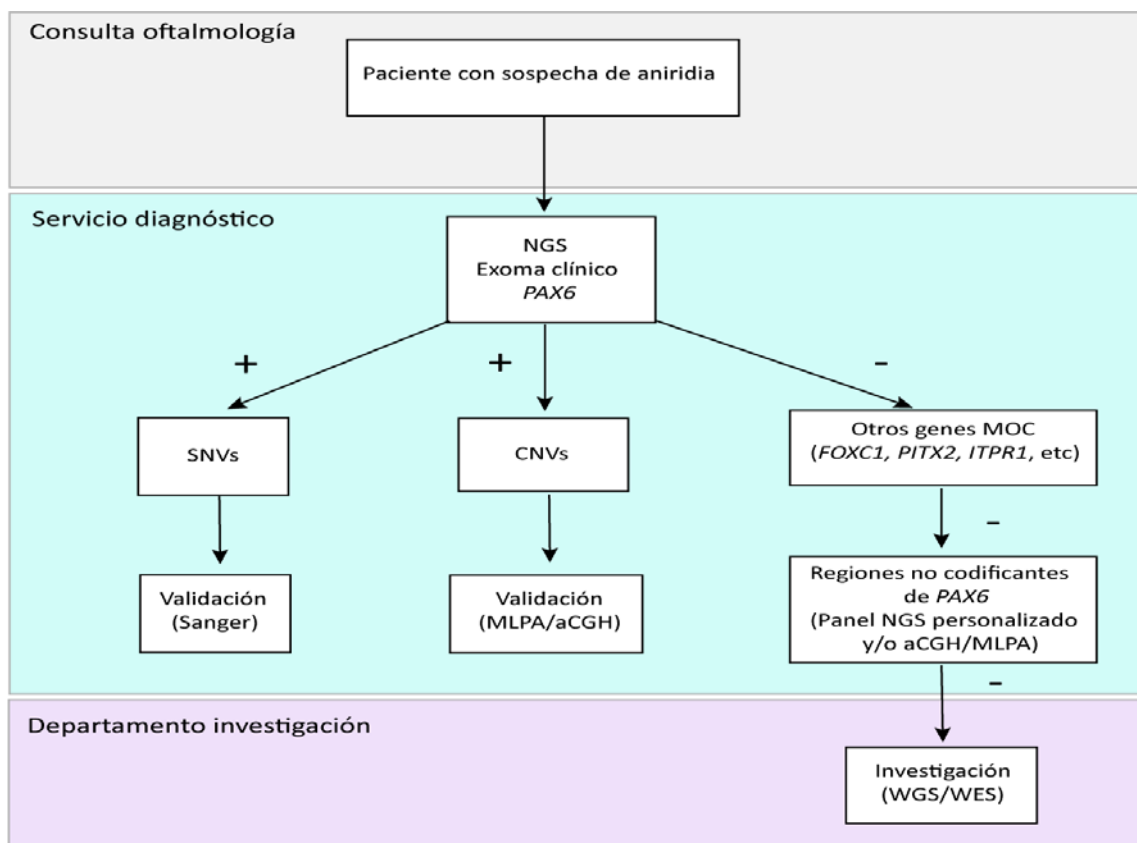


Figura D1. Nuevo algoritmo de diagnóstico genético de la aniridia propuesto tras los resultados obtenidos en el presente trabajo.

Conclusiones

- 1- El estudio de nuevos mecanismos que llevan a la aniridia mediante la incorporación de tecnología de última generación, aCGH y NGS, nos ha permitido **identificar variantes en las regiones no codificantes y cis-reguladoras de PAX6** en un 10% de nuestra cohorte.
- 2- El uso de aCGH personalizados para el *locus* 11p13 permite acotar los puntos de corte de las microdeleciones afectando a los enhancers situados a 3' de *PAX6* y con ello, hemos delimitado una **nueva región crítica para la regulación transcripcional de PAX6** de solo 18 Kb. Esta incluye al *cis*-regulador E180, recientemente identificado, lo que pone en evidencia su posible implicación en la patogénesis de la aniridia. Por el contrario, esta región excluye al *cis*-regulador SIMO, previamente considerado como el principal responsable de la haploinsuficiencia de *PAX6* asociada a estas microdeleciones de los reguladores 3'.
- 3- Este estudio ha **ampliado el espectro mutacional asociado a la región 5'UTR de PAX6** gracias a la identificación y estudio de 5 nuevas variantes. Entre ellas, aparece sobrerrepresentada la nueva variante c.-129+1G>A identificada en 4 familias no relacionadas. Este trabajo no ha detectado variantes potencialmente patogénicas en el 3'UTR de *PAX6* en nuestra cohorte.
- 4- Estudios de **caracterización funcional** realizados mediante una estrategia de midigenes para 12 variantes de significado incierto localizadas en el 5'UTR o en el exón 6 de *PAX6* han confirmado la implicación de 10 de ellas en la desregulación del *splicing* y con ello, han permitido su **reclasificación como variantes probablemente patogénicas**. Las dos variantes restantes, ambas *missense*, modifican el plegamiento 3D de la proteína en modelizaciones *in silico* y, por tanto, también fueron consideradas causales.
- 5- Estos **estudios de expresión *in vitro*** nos han permitido establecer que existe una **compleja regulación del *splicing* en el exón 6** que estaría mediada por al menos 6 donadores alternativos, exónicos e intrónicos, que compiten con el canónico, así como por elementos reguladores del *splicing*. Nuestro trabajo indica que todos, excepto el donador en posición c.157, participan en mecanismos patogénicos asociados a 5 variantes exónicas o del sitio donador de *splicing*. Además hemos confirmado la existencia de 3 donadores críticos

- en posiciones exónicas, c.157, c.173 y c.250, y hemos identificado por primera vez la implicación de otros 3 donadores crípticos, c.177, c.331 y c.357+6, en la etiopatología de la aniridia.
- 6- Nuestro trabajo ha **confirmado por primera vez el mosaicismo somático y/o gonosomal en PAX6** como causa de recurrencia y/o variabilidad intrafamiliar en patologías asociadas a *PAX6*. Además, en una de las familias, el análisis del ADN seminal del individuo mosaico permitió calcular que el riesgo real de transmisión de la variante su descendencia era mucho mayor de lo que se hubiese esperado para un caso esporádico. Todo ello presenta importantes implicaciones en el consejo genético de las familias. En un futuro, a medida que se implementen nuevas técnicas de mayor sensibilidad para la detección de alelos de baja frecuencia, la posibilidad de mosaicismo parental deberá ser tomada en cuenta en el diagnóstico genético de la aniridia en casos esporádicos con mutaciones *de novo* para calcular el riesgo de recurrencia en padres asintomáticos.
 - 7- Nuestro trabajo ha **identificado variantes patogénicas en los genes FOXC1 y PITX2** en el 4% de nuestra cohorte, principalmente **en casos con fenotipo de aniridia-like**. Estos fenotipos atípicos también están presentes en portadores de variantes *missense* en *PAX6*, así como de variantes localizadas en el 5'UTR y/o en sitios no canónicos de *splicing*. Nuestros hallazgos permiten concluir que **PAX6 es el principal gen asociado al fenotipo clásico de aniridia**.
 - 8- En nuestra cohorte **se detecta un número considerable de VUS**, 10%, por lo que es necesario realizar **posteriores estudios funcionales** para valorar su posible patogenicidad y con ello, **establecer una correcta interpretación clínica de las mismas**. Nuestros resultados demuestran que algunas variantes consideradas VUS según los criterios actuales de clasificación, incluidas variantes sinónimas, pueden no ser neutras y se debería valorar en ellas una posible modificación del proceso de *splicing*.
 - 9- En resumen, nuestros estudios indican que el uso de NGS y arrays de CGH con diseños personalizados, junto con el estudio funcional de variantes de significado incierto, **augmenta el rendimiento diagnóstico en un 20%**, pasando del 63% tras el uso de algoritmos clásicos al 84% tras el cribado de las regiones

no codificantes y otros genes. Tras nuestros resultados **hemos propuesto un nuevo algoritmo de diagnóstico** en donde se propone evaluar de manera simultánea SNVs y CNVs mediante estrategias de NGS **con el fin de agilizar el diagnóstico molecular** de la aniridia y del síndrome de WAGR.

Bibliografía

A

- Acedo A, Hernandez-Moro C, Curiel-Garcia A, Diez-Gomez B, Velasco EA. Functional classification of BRCA2 DNA variants by splicing assays in a large minigene with 9 exons. *Hum Mutat.* Feb 2015;36(2):210-221.
- Acuna-Hidalgo R, Bo T, Kwint MP, et al. Post-zygotic Point Mutations Are an Underrecognized Source of De Novo Genomic Variation. *Am J Hum Genet.* Jul 02 2015;97(1):67-74.
- Acuna-Hidalgo R, Veltman JA, Hoischen A. New insights into the generation and role of de novo mutations in health and disease. *Genome Biol.* Nov 28 2016;17(1):241.
- Addis L, Ahn JW, Dobson R, et al. Microdeletions of ELP4 Are Associated with Language Impairment, Autism Spectrum Disorder, and Mental Retardation. *Hum Mutat.* Sep 2015;36(9):842-850.
- Albert S, Garanto A, Sangermano R, et al. Identification and Rescue of Splice Defects Caused by Two Neighboring Deep-Intronic ABCA4 Mutations Underlying Stargardt Disease. *Am J Hum Genet.* Apr 5 2018;102(4):517-527.
- Alkemade PP, van Balen AT. Hereditary epithelial dystrophy of the cornea: Meesmann type. *Br J Ophthalmol.* Oct 1966;50(10):603-605.
- Ansari M, Rainger J, Hanson IM, et al. Genetic Analysis of 'PAX6-Negative' Individuals with Aniridia or Gillespie Syndrome. *PLoS One.* 2016;11(4):e0153757.
- Axton RA, Hanson IM, Love J, Seawright A, Prosser J, van Heyningen V. Combined SSCP/heteroduplex analysis in the screening for PAX6 mutations. *Mol Cell Probes.* Aug 1997;11(4):287-292.
- Azuma N, Nishina S, Yanagisawa H, Okuyama T, Yamada M. PAX6 missense mutation in isolated foveal hypoplasia. *Nat Genet.* Jun 1996;13(2):141-142.
- Azuma N, Yamaguchi Y, Handa H, Hayakawa M, Kanai A, Yamada M. Missense mutation in the alternative splice region of the PAX6 gene in eye anomalies. *Am J Hum Genet.* Sep 1999;65(3):656-663.
- Azuma N, Yamaguchi Y, Handa H, et al. Mutations of the PAX6 gene detected in patients with a variety of optic-nerve malformations. *Am J Hum Genet.* Jun 2003;72(6):1565-1570.

B

- Bamiou DE, Campbell NG, Musiek FE, et al. Auditory and verbal working memory deficits in a child with congenital aniridia due to a PAX6 mutation. *Int J Audiol.* Apr 2007;46(4):196-202.
- Bayrakli F, Guney I, Bayri Y, et al. A novel heterozygous deletion within the 3' region of the PAX6 gene causing isolated aniridia in a large family group. *J Clin Neurosci.* Dec 2009;16(12):1610-1614.
- Beccari L, Marco-Ferreres R, Bovolenta P. The logic of gene regulatory networks in early vertebrate forebrain patterning. *Mech Dev.* Feb 2013;130(2-3):95-111.
- Bentley CA, Zidehsarai MP, Grindley JC, Parlow AF, Barth-Hall S, Roberts VJ. Pax6 is implicated in murine pituitary endocrine function. *Endocrine.* Apr 1999;10(2):171-177.
- Beryozkin A, Levy G, Blumenfeld A, et al. Genetic Analysis of the Rhodopsin Gene Identifies a Mosaic Dominant Retinitis Pigmentosa Mutation in a Healthy Individual. *Invest Ophthalmol Vis Sci.* Mar 2016;57(3):940-947.
- Bhatia S, Bengani H, Fish M, et al. Disruption of autoregulatory feedback by a mutation in a remote, ultraconserved PAX6 enhancer causes aniridia. *Am J Hum Genet.* Dec 05 2013;93(6):1126-1134.
- Bhatia S, Monahan J, Ravi V, et al. A survey of ancient conserved non-coding elements in the PAX6 locus reveals a landscape of interdigitated cis-regulatory archipelagos. *Dev Biol.* Mar 15 2014;387(2):214-228.
- Bhatia S, Kleinjan DA. Disruption of long-range gene regulation in human genetic disease: a kaleidoscope of general principles, diverse mechanisms and unique phenotypic consequences. *Hum Genet.* Feb 05 2014b:133:815-845.
- Blanco-Kelly F, Palomares M, Vallespin E, et al. Improving molecular diagnosis of aniridia and WAGR syndrome using customized targeted array-based CGH. *PLoS One.* 2017;12(2):e0172363.
- Blanco-Kelly F, Villaverde C, Lorda I, et al. Chapter 3. Genetic implications of aniridia. Spanish guidelines for the management of congenital aniridia. <http://s1.aniridia.eu/wordpress/wp-content/uploads/2014/10/Spanish-Guidelines-for-the-Management-of-Congenital-Aniridia.pdf>. 2016.

- Blanco-Kelly F, Villaverde-Montero C, Lorda-Sanchez I, Millan JM, Trujillo-Tiebas MJ, Ayuso C. Guidelines for genetic study of aniridia. *Arch Soc Esp Oftalmol*. Apr 2013;88(4):145-152.
- Bobilev AM, McDougal ME, Taylor WL, Geisert EE, Netland PA, Lauderdale JD. Assessment of PAX6 alleles in 66 families with aniridia. *Clin Genet*. Jun 2016;89(6):669-677.
- Brillen AL, Schoneweis K, Walotka L, et al. Succession of splicing regulatory elements determines cryptic 5ss functionality. *Nucleic Acids Res*. Apr 20 2017;45(7):4202-4216.
- Brown A, McKie M, van Heyningen V, Prosser J. The Human PAX6 Mutation Database. *Nucleic Acids Res*. Jan 01 1998;26(1):259-264.
- Buratti E, Chivers M, Hwang G, Vorechovsky I. DBASS3 and DBASS5: databases of aberrant 3'- and 5'-splice sites. *Nucleic Acids Res*. Jan 2011;39(Database issue):D86-91.
- Byers PH. Killing the messenger: new insights into nonsense-mediated mRNA decay. *J Clin Invest*. Jan 2002;109(1):3-6.

C

- Carvalho DR, Medeiros JEG, Ribeiro DSM, Martins B, Sobreira NLM. Additional features of Gillespie syndrome in two Brazilian siblings with a novel ITPR1 homozygous pathogenic variant. *Eur J Med Genet*. Mar 2017;61(3):134-138.
- Casas-Llera P, Siverio A, Esquivel G, Bautista C, Alio JL. Spectral-domain optical coherence tomography foveal morphology as a prognostic factor for vision performance in congenital aniridia. *Eur J Ophthalmol*. Dec 17 2018;1120672118818352.
- Ceroni F, Aguilera-Garcia D, Chassaing N, et al. New GJA8 variants and phenotypes highlight its critical role in a broad spectrum of eye anomalies. *Hum Genet*. Feb 20 2018.
- Chao LY, Mishra R, Strong LC, Saunders GF. Missense mutations in the DNA-binding region and termination codon in PAX6. *Hum Mutat*. Feb 2003;21(2):138-145.
- Cheng F, Song W, Kang Y, Yu S, Yuan H. A 556 kb deletion in the downstream region of the PAX6 gene causes familial aniridia and other eye anomalies in a Chinese family. *Mol Vis*. 2011;17:448-455.

- Chou CM, Nelson C, Tarle SA, et al. Biochemical Basis for Dominant Inheritance, Variable Penetrance, and Maternal Effects in RBP4 Congenital Eye Disease. *Cell*. Apr 23 2015;161(3):634-646.
- Clericuzio C, Hingorani M, Crolla JA, van Heyningen V, Verloes A. Clinical utility gene card for: WAGR syndrome. *Eur J Hum Genet*. Apr 2011;19(4).
- Corso-Diaz X, Borrie AE, Bonaguro R, et al. Absence of NR2E1 mutations in patients with aniridia. *Mol Vis*. 2012;18:2770-2782.
- Coutinho P, Pavlou S, Bhatia S, Chalmers KJ, Kleinjan DA, van Heyningen V. Discovery and assessment of conserved Pax6 target genes and enhancers. *Genome Res*. Aug 2011;21(8):1349-1359.
- Crolla JA, Cawdery JE, Oley CA, et al. A FISH approach to defining the extent and possible clinical significance of deletions at the WAGR locus. *J Med Genet*. Mar 1997;34(3):207-212.
- Crolla JA, van Heyningen V. Frequent chromosome aberrations revealed by molecular cytogenetic studies in patients with aniridia. *Am J Hum Genet*. Nov 2002;71(5):1138-1149.
- Cvekl A, Callaerts P. PAX6: 25th anniversary and more to learn. *Exp Eye Res*. Mar 2017;156:10-21.

D

- Davis LK, Meyer KJ, Rudd DS, et al. Pax6 3' deletion results in aniridia, autism and mental retardation. *Hum Genet*. May 2008;123(4):371-378.
- Davis N, Yoffe C, Raviv S, et al. Pax6 dosage requirements in iris and ciliary body differentiation. *Dev Biol*. Sep 1 2009;333(1):132-142.
- Davis-Silberman N, Kalich T, Oron-Karni V, et al. Genetic dissection of Pax6 dosage requirements in the developing mouse eye. *Hum Mol Genet*. Aug 1 2005;14(15):2265-2276.
- de la Hoya M, Soukariéh O, Lopez-Perolio I, et al. Combined genetic and splicing analysis of BRCA1 c.[594-2A>C; 641A>G] highlights the relevance of naturally occurring in-frame transcripts for developing disease gene variant classification algorithms. *Hum Mol Genet*. Jun 1 2016;25(11):2256-2268.
- DeLano WL, The PyMOL Molecular Graphics System. Schrödinger LLC. www.pymol.org. version 1.7; <http://www.pymol.org> 2002.

- D'Elia AV, Pellizzari L, Fabbro D, et al. A deletion 3' to the PAX6 gene in familial aniridia cases. *Mol Vis.* 2007;13:1245-1250.
- D'Elia AV, Pellizzari L, Fabbro D, et al. A deletion 3' to the PAX6 gene in familial aniridia cases. *Mol Vis.* 2007;13:1245-1250.
- Deml B, Reis LM, Lemyre E, Clark RD, Kariminejad A, Semina EV. Novel mutations in PAX6, OTX2 and NDP in anophthalmia, microphthalmia and coloboma. *Eur J Hum Genet.* Apr 2016;24(4):535-541.
- Dentici ML, Barresi S, Nardella M, et al. Identification of novel and hotspot mutations in the channel domain of ITPR1 in two patients with Gillespie syndrome. *Gene.* Sep 10 2017;628:141-145.
- Desplan C. Eye development: governed by a dictator or a junta? *Cell.* Dec 26 1997;91(7):861-864.
- Desviat LR, Perez B, Ugarte M. Minigenes to confirm exon skipping mutations. *Methods Mol Biol.* 2012;867:37-47.
- Dora N, Ou J, Kucerova R, Parisi I, West JD, Collinson JM. PAX6 dosage effects on corneal development, growth, and wound healing. *Dev Dyn.* May 2008;237(5):1295-1306.
- Drivas TG, Wojno AP, Tucker BA, Stone EM, Bennett J. Basal exon skipping and genetic pleiotropy: A predictive model of disease pathogenesis. *Sci Transl Med.* Jun 10 2015;7(291):291ra297.
- Dubey SK, Mahalaxmi N, Vijayalakshmi P, Sundaresan P. Mutational analysis and genotype-phenotype correlations in southern Indian patients with sporadic and familial aniridia. *Mol Vis.* 2015;21:88-97.
- Duncan MK, Kozmik Z, Cveklova K, Piatigorsky J, Cvekl A. Overexpression of PAX6(5a) in lens fiber cells results in cataract and upregulation of (alpha)5(beta)1 integrin expression. *J Cell Sci.* Sep 2000;113 (Pt 18):3173-3185.

E

- Epstein JA, Glaser T, Cai J, Jepeal L, Walton DS, Maas RL. Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing. *Genes Dev.* Sep 1 1994;8(17):2022-2034.

Erez A, Li J, Geraghty MT, et al. Mosaic deletion 11p13 in a child with dopamine beta-hydroxylase deficiency--case report and review of the literature. *Am J Med Genet A*. Mar 2010;152A(3):732-736.

F

Faivre L, Williamson KA, Faber V, et al. Recurrence of SOX2 anophthalmia syndrome with gonosomal mosaicism in a phenotypically normal mother. *Am J Med Genet A*. Mar 15 2006;140(6):636-639.

Fantes J, Redeker B, Breen M, et al. Aniridia-associated cytogenetic rearrangements suggest that a position effect may cause the mutant phenotype. *Hum Mol Genet*. Mar 1995;4(3):415-422.

Fantes J, Redeker B, Breen M, et al. Aniridia-associated cytogenetic rearrangements suggest that a position effect may cause the mutant phenotype. *Hum Mol Genet*. Mar 1995;4(3):415-422.

Favor J, Peters H, Hermann T, et al. Molecular characterization of Pax6(2Neu) through Pax6(10Neu): an extension of the Pax6 allelic series and the identification of two possible hypomorph alleles in the mouse *Mus musculus*. *Genetics*. Dec 2001;159(4):1689-1700.

Fernandez-San Jose P, Blanco-Kelly F, Corton M, et al. Prevalence of Rhodopsin mutations in autosomal dominant Retinitis Pigmentosa in Spain: clinical and analytical review in 200 families. *Acta Ophthalmol*. Feb 2015;93(1):e38-44.

Filatova AY, Vasilyeva TA, Marakhonov AV, Voskresenskaya AA, Zinchenko RA, Skoblov MY. Functional reassessment of PAX6 single nucleotide variants by in vitro splicing assay. *Eur J Hum Genet*. Mar 2019;27(3):488-493.

Firth LC, Baker NE. Retinal determination genes as targets and possible effectors of extracellular signals. *Dev Biol*. Mar 15 2009;327(2):366-375.

Fischbach BV, Trout KL, Lewis J, Luis CA, Sika M. WAGR syndrome: a clinical review of 54 cases. *Pediatrics*. Oct 2005;116(4):984-988.

Fraile-Bethencourt E, Diez-Gomez B, Velasquez-Zapata V, Acedo A, Sanz DJ, Velasco EA. Functional classification of DNA variants by hybrid minigenes: Identification of 30 spliceogenic variants of BRCA2 exons 17 and 18. *PLoS Genet*. Mar 2017;13(3):e1006691.

- Fraile-Bethencourt E, Valenzuela-Palomo A, Diez-Gomez B, Caloca MJ, Gomez-Barrero S, Velasco EA. Minigene Splicing Assays Identify 12 Spliceogenic Variants of BRCA2 Exons 14 and 15. *Front Genet.* 2019;10:503.
- Franzoni A, Russo PD, Baldan F, et al. A CGH array procedure to detect PAX6 gene structural defects. *Mol Cell Probes.* Apr 2017;32:65-68.
- Fujimura N, Klimova L, Antosova B, Smolikova J, Machon O, Kozmik Z. Genetic interaction between Pax6 and beta-catenin in the developing retinal pigment epithelium. *Dev Genes Evol.* Apr 2015;225(2):121-128.
- Furuta Y, Hogan BL. BMP4 is essential for lens induction in the mouse embryo. *Genes Dev.* Dec 1 1998;12(23):3764-3775.

G

- Garcia CM, Huang J, Madakashira BP, et al. The function of FGF signaling in the lens placode. *Dev Biol.* Mar 1 2011;351(1):176-185.
- Gerber S, Alzayady KJ, Burglen L, et al. Recessive and Dominant De Novo ITPR1 Mutations Cause Gillespie Syndrome. *Am J Hum Genet.* May 05 2016;98(5):971-980.
- Glaser T, Jepeal L, Edwards JG, Young SR, Favor J, Maas RL. PAX6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. *Nat Genet.* Aug 1994;7(4):463-471.
- Glaser T, Walton DS, Maas RL. Genomic structure, evolutionary conservation and aniridia mutations in the human PAX6 gene. *Nat Genet.* Nov 1992;2(3):232-239.
- Graziano C, D'Elia AV, Mazzanti L, et al. A de novo nonsense mutation of PAX6 gene in a patient with aniridia, ataxia, and mental retardation. *Am J Med Genet A.* Aug 01 2007;143A(15):1802-1805.
- Gregory-Evans CY, Wallace VA, Gregory-Evans K. Gene networks: dissecting pathways in retinal development and disease. *Prog Retin Eye Res.* Mar 2013;33:40-66.
- Gregory-Evans CY, Wang X, Wasan KM, Zhao J, Metcalfe AL, Gregory-Evans K. Postnatal manipulation of Pax6 dosage reverses congenital tissue malformation defects. *J Clin Invest.* Jan 2014;124(1):111-116.

- Griffin C, Kleinjan DA, Doe B, van Heyningen V. New 3' elements control Pax6 expression in the developing pretectum, neural retina and olfactory region. *Mech Dev.* Mar 2002;112(1-2):89-100.
- Grocott T, Johnson S, Bailey AP, Streit A. Neural crest cells organize the eye via TGF-beta and canonical Wnt signalling. *Nat Commun.* 2011;2:265.
- Gronskov K, Olsen JH, Sand A, et al. Population-based risk estimates of Wilms tumor in sporadic aniridia. A comprehensive mutation screening procedure of PAX6 identifies 80% of mutations in aniridia. *Hum Genet.* Jul 2001;109(1):11-18.
- Gronskov K, Olsen JH, Sand A, et al. Population-based risk estimates of Wilms tumor in sporadic aniridia. A comprehensive mutation screening procedure of PAX6 identifies 80% of mutations in aniridia. *Hum Genet.* Jul 2001;109(1):11-18.
- Gronskov K, Rosenberg T, Sand A, Brondum-Nielsen K. Mutational analysis of PAX6: 16 novel mutations including 5 missense mutations with a mild aniridia phenotype. *Eur J Hum Genet.* Apr 1999;7(3):274-286.
- Gupta SK, Orr A, Bulman D, De Becker I, Guernsey DL, Neumann PE. A novel PAX6 frameshift mutation in a kindred from Atlantic Canada with familial aniridia. *Can J Ophthalmol.* Oct 1999;34(6):330-334.

H

- Hall HN, Williamson KA, FitzPatrick DR. The genetic architecture of aniridia and Gillespie syndrome. *Hum Genet.* Sep 2019;138(8-9):881-898.
- Hanson I, Churchill A, Love J, et al. Missense mutations in the most ancient residues of the PAX6 paired domain underlie a spectrum of human congenital eye malformations. *Hum Mol Genet.* Feb 1999;8(2):165-172.
- Hanson IM, Fletcher JM, Jordan T, et al. Mutations at the PAX6 locus are found in heterogeneous anterior segment malformations including Peters' anomaly. *Nat Genet.* Feb 1994;6(2):168-173.
- Hanson IM, Seawright A, Hardman K, et al. PAX6 mutations in aniridia. *Hum Mol Genet.* Jul 1993;2(7):915-920.
- Hart AW, Mella S, Mendrychowski J, van Heyningen V, Kleinjan DA. The developmental regulator Pax6 is essential for maintenance of islet cell function in the adult mouse pancreas. *PLoS One.* 2013;8(1):e54173.

- Hentze MW, Kulozik AE. A perfect message: RNA surveillance and nonsense-mediated decay. *Cell*. Feb 5 1999;96(3):307-310.
- Hingorani M, Hanson I, van Heyningen V. Aniridia. *Eur J Hum Genet*. Oct 2012;20(10):1011-1017.
- Hingorani M, Williamson KA, Moore AT, van Heyningen V. Detailed ophthalmologic evaluation of 43 individuals with PAX6 mutations. *Invest Ophthalmol Vis Sci*. Jun 2009;50(6):2581-2590.
- Hu P, Meng L, Ma D, et al. A novel 11p13 microdeletion encompassing PAX6 in a Chinese Han family with aniridia, ptosis and mental retardation. *Mol Cytogenet*. 2015;8(1):3.
- Huang J, Liu Y, Filas B, Gunhaga L, Beebe DC. Negative and positive auto-regulation of BMP expression in early eye development. *Dev Biol*. Nov 15 2015;407(2):256-264.
- Huynh MT, Boudry-Labis E, Duban B, et al. WAGR syndrome and congenital hypothyroidism in a child with a Mosaic 11p13 deletion. *Am J Med Genet A*. Jun 2017;173(6):1690-1693.

I

- Ito YA, Footz TK, Berry FB, et al. Severe molecular defects of a novel FOXC1 W152G mutation result in aniridia. *Invest Ophthalmol Vis Sci*. Aug 2009;50(8):3573-3579.

J

- Jidigam VK, Srinivasan RC, Patthey C, Gunhaga L. Apical constriction and epithelial invagination are regulated by BMP activity. *Biol Open*. Nov 30 2015;4(12):1782-1791.
- Johansson LF, van Dijk F, de Boer EN, et al. CoNVaDING: Single Exon Variation Detection in Targeted NGS Data. *Hum Mutat*. May 2016;37(5):457-464.

K

- Kammandel B, Chowdhury K, Stoykova A, Aparicio S, Brenner S, Gruss P. Distinct cis-essential modules direct the time-space pattern of the Pax6 gene activity. *Dev Biol*. Jan 1 1999;205(1):79-97.

- Kelberman D, Islam L, Holder SE, et al. Digenic inheritance of mutations in FOXC1 and PITX2 : correlating transcription factor function and Axenfeld-Rieger disease severity. *Hum Mutat.* Oct 2011;32(10):1144-1152.
- Kerem E, Konstan MW, De Boeck K, et al. Ataluren for the treatment of nonsense-mutation cystic fibrosis: a randomised, double-blind, placebo-controlled phase 3 trial. *Lancet Respir Med.* Jul 2014;2(7):539-547.
- Khan AO, Aldahmesh MA, Al-Amri A. Heterozygous FOXC1 mutation (M161K) associated with congenital glaucoma and aniridia in an infant and a milder phenotype in her mother. *Ophthalmic Genet.* Jun 2008;29(2):67-71.
- Kioussi C, O'Connell S, St-Onge L, et al. Pax6 is essential for establishing ventral-dorsal cell boundaries in pituitary gland development. *Proc Natl Acad Sci U S A.* Dec 7 1999;96(25):14378-14382.
- Kleinjan DA, Seawright A, Childs AJ, van Heyningen V. Conserved elements in Pax6 intron 7 involved in (auto)regulation and alternative transcription. *Dev Biol.* Jan 15 2004;265(2):462-477.
- Kleinjan DA, Seawright A, Mella S, et al. Long-range downstream enhancers are essential for Pax6 expression. *Dev Biol.* Nov 15 2006;299(2):563-581.
- Kleinjan DA, Seawright A, Schedl A, Quinlan RA, Danes S, van Heyningen V. Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6. *Hum Mol Genet.* Sep 15 2001;10(19):2049-2059.
- Kloss BA, Reis LM, Bremond-Gignac D, Glaser T, Semina EV. Analysis of FOXD3 sequence variation in human ocular disease. *Mol Vis.* 2012;18:1740-1749.
- Kopanos C, Tsiolkas V, Kouris A, et al. VarSome: The Human Genomic Variant Search Engine. *Oxford Bioinformatics.* 2018/10/30 2018.
- Krizman DB. Gene Isolation by Exon Trapping. In: Rapley R. (eds) *The Nucleic Acid Protocols Handbook.* Humana Press, Totowa, NJ. 2000.
- Kumar SU, Priyanka N, Sneha P, Doss CGP. Functional and structural characterization of missense mutations in PAX6 gene. *Front Biol.* Dec 2015:e1-9.

L

- Lakowski J, Majumder A, Lauderdale JD. Mechanisms controlling Pax6 isoform expression in the retina have been conserved between teleosts and mammals. *Dev Biol.* Jul 15 2007;307(2):498-520.
- Lauderdale JD, Wilensky JS, Oliver ER, Walton DS, Glaser T. 3' deletions cause aniridia by preventing PAX6 gene expression. *Proc Natl Acad Sci U S A.* Dec 5 2000;97(25):13755-13759.
- Lee HJ, Colby KA. A review of the clinical and genetic aspects of aniridia. *Semin Ophthalmol.* Sep-Nov 2013;28(5-6):306-312.
- Levin AA. Treating Disease at the RNA Level with Oligonucleotides. *N Engl J Med.* Jan 3 2019;380(1):57-70.
- Li PC, Yao Q, Ren X, et al. Analysis of PAX6 gene in a Chinese family with congenital aniridia. *Zhonghua Yan Ke Za Zhi.* Oct 2009;45(10):931-934.

M

- Machon O, Kreslova J, Ruzickova J, et al. Lens morphogenesis is dependent on Pax6-mediated inhibition of the canonical Wnt/beta-catenin signaling in the lens surface ectoderm. *Genesis.* Feb 2010;48(2):86-95.
- Malandrini A, Mari F, Palmeri S, et al. PAX6 mutation in a family with aniridia, congenital ptosis, and mental retardation. *Clin Genet.* Aug 2001;60(2):151-154.
- Mann IC. The Development of the Human Iris. *Br J Ophthalmol.* Oct 1925;9(10):495-512.
- Mathers PH, Grinberg A, Mahon KA, Jamrich M. The Rx homeobox gene is essential for vertebrate eye development. *Nature.* Jun 5 1997;387(6633):603-607.
- Maurya SK, Mishra R. Pax6 Binds to Promoter Sequence Elements Associated with Immunological Surveillance and Energy Homeostasis in Brain of Aging Mice. *Ann Neurosci.* May 2017;24(1):20-25.
- McBride DJ, Buckle A, van Heyningen V, Kleinjan DA. DNase hypersensitivity and ultraconservation reveal novel, interdependent long-range enhancers at the complex Pax6 cis-regulatory region. *PLoS One.* 2011;6(12):e28616.
- McDonald CM, Henricson EK, Abresch RT, et al. Long-term effects of glucocorticoids on function, quality of life, and survival in patients with Duchenne muscular dystrophy: a prospective cohort study. *Lancet.* Feb 3 2017;391(10119):451-461.

- McKenna A, Hanna M, Banks E, et al. The Genome Analysis Toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* Sep 2010;20(9):1297-1303.
- Miao Q, Ping X, Tang X, et al. Experimental assessment of novel PAX6 splicing mutations in two Chinese families with aniridia. *Gene.* Sep 30 2017;630:44-48.
- Mic FA, Molotkov A, Molotkova N, Duester G. Raldh2 expression in optic vesicle generates a retinoic acid signal needed for invagination of retina during optic cup formation. *Dev Dyn.* Oct 2004;231(2):270-277.
- Mirzayans F, Pearce WG, MacDonald IM, Walter MA. Mutation of the PAX6 gene in patients with autosomal dominant keratitis. *Am J Hum Genet.* Sep 1995;57(3):539-548.
- Mitchell TN, Free SL, Williamson KA, et al. Polymicrogyria and absence of pineal gland due to PAX6 mutation. *Ann Neurol.* May 2003;53(5):658-663.
- Morcillo J, Martinez-Morales JR, Trousse F, Fermin Y, Sowden JC, Bovolenta P. Proper patterning of the optic fissure requires the sequential activity of BMP7 and SHH. *Development.* Aug 2006;133(16):3179-3190.
- Murton NJ, Rehman I, Black GC, Inglehearn CF, Churchill AJ. A novel deletion (IVS11+3del4) identified in the human PAX6 gene in a patient with aniridia. *Hum Mutat.* Jun 2000;15(6):582.
- Myers CT, Hollingsworth G, Muir AM, et al. Parental Mosaicism in "De Novo" Epileptic Encephalopathies. *N Engl J Med.* Apr 26 2018;378(17):1646-1648.

N

- Nakayama T, Fisher M, Nakajima K, et al. *Xenopus* pax6 mutants affect eye development and other organ systems, and have phenotypic similarities to human aniridia patients. *Dev Biol.* Dec 15 2015;408(2):328-344.
- Nallathambi J, Neethirajan G, Shashikant S, Vijayalakshmi P, Sundaresan P. PAX6 missense mutations associated in patients with optic nerve malformation. *Mol Vis.* Mar 30 2006;12:236-242.
- Nanan KK, Ocheltree C, Sturgill D, et al. Independence between pre-mRNA splicing and DNA methylation in an isogenic minigene resource. *Nucleic Acids Res.* Dec 15 2017;45(22):12780-12797.

Navratilova P, Fredman D, Hawkins TA, Turner K, Lenhard B, Becker TS. Systematic human/zebrafish comparative identification of cis-regulatory activity around vertebrate developmental transcription factor genes. *Dev Biol.* Mar 15 2009;327(2):526-540.

Neethirajan G, Krishnadas SR, Vijayalakshmi P, Shashikant S, Sundaresan P. PAX6 gene variations associated with aniridia in south India. *BMC Med Genet.* Apr 16 2004;5:9.

Nelson LB, Spaeth GL, Nowinski TS, Margo CE, Jackson L. Aniridia. A review. *Surv Ophthalmol.* May-Jun 1984;28(6):621-642.

Netland PA, Scott ML, Boyle JWT, Lauderdale JD. Ocular and systemic findings in a survey of aniridia subjects. *J AAPOS.* Dec 2011;15(6):562-566.

Netter FH. *Atlas de Anatomía Humana.* Elsevier Inc. 2012;4ª edición:107, lámina 187.

O

Ohuchi H, Sato K, Habuta M, Fujita H, Bando T. Congenital eye anomalies: More mosaic than thought? *Congenit Anom (Kyoto).* May 2018;59(3):56-73.

P

Perez-Solorzano S, Chacon-Camacho OF, Astiazaran MC, Ledesma-Gil G, Zenteno JC. PAX6 allelic heterogeneity in Mexican congenital aniridia patients: expanding the mutational spectrum with seven novel pathogenic variants. *Clin Exp Ophthalmol.* Dec 2017;45(9):875-883.

Plaisancie J, Tarilonte M, Ramos P, et al. Implication of non-coding PAX6 mutations in aniridia. *Hum Genet.* Oct 2018;137(10):831-846.

Ploski JE, Shamsheer MK, Radu A. Paired-type homeodomain transcription factors are imported into the nucleus by karyopherin 13. *Mol Cell Biol.* Jun 2004;24(11):4824-4834.

Prosser J, van Heyningen V. PAX6 mutations reviewed. *Hum Mutat.* 1998;11(2):93-108.

Q

Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics*. Mar 1 2015;31(5):761-763.

Quinlan AR, Hall IM. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics*. Mar 15 2010;26(6):841-842.

R

Rajagopal R, Huang J, Dattilo LK, et al. The type I BMP receptors, *Bmpr1a* and *Acvr1*, activate multiple signaling pathways to regulate lens formation. *Dev Biol*. Nov 15 2009;335(2):305-316.

Ravi V, Bhatia S, Gautier P, et al. Sequencing of *Pax6* loci from the elephant shark reveals a family of *Pax6* genes in vertebrate genomes, forged by ancient duplications and divergences. *PLoS Genet*. 2013;9(1):e1003177.

Raviv S, Bharti K, Rencus-Lazar S, et al. PAX6 regulates melanogenesis in the retinal pigmented epithelium through feed-forward regulatory interactions with MITF. *PLoS Genet*. 2014;10(5):e1004360.

Redeker EJ, de Visser AS, Bergen AA, Mannens MM. Multiplex ligation-dependent probe amplification (MLPA) enhances the molecular diagnosis of aniridia and related disorders. *Mol Vis*. 2008;14:836-840.

Reed TE, Falls HF. A pedigree of aniridia with a discussion of germinal mosaicism in man. *Am J Hum Genet*. Mar 1955;7(1):28-38.

Remington LA. *Clinical Anatomy and Physiology of the Visual System*. Elsevier Inc. 2012;3rd edition(Chapters 7 and 13):124, 125 and 234; figures 127-121, 127-122 and 113-121.

Rentzsch P, Witten D, Cooper GM, Shendure J, Kircher M. CADD: predicting the deleteriousness of variants throughout the human genome. *Nucleic Acids Res*. Jan 8 2019;47(D1):D886-D894.

Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. May 2015;17(5):405-424.

Richardson R, Hingorani M, Van Heyningen V, Gregory-Evans C, Moosajee M. Clinical utility gene card for: Aniridia. *Eur J Hum Genet*. Nov 2016;24(11).

- Richardson R, Smart M, Tracey-White D, Webster AR, Moosajee M. Mechanism and evidence of nonsense suppression therapy for genetic eye disorders. *Exp Eye Res.* Feb 2017;155:24-37.
- Riera M, Wert A, Nieto I, Pomares E. Panel-based whole exome sequencing identifies novel mutations in microphthalmia and anophthalmia patients showing complex Mendelian inheritance patterns. *Mol Genet Genomic Med.* Nov 2017;5(6):709-719.
- Robinson DO, Howarth RJ, Williamson KA, van Heyningen V, Beal SJ, Crolla JA. Genetic analysis of chromosome 11p13 and the PAX6 gene in a series of 125 cases referred with aniridia. *Am J Med Genet A.* Mar 1 2008;146A(5):558-569.
- Roux LN, Petit I, Domart R, et al. Modeling of Aniridia-Related Keratopathy by CRISPR/Cas9 Genome Editing of Human Limbal Epithelial Cells and Rescue by Recombinant PAX6 Protein. *Stem Cells.* Sep 2018;36(9):1421-1429.
- Rush CC. Congenital Aniridia. *Trans Am Ophthalmol Soc.* 1926;24:332-341.

S

- Sadagopan KA, Liu GT, Capasso JE, Wuthisiri W, Keep RB, Levin AV. Aniridia-like phenotype caused by 6p25 dosage aberrations. *Am J Med Genet A.* Mar 2015;167A(3):524-528.
- Sahel JA, Marazova K. Toward postnatal reversal of ocular congenital malformations. *J Clin Invest.* Jan 2014;124(1):81-84.
- Sangermano R, Khan M, Cornelis SS, et al. ABCA4 midigenes reveal the full splice spectrum of all reported noncanonical splice site variants in Stargardt disease. *Genome Res.* Jan 2017;28(1):100-110.
- Sannan NS, Gregory-Evans CY, Lyons CJ, et al. Correlation of novel PAX6 gene abnormalities in aniridia and clinical presentation. *Can J Ophthalmol.* Dec 2017;52(6):570-577.
- Sanz DJ, Acedo A, Infante M, et al. A high proportion of DNA variants of BRCA1 and BRCA2 is associated with aberrant splicing in breast/ovarian cancer patients. *Clin Cancer Res.* Mar 15 2010;16(6):1957-1967.
- Schedl A, Ross A, Lee M, et al. Influence of PAX6 gene dosage on development: overexpression causes severe eye abnormalities. *Cell.* Jul 12 1996;86(1):71-82.

- Schedl A, Ross A, Lee M, et al. Influence of PAX6 gene dosage on development: overexpression causes severe eye abnormalities. *Cell*. Jul 12 1996;86(1):71-82.
- Schilter KF, Reis LM, Schneider A, et al. Whole-genome copy number variation analysis in anophthalmia and microphthalmia. *Clin Genet*. Nov 2013;84(5):473-481.
- Schmidt-Sidor B, Szymanska K, Williamson K, et al. Malformations of the brain in two fetuses with a compound heterozygosity for two PAX6 mutations. *Folia Neuropathol*. 2009;47(4):372-382.
- Schwarz M, Cecconi F, Bernier G, et al. Spatial specification of mammalian eye territories by reciprocal transcriptional repression of Pax2 and Pax6. *Development*. Oct 2000;127(20):4325-4334.
- Shaham O, Menuchin Y, Farhy C, Ashery-Padan R. Pax6: A multi-level regulator of ocular development. *Prog Retin Eye Res*. May 3 2012.
- Shen MW, Arbab M, Hsu JY, et al. Predictable and precise template-free CRISPR editing of pathogenic variants. *Nature*. Nov 2018;563(7733):646-651.
- Shi Z, Xin H, Tian D, et al. Modeling human point mutation diseases in *Xenopus tropicalis* with a modified CRISPR/Cas9 system. *FASEB J*. Jun 2019;33(6):6962-6968.
- Shortt AJ, Tuft SJ. Ocular surface reconstruction. *Br J Ophthalmol*. Jul 2011;95(7):901-902.
- Simpson TI, Price DJ. Pax6; a pleiotropic player in development. *Bioessays*. Nov 2002;24(11):1041-1051.
- Singh S, Mishra R, Arango NA, Deng JM, Behringer RR, Saunders GF. Iris hypoplasia in mice that lack the alternatively spliced Pax6(5a) isoform. *Proc Natl Acad Sci U S A*. May 14 2002;99(10):6812-6815.
- Singh S, Tang HK, Lee JY, Saunders GF. Truncation mutations in the transactivation region of PAX6 result in dominant-negative mutants. *J Biol Chem*. Aug 21 1998;273(34):21531-21541.
- Sisodiya SM, Free SL, Williamson KA, et al. PAX6 haploinsufficiency causes cerebral malformation and olfactory dysfunction in humans. *Nat Genet*. Jul 2001;28(3):214-216.

- Skeens HM, Brooks BP, Holland EJ. Congenital aniridia variant: minimally abnormal irides with severe limbal stem cell deficiency. *Ophthalmology*. Jul 2011;118(7):1260-1264.
- Smith AN, Miller LA, Song N, Taketo MM, Lang RA. The duality of beta-catenin function: a requirement in lens morphogenesis and signaling suppression of lens fate in periocular ectoderm. *Dev Biol*. Sep 15 2005;285(2):477-489.
- Solomon BD, Pineda-Alvarez DE, Balog JZ, et al. Compound heterozygosity for mutations in PAX6 in a patient with complex brain anomaly, neonatal diabetes mellitus, and microphthalmia. *Am J Med Genet A*. Nov 2009;149A(11):2543-2546.
- Souzeau E, Rudkin AK, Dubowsky A, et al. PAX6 molecular analysis and genotype-phenotype correlations in families with aniridia from Australasia and Southeast Asia. *Mol Vis*. 2018;24:261-273.
- Staden, P. RJD, K. BJ. *Analysing Sequences Using the Staden Package and EMBOSS. Introduction to Bioinformatics. A Theoretical and Practical Approach*. Eds. Stephen A. Krawetz and David D. Womble. Human Press Inc. Totawa, NJ 07512 2003.
- Stosser MB, Lindy AS, Butler E, et al. High frequency of mosaic pathogenic variants in genes causing epilepsy-related neurodevelopmental disorders. *Genet Med*. Apr 2017;20(4):403-410.
- Sun J, Rockowitz S, Xie Q, Ashery-Padan R, Zheng D, Cvekl A. Identification of in vivo DNA-binding mechanisms of Pax6 and reconstruction of Pax6-dependent gene regulatory networks during forebrain and lens development. *Nucleic Acids Res*. Aug 18 2015;43(14):6827-6846.

T

- Tabata H, Koinui A, Ogura A, Nishihara D, Yamamoto H. A novel nuclear localization signal spans the linker of the two DNA-binding subdomains in the conserved paired domain of Pax6. *Genes Genet Syst*. Sep 15 2018;93(2):75-81.
- Tang HK, Chao LY, Saunders GF. Functional analysis of paired box missense mutations in the PAX6 gene. *Hum Mol Genet*. Mar 1997;6(3):381-386.
- Tarilonte M, Morin M, Ramos P, et al. Parental Mosaicism in PAX6 Causes Intra-Familial Variability: Implications for Genetic Counseling of Congenital Aniridia and Microphthalmia. *Front Genet*. 2018;9:479.

- Terzic J, Saraga-Babic M. Expression pattern of PAX3 and PAX6 genes during human embryogenesis. *Int J Dev Biol.* Sep 1999;43(6):501-508.
- Thomas S, Thomas MG, Andrews C, et al. Autosomal-dominant nystagmus, foveal hypoplasia and presenile cataract associated with a novel PAX6 mutation. *Eur J Hum Genet.* Mar 2014;22(3):344-349.
- Thompson PJ, Mitchell TN, Free SL, et al. Cognitive functioning in humans with mutations of the PAX6 gene. *Neurology.* Apr 13 2004;62(7):1216-1218.
- Ticho BH, Hilchie-Schmidt C, Egel RT, Traboulsi EI, Howarth RJ, Robinson D. Ocular findings in Gillespie-like syndrome: association with a new PAX6 mutation. *Ophthalmic Genet.* Dec 2006;27(4):145-149.
- Ton CC, Hirvonen H, Miwa H, et al. Positional cloning and characterization of a paired box- and homeobox-containing gene from the aniridia region. *Cell.* Dec 20 1991;67(6):1059-1074.
- Tzoulaki I, White IM, Hanson IM. PAX6 mutations: genotype-phenotype correlations. *BMC Genet.* May 26 2005;6:27.

V

- van der Klift HM, Jansen AM, van der Steenstraten N, et al. Splicing analysis for exonic and intronic mismatch repair gene variants associated with Lynch syndrome confirms high concordance between minigene assays and patient RNA analyses. *Mol Genet Genomic Med.* Jul 2015;3(4):327-345.
- van Heyningen V, Williamson KA. PAX6 in sensory development. *Hum Mol Genet.* May 15 2002;11(10):1161-1167.
- van Raamsdonk CD, Tilghman SM. Dosage requirement and allelic expression of PAX6 during lens placode formation. *Development.* Dec 2000;127(24):5439-5448.
- Vasilyeva TA, Voskresenskaya AA, Kasmann-Kellner B, et al. Molecular analysis of patients with aniridia in Russian Federation broadens the spectrum of PAX6 mutations. *Clin Genet.* Dec 2017;92(6):639-644.
- Vincent MC, Pujo AL, Olivier D, Calvas P. Screening for PAX6 gene mutations is consistent with haploinsufficiency as the main mechanism leading to various ocular defects. *Eur J Hum Genet.* Feb 2003;11(2):163-169.
- Voskresenskaya A, Pozdeyeva N, Vasilyeva T, et al. Clinical and morphological manifestations of aniridia-associated keratopathy on anterior segment

optical coherence tomography and in vivo confocal microscopy. *Ocul Surf.* Oct 2017;15(4):759-769.

W

Wang K, Li M, Hakonarson H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* Sep 2010;38(16):e164.

Wang P, Sun W, Li S, Xiao X, Guo X, Zhang Q. PAX6 mutations identified in 4 of 35 families with microcornea. *Invest Ophthalmol Vis Sci.* Sep 19 2012;53(10):6338-6342.

Wang X, Gregory-Evans CY. Nonsense suppression therapies in ocular genetic diseases. *Cell Mol Life Sci.* May 2015;72(10):1931-1938.

Wang X, Gregory-Evans K, Wasan KM, Sivak O, Shan X, Gregory-Evans CY. Efficacy of Postnatal In Vivo Nonsense Suppression Therapy in a Pax6 Mouse Model of Aniridia. *Mol Ther Nucleic Acids.* Jun 16 2017;7:417-428.

Wawersik S, Purcell P, Rauchman M, Dudley AT, Robertson EJ, Maas R. BMP7 acts in murine lens placode development. *Dev Biol.* Mar 1 1999;207(1):176-188.

Wawrocka A, Budny B, Debicki S, Jamsheer A, Sowinska A, Krawczynski MR. PAX6 3' deletion in a family with aniridia. *Ophthalmic Genet.* Mar 2012;33(1):44-48.

Wawrocka A, Krawczynski MR. The genetics of aniridia - simple things become complicated. *J Appl Genet.* May 2018;59(2):151-159.

Weisschuh N, Wissinger B, Gramer E. A splice site mutation in the PAX6 gene which induces exon skipping causes autosomal dominant inherited aniridia. *Mol Vis.* 2012;18:751-757.

Williams AL, Bohnsack BL. Neural crest derivatives in ocular development: discerning the eye of the storm. *Birth Defects Res C Embryo Today.* Jun 2015;105(2):87-95.

Williams SC, Altmann CR, Chow RL, Hemmati-Brivanlou A, Lang RA. A highly conserved lens transcriptional control element from the Pax-6 gene. *Mech Dev.* May 1998;73(2):225-229.

X

Xie Q, Cvekl A. The orchestration of mammalian tissue morphogenesis through a series of coherent feed-forward loops. *J Biol Chem*. Dec 16 2011;286(50):43259-43271.

Xu HE, Rould MA, Xu W, Epstein JA, Maas RL, Pabo CO. Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. *Genes Dev*. May 15 1999;13(10):1263-1275.

Y

Yahalom C, Blumenfeld A, Hendler K, et al. Mild aniridia phenotype: an under-recognized diagnosis of a severe inherited ocular disease. *Graefes Arch Clin Exp Ophthalmol*. Nov 2018;256(11):2157-2164.

Yasuda T, Kajimoto Y, Fujitani Y, et al. PAX6 mutation as a genetic factor common to aniridia and glucose intolerance. *Diabetes*. Jan 2002;51(1):224-230.

Yasue A, Kono H, Habuta M, et al. Relationship between somatic mosaicism of Pax6 mutation and variable developmental eye abnormalities-an analysis of CRISPR genome-edited mouse embryos. *Sci Rep*. Dec 2017;7(1):53.

Yogarajah M, Matarin M, Vollmar C, et al. PAX6, brain structure and function in human adults: advanced MRI in aniridia. *Ann Clin Transl Neurol*. May 2016;3(5):314-330.

Z

Zambonin JL, Bellomo A, Ben-Pazi H, et al. Spinocerebellar ataxia type 29 due to mutations in ITPR1: a case series and review of this emerging congenital ataxia. *Orphanet J Rare Dis*. Jun 28 2017;12(1):121.

Zhang R, Linpeng S, Wei X, et al. Novel variants in PAX6 gene caused congenital aniridia in two Chinese families. *Eye (Lond)*. Jun 2017;31(6):956-961.

Zhang X, Qin G, Chen G, et al. Variants in TRIM44 Cause Aniridia by Impairing PAX6 Expression. *Hum Mutat*. Dec 2015;36(12):1164-1167.

Zhang X, Wang P, Li S, Xiao X, Guo X, Zhang Q. Mutation spectrum of PAX6 in Chinese patients with aniridia. *Mol Vis*. 2011;17:2139-2147.

Zhang X, Zhang Q, Tong Y, et al. Large novel deletions detected in Chinese families with aniridia: correlation between genotype and phenotype. *Mol Vis*. Feb 19 2011;17:548-557.

Anexos

Elemento regulador	Posición (hg19)	Tamaño (pb)	Distancia a PAX6 (Kb)	Localización	Expresión	Referencia
HS8B	chr11:31662110-31662778	668	-144	ELP4		Kleinjan, 2001
HS8	chr11:31662868-31663475	607	-143	ELP4		Kleinjan, 2001
HS5	chr11:31671019-31671690	671	-135	ELP4	Copa óptica, neuroretina y diencefalo	McBride DJ, 2011
HS3	chr11:31676764-31677347	583	-129	ELP4	Neuroretina	Kleinjan DA, 2001
HS2	chr11:31677439-31677991	552	-128	ELP4	Neuroretina	Kleinjan DA, 2001
SIMO	chr11:31685581-31686174	593	-120	ELP4	Cristalino, neuroretina, diencefalo, rombencéfalo	Kleinjan DA, 2001
E100	chr11:31734245-31734929	684	-71	ELP4	Neuroretina, diencefalo, bulbos olfatorios	Griffin C, 2002
E60A	chr11:31784779-31785426	647	-21	ELP4	Copa óptica, neuroretina y estructuras neurales	McBride DJ, 2011
7CE1	chr11:31820755-31821453	698	15	intrón PAX6	Ojo	Kleinjan DA, 2004
NRE	chr11:31825361-31826132	771	20	intrón PAX6	Neuroretina	Kammandel B, 1999
OCE1	chr11:31837809-31838293	484	32	5'UTR-intrónica		Bathia, 2013
agCNE13 (EE)	chr11:31843104-31843692	588	37	RNC1	Cristalino	Williams SC, 1998
agCNE12 (P)	chr11:31843629-31844067	438	38	RNC1	Páncreas	Kammandel B, 1999
agCNE9 (Up-10)	chr11:31847850-31848450	600	42	RNC1	Glándula pineal	Bhatia S, 2014
Mouse-like Paupar	chr11:31847922-31850844	2922	45	RNC1		Vance_2014
CTCF6 elemento	chr11:31892452-31893062	610	87	RNC1		Bathia, 2013
agCNE2 (E-200)	chr11:32052624-32053243	619	247	RNC1	Bulbos olfatorios, cerebelo	Ravi V, 2013
E-200 proximal	chr11:32052714-32053362	648	247	RNC1		Ravi V, 2013
E-200 distal	chr11:32053010-32053783	773	247	RNC1		Ravi V, 2013

Anexo 1: Listado de los 20 elementos cis-reguladores de PAX6 incluidos en el panel de NGS personalizado. Aparece el nombre, el tipo de variante detectada, su localización y su distancia al promotor de PAX6.

MUTATION NAME	PRIMER NAME	SECUENCE (5'-3')	TM (°C)	PRIMER SIZE
MINIGEN WT				
Minigen PAX6_ Ex1-2	tail- NotI -primer1F	ATT GCGGCCGC GGAGAGGGAGCATCCAATC	60	30
	tail- BamHI-Mrel - primer2R	ATAGGATCCC GCCGGCG GGCTTTCGCTGGAAGTAGAA	60	37
Minigen PAX6_ Ex3-4	tail- NotI-Mrel -primer3F	AAT GCGGCCGC CGCCGGCGCTCCCAAGCCAGACCAGAG	61	38
	tail- BamHI -primer4R	ATAGGATCCTACCGCAGCTTCGAAAATC	60	29
DIRECTED MUTATION				
c.-129+1G>A	c.-129+1G>A_F	GAAAGAAGCGGACTCATCTTTATGAGGCATCCTTTCT	78,28	37
	c.-129+1G>A_R	AGAAAGGATGCCTCATAAAGATGAGTCCGCTTCTTTC	78,28	37
c.-117_-118delTT	c.-117_-118delTT_F	GCGCGCCCCTAGTT--AGTCTTCCCCCTAA	78,71	28
	c.-117_-118delTT_R	TTAGGGGGAAGACT--AACTAGGGGCGCGC	78,71	28
c.-128-2delA	c.-128-2delA_F	CCTTTTTCTCCTTCTGTTTTGTCTT-GGGGGAAGACTTTA	78,45	39
	c.-128-2delA_R	TAAAGTCTTCCCCC-AAGACAAAACAGAAGGAGAAAAAGG	78,45	39
c.-52+1G>A	c.-52+1G>A_F	AGCCACCACAGA AACTTGTCTGAAATCTCGGATGTC	79,27	35
	c.-52+1G>A_R	GACATCCGAGATTT CAGACAAGTTCTGTGGTGGCT	79,27	35
c.-52+3_-52+6delinsTG	c.-52+3_-52+6delinsTG_F	CCGAGATTT CAGGCTTGTCTGTGGTGGCTGC	78,71	31
	c.-52+3_-52+6delinsTG_R	GCAGCCACCACAGACAAGCCTGAAATCTCGG	78,71	31

c.-52+5delG	c.-52+5delG_F	CCGAGATTTTCAGGCAA-TTCTGTGGTGGCTGC	78,98	31
	c.-52+5delG_R	GCAGCCACCACAGAA-TTGCCTGAAATCTCGG	78,98	31
RT-PCR				
pSPL3_SD6+FAM	SD6_F	TCTGAGTCACCTGGACAACC	60	20
pSPL3_SA2	SA2_R	ATCTCAGTGGTATTTGTGAGC	60	21

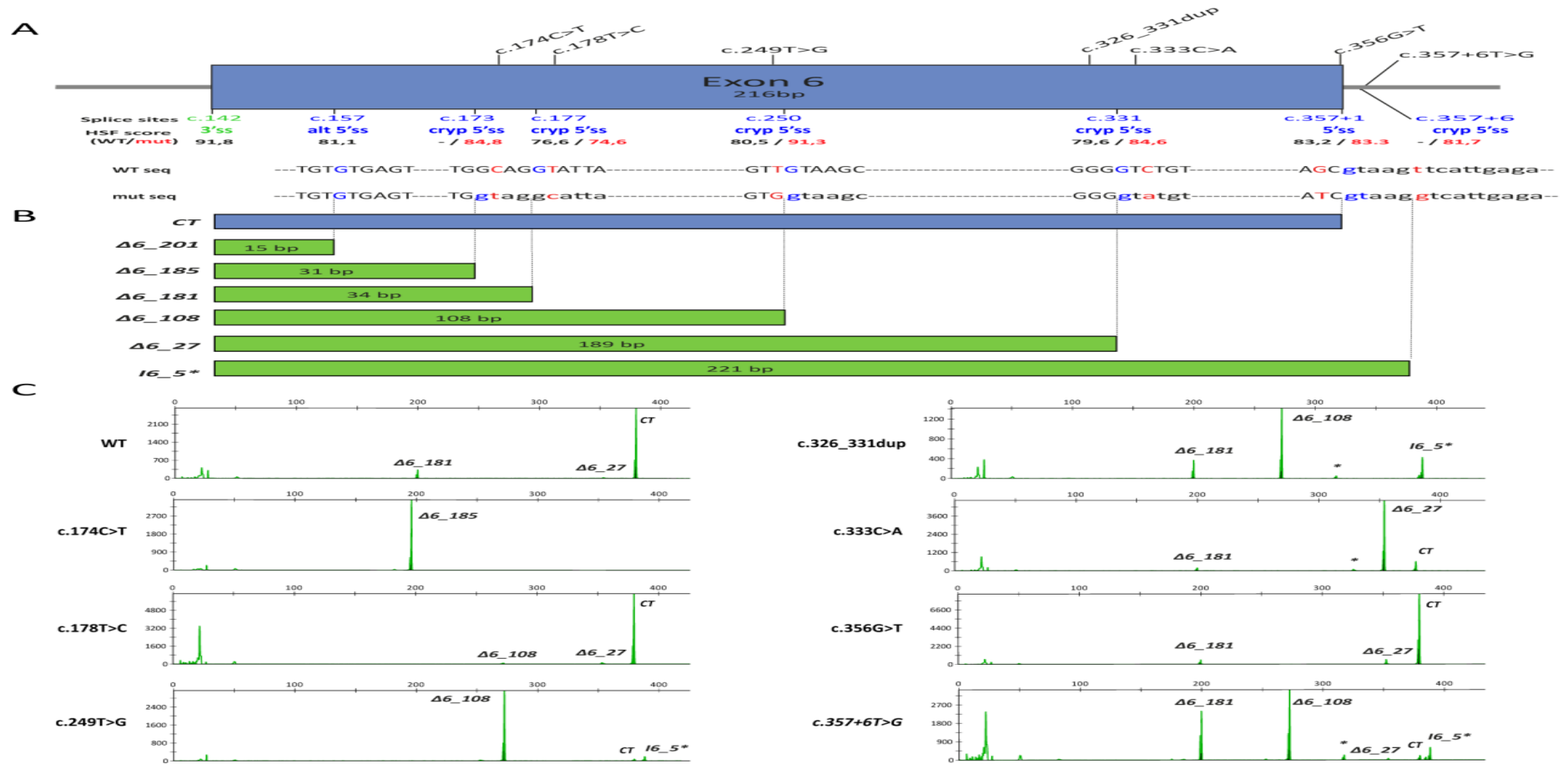
Anexo 2: Lista de oligonucleótidos utilizados en el apartado 4.1.

Nombre	Secuencia	Tm
Validación por Sanger		
PAX6_6F	GTGGTTTTCTGTCCACTTCC	57.3
PAX6_6R	AGGAGAGAGCATTGGGCTTA	57.3
Construcción Minigen "PAX6_exón 6"		
PAX6_int5_F	AATGCGGCCGCTGTAGTTTTAGGGCTACAAATGTAAT	75.6
PAX6_int6_R	ATAGGATCCTTGCAGCATGCAAATGAAGT	69.9
PAX6_c.174C>T_F	GCCAGTCTCGTAATACCTACCCAGAATTTACTCACA	71.7
PAX6_c.174C>T_R	TGTGAGTAAAATTCTGGGTAGGTATTACGAGACTGGC	71.7
PAX6_c.178T>C_F	GGAGCCAGTCTCGTAATGCCTGCCAGAATTTTAC	76.8
PAX6_c.178T>C_R	GTAAAATTCTGGGCAGGTATTACGAGACTGGCTCC	73.6
PAX6_c.249T>G_F	TGGGCTATTTTGCTTACCACTTCTGGAGTCGCTAC	74.7
PAX6_c.249T>G_R	GTAGCGACTCCAGAAGTGGTAAGCAAATAGCCCA	74.7
PAX6_c.326_331dup_F	GTTATCGTTGGTACAGACCCCTCCCCCTCGGACAGTAAT	80.0
PAX6_c.326_331dup_R	ATTACTGTCCGAGGGGGAGGGGGTCTGTACCAACGATAAC	80.0
PAX6_c.333C>A_F	GTTATCGTTGGTACATACCCCTCGGACAGTAAT	71.8
PAX6_c.333C>A_R	ATTACTGTCCGAGGGGGTATGTACCAACGATAAC	71.8
PAX6_c.356G>T_F	GTTCTCAATGAACTTACGATTGGTATGTTATCGTTGGTACAGAC	73.8
PAX6_c.356G>T_R	GTCTGTACCAACGATAACATACCAATCGTAAGTTCATTGAGAAC	73.8
PAX6_c.357+6T>G_F	GGGCAGATGTTCTCAATGCACCTTACGCTTGGTATGTT	76.8
PAX6_c.357+6T>G_R	AACATACCAAGCGTAAGGTCATTGAGAACATCTGCCC	76.1
Análisis ARN		
pSPL3_SD6_F	TCTGAGTCACCTTGGACAACC	60.1
pSPL3_SA2_R	ATCTCAGTGGTATTTGTGAGC	54.4
PAX6_c.148-171_F	AACGGATGTGTGAGTAAAATTCTG	59.5
PAX6_c.357+8_SA2_R	TCCAGGTGAACTTACGCTTGG	62.5
PAX6_10R	CATTTGGCCCTTCGATTAGA	60.0
GAPDH_F	CGACCACTTTGTCAAGCTCA	60.0
GAPDH_R	AGGGGAGATTCAGTGTGGTG	60.0

Anexo 3: Lista de oligonucleótidos utilizados en el apartado 4.2.

Microsatélite	Primer Forward	Primer Reverse	Posición	Tamaño PCR	Color	Repetición
D11S929	AGGCCCTTCCAAGATCAG	CCCAGTTGCCGAACCTACC	chr11:25852052_25852283	230pb	PET	dinucleótido
D11S914	ATCTCATGGGAGTACCGT	GACCCACATCACCATTAC	chr11:31363290_31363565	270pb	VIC	dinucleótido
D11S995	GATTTTCCCACATCTATAATTATA	CTCTCAGAGAACTGAACCAA	chr11:31575212_31575212	120pb	NED	dinucleótido
PAX6	AATGCCACATCTTCAGTAC	TTGGAATGGCATTTCAGTGAC	chr11:31815845_31816135	290pb	NED	dinucleótido
GDB250586	AATGAGACTTACTGGGTGAGG	TTACACAGTAATTTCAAGCAACGG	chr11:32387938_32388081	140pb	NED	dinucleótido
D11S935	TACTAACCAAAAGAGTTGGGG	CTATCATTTCAGAAAATGTTGGC	chr11:36023259_36023462	200pb	VIC	dinucleótido
D11S1883	TTCAGTAACAGGAGACAAAAGG	GTTTCTTTGGTTTCGGATCTCTTCTCA	chr11:63373733_63373991	250pb	NED	dinucleótido

Anexo 4: Microsatélites utilizados en el análisis de haplotipos y sus características.

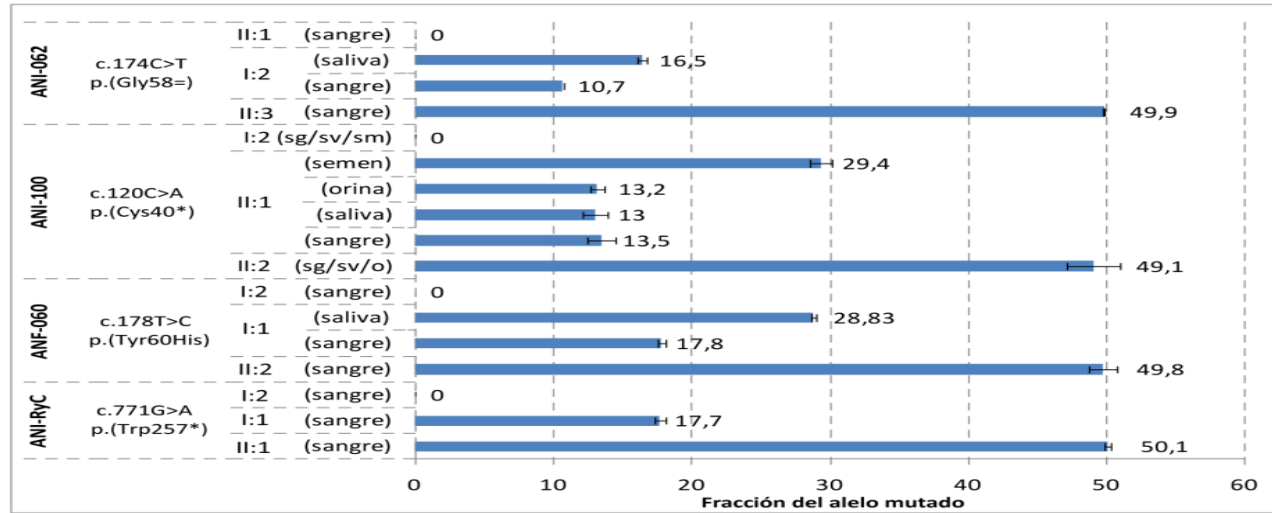


Anexo 5: Resumen de los resultados obtenidos mediante PCR semi-cuantitativa en las siete variantes estudiadas: c.174C>T, c.178T>C, c.249T>G, c.326_331dup, c.333C>A, c.356G>T, c.357+6T>G. A) Esquema de los distintos donadores detectados a lo largo del exón 6 mediante minigenes y el potencial spliceogénico de cada uno de ellos en la secuencia silvestre y en la mutada. B) Representación de las distintas isoformas resultantes del uso de los distintos donadores por el mecanismo de splicing. C) Resultados de la PCR semi-cuantitativa.

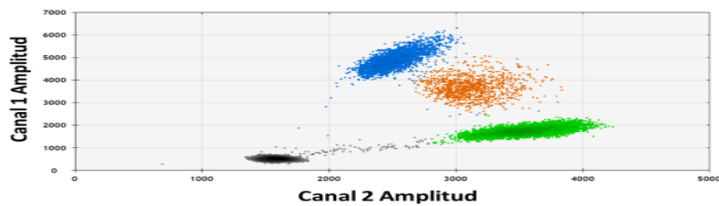
	Isoformas observadas							Isoformas Frameshift (%)	Isoformas In-frame (%)	Clasificación severidad	Fenotipo iris
	CT (%)	$\Delta 6_{185}$ (%)	$\Delta 6_{181}$ (%)	$\Delta 6_{108}$ (%)	$\Delta 6_{27}$ (%)	I6_5 (%)	Otra				
Minigenes											
<i>WT</i>	86		12		2			12	2		
<i>c.174C>T</i>	0	100						100	0	Muy severa (1)	Iris asimétricos
<i>c.178T>C</i>	93			1	6			0	7		Corectopia ^{&}
<i>c.249T>G</i>	3			91				6	91	Baja (4)	Variabilidad intra- individual
<i>c.326_331dup</i>	0		14*	66				33	66	Severa (2)	Aniridia completa
<i>c.333C>A</i>	13		3		84			3	84	Baja (5)	N.D.
<i>c.356G>T</i>	88		4		8			4	8		Membrana pupilar y corectopia ^{&}
<i>c.357+6T>G</i>	3		33	49	1*	10	4	43	51	Severa (3)	Variabilidad interindividual
Tamaño pico (pb)	381	196	199	273	354	385	317				
Posición donador	c.357+1	c.173	c.177	c.250	c.331	c.357+6					
Consecuencia proteica	---	p.(R59Vfs*12)	p.(R59Sfs*13)	p.(K86_S121del)	p.(C112_V120del)	p.(S121Rfs*5)					

Anexo 6: Resumen resultados minigenes obtenidos tras semi-cuantificación por electroforesis capilar. Distribución, en porcentaje, de las isoformas encontradas para cada variante. La posición del donador refleja los dinucleótidos GT del 5'SS. Se detallan las consecuencias proteicas de cada isoforma. CT indica el transcrito canónico. * indica niveles similares al WT. $\Delta 6$ indica delección del exón 6 (ejemplo: $\Delta 6_{185}$ indica delección parcial de 185 pb). I6_5 indica inserción de los primeros 5 nucleótidos del intrón 6. [&]: indica microftalmia.

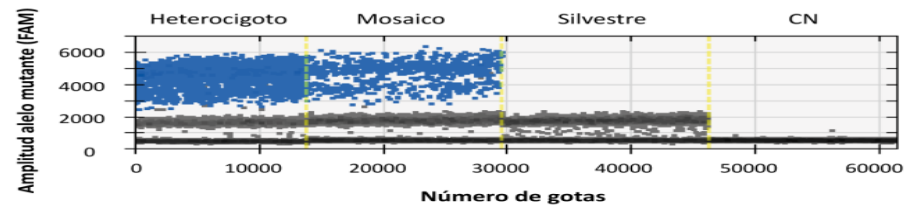
A



B



C



Anexo 7: Estudio de mosaicismo ligado a variantes en PAX6 mediante ddPCR. A) Representación en porcentaje de la presencia del alelo mutado dividido entre el total de alelos. **B)** Representación de los alelos amplificadas, cada punto es una gota. Azul (FAM): amplificación de un alelo mutante. Verde (VIC): amplificación de un alelo silvestre. Naranja: amplificación de un alelo mutado y uno silvestre en la misma gota. Gris: gotas sin amplificación. **C)** Representación del patrón del alelo mutado en los distintos individuos estudiados. Puntos azules (FAM): gotas con amplificación del alelo mutante. Puntos grises: gotas sin amplificación. Véase las dos poblaciones de puntos azules en el individuo mosaico.

Publicaciones

Resumen de las publicaciones derivadas de la presente Tesis Doctoral:

Plaisancié J[#], **Tarilonte M[#]**, Ramos P, Jeanton-Scaramouche C, Gaston V, Dollfus H, Aguilera D, Kaplan J, Fares-Taie L, Blanco-Kelly F, Villaverde C, Francannet C, Goldenberg A, Arroyo I, Rozet JM, Ayuso C, Chassaing N, Calvas P, Corton M. [#] Equally contribution. Implication of non-coding PAX6 mutations in aniridia. **Hum Genet.** 2018 Oct;137(10):831-846. doi: 10.1007/s00439-018-1940-x. Epub 2018 Oct 5. PMID: 30291432

Tarilonte M[#], Morín M[#], Ramos P, Galdós M, Blanco-Kelly F, Villaverde C, Rey-Zamora D, Rebolleda G, Muñoz-Negrete FJ, Tahsin-Swafiri S, Gener B, Moreno-Pelayo MA, Ayuso C, Villamar M, Corton M. [#] Equally contribution. Parental Mosaicism in PAX6 Causes Intra-Familial Variability: Implications for Genetic Counseling of Congenital Aniridia and Microphthalmia. **Front Genet.** 2018 Oct 17;9:479. doi: 10.3389/fgene.2018.00479. eCollection 2018.

María Tarilonte, MsC; Patricia Ramos, Jennifer Moya, Guillermo Sanz, MD; Fiona Blanco-Kelly, MD, PhD; Saoud T Swafiri, MD; Cristina Villaverde; Raquel Romero, PhD; Blanca Gener, MD; Patrick Calvas, MD, PhD; Carmen Ayuso, MD, PhD; Marta Corton, PhD. Activation of cryptic donor splice sites by non-coding PAX6 variants contributes to congenital aniridia. En proceso de publicación.

Otras publicaciones:

Ceroni F, Aguilera-Garcia D, Chassaing N, Bax DA, Blanco-Kelly F, Ramos P, **Tarilonte M**, Villaverde C, da Silva LRJ, Ballesta-Martínez MJ, Sanchez-Soler MJ, Holt RJ, Cooper-Charles L, Bruty J, Wallis Y, McMullan D, Hoffman J, Bunyan D, Stewart A, Stewart H, Lachlan K; DDD Study, Fryer A, McKay V, Roume J, Dureau P, Saggat A, Griffiths M, Calvas P, Ayuso C, Corton M, Ragge NK. New GJA8 variants and phenotypes highlight its critical role in a broad spectrum of eye anomalies. **Hum Genet.** 2019 Sep;138(8-9):1027-1042. doi: 10.1007/s00439-018-1875-2. Epub 2018 Feb 20. PMID: 29464339



Implication of non-coding *PAX6* mutations in aniridia

Julie Plaisancié^{1,2} · M. Tarilonte^{3,4} · P. Ramos^{3,4} · C. Jeanton-Scaramouche¹ · V. Gaston¹ · H. Dollfus⁵ · D. Aguilera^{3,4} · J. Kaplan⁶ · L. Fares-Taie⁶ · F. Blanco-Kelly^{3,4} · C. Villaverde^{3,4} · C. Francannet⁷ · A. Goldenberg⁸ · I. Arroyo^{4,9} · J. M. Rozet⁶ · C. Ayuso^{3,4} · N. Chassaing^{1,2} · P. Calvas^{1,2} · M. Corton^{3,4}

Received: 9 July 2018 / Accepted: 23 September 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

There is an increasing implication of non-coding regions in pathological processes of genetic origin. This is partly due to the emergence of sophisticated techniques that have transformed research into gene expression by allowing a more global understanding of the genome, both at the genomic, epigenomic and chromatin levels. Here, we implemented the analysis of *PAX6*, whose coding loss-of-function variants are mainly implied in aniridia, by studying its non-coding regions (untranslated regions, introns and *cis*-regulatory sequences). In particular, we have taken advantage of the development of high-throughput approaches to screen the upstream and downstream regulatory regions of *PAX6* in 47 aniridia patients without identified mutation in the coding sequence. This was made possible through the use of custom targeted resequencing and/or CGH array to analyze the entire *PAX6* locus on 11p13. We found candidate variants in 30 of the 47 patients. 9/30 correspond to the well-known described 3' deletions encompassing SIMO and other enhancer elements. In addition, we identified numerous different variants in various non-coding regions, in particular untranslated regions. Among these latter, most of them demonstrated an *in vitro* functional effect using a minigene strategy, and 12/21 are thus considered as causative mutations or very likely to explain the phenotypes. This new analysis strategy brings molecular diagnosis to more than 90% of our aniridia patients. This study revealed an outstanding mutation pattern in non-coding *PAX6* regions confirming that *PAX6* remains the major gene for aniridia.

Keywords Aniridia · Eye development · *PAX6* · Non-coding mutation · *Cis*-regulatory region · 5'UTR · Minigene assay

J. Plaisancié and M. Tarilonte contributed equally to this work.

P. Calvas and M. Corton contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s00439-018-1940-x>) contains supplementary material, which is available to authorized users.

✉ Julie Plaisancié
plaisancie.j@chu-toulouse.fr

- ¹ Service de Génétique Médicale, Pavillon Lefebvre, Hôpital Purpan, CHU Toulouse, Place du Dr Baylac, 31059 Toulouse Cedex 9, France
- ² INSERM U1056, Université Toulouse III, Toulouse, France
- ³ Department of Genetics, Instituto de Investigacion Sanitaria-Fundacion Jimenez Diaz University Hospital-Universidad Autónoma de Madrid (IIS-FJD, UAM), Madrid, Spain
- ⁴ Center for Biomedical Network Research on Rare Diseases (CIBERER), ISCIII, Madrid, Spain

Introduction

Aniridia (MIM#106210) is a panocular developmental anomaly that takes its name from a characteristic iris hypoplasia, visible in most cases to the naked eye. This hypoplasia not only manifests as an easily recognizable complete

- ⁵ Centre de Référence pour les affections rares en génétique ophtalmologique, CARGO, Filière SENSGENE, Hôpitaux Universitaires de Strasbourg, Strasbourg, France
- ⁶ Laboratoire de Génétique Ophtalmologique INSERM U1163, Institut Imagine, Paris, France
- ⁷ Service de Génétique Médicale, CHU Estaing, Clermont-Ferrand, France
- ⁸ Service de Génétique, CHU de Rouen, Centre Normand de Génomique Médicale et Médecine Personnalisée, Rouen, France
- ⁹ Department of Genetics, Hospital of Cáceres, Cáceres, Spain

absence of iris (“aniridia”), it may also take the form of pupil irregularities mimicking iris coloboma or simple defects in the iris architecture (slots in the iris anterior layer) visible at the slit lamp (Jordan et al. 1992; Nelson et al. 1984). Visual acuity may be reduced in these individuals due to foveal hypoplasia, cataract, corneal dystrophy, as well as abnormalities of the optic nerve. Visual outcome of patients is worsened by glaucoma in half of the cases, often in the second decade of life (Nelson et al. 1984).

Aniridia is a rare genetic disease with an incidence estimated between 1 per 50,000 and 96,000 births (Boonstra et al. 2012; Shaw et al. 1960). In more than 80% of the cases, this affection is due to monoallelic *PAX6* (Paired box 6) mutations and 70% of patients with isolated aniridia have an affected relative. About 1/3 of the cases carry de novo variants (Vincent et al. 2003). Aniridia can also be part of rare syndromes: the WAGR syndrome (MIM#194072) and the Gillespie syndrome (MIM#206700). In addition, mutations in the *PITX2* and *FOXC1* genes are occasionally responsible for severe iris hypoplasia mimicking aniridia (Ansari et al. 2016; Perveen et al. 2000).

The *PAX6* gene is known to play a major role in normal ocular development processes (Glaser et al. 1992; Hever et al. 2006). It is expressed during early and late ocular development (cornea, iris, lens and retina) but also in olfactory (Simpson and Price 2002) and central nervous systems (Simpson and Price 2002) and the endocrine pancreas (Simpson and Price 2002, St-Onge et al. 1997). It encodes a transcription factor whose function has been extremely conserved across evolution of species that use vision, such as evidenced by its loss-of-function in the mouse *small eye* (Sey) (Hill et al. 1991) or the *Drosophila eyeless* (*ey*) (Quiring et al. 1994). This transcription factor exerts its functions through the presence of three functional domains: the paired domain (PD) and the homeodomain (HD), allowing the protein to bind to DNA, and a proline, serine and threonine-rich transactivation (PST) enriched domain, which regulates the transcription of target genes. In humans, this gene uses at least three promoters (Azuma et al. 1999; Singh et al. 2002): P0 (exon 0), P1 (exon 1) and P α (intron 4) allowing, respectively, the production of at least three *PAX6* isoforms: (P0) the canonical *PAX6* isoform of 422 amino acids (NM_000280) contains all the functional domains mentioned above; (P1) the inclusion of an alternative exon (exon 6, formerly named exon 5a) leads to the synthesis of a second isoform of 436 amino acids (NM_001604) with an interrupted PD which recognizes a different DNA sequence and finally (P α), a last isoform in which PD is lost.

To date, more than 400 unique *PAX6* mutations are described, located across the entire coding sequence (Human *PAX6* Mutation Database). Two thirds are single nucleotide substitutions (SNVs) predominantly private; small deletions or insertions (indels) account for the

remaining third. Less than 2% of mutational events are large exonic rearrangements. Complete *PAX6* gene deletions that may also extend to neighboring genes at 11p13 locus represent also a very frequent cause of aniridia, explaining up to 30% of cases.

Most of the mutations observed in aniridia are truncating variants (nonsense, frameshift or splicing) leading to the introduction of premature termination codons (PTCs). It is hypothesized that aberrant *PAX6*-derived mRNAs are degraded by nonsense-mediated mRNA decay (NMD) (Bhuvanagiri et al. 2010; Vincent et al. 2003). Truncating variants and complete *PAX6* deletions lead to a similar severe effect, due to haploinsufficiency. Concerning missense variants, around four fifth of them are located in the PD of *PAX6*, thus potentially affecting its DNA binding activity. More than half of the missense mutations are associated with an atypical phenotype, ranging from defects in iris architecture with preserved visual acuity (Hanson et al. 1999) to phenotypes more severe than classical aniridia, including optic nerve malformations (Azuma et al. 2003), Peters anomaly (Hanson et al. 1994) and microphthalmia (Chassaing et al. 2013). There is no correlation between the nature and position of the mutations in functional regions and the observed phenotypes (Aggarwal et al. 2011; Gronskov et al. 1999; Gupta et al. 1998; Xie et al. 2014).

PAX6 is the master gene of eye development (Glaser et al. 1992; Hever et al. 2006), which quantitative and spatio-temporal expression must be finely controlled during embryonic development. This is mediated through a complex and extended *cis*-regulating domain comprising a large number of tissue-specific regulators, most of them distributed over hundreds of kilobases between the neighboring *RCN1* and *ELP4* genes (Bhatia et al. 2013; Kammandel et al. 1999; Kleinjan et al. 2001, 2004, 2006; McBride et al. 2011; Navratilova and Becker 2009; Ravi et al. 2013). The entire *cis*-regulatory domain of *PAX6* is highly conserved across vertebrate species. There are sequences that have existed for 450 million years in ancient vertebrates such as chimeric elephant (*Callorhinchus milii*). Additional sequences in that conserved region have been recently identified as “agCNE” for “ancient gnathostomes non-coding element” (Bhatia et al. 2014).

The regulatory role of these *cis*-elements was suggested by the existence of aniridia in patients with different 11p13 chromosome rearrangements affecting the downstream elements while preserving the coding *PAX6* sequence (Fantes et al. 1995). One of them, called the SIMO element, was particularly well characterized and has been involved in the self-amplifying loop of *PAX6* (Bhatia et al. 2013). A de novo point variant in SIMO was recently described as responsible for aniridia (Bhatia et al. 2013), demonstrating the importance of the sequence integrity of this conserved element. Thus, variants in these *cis*-elements may affect the *PAX6*

dosage in a given tissue or disrupt its ability to respond to external signals.

Thus, to better understand the frequency of non-coding variants in aniridia phenotype spectrum, we performed targeted Next Generation Sequencing (NGS) of non-coding *PAX6* sequences in a cohort of aniridia patients without any mutation or rearrangement in the coding *PAX6* sequence.

Materials and methods

Cohort description

This study was designed in compliance with the tenets of the Helsinki Declaration and patient enrollment had approval by the regional Ethics Committee of the University Hospital of Toulouse (CCPSOM2) and University Hospital Fundación Jiménez Díaz (Madrid, Spain). Informed consent was obtained from all participants included in this study.

A cohort of 27 French and 20 Spanish families with clinical diagnosis of isolated ($n=44$) or associated aniridia ($n=3$) (additional features are detailed in Table 2) in a familial ($n=19$) or sporadic pattern ($n=28$) was analyzed for non-coding *PAX6* variants. There was no phenotype suggestive of Gillespie syndrome. None of the patient had an intragenic *PAX6* coding mutation as determined by a combination of Sanger sequencing of the coding exons and splice sites, semi-quantitative methods and/or MLPA for the screening of exonic rearrangements, following a previously reported genetic diagnosis algorithm (Blanco-Kelly et al. 2017). Mutations in exons or intron–exon boundaries of *FOXC1* and *PITX2* were excluded as well.

PAX6 locus analysis

Different targeted resequencing approaches focused on aniridia, thereafter globally called “aniridia panel”, were developed, allowing to routinely analyze the entire *PAX6* locus: all coding and non-coding exons, intronic and 3'UTR regions for all the different RefSeq *PAX6* isoforms (NM_000280, NM_001258462 and NM_001604) as well as the *WT1* gene and 31 sequences conserved at the *PAX6* locus (NC000011.9, Online resource 1) between the *RCN1* and *ELP4* genes (580 kb). The selection of these conserved sequences was based on the results of previous works (Bhatia et al. 2014; Griffin et al. 2002; Kammandel et al. 1999; Kleinjan et al. 2001, 2004, 2006; McBride et al. 2011; Ravi et al. 2013; Williams et al. 1998) in which these different elements were studied with reporter transgenic mouse and zebrafish models.

The French cohort was screened by means of an *Ampliseq*-based multiplex PCR approach (Thermo Fisher Scientific, USA). Libraries preparation and further NGS

sequencing on an *Ion Torrent PGM*TM sequencer (Thermo Fisher Scientific, USA) were performed following the manufacturer's protocols. This panel has a size of 44.15 kb with a total of 198 amplicons distributed over 2 pools. The real coverage approaches the theoretical one (99.86%), since it lacks about 225 bp in the NRE element (intron 4) and 70 bp in the element 7CE1 (intron 6) (Online resource 1). The average depth of this panel for each patient is 1500X with a minimum of 800X. Under these conditions, all the regions are covered by a minimum depth of 50X. For the analysis of point mutations (SNV and Indel), the bioinformatics analysis pipelines, *Variant Caller v4.4.0.6* (Torrent Suite, <https://www.thermofisher.com>) and *NextGENe v.2.3.4* (https://softgenetics.com/NextGENe_01.php) were used. Concerning the analysis of large rearrangements (> 0.1 kb), thereafter called CNV, the results were obtained from a custom in-house pipeline adapted from Charbonnier et al. (Charbonnier et al. 2000). It consists in an analysis of reading depths, based on the comparison of a sample, region by region, to a set of control samples in the same run. A minimum depth of 100X was required with the need to use at least 8 “controls” by run.

The Spanish cohort was screened by means of customized oligonucleotide CGH array and/or targeted NGS panels. First, the previously reported WAGR array (Agilent 15K CGH) (Blanco-Kelly et al. 2017) and/or a customized design targeting 150 eye developmental-related genes (Agilent 4×180 K CGH arrays) (Ceroni et al. 2018), including the 5 Mb of the 11p13 locus, were used for CNV analysis in regulatory regions of *PAX6*. Additionally, 15 index cases were analyzed using two different customized targeted NGS panels for eye development genes, as previously described (Ceroni et al. 2018), both including the entire *PAX6* locus and regulatory regions, as above indicated. Briefly, target enrichment was performed using HaloPlex or SureSelect QXT technologies (Agilent Technologies, Santa Clara, CA) on Illumina MiSeq or Nextseq 500 platforms running on paired-end mode at a minimum of 400X. The family Aniridia-0006 was analyzed by whole exome sequencing (WES), as previously described (Ceroni et al. 2018). Bioinformatic analysis for mapping, variant calling and annotation for SNVs analysis and read depth comparison for CNV analysis were performed using standard procedures and custom in-house pipelines (Sanchez-Navarro et al. 2018).

All the detected variants were further validated and segregated in families by Sanger sequencing for punctual variants or in case of NGS-detected CNVs, by commercial CGH arrays (Illumina technologies, 60 K and/or 180 K formats) or by commercial MLPA (kit P219 v.B2, MRC Holland) when CGH arrays were used as first-tier approach, as previously described (Blanco-Kelly et al. 2017).

To investigate a possible founder event in three Spanish families, 8 highly polymorphic microsatellite markers

on 11p13 (Online resource 2) flanking around 34.5 Mb (11p15.5, chr11:1566686 to 11p13, chr11:36023782) of the *PAX6* gene were used. For the STRs genotyping, PCR products were analyzed on the ABI 3130xl Genetic Analyzer (ThermoFisher) and further analyzed with the GeneMapper v3.5 software (ThermoFisher). Haplotype reconstruction was performed using the software Cyrillic v2.1 (Cyrillic Software, Wallingford, UK).

Assessment of pathogenicity by in silico analysis

Pathogenicity of detected SNVs and small indels was sequentially assessed based on the following criteria: (1) variants previously reported in mutation databases, including HGMD Professional and the *PAX6* Mutation Database, (2) novel or rare variants with a minor allele frequency (MAF) < 0.002% (i.e., 1 per 50 000) and (3) de novo variants for the sporadic cases and variants showing dominant inheritance and co-segregation in all affected relatives for familial forms.

Population frequencies and pathogenicity analysis of the detected variants were assessed using public databases (gnomAD and CIBERER Spanish Variant Server) and in silico prediction tools for putative splicing misregulation (Human Splicing Finder, Splice Site Finder-like, MaxEntScan, NNSPLICE, GeneSplicer, ESEfinder, Ex-skip on Alamut software and SPARN software) and for open reading frames disruption (ORFfinder and OrfPredictor). All variants were classified according to the ACMG (American College of Medical Genetics and Genomics) recommendations (Richards et al. 2015) as pathogenic, likely pathogenic or of unknown significance (VUS). The nomenclature of intragenic variants was referred for the canonical RefSeq *PAX6* isoform NM_000280.

In vitro splicing assays

An in vitro minigene approach was used to assess the impact on the splicing process of 6 variants located in the 5' UTR of *PAX6* at exon 2 and 3. A wild-type minigene for exons 1–4 of the RefSeq *PAX6* isoform NM_000280 (*PAX6_ex1-4*, Online resource 3A) was assembled by using a sequential classical restriction digestion/ligation cloning strategy with two intermediate constructs (*PAX6_ex1-2* and *PAX6_ex3-4*). The final construct included the complete introns 1 and 3, and around 200 bp of surrounding sequences for introns 2 and 4. Minigenes were generated following the recommendations of a previously described protocol (Desviat et al. 2012). The primers used are detailed in Online resource 4. Variants were introduced in the wild-type *PAX6_ex1-4* minigene construct by site-directed mutagenesis using the QuikChange Lightning Site-Directed Mutagenesis kit

(Agilent) with specifically designed sense and antisense primers for each variant (Online resource 4).

To assess if there were tissue-related alternative splicing patterns, the minigene assays were performed using two different human cell lines in duplicate: (1) HEK-293 (Human Embryonic Kidney) cell line due to its well-known high transfection efficiency and (2) ARPE-19 (Adult Retinal Pigmented Epithelium) cell line, a model of pigmented retinal cells. Both cell lines were grown at 37 °C and 5% CO₂ atmosphere in DMEM with ultraglucutamine and 4.5 g/l glucose (Lonza, Belgium) or DMEM F-12 (Ham) (1:1) with L-glutamine and 15 mM HEPES (Lonza, Belgium) for HEK-293 or ARPE-19, respectively, supplemented with 10% fetal bovine serum and 1% penicillin/streptomycin (Gibco, Thermo Fisher Scientific). A total of 4 × 10⁵ cells grown in 6-well plates were transfected with 2 µg or 3 µg of each minigene constructions (empty vector, wild-type and mutants *PAX6_ex1-4*) for HEK-293 and ARPE-19, respectively, using JetPEI reagent (Polyplus Transfection, New York, NY, USA). At 24-h post-transfection, the cells were harvested and total RNA was extracted using Trizol reagent (Thermo Fisher Scientific) in accordance with the manufacturer's instructions. Reverse transcription and PCR (RT-PCR) were performed with Superscript IV First-Strand Synthesis System (Invitrogen, Thermo Fisher Scientific) for cDNA synthesis and FastStart Taq DNA Polymerase (Roche Applied Science, Germany) using vector-specific primers (Online resource 4). Amplified products were separated by agarose gel electrophoresis and the excised bands analyzed by Sanger sequencing after extraction with NucleoSpin Gel and PCR Clean-up (Macherey–Nagel, Düren, Germany). Quantification of all transcripts was performed by semi-quantitative fluorescent PCRs. FAM-labeled products were run with ROX-1000 Size Standard on the ABI3130xl Genetic Analyzer (Thermo Fisher Scientific) and further analyzed with the GeneMapper v3.5 software. Peak areas were used to quantify the relative abundance of each transcript that was the average of at least three replicates.

Patient's RNA analysis by RT-PCR

RT-PCR was performed to determine effects of 5'UTR sequence variants on *PAX6* pre-mRNA processing from lymphoblastoid cell lines (LCLs) from patients and control individuals. LCLs were established by Epstein Barr virus (EBV) transformation of blood lymphocytes from patients carrying the **c.-129+1G>A** variant in exon 2 (ANI-006 and ANI-069), patient SG140776 carrying the **c.-118_-117delTT** variant and patient SG130882 carrying the **c.-7421C>T** variant (located within the promoter 0 region). Cell lines were cultured in RPMI-1640 media supplemented with 10% fetal bovine serum and 1% (v/v) penicillin–streptomycin (Gibco, Thermo Fisher Scientific).

To inhibit nonsense-mediated decay, LCLs were incubated with different NMD inhibitors: (1) cycloheximide (Sigma-Aldrich, St. Louis, MO) 0.5 mg/mL for 4 h, (2) cycloheximide (Sigma-Aldrich, St. Louis, MO, USA) 0.75 mg/mL for 4 h, (3) emetine 300 µg/mL for 4 h. Total RNAs were reversely transcribed into cDNA with Superscript IV First-Strand Synthesis System (Invitrogen, ThermoFisher Scientific) and using random primers. RT-PCR experiments were performed using different *PAX6* exonic primer pairs spanning exon 1–5 and exon 1–8, including specific primers senses for amplification of the different alternative exons 1 for *PAX6* (Online resource 4). The housekeeping *GAPDH* gene was used as internal control. Amplified products were further analyzed by Sanger sequencing.

Results

We studied 47 index cases suffering from aniridia, without point mutation or exonic rearrangement in the *PAX6* coding sequence, using a comprehensive screening of non-coding regions of the *PAX6* locus, based mainly on custom targeted NGS and CGH arrays allowing both SNVs and CNVs identification.

Microdeletions affecting 3' *PAX6* regulatory regions at the 11p13 locus

Comparative analysis of reading depths obtained by NGS and/or custom CGH arrays analysis revealed that 9 out

of the 47 patients were carrying microdeletions of the downstream region of *PAX6* (Table 1, Online resource 5). All rearrangements were confirmed by a secondary approach prior to refinement of boundaries using aCGH. The deletions ranged from 150 to 1500 kb in size and involved the 3' regulatory sequences of *PAX6* located in the intronic regions of *ELP4* (Table 1, Online resource 5). Neither *PAX6* nor *WT1* were involved in deletions (Online resource 5). Accordingly, the observed phenotype of all these patients was isolated aniridia without any extra-ocular manifestation. In the 5 out of the 19 familial cases of aniridia, 3' microdeletions segregated with the phenotype in an autosomal dominant pattern (Table 1, Online resources 2 and 5).

SNV or Indel in non-coding regions of *PAX6*

In the remaining 38 patients, we looked for point mutations in the non-coding *PAX6* sequences using customized “aniridia panels” that allow the complete capture of the entire *PAX6* locus. All variants with a MAF < 1% were reported, to be considered in case of complex alleles. A comprehensive NGS analysis of the UTRs, intronic and regulatory regions revealed a total of 7 pathogenic or likely pathogenic variants in 5' UTR in 12 index cases (12/47, 25% of patients) and also 15 VUS located in introns, 3' UTR and *cis*-elements in 11 patients (Tables 2, 3). For 36% (17/47) patients (6 French and 11 Spanish patients), only benign and/or likely benign variants were identified at the *PAX6* locus.

Table 1 Summary of 11p13 microdeletions affecting 3' regulatory *PAX6* regions found in 9 patients with isolated aniridia

Patient	Sex	Age (years)	Phenotype	Case	CGH array	Size of the microdeletion (kb)	Distance from the last exon of <i>PAX6</i> (kb)	Inheritance
SG112219	M	6	Isolated aniridia	Sporadic	arr[hg19]11p14.1p13 (30,432,767–31,735,748)x1	1303	± 70	De novo
SG130616	M	4	Isolated aniridia	Familial	arr[hg19]11p14.1p13 (30,741,612–31,775,457)x1	1034	± 31	Inherited
ADN170025	M	4	Isolated aniridia	Familial	arr[hg19] 11p14.1p13 (30,921,059–31,696,685)x1	776	± 110	Inherited
ANI-067	F	10	Isolated aniridia	Familial	arr[hg19]11p13 (31,083,877–31,704,548)x1	621	± 102	Inherited
ADN100128	M	6	Isolated aniridia	Sporadic	arr[hg19]11p13 (31,172,440–31,747,602)x1	575	± 59	De novo
ANI-008	F	61	Isolated aniridia	Familial	arr[hg19]11p13 (31,147,306–31,714,853)x1	568	± 91	Inherited
ANI-021	F	15	Isolated aniridia	Sporadic	arr[hg19]11p13 (31,186,493–31,698,208)x1	512	± 108	De novo
ANI-105	F	10	Isolated aniridia	Sporadic	arr[hg19]11p13 (31,184,376–31,686,880)x1	502	± 119	De novo
SG101836	M	41	Isolated aniridia	Familial	arr[hg19]11p13 (31,648,248–31,804,088)x1	155	± 2	Inherited

Table 2 Summary of non-coding variants identified in our cohort and related phenotypes

Patient	Sex	Age (years)	Ocular phenotype	Other signs	SNV	Location	Segregation analysis	Allele frequency (gnomAD)	References
5' and 3' UTR									
SG071509	F	32	Isolated aniridia		c.[-129+1G>A];[=]	Intron 2	De novo	N.D	This study
ANI-006	M	33	Isolated aniridia		c.[-129+1G>A];[=]	Intron 2	In affected mother	N.D	This study
ANI-069	M	34	Isolated aniridia		c.[-129+1G>A];[=]	Intron 2	In affected mother	N.D	This study
ANI-088	M	6	Isolated aniridia		c.[-129+1G>A];[=]	Intron 2	In affected father	N.D	This study
SG080661	F	9	Isolated aniridia		c.[-128-2delA];[=]	Intron 2	De novo	N.D	Axton (1997), Robinson (2008) and Vasilyeva (2017)
SG130068 ^a	M	16	Isolated aniridia		c.[-128-2delA];[=]	Intron 2	N.A Sporadic case	N.D	Axton (1997), Robinson (2008) and Vasilyeva (2017)
ANI-057	M	51	Isolated incomplete aniridia with Peters anomaly		c.[-128-2delA];[=]	Intron 2	De novo	N.D	Axton (1997), Robinson (2008) and Vasilyeva (2017)
SG140776 ^a	F	10	Isolated aniridia		c.[-118_-117delTT];[=]	Exon 3	In affected mother	N.D	This study
SG120216	F	5	Isolated aniridia		c.[-52+1G>A];[=]	Intron 3	In affected mother and brother	N.D	Gronskov (1999) and Redeker (2008)
SG130439	F	5	Isolated aniridia		c.[-52+3_-52+6delin-sTG];[=]	Intron 3	In affected mother	N.D	This study
SG070836	M	14	Isolated aniridia		c.[-52+3_-52+4delAA];[=]	Intron 3	In affected mother	N.D	This study
ANI-012	M	46	Isolated aniridia		c.[-52+5delG];[=]	Intron 3	In affected father and sister	N.D	This study
SG122301	F	36	Aniridia	MO, MS	c.[*4389_*4390del];[=]	3'UTR	In affected mother	rs777790874 0.02%	This study
Deep intronic									
SG102229	M	7	Isolated aniridia		c.[-357+136G>A];[=]	Intron 6	De novo	ND	This study
SG132488 ^a	F	31	Isolated aniridia		c.[357+334G>A];[=]	Intron 6	NA	ND	This study
ADN130014	F	4	Aniridia	MO, MS	c.[1033-275A>C];[=]	Intron 11	NA	rs3026392 0.08%	This study
3' regulatory regions									
SG132488	F	31	Isolated aniridia		g.[31667759T>C];[=] g.[31845390G>A];[=]	HS6 agCNE11	NA	0.4% 0.003%	This study
SG140776	F	10	Isolated aniridia		g.[32062883G>A];[=]	agCNE1	In affected mother	0.04%	This study
SG130068	M	16	Isolated aniridia		g.[31847009G>A];[=]	agCNE10	NA Sporadic case	0.9%	This study
SG130882	F	7	Isolated aniridia		c.[-742]C>T];[=]	Upstream exon 0	Four affected relatives	0.3%	This study

Table 2 (continued)

Patient	Sex	Age (years)	Ocular phenotype	Other signs	SNV	Location	Segregation analysis	Allele frequency (gnomAD)	References
SG160682	F	31	Isolated aniridia		g.[32063020A>G];[=] g.[31671543A>G];[=]	agCNE1 HS5	NA	0.6% 0.8%	This study
SG132067	M	36	Isolated aniridia	PMG, SE, RF	g.[32063079C>A];[=]	agCNE1	NA	0.02%	This study
SG150428	F	2	Isolated aniridia		g.[31908528C>T];[=]	E72	NA	0.5%	This study
SG080877	F	9	Isolated aniridia		g.[31623206G>A];[=] g.[31671543A>G];[=]	RB HS5	NA	0.003% 0.8%	This study

MO microphthalmia, MS microspherophakia, PMG polymicrogyria, SE schizencephaly, RF renal failure, NA not applicable, ND not displayed

^aFor these patients, additional rare variants were also found in 3' regulatory regions (indicated the dedicated section)

Likely pathogenic variants in 5'UTR

Twelve out of the 38 index cases (5 Spanish and 7 French families) tested with targeted NGS approaches harbored unreferenced or extremely rare variants (MAF < 0.002%) in the 5'UTR of *PAX6* (Fig. 1; Tables 2, 3).

Seven 5'UTR variants, located in canonical splicing sequences surrounding non-coding exon 2 and 3 (Fig. 1; Tables 2, 3), were considered pathogenic or likely pathogenic based on in silico predictions and disease co-segregation. These include three variants reported in more than two unrelated families either in this work (**c.-129+1G>A** [SCV000803674]) or in the literature in patients with isolated aniridia (**c.-128-2delA** and **c.-52+1G>A**) (Axton et al. 1997; Gronskov et al. 1999; Redeker et al. 2008; Robinson et al. 2008). Four unreferenced variants: **c.-118_-117delTT** [SCV000803675], **c.-52+3_-52+4del** [SCV000803676], **c.-52+3_-52+6delinsTG** [SCV000803677] and **c.-52+5del** [SCV000803678] were identified in four distinct unrelated families (Fig. 1; Tables 2, 3). In silico analyses predict a splicing defect of pre-mRNAs carrying the three latter variations (Table 3). Segregation of all these 5'UTR variants was consistent with the phenotype in an autosomal dominant pattern in the eight familial cases (Table 2, Online resource 2), while they likely occurred as de novo events in the four remaining sporadic cases (Table 2).

Variants of unknown significance (VUS) in 3'UTR and intronic regions of *PAX6*

Four VUS were identified, one of which was located in the 3'UTR of *PAX6*: **c.*4389_*4390delAA** (rs777790874). This variant, found in an index case (SG122301) and her affected mother, had a reported MAF of 0.0001941 (gnomAD). The three others VUS consisted in deep intronic changes (Tables 2, 3). The unreferenced **c.357+136G>A** [SCV000803679] change located in intron 6 has been identified in a sporadic case (SG102229). None of his parents carried the mutation suggesting a de novo event. It is expected to create a new cryptic acceptor splicing site according to the prediction tools (Table 3).

A second unreported variant has been detected in intron 6, the **c.357+334G>A** [SCV000803680], that could strengthen a putative cryptic donor site (Table 3). This variant was identified in the patient SG132488 who also carried rare variants in the agCNE11 and the HS6 elements (Table 2). Due to unavailability of parental DNA, segregation analysis could not be performed in this family.

Finally, the third deep intronic variant, **c.1033-275A>C**, located in intron 11 had a MAF of 0.08% (rs3026392) and is predicted to affect splicing (Tables 2, 3). This variant was found in the patient ADN130014, whom parent's DNAs were unavailable, hampering segregation analysis.

Table 3 Splicing predictions and in vitro outcomes of 5'UTR and deep intronic variants

Variant	No. of patients	No. of families	Localization ^a	In silico splicing scores ^b	In silico prediction outcome	Splicing defect in LCLs	Splicing defect in minigenes ^c
c.-129+1G>A	7	4	Canonical 5'SS Exon 2	- 100% (4/4 tools) dPSI (- 13.17)	Disruption of 5'SS and one ESE site	No observed aberrant splicing Only CT	Six different products observed in in ARPE-19 line: E2 skipping (9%), E2-E3 skipping (12%), E2-4 skipping (32%), 406 bp intron 2 retention (6%), 254 bp intron 2 retention (19%) and CT (4%)
c.-128-2delA	9	6	Canonical 3'SS Exon 3	- 100% (4/4 tools)	Disruption of 3'SS and two ESEs	NA	E3 skipping (96%) and CT (4%)
c.-118_-117delTT	2	1	Exon 3	No alteration	No alteration	No observed aberrant splicing Only CT	Assay failure
c.-52+1G>A	6	4	Canonical 5'SS Exon 3	- 100% on SSF, no variation for HSF dPSI (- 23.00)	Disruption of 5'SS and creation of new ESE sites	E3-6 skipping [Gronskov, 1999]	E3 skipping (100%)
c.-52+3_-52+4delAA	2	1	Canonical 5'SS Exon 3	- 100% on SSF, no variation for HSF	Disruption of 5'SS and creation of a new ESE site	NA	Not done
c.-52+3_-52+6delinsTG	2	1	Canonical 5'SS Exon 3	- 100% on SSF, no variation for HSF	Disruption of 5'SS and ESE site	NA	E3 skipping (100%)
c.-52+5delG	2	1	Canonical 5'SS Exon 3	- 12.8% on SSF, no variation for HSF	Altering 5'SS and disruption of ESE site	NA	E3 skipping (96%) and CT (4%)
c.-7421C>T	5	1	Promoter 0 region	No alteration	None	No observed aberrant splicing Only CT	NA
c.357+136G>A	1	1	Deep Intron 6	+ 100% (4/4 tools)	New cryptic 3'SS in c.357+138	NA	NA
c.357+334G>A	1	1	Deep Intron 6	+ 100% (3/4 tools)	Creation or strengthening of 5'SS in c.357+331	NA	NA
c.1033-275A>C	1	1	Deep Intron 11	+ 100% (2/4 tools)	Creation or strengthening of 3'SS in c.1033-262	NA	NA

^aAffected motifs: 3'SS = 3' acceptor splice site; 5'SS = 5' donor splice site; *ESE* exonic splicing enhancer, *ESS* exonic splicing silencer

^bIn silico splicing scores: *SSF* splice site finder, *HSF* human splicing finder; dPSI = changes of percent splicing inclusion exon, calculated by SPANR predictor for SNVs

^cThe proportion of each transcript is indicated between parentheses. *CT* canonical transcript, *NA* not applicable

Variants of unknown significance in the cis-regulatory *PAX6* elements

Finally, screening of the *cis*-regulatory regions of *PAX6* detected rare sequence variants in five index cases (Table 2).

The index case SG130882 (Table 2 and Online resource 2) presenting with an isolated aniridia, was carrying the g.31839767C>T (hg37) corresponding to the variation **c.-7421C>T** (NM_000280) [SCV000803681] equivalent to the variation c.-727C>T (NM_001258462) positioned just

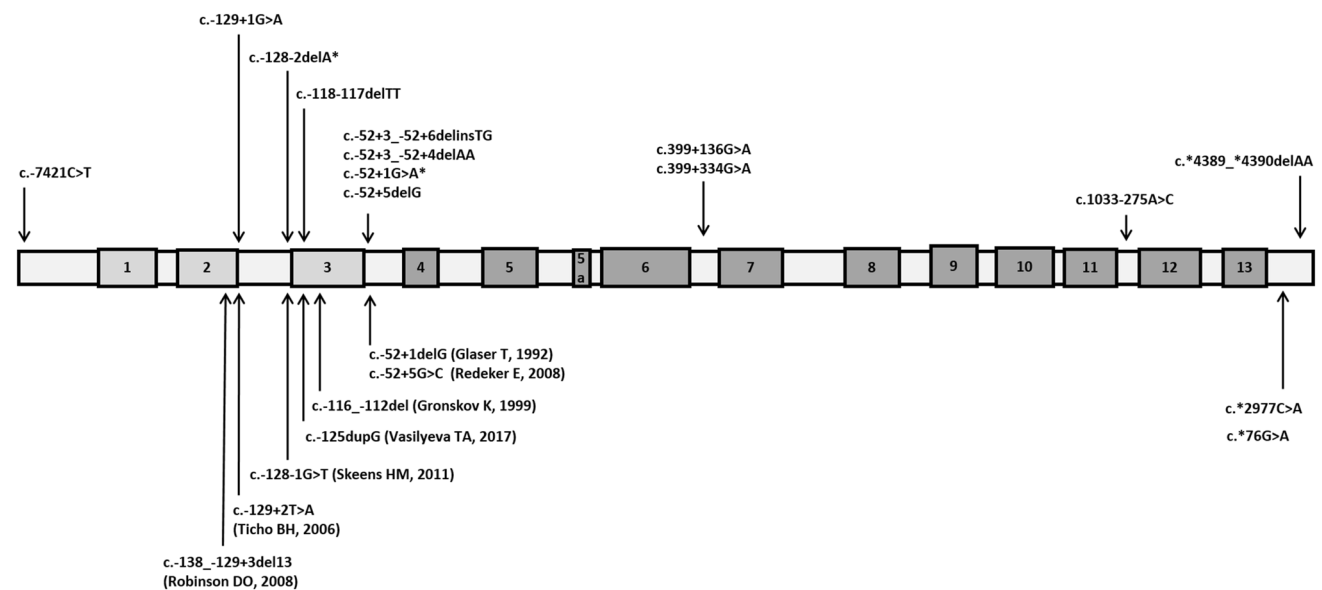


Fig. 1 Schematic representation of the *PAX6* gene. The dark gray boxes correspond to the 11 coding exons, the pale gray ones to the 3 non-coding exons. The upper section shows the position of the main variants identified in this study (*indicates if the variation has already

been described in the literature). The lower section shows the position of the variants reported in the literature and/or in the Human *PAX6* mutation Database (LOVD)

upstream of the promoter P0. This variation segregated in all the other four affected relatives and was absent from an unaffected individual tested in this family (Online resource 2). However, it is listed in the dbSNP database (rs537446654) with an allelic frequency around 0.3%, making it a relatively frequent variant compared to the frequency of the aniridia phenotype. In addition, in silico tools do not make it possible to precisely predict the effect of this type of variant. Interestingly, linkage analyses were also performed in this family and have shown that the phenotype segregated with *PAX6*. Thus, either this variant is effectively pathogenic with incomplete penetrance (but this is not the case in this family) or, another mutation exists on the same allele at the *PAX6* locus that has not been identified here despite an extensive analysis of the locus. Finally, one can also imagine that another variation (miRNA, modifying gene) is necessary for the emergence of the phenotype.

Finally, four index cases were carrying variations with an allele frequency < 1% in different conserved elements at the *PAX6* locus (Table 2). However, segregation analysis of these variants was not possible without parental DNA.

Functional analysis of splicing for 5'UTR variants

All the 5'UTR variants identified in this work were located in canonical splicing sites of non-coding exons 2 and 3 (Fig. 1; Tables 2, 3). To evaluate their possible effect on splicing and provide evidence of their pathogenicity, we functionally tested these variants by hybrid minigene assays

and/or RT-PCR using total RNA from LCLs when available. RT-PCR experiment was also performed for studying the variant **c.-7421C>T** (promoter 0 region) to be sure it had no effect on splicing.

The **c.-129+1G>A**, **c.-128-2delA**, **c.-52+1G>A**, **c.-52+3_-52+6delinsTG** and **c.-52+5del** variants were introduced into the wild-type *PAX6*_ex1-4 construct by site-directed mutagenesis (Table 3). In contrast, we failed to produce the indel **c.-118_-117delTT** (g.31828462_31828463del). The **c.-52+3_-52+4del** was not tested by hybrid minigene assay given its probable similar effect on splicing than the **c.-52+1G>A**, **c.-52+3_-52+6delinsTG** and **c.-52+5del**.

Minigene transfection assays were performed in two different human cell lines, HEK293 and the ARPE-19. Different spliced transcripts were evaluated by sequencing and also quantified to determine their precise contribution on overall splicing of involved exons (Fig. 2 and Online resources 3, 6 and 7). RT-PCR of the wild-type construct for exons 1–4 resulted in a single mRNA isoform of the expected size, showing a correct splicing by sequencing (Fig. 2d). In contrast, minigene assays for the five analyzed variants revealed an aberrant pre-mRNA processing.

In case of the novel **c.-129+1G>A** variant on exon 2, a complex pattern of at least five different aberrant splicing isoforms (Fig. 2E and Online resource 3 and 6) that included, not only the predicted skipping of exon 2, either alone or accompanied by exon 3 (ex2-3 skipping) and exon 4 (ex2-4 skipping), but also two longer transcripts due to

partial retention of intron 2 using different alternative deep-intronic donors. In addition, the wild-type isoform was also observed but only representing <5% of the total transcribed transcripts (Table 3). Overall, despite some differences in terms of the amount of each aberrant transcript (Table 3 and Online resource 6), the same pattern of aberrant and normal splicing was observed in both cell lines. The most abundant splicing event, in both cell lines, was the partial retention of intron 2. Unlike in the HEK293 cells, the aberrant transcript with the exon2-4 skipping was relatively abundant in the ARPE-19 cells (>32%). In LCLs from two carriers of this variant (ANI-006 and ANI-069), we could not detect any exon skipping by RT-PCR in comparison with controls (data not shown). This absence of aberrant transcripts could be explained by NMD in contrast to minigene experiments. To discard that aberrant transcripts could be degraded by the NMD mechanism, LCLs were thus treated with different NMD inhibitors yielding a similar pattern of canonical transcription for the main different *PAX6* isoforms (data not shown). These findings could support the observed cell-specific differences in the splicing processing of this variant on minigene assays.

Four different variants (Table 3) located either on acceptor (**c.-128-2delA**) or donor (**c.-52+1G>A**; **c.-52+3_-52+6delinsTG** and **c.-52+5del**) splicing sites of exon 3 were also tested by hybrid minigene analysis. All of these variants lead to exon 3 skipping (Fig. 2F and Online resources 3 and 7). The normally spliced transcript was also marginally observed for the **c.-128-2delA** and **c.-52+5del** (<5%, Table 3).

In LCLs, respectively, derived from patient SG140776 carrying the **c.-118_-117delTT** and patient SG130882 carrying the **c.-7421C>T**, we could not detect any aberrant transcript by RT-PCR (data not shown). Sanger sequencing of the unique band observed in patient SG140776 showed heterozygosity for the **c.-118_-117delTT**. Concerning the **c.-7421C>T** variant, the absence of apparent splicing effect is an additional argument, besides its frequency, to think it is more a genetic mark (in linkage disequilibrium with an undetected variant) associated with the disease in this family than the reflect of a pathological mechanism.

Discussion

85 to 95% of patients with aniridia carry mutations involving the coding sequence of *PAX6*. Thus, 5 to 15% of patients remain without molecular diagnosis. Aniridia is a very rare disease (incidence between 1 per 50.000 and 96.000 births), and therefore, patients without molecular diagnosis are, in the end, very few (Vasilyeva et al. 2017). To uncover new genes or pathophysiological mechanisms that cause rare diseases, it is necessary to study together several patients to be

able to reveal common processes. Here, we have constituted an exceptional cohort of 47 patients with aniridia without mutation in the coding sequence of *PAX6*. To bring a genetic diagnosis to these patients, we developed targeted tools to analyze the entire *PAX6* locus to search for mutations affecting regulatory or untranslated regions of *PAX6*.

Using NGS and aCGH approaches, we were able to identify CNVs of less than 1500 kb in the *PAX6* downstream region in 9 patients. Interestingly, all the deletions we identified did not include *PAX6* but the SIMO element, known to be involved in aniridia phenotype (Bhatia et al. 2013). Nevertheless, we found a minimum shared region of 18 kb (chr11:31648248–31666340, Online resource 5) for all the 3' downstream deletions except one, reported by Simioni et al. (Simioni et al. 2012). This shared region did not contain SIMO but the E180 element (Online resource 1 and 5). There are previously three published cases carrying 3' deletions that did not include the SIMO element (Bayrakli et al. 2009; Cheng et al. 2011; Simioni et al. 2012). This comparison highlights additional enhancers that could be important in aniridia, besides the SIMO element. Of note, no chromosomal aberration was detected in the upstream non-coding region between *RCNI* and *PAX6* both in this work, in the literature or in the public databases.

We found SNVs or small indels in 21 out of the 38 remaining patients tested on “aniridia panel” by NGS. Over half of them (12/21) was carrying variants in the 5'UTR of *PAX6*. A total of 7 extremely rare and novel variations in the 5'UTR were identified in these 12 unrelated patients: **c.-129+1G>A**, **c.-128-2delA**, **c.-118_-117delTT**, **c.-52+1G>A**, **c.-52+3_-52+4delAA**, **c.-52+3_-52+6delinsTG** and **c.-52+5delG**. All these variants were located in canonical donor or acceptor splice sites of non-coding exons (or very close to, as for the **c.-118_-117delTT** located in exon 3). Thus, it was supposed to affect splicing processes. Variants in 5'UTR, with potentially similar effects, have been previously described in patients with iris anomalies as detailed in Fig. 1: **c.-129+2G>A** (Ticho et al. 2006), **c.-138_-129+3del13** (Robinson et al. 2008), **c.-128-1G>T** (Skeens et al. 2011), **c.-116_-112del** (Gronskov et al. 1999), **c.-52+5G>C** (Redeker et al. 2008) and **c.-52+1delG** (Glaser et al. 1992).

According to the ACMG guidelines (Richards et al. 2015), the variations **c.-52+1G>A**, **c.-128-2delA** and **c.-129+1G>A** are considered as pathogenic since (1) they are each found in at least two affected families, (2) they affect canonical splice sites, (3) they have been predicted to be pathogenic by in silico analyses and (4) in vivo and/or in vitro functional studies have demonstrated potentially pathogenic effects. The **c.-52+1G>A** mutation led to skipping of exons 3–6 of the *PAX6* gene as Gronskov K. et al. demonstrated by RT-PCR on LCLs from carriers (Gronskov et al. 1999). Accordingly, our minigene splicing assay

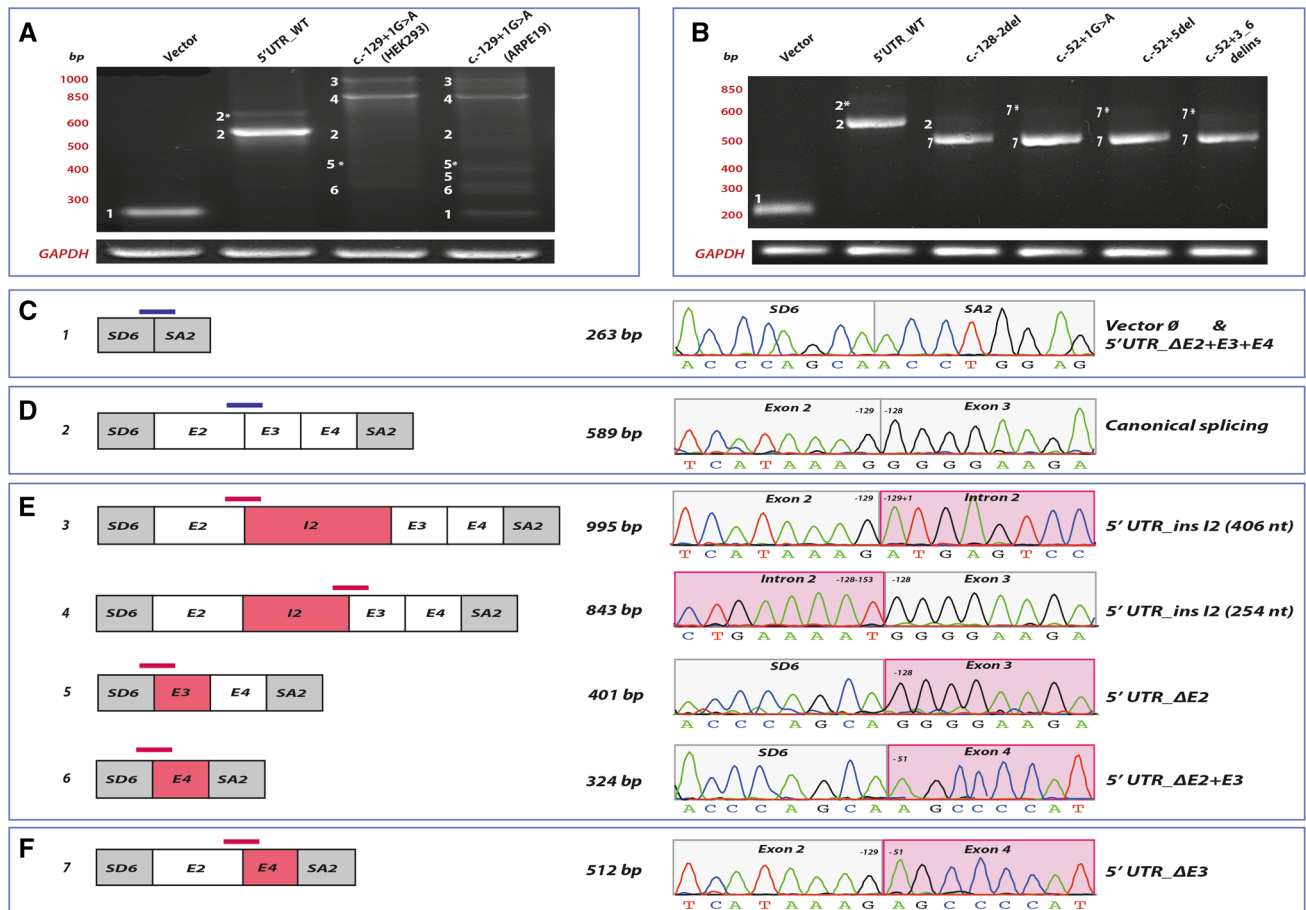


Fig. 2 In vitro hybrid minigene assays for the splicing analysis of five variants located in 5'UTR of *PAX6*. Experiments have been conducted in two different human cell lines (HEK293T and ARPE-19) that were transfected with the mutant constructs and compared with the wild-type construct and the empty vector used as controls. RT-PCRs using specific vector primers were performed and analyzed by gel electrophoresis. *GAPDH* expression was also analyzed as internal control. **a, b** Splicing outcome for the five mutants versus wild-type minigene constructs in the two analyzed cell lines, the kidney-derived HEK293T and the retinal-derived ARPE-19. Some low abundant artefactual isoforms were also obtained due to the alternative use of a cryptic acceptor site on the pSPL3, located downstream of SA2, leading to the inclusion of additional 114 nucleotides of vector sequence (represented by asterisk). **c–f** Schematic representation of splicing events and transcripts generated by different minigene constructs: **c** the empty vector, **d** wild type, **e** the **c.-129+1G>A** on exon 2 and **f** 4 variants on exon 3. The obtained sequence for each canonical or

aberrant splicing junctions (blue or red bars, respectively) was also shown. *PAX6* exons and pSPL3 splicing sites are represented by white and gray boxes, respectively. Anomalous spliced exons and/or intron retentions are shown in red. **a, e** Different isoforms were obtained for the novel variant **c.-129+1G>A** in the two transfected cell lines: (1) two aberrant larger transcripts due to different partial retentions of intron 2 by the disruption of the 5'SS and the use of different deep intronic cryptic donors, (2) two aberrant shorter isoforms for the exon skipping of exon 2 and exon 2–3 and (3) the canonical *PAX6* transcript. An additional isoform resulting from exon skipping of exon 2–4 was observed but only in the ARPE-19 line. **b, e** One single high-abundant aberrant isoform was observed in ARPE-19 cells for the variants **c.-52+1G>A**, **c.-52+3_-52+6delinsTG** and **c.-52+5delG**, disrupting the 5'SS of exon 3. In case of the variant **c.-128-2delA**, located on the 3'SS of exon 3, a similar pattern was obtained, besides the presence of the residual canonical transcript. Similar findings were obtained in HEK-293 cells (data not shown)

using an exons 1–4 construction for the 5'UTR of *PAX6*, showed a similar aberrant skipping of exon 3 for this variant. We have observed a similar in vitro effect for the previously reported **c.-128-2delA** (Axton et al. 1997; Robinson et al. 2008; Vasilyeva et al. 2017) that has been also identified in two unrelated French and Spanish families of our cohort. Concerning the novel variant **c.-129+1G>A**, that was identified in a total of seven patients with aniridia from four unrelated Spanish and French families, the

minigene approach showed complex splicing events. For this variant, partial retention of intron 2 was the main event observed in two different cell lines. Skipping of exon 2, alone or accompanied by the exons 3 and 4, was also detected in a lesser extent, mainly in retinal-derived cells. From this body of evidence, all these variants have been considered as causative variants in our cohort. *In fine*, three additional variants, absent from public databases, **c.-52+3_-52+4delAA**, **c.-52+3_-52+6delinsTG**

and **c.-52+5delIG**, were identified here for the first time, showing a similar effect on the canonical donor splicing site of exon 3 when using the above minigene constructions. Thus, as they are likely to mimic *in vivo* the effect of the previously reported pathogenic **c.-52+1G>A** variant (Gronskov et al. 1999), the above-mentioned new variants can reasonably be considered also as causative. Additionally, all these pathogenic variants segregated with the phenotype following a dominant inheritance in the families.

In general, disruption of consensus splicing regions may be relatively well predicted by *in silico* prediction tools. However, the specific nature of the resulting spliced transcripts and even their clinical significance is not always straightforward and it cannot be accurately defined without *in vivo* or *in vitro* testing (Spurdle et al. 2008). Thus, *in vitro* splicing assays represented a convenient approach to provide evidence they hamper normal splicing. This is likely to provide useful functional assays in all patients whose RNA is not available. In this work, all the analyzed variants were shown to have a high impact on splicing of non-coding exons of *PAX6* using one single minigene construct. Our strategy of *in vitro* splicing assays allows analyzing multiple variants in the 5'UTR of *PAX6* that included not only the three non-coding exons (exons 1 to 3) but also, the first coding exon (exon 4) and most of the intronic regions. Thus, this approach allows improved reproducibility of physiological splicing patterns by keeping the genomic context. However, it should be considered only as a supporting test, since minigene findings should be validated in RNA from patients whenever possible. We had the opportunity to derive LCLs from carriers of the novel **c.-129+1G>A** variant. We noticed in previous attempts that, in most cases, canonical splice site variants allowed only the recovery of mRNA produced by the normal allele and that it was likely that the mutant allele led to aberrant transcripts carrying premature stop codons and was then sensitive to NMD (Vincent et al. 2003). For this mutant allele, we thus conducted cell culture assays in the presence of different inhibitors of translation in an attempt to recover all the processed *PAX6* transcripts before they could be degraded. Despite these handling precautions, we failed to detect any aberrant isoforms in cultured LCLs. It is well-known that outcome of splicing assays may be biased by experimental and technical limitations, such as those related with analytical sensitivity to detect low abundance transcripts (Whiley et al. 2014). However, RT-PCR performed on LCLs usually allows the detection of numerous splice forms of *PAX6* (Vincent et al. 2003). Thus, total NMD might probably explain the absence of detection of any aberrant transcripts for the **c.-129+1G>A** variant as it was shown for other splicing mutations (Vincent et al. 2003).

It is worth raising the question whether the observed aberrant transcripts are pathogenic *in vivo*. In general, it is well accepted that transcripts lacking an open reading frame (ORF),

as those involving skipping of exon 4 of *PAX6*, will be deleterious, having a direct impact on the protein function or on mRNA expression level. In some instances, canonical splicing variants lead to in-frame single exon deletions affecting domains of unknown function, rendering pathological consequences assessment difficult (Spurdle et al. 2008). As for the *PAX6* aberrant transcripts only involving the non-coding exons of *PAX6*, the translation of full-length proteins is expected. In that case, the pathological impact, which is that of a haploinsufficiency, could be related to an imbalance in the relative amounts of endogenous transcripts due to either a decreased stability of aberrant splice products, or a defect in regulatory regions controlling the *PAX6* expression. Moreover, major changes occur in the mRNA structure and length that could be able to modify the secondary structure of the mature mRNA and its stability (Mignone et al. 2002). These are plausible causes for the apparently lack of *in vivo* expression of aberrant isoforms in LCLs for the **c.-129+G>A** variant.

With respect to the variation **c.-118_-117delTT**, a study of transcripts from patient samples does not reveal any effect on splicing (both alleles were detected by Sanger sequencing) and a minigene strategy could not be performed for this mutation. Thus, this variation does not seem to affect normal splicing. As a similar variant **c.-116_-112del** has been previously identified (Gronskov et al. 1999), other pathological mechanisms could then be proposed for these closely located mutations. For example, these two variants could define or modify additional upstream Open Reading Frame (uORF). In most cases, translation of uORF is associated with a reduction in downstream main ORF expression because, on one hand, reinitiation from the ORF is generally inefficient, and on the other hand, the ribosome can start to translate uORF and stall, triggering the degradation of mRNA (Hinnebusch et al. 2016). This is also illustrated by the demonstration that, in 100 eukaryotic genes, these uORFs were associated with a reduction in the level of synthesized proteins (Calvo et al. 2009). The reduction of ORF expression, secondary to the use of uORF, is correlated with the degree of conservation of the alternative codon during evolution, its integration within a Kozak sequence, its remoteness from the cap and the existence of multiple uORFs in the 5'UTR (Calvo et al. 2009). Nevertheless, although the role of these uORFs in translational regulation is still poorly understood, their importance is suggested by different findings: (1) nearly half of human transcripts have at least one uORF, (2) their important degree of conservation within the 5'UTR during the evolution of the species and finally (3) their implication in human pathologies (Somers et al. 2015; Wiestner et al. 1998). In the case of the variants **c.-118_-117del** and **c.-116_-112del**, *in silico* predictions found a similar effect consisting in the appearance of a novel ORF upstream of the main ORF. Thus, despite being speculative, this enlightens a novel possible pathogenic effect to be further studied.

In addition, a likely splicing defect has also been predicted for a deep intronic variation **c.357+136G>A** in the intron 6 of the *PAX6* gene, never reported previously, that was found de novo in a sporadic case of this cohort. Similarly, in a microphthalmic mouse, Puk O et al. (Puk et al. 2013) described a profound mutation in intron 6 (+1 kb) leading to the insertion of a novel cryptic pseudoexon introducing a premature stop codon. The intronic variant we report here is predicted to introduce a cryptic splice acceptor site at position c.357+138 that could potentially cause the aberrant splicing of exon 7 by partial retention of intron 6. However, RNA expression studies have not been carried out due to the impossibility of establishing LCL in this patient. Thus, it remains to elucidate if this variant could lead to the partial intron retention or the inclusion of an alternative cryptic exon.

Other variants of unknown significance have been also identified in our study, for which there are some difficulties to demonstrate their pathogenic effects and then their causality. This is the case of a rare variation **c.*4389_4390delAA** (rs777790874) found in the 3'UTR segregating with aniridia in a familial case. Even if this variant is predicted to affect miRNA recruitment by in silico predictions, it is not easy to demonstrate its real pathogenicity.

A more precise interpretation of the 13 remaining variations (Table 2) of unknown significance was partly impaired by the lack of segregation analysis. Indeed, these variants are, for the most part (11/13) located in the remote *cis*-regulatory elements of *PAX6*. Interestingly, none of these identified variants was located in the SIMO enhancer as reported by Bhatia et al. (Bhatia et al. 2013). Therefore, due to the recent report from Bhatia et al. (Bhatia et al. 2013), one can consider these variants as strong candidates. However, segregation analysis is not possible due to unavailability of parent's DNA. Nevertheless, we decided to expose the results of the entire series, as the number of candidate hits in the non-coding region of *PAX6* was important information that might question the routine procedure of laboratories involved in aniridia diagnosis.

We confirm here that CNVs involving the *PAX6* downstream region are a frequent cause of aniridia, representing up to 20% (9/47) of patients without *PAX6* coding variants in our cohort. Our work highlights the need to implement the study of these regulatory regions by CNV analysis in the routine genetic study of aniridia, which can be easily searched as a first-line analysis using CGH array. Here, we could be able to detect them using targeted strategies not only by CGH array, as we previously published (Blanco-Kelly et al. 2017), but also using NGS data through a novel comparative read-depth strategy by capture of 3' regulatory enhancers. In such a combined strategy, SNVs and CNVs could be straightforward analyzed, increasing the diagnostic rate for aniridia. Therefore, our targeting NGS strategy allows obtaining earlier molecular diagnoses for aniridia and

WAGR syndrome by a relatively low-cost, robust and clinically useful approach. This is especially important for clinical follow-up of patients, discarding the medically actionable Wilms tumor.

Our exhaustive analysis of non-coding regions enables to reveal the astonishing frequency of mutations located in the *PAX6* untranslated regions in patients with aniridia. In general, even if we know that variants in the untranslated regions can be pathogenic, for instance through altering splicing, creating or disrupting open reading frames or recruitment of miRNAs (Chatterjee and Pal 2009), they are finally not commonly searched in routine diagnosis analyses. Until now, only 11 variants of *PAX6* were reported in these untranslated regions (the Human *PAX6* mutation Database) (Fig. 1), most of them were reported once and just few of them were considered as causative mutations. With this work, we can realize that the splicing errors of the non-coding exons constitute a non-negligible cause of aniridia, and even if the reported variants are, for most of them, private mutations, they affect almost all the same canonical splicing regions suggesting a similar and common pathogenic effect. This work illustrates the need for searching such variants in untranslated regions in routine diagnostic procedures. Unlike variants identified in the remote *cis*-regulatory sequences, the accountability of such variants in the phenotype is, for most of them, indisputable or easily demonstrable.

Thus, among the patient's population studied in this work, that is, patients with aniridia without any mutation in the *PAX6* coding sequence, around 45% (21/47) benefit now from a certain diagnosis and around 55% (26/47) remain undiagnosed with, for one third of them (9/26), candidate variants whom pathogenicity might be confirmed to enable a definitive diagnosis. We may emphasize the absence of detected point mutation in the SIMO element in this cohort, which appears to be a very rare event.

Finally, around 36% of this cohort (17/47) did not show any (likely) pathogenic changes involving the *PAX6* locus. Several assumptions can be made: (1) variants could be present in incompletely covered areas, but it is unlikely as the actual coverage of our targeting approaches 100%, (2) it could be the involvement of still unknown *PAX6* regulatory elements not previously described, (3) it could be the involvement of another gene not yet discovered and finally (4) it might be the undetected presence of copy neutral structural variants (f.i. balanced translocation). However, 12 out of these 17 patients benefited from conventional karyotyping, excluding balanced rearrangements such as reciprocal translocations or inversions. Thus, an exome analysis strategy could be applied in these undiagnosed patients, possibly supplemented with an analysis of the entire genome. In light of this work, WGS would be the most appropriate strategy for searching variants both in new genes or regulatory regions and even copy neutral structural variants.

In brief, our targeted NGS approach allows the analysis of the entire *PAX6* locus, which we now routinely use as a first-line technique in all patients with aniridia. The “aniridia panels” we presented hereby show a high detection rate of variants in the non-coding sequences of *PAX6*. This new analysis strategy brings molecular diagnosis to more than 90% of aniridia patients and allowed us to confirm that *PAX6* remains the major gene for aniridia.

Acknowledgements We acknowledge generous support from the families published in this article. We thank Salvador Marti and Virginia Corrochano from CIBERER Biobank (Valencia, Spain) for their help in the generation of LCLs. This work was supported by *Géniris* funding awarded to Dr Julie Plaisancié, *Retina France* funding awarded to Jean-Michel Rozet and Patrick Calvas, Spanish Institute of Health Carlos III (ISCIII)/European Regional Development Fund (ERDF) (PI17_01164 and CP12/03256), Spanish Ministry of Economy and Competitiveness/ERDF (MINECO, SAF2013-46943-R), Spanish Federation of Rare Diseases (FEDER) and Mutua Madrileña Foundation, funding awarded to Marta Cortón, CIBERER (06/07/0036), the University Chair UAM-IIS-FJD of Genomic Medicine, the Ramon Areces Foundation and Regional Government of Madrid (CAM, B2017/BMD3721), funding awarded to Carmen Ayuso. Marta Cortón is sponsored by the Miguel Servet Program (CP12/03256 and CPII17_00006) from ISCIII and Maria Tarilonte received a Conchita Rabago PhD fellowship.

Compliance with ethical standards

Conflict of interest The authors declare that they have no competing interests (Note: Nucleotide sequence data reported are available in the ClinVar database under the accession numbers: SCV000803674 to SCV000803681).

Ethical approval All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

URLs *Human *PAX6* Mutation Database: http://lsdb.hgu.mrc.ac.uk/home.php?select_db=PAX6. *HGMD Professional: <http://www.biobase-international.com/product/hgmd>. *GnomAD: <http://gnomad.broadinstitute.org>. *Decipher: <https://decipher.sanger.ac.uk>. *ESEfinder: <http://kraimer01.cshl.edu/cgi-bin/tools/ESE3/esefinder.cgi?process=home>. *Ex-skip: <http://ex-skip.img.cas.cz>. *Human Slicing Finder: <http://www.umd.be/HSF>. *SPARN software: <http://tools.genes.toronto.edu>. *The Human *PAX6* mutation Database: http://lsdb.hgu.mrc.ac.uk/home.php?select_db=PAX6. *CIBERER Spanish Variant Server: <http://cvs.babelomics.org>. *ORFfinder: <https://www.ncbi.nlm.nih.gov/orffinder>. *OrfPredictor: <https://bioinformatics.yu.edu/tools/OrfPredictor.html>.

References

- Aggarwal S, Jinda W, Limwongse C, Atchaneeyasakul LO, Phadke SR (2011) Run-on mutation in the *PAX6* gene and chorioretinal degeneration in autosomal dominant aniridia. *Mol Vis* 17:1305–1309
- Ansari M, Rainger J, Hanson IM, Williamson KA, Sharkey F, Harewood L, Sandilands A, Clayton-Smith J, Dollfus H, Bitoun P, Meire F, Fantes J, Franco B, Lorenz B, Taylor DS, Stewart F, Willoughby CE, McEntagart M, Khaw PT, Clericuzio C, Van Maldergem L, Williams D, Newbury-Ecob R, Traboulsi EI, Silva ED, Madlom MM, Goudie DR, Fleck BW, Wieczorek D, Kohlhase J, McTrusty AD, Gardiner C, Yale C, Moore AT, Russell-Eggitt I, Islam L, Lees M, Beales PL, Tuft SJ, Solano JB, Splitt M, Hertz JM, Prescott TE, Shears DJ, Nischal KK, Doco-Fenzy M, Prieur F, Temple IK, Lachlan KL, Damante G, Morrison DA, van Heyningen V, FitzPatrick DR (2016) Genetic analysis of ‘PAX6-negative’ individuals with aniridia or gillespie syndrome. *PLoS One* 11:e0153757. <https://doi.org/10.1371/journal.pone.0153757>
- Axton RA, Hanson IM, Love J, Seawright A, Prosser J, van Heyningen V (1997) Combined SSCP/heteroduplex analysis in the screening for *PAX6* mutations. *Mol Cell Probes* 11:287–292. <https://doi.org/10.1006/mcpr.1997.0117>
- Azuma N, Yamaguchi Y, Handa H, Hayakawa M, Kanai A, Yamada M (1999) Missense mutation in the alternative splice region of the *PAX6* gene in eye anomalies. *Am J Hum Genet* 65:656–663. <https://doi.org/10.1086/302529>
- Azuma N, Yamaguchi Y, Handa H, Tadokoro K, Asaka A, Kawase E, Yamada M (2003) Mutations of the *PAX6* gene detected in patients with a variety of optic-nerve malformations. *Am J Hum Genet* 72:1565–1570. <https://doi.org/10.1086/375555>
- Bayrakli F, Guney I, Bayri Y, Ercan-Sencicek AG, Ceyhan D, Canakaya T, Mason C, Bilguvar K, Bayrakli S, Mane SM, State MW, Gunel M (2009) A novel heterozygous deletion within the 3’ region of the *PAX6* gene causing isolated aniridia in a large family group. *J Clin Neurosci* 16:1610–1614. <https://doi.org/10.1016/j.jocn.2009.03.022>
- Bhatia S, Bengani H, Fish M, Brown A, Divizia MT, de Marco R, Damante G, Grainger R, van Heyningen V, Kleinjan DA (2013) Disruption of autoregulatory feedback by a mutation in a remote, ultraconserved *PAX6* enhancer causes aniridia. *Am J Hum Genet* 93:1126–1134. <https://doi.org/10.1016/j.ajhg.2013.10.028>
- Bhatia S, Monahan J, Ravi V, Gautier P, Murdoch E, Brenner S, van Heyningen V, Venkatesh B, Kleinjan DA (2014) A survey of ancient conserved non-coding elements in the *PAX6* locus reveals a landscape of interdigitated cis-regulatory archipelagos. *Dev Biol* 387:214–228. <https://doi.org/10.1016/j.ydbio.2014.01.007>
- Bhuvanagiri M, Schlitter AM, Hentze MW, Kulozik AE (2010) NMD: RNA biology meets human genetic medicine. *Biochem J* 430:365–377. <https://doi.org/10.1042/BJ20100699>
- Blanco-Kelly F, Palomares M, Vallespin E, Villaverde C, Martin-Arenas R, Velez-Monsalve C, Lorda-Sanchez I, Nevado J, Trujillo-Tiebas MJ, Lapunzina P, Ayuso C, Corton M (2017) Improving molecular diagnosis of aniridia and WAGR syndrome using customized targeted array-based CGH. *PLoS One* 12:e0172363. <https://doi.org/10.1371/journal.pone.0172363>
- Boonstra N, Limburg H, Tijmes N, van Genderen M, Schuil J, van Nispen R (2012) Changes in causes of low vision between 1988 and 2009 in a Dutch population of children. *Acta Ophthalmol* 90:277–286. <https://doi.org/10.1111/j.1755-3768.2011.02205.x>
- Calvo SE, Pagliarini DJ, Mootha VK (2009) Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc Natl Acad Sci USA* 106:7507–7512. <https://doi.org/10.1073/pnas.0810916106>
- Ceroni F, Aguilera-Garcia D, Chassaing N, Bax DA, Blanco-Kelly F, Ramos P, Tarilonte M, Villaverde C, da Silva LRJ, Ballesta-Martinez MJ, Sanchez-Soler MJ, Holt RJ, Cooper-Charles L, Bruty J, Wallis Y, McMullan D, Hoffman J, Bunyan D, Stewart A, Stewart H, Lachlan K, Study DDD, Fryer A, McKay V, Roume J, Dureau P, Saggat A, Griffiths M, Calvas P, Ayuso C, Corton M, Ragge NK (2018) New *GJA8* variants and phenotypes highlight its critical role in a broad spectrum of eye anomalies. *Hum Genet*. <https://doi.org/10.1007/s00439-018-1875-2>
- Charbonnier F, Raux G, Wang Q, Drouot N, Cordier F, Limacher JM, Saurin JC, Puisieux A, Olschwang S, Frebourg T (2000) Detection of exon deletions and duplications of the mismatch repair

- genes in hereditary nonpolyposis colorectal cancer families using multiplex polymerase chain reaction of short fluorescent fragments. *Cancer Res* 60:2760–2763
- Chassaing N, Causse A, Vigouroux A, Delahaye A, Alessandri JL, Boespflug-Tanguy O, Boute-Benejean O, Dollfus H, Duban-Bedu B, Gilbert-Dussardier B, Giuliano F, Gonzales M, Holder-Espinasse M, Isidor B, Jacquemont ML, Lacombe D, Martin-Coignard D, Mathieu-Dramard M, Odent S, Picone O, Pinson L, Quelin C, Sigaudy S, Toutain A, Thauvin-Robinet C, Kaplan J, Calvas P (2013) Molecular findings and clinical data in a cohort of 150 patients with anophthalmia/microphthalmia. *Clin Genet*
- Chatterjee S, Pal JK (2009) Role of 5'- and 3'-untranslated regions of mRNAs in human diseases. *Biol Cell* 101:251–262. <https://doi.org/10.1042/BC20080104>
- Cheng F, Song W, Kang Y, Yu S, Yuan H (2011) A 556 kb deletion in the downstream region of the PAX6 gene causes familial aniridia and other eye anomalies in a Chinese family. *Mol Vis* 17:448–455
- Desviat LR, Perez B, Ugarte M (2012) Minigenes to confirm exon skipping mutations. *Methods Mol Biol* 867:37–47. https://doi.org/10.1007/978-1-61779-767-5_3
- Fantes J, Redeker B, Breen M, Boyle S, Brown J, Fletcher J, Jones S, Bickmore W, Fukushima Y, Mannens M et al (1995) Aniridia-associated cytogenetic rearrangements suggest that a position effect may cause the mutant phenotype. *Hum Mol Genet* 4:415–422
- Glaser T, Walton DS, Maas RL (1992) Genomic structure, evolutionary conservation and aniridia mutations in the human PAX6 gene. *Nat Genet* 2:232–239. <https://doi.org/10.1038/ng1192-232>
- Griffin C, Kleinjan DA, Doe B, van Heyningen V (2002) New 3' elements control Pax6 expression in the developing pretectum, neural retina and olfactory region. *Mech Dev* 112:89–100
- Gronskov K, Rosenberg T, Sand A, Brondum-Nielsen K (1999) Mutational analysis of PAX6: 16 novel mutations including 5 missense mutations with a mild aniridia phenotype. *Eur J Hum Genet* 7:274–286. <https://doi.org/10.1038/sj.ejhg.5200308>
- Gupta SK, De Becker I, Tremblay F, Guernsey DL, Neumann PE (1998) Genotype/phenotype correlations in aniridia. *Am J Ophthalmol* 126:203–210
- Hanson IM, Fletcher JM, Jordan T, Brown A, Taylor D, Adams RJ, Punnett HH, van Heyningen V (1994) Mutations at the PAX6 locus are found in heterogeneous anterior segment malformations including Peters' anomaly. *Nat Genet* 6:168–173. <https://doi.org/10.1038/ng0294-168>
- Hanson I, Churchill A, Love J, Axton R, Moore T, Clarke M, Meire F, van Heyningen V (1999) Missense mutations in the most ancient residues of the PAX6 paired domain underlie a spectrum of human congenital eye malformations. *Hum Mol Genet* 8:165–172
- Hever AM, Williamson KA, van Heyningen V (2006) Developmental malformations of the eye: the role of PAX6, SOX2 and OTX2. *Clin Genet* 69:459–470
- Hill RE, Favor J, Hogan BL, Ton CC, Saunders GF, Hanson IM, Prosser J, Jordan T, Hastie ND, van Heyningen V (1991) Mouse small eye results from mutations in a paired-like homeobox-containing gene. *Nature* 354:522–525
- Hinnebusch AG, Ivanov IP, Sonenberg N (2016) Translational control by 5'-untranslated regions of eukaryotic mRNAs. *Science* 352:1413–1416. <https://doi.org/10.1126/science.aad9868>
- Jordan T, Hanson I, Zaletayev D, Hodgson S, Prosser J, Seawright A, Hastie N, van Heyningen V (1992) The human PAX6 gene is mutated in two patients with aniridia. *Nat Genet* 1:328–332. <https://doi.org/10.1038/ng0892-328>
- Kammandel B, Chowdhury K, Stoykova A, Aparicio S, Brenner S, Gruss P (1999) Distinct cis-essential modules direct the time-space pattern of the Pax6 gene activity. *Dev Biol* 205:79–97. <https://doi.org/10.1006/dbio.1998.9128>
- Kleinjan DA, Seawright A, Schedl A, Quinlan RA, Danes S, van Heyningen V (2001) Aniridia-associated translocations, DNase hypersensitivity, sequence comparison and transgenic analysis redefine the functional domain of PAX6. *Hum Mol Genet* 10:2049–2059
- Kleinjan DA, Seawright A, Childs AJ, van Heyningen V (2004) Conserved elements in Pax6 intron 7 involved in (auto)regulation and alternative transcription. *Dev Biol* 265:462–477
- Kleinjan DA, Seawright A, Mella S, Carr CB, Tyas DA, Simpson TI, Mason JO, Price DJ, van Heyningen V (2006) Long-range downstream enhancers are essential for Pax6 expression. *Dev Biol* 299:563–581. <https://doi.org/10.1016/j.ydbio.2006.08.060>
- McBride DJ, Buckle A, van Heyningen V, Kleinjan DA (2011) DNaseI hypersensitivity and ultraconservation reveal novel, interdependent long-range enhancers at the complex Pax6 cis-regulatory region. *PLoS One* 6:e28616. <https://doi.org/10.1371/journal.pone.0028616>
- Mignone F, Gissi C, Liuni S, Pesole G (2002) Untranslated regions of mRNAs. *Genome Biol* 3:REVIEWS0004
- Navratilova P, Becker TS (2009) Genomic regulatory blocks in vertebrates and implications in human disease. *Brief Funct Genomic Proteomic* 8:333–342. <https://doi.org/10.1093/bfgp/elp019>
- Nelson LB, Spaeth GL, Nowinski TS, Margo CE, Jackson L (1984) Aniridia. A review. *Surv Ophthalmol* 28:621–642
- Perveen R, Lloyd IC, Clayton-Smith J, Churchill A, van Heyningen V, Hanson I, Taylor D, McKeown C, Super M, Kerr B, Winter R, Black GC (2000) Phenotypic variability and asymmetry of Rieger syndrome associated with PITX2 mutations. *Invest Ophthalmol Vis Sci* 41:2456–2460
- Puk O, Yan X, Sabrautzki S, Fuchs H, Gailus-Durner V, Hrabe de Angelis M, Graw J (2013) Novel small-eye allele in paired box gene 6 (Pax6) is caused by a point mutation in intron 7 and creates a new exon. *Mol Vis* 19:877–884
- Quiring R, Walldorf U, Kloter U, Gehring WJ (1994) Homology of the eyeless gene of Drosophila to the Small eye gene in mice and Aniridia in humans. *Science* 265:785–789
- Ravi V, Bhatia S, Gautier P, Loosli F, Tay BH, Tay A, Murdoch E, Coutinho P, van Heyningen V, Brenner S, Venkatesh B, Kleinjan DA (2013) Sequencing of Pax6 loci from the elephant shark reveals a family of Pax6 genes in vertebrate genomes, forged by ancient duplications and divergences. *PLoS Genet* 9:e1003177. <https://doi.org/10.1371/journal.pgen.1003177>
- Redeker EJ, de Visser AS, Bergen AA, Mannens MM (2008) Multiplex ligation-dependent probe amplification (MLPA) enhances the molecular diagnosis of aniridia and related disorders. *Mol Vis* 14:836–840
- Richards S, Aziz N, Bale S, Bick D, Das S, Gastier-Foster J, Grody WW, Hegde M, Lyon E, Spector E, Voelkerding K, Rehml HL, Committee ALQA (2015) Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med* 17:405–424. <https://doi.org/10.1038/gim.2015.30>
- Robinson DO, Howarth RJ, Williamson KA, van Heyningen V, Beal SJ, Crolla JA (2008) Genetic analysis of chromosome 11p13 and the PAX6 gene in a series of 125 cases referred with aniridia. *Am J Med Genet A* 146A:558–569. <https://doi.org/10.1002/ajmg.a.32209>
- Sanchez-Navarro I, Blanco-Kelly LRJdS, Zurita F, Sanchez-Bolivar O, Villaverde N, Lopez-Molina C, Garcia-Sandoval MI, Tahsin-Swafiri B, Minguez S, Riveiro-Alvarez P, Lorda R, Sanchez-Alcudia I, Perez-Carro R, Valverde R, Liu D, Tian Y, Hakonarson L, Avila-Fernandez H, Corton A, Ayuso M C (2018) Combining targeted panel-based resequencing and copy-number variation analysis for the diagnosis of inherited syndromic retinopathies and associated ciliopathies. *Sci Rep* 8:5285. <https://doi.org/10.1038/s41598-018-23520-1>

- Shaw MW, Falls HF, Neel JV (1960) Congenital Aniridia. *Am J Hum Genet* 12:389–415
- Simioni M, Vieira TP, Sgardoli IC, Freitas EL, Rosenberg C, Maurer-Morelli CV, Lopes-Cendes I, Fett-Conte AC, Gil-da-Silva-Lopes VL (2012) Insertional translocation of 15q25-q26 into 11p13 and duplication at 8p23.1 characterized by high resolution arrays in a boy with congenital malformations and aniridia. *Am J Med Genet A* 158A:2905–2910. <https://doi.org/10.1002/ajmg.a.35603>
- Simpson TI, Price DJ (2002) Pax6; a pleiotropic player in development. *Bioessays* 24:1041–1051. <https://doi.org/10.1002/bies.10174>
- Singh S, Mishra R, Arango NA, Deng JM, Behringer RR, Saunders GF (2002) Iris hypoplasia in mice that lack the alternatively spliced Pax6(5a) isoform. *Proc Natl Acad Sci USA* 99:6812–6815. <https://doi.org/10.1073/pnas.102691299>
- Skeens HM, Brooks BP, Holland EJ (2011) Congenital aniridia variant: minimally abnormal irides with severe limbal stem cell deficiency. *Ophthalmology* 118:1260–1264. <https://doi.org/10.1016/j.ophtha.2010.11.021>
- Somers J, Wilson LA, Kilday JP, Horvilleur E, Cannell IG, Poyry TA, Cobbold LC, Kondrashov A, Knight JR, Puget S, Grill J, Grundy RG, Bushell M, Willis AE (2015) A common polymorphism in the 5' UTR of ERCC5 creates an upstream ORF that confers resistance to platinum-based chemotherapy. *Genes Dev* 29:1891–1896. <https://doi.org/10.1101/gad.261867.115>
- Spurdle AB, Couch FJ, Hogervorst FB, Radice P, Sinilnikova OM, Group IUGVW (2008) Prediction and assessment of splicing alterations: implications for clinical testing. *Hum Mutat* 29:1304–1313. <https://doi.org/10.1002/humu.20901>
- St-Onge L, Sosa-Pineda B, Chowdhury K, Mansouri A, Gruss P (1997) Pax6 is required for differentiation of glucagon-producing alpha-cells in mouse pancreas. *Nature* 387:406–409. <https://doi.org/10.1038/387406a0>
- Ticho BH, Hilchie-Schmidt C, Egel RT, Traboulsi EI, Howarth RJ, Robinson D (2006) Ocular findings in Gillespie-like syndrome: association with a new PAX6 mutation. *Ophthalmic Genet* 27:145–149. <https://doi.org/10.1080/13816810600976897>
- Vasilyeva TA, Voskresenskaya AA, Kasmann-Kellner B, Khlebnikova OV, Pozdeyeva NA, Bayazutdinova GM, Kutsev SI, Ginter EK, Semina EV, Marakhonov AV, Zinchenko RA (2017) Molecular analysis of patients with aniridia in Russian Federation broadens the spectrum of PAX6 mutations. *Clin Genet* 92:639–644. <https://doi.org/10.1111/cge.13019>
- Vincent MC, Pujo AL, Olivier D, Calvas P (2003) Screening for PAX6 gene mutations is consistent with haploinsufficiency as the main mechanism leading to various ocular defects. *Eur J Hum Genet* 11:163–169
- Whiley PJ, de la Hoya M, Thomassen M, Becker A, Brandao R, Pedersen IS, Montagna M, Menendez M, Quiles F, Gutierrez-Enriquez S, De Leeneer K, Tenes A, Montalban G, Tserpelis D, Yoshimatsu T, Tirapo C, Raponi M, Caldes T, Blanco A, Santamarina M, Guidugli L, de Garibay GR, Wong M, Tancredi M, Fachal L, Ding YC, Kruse T, Lattimore V, Kwong A, Chan TL, Colombo M, De Vecchi G, Caligo M, Baralle D, Lazaro C, Couch F, Radice P, Southey MC, Neuhausen S, Houdayer C, Fackenthal J, Hansen TV, Vega A, Diez O, Blok R, Claes K, Wappenschmidt B, Walker L, Spurdle AB, Brown MA, consortium E (2014) Comparison of mRNA splicing assay protocols across multiple laboratories: recommendations for best practice in standardized clinical testing. *Clin Chem* 60:341–352. <https://doi.org/10.1373/clinchem.2013.210658>
- Wiestner A, Schlemper RJ, van der Maas AP, Skoda RC (1998) An activating splice donor mutation in the thrombopoietin gene causes hereditary thrombocythaemia. *Nat Genet* 18:49–52. <https://doi.org/10.1038/ng0198-49>
- Williams SC, Altmann CR, Chow RL, Hemmati-Brivanlou A, Lang RA (1998) A highly conserved lens transcriptional control element from the Pax-6 gene. *Mech Dev* 73:225–229
- Xie Q, Ung D, Khafizov K, Fiser A, Cvekl A (2014) Gene regulation by PAX6: structural-functional correlations of missense mutants and transcriptional control of Trpm3/miR-204. *Mol Vis* 20:270–282



Parental Mosaicism in *PAX6* Causes Intra-Familial Variability: Implications for Genetic Counseling of Congenital Aniridia and Microphthalmia

María Tarilonte^{1†}, Matías Morín^{2,3†}, Patricia Ramos¹, Marta Galdós⁴, Fiona Blanco-Kelly^{1,3}, Cristina Villaverde^{1,3}, Dolores Rey-Zamora², Gema Rebolleda⁵, Francisco J. Muñoz-Negrete⁵, Saoud Tahsin-Swafiri^{1,3}, Blanca Gener^{3,6}, Miguel-Angel Moreno-Pelayo^{2,3}, Carmen Ayuso^{1,3}, Manuela Villamar^{2,3*} and Marta Corton^{1,3*}

OPEN ACCESS

Edited by:

Yuk Fai Leung,
Purdue University, United States

Reviewed by:

Zi-Bing Jin,
Wenzhou Medical University, China
Larry Baum,
The University of Hong Kong,
Hong Kong

*Correspondence:

Manuela Villamar
taugen.hrc@salud.madrid.org
Marta Corton
mcorton@fjd.es

† These authors have contributed
equally to this work

Specialty section:

This article was submitted to
Genetic Disorders,
a section of the journal
Frontiers in Genetics

Received: 10 August 2018

Accepted: 26 September 2018

Published: 17 October 2018

Citation:

Tarilonte M, Morín M, Ramos P,
Galdós M, Blanco-Kelly F,
Villaverde C, Rey-Zamora D,
Rebolleda G, Muñoz-Negrete FJ,
Tahsin-Swafiri S, Gener B,
Moreno-Pelayo M-A, Ayuso C,
Villamar M and Corton M (2018)
Parental Mosaicism in *PAX6* Causes
Intra-Familial Variability: Implications
for Genetic Counseling of Congenital
Aniridia and Microphthalmia.
Front. Genet. 9:479.
doi: 10.3389/fgene.2018.00479

¹ Department of Genetics and Genomics, Instituto de Investigación Sanitaria de la Fundación Jiménez Díaz, University Hospital – Universidad Autónoma de Madrid, Madrid, Spain, ² Servicio de Genética, Instituto Ramón y Cajal de Investigación Sanitaria, Hospital Universitario Ramón y Cajal, Madrid, Spain, ³ Centre for Biomedical Network Research on Rare Diseases, Instituto de Salud Carlos III, Madrid, Spain, ⁴ Department of Ophthalmology, Cruces University Hospital, Bilbao, Spain, ⁵ Department of Glaucoma, Instituto Ramón y Cajal de Investigación Sanitaria, Hospital Universitario Ramón y Cajal, Madrid, Spain, ⁶ Department of Genetics, BioCruces Health Research Institute, Cruces University Hospital, Bilbao, Spain

Mutations in *PAX6* are involved in several developmental eye disorders. These disorders have considerable phenotypic variability, ranging from panocular forms of congenital aniridia and microphthalmia to isolated anomalies of the anterior or posterior segment. Here, we describe 3 families with variable inter-generational ocular expression of aniridia, iris coloboma, or microphthalmia, and an unusual transmission of *PAX6* mutations from an unaffected or mildly affected parent; all of which raised suspicion of gonosomal mosaicism. We first identified two previously known nonsense mutations and one novel likely pathogenic missense variant in *PAX6* in probands by means of targeted NGS. The subsequent segregation analysis by Sanger sequencing evidenced the presence of highly probable mosaic events in paternal blood samples. Mosaicism was further confirmed by droplet digital PCR analysis in several somatic tissues of mosaic fathers. Quantification of the mutant allele fraction in parental samples showed a marked deviation from 50%, with a range between 12 and 29% depending on cell type. Gonosomal mosaicism was definitively confirmed in one of the families thanks to the availability of a sperm sample from the mosaic father. Thus, the recurrence risk in this family was estimated to be about one-third. This is the first report confirming parental *PAX6* mosaicism as a cause of disease recurrence in aniridia and other related phenotypes. In addition, we demonstrated that post-zygotic mosaicism is a frequent and underestimated pathogenic mechanism in aniridia, explaining intra-familial phenotypic variability in many cases. Our findings may have substantial implications for genetic counseling in congenital aniridia. Thus, we also highlight the importance of comprehensive genetic screening of parents for new sporadic cases with aniridia or related developmental eye disease to more accurately assess recurrence risk.

In conclusion, somatic and/or gonosomal mosaicism should be taken into consideration as a genetic factor to explain not only families with unaffected parents despite multiple affected children but also variable expressivity, apparent *de novo* cases, and even uncharacterized cases of aniridia and related developmental eye disorders, apparently lacking *PAX6* mutations.

Keywords: parental mosaicism, *PAX6*, aniridia, variable expressivity, microphthalmia, post-zygotic variants

INTRODUCTION

PAX6 encodes a highly conserved homeodomain-containing transcription factor that plays pivotal roles in normal ocular and neural development (Cvekl and Callaerts, 2017). Dominant *PAX6* mutations lead to a spectrum of ocular developmental anomalies (ODAs) depending on the mutation type and gene dosage (van Heyningen and Williamson, 2002). *PAX6* haploinsufficiency, which results from loss-of-function variants or 11p13 microdeletions involving this gene or their 5' regulatory regions, is the major cause of congenital aniridia (MIM# 106210) (Hingorani et al., 2012). By contrast, missense mutations usually exhibit a moderate impact on *PAX6* functionality and are often associated with some atypical *PAX6*-associated phenotypes, such as mild forms of iris coloboma or isolated foveal hypoplasia, or more severe phenotypes of Peter's anomaly and microphthalmia (Hanson et al., 1999; Azuma et al., 2003; Nallathambi et al., 2006; Jia et al., 2010; Thomas et al., 2014). However, no obvious genotype-phenotype correlations have been established to date (Gupta et al., 1998; Gronskov et al., 1999; Dubey et al., 2015; Sannan et al., 2017).

Aniridia, characterized by the incomplete development of the iris and fovea, is the most frequent *PAX6*-related condition, with a worldwide incidence of 1:50,000–100,000 births (Hingorani et al., 2012). Patients exhibit photophobia, low visual acuity, and nystagmus. Phenotypic variability is commonly observed (Lee and Colby, 2013). Iris hypoplasia may manifest as a complete absence of the iris (or “aniridia”) but also may present with minor structural iris defects, leading to atypical iris coloboma, iris holes, or stromal hypoplasia (Skeens et al., 2011), symptoms which can be only recognizable on slit-lamp examination. Additionally, a wide range of abnormalities in the cornea, anterior chamber, lens, and optic nerve have been reported conferring a higher risk in patients for secondary glaucoma, cataracts, and aniridia-associated keratopathy (AAK), which further worsen the visual outcome (Lee and Colby, 2013). Inter-familial and intra-familial variability on disease onset and severity of these secondary symptoms manifesting as congenital, childhood, or even adult forms are also observed (Hingorani et al., 2012; Vasilyeva et al., 2017). However, the molecular mechanisms underlying the variable expressivity of the *PAX6* mutations have yet to be elucidated. Identifying and understanding the genetic mechanisms that affect the severity of the disease is essential to provide a more accurate diagnosis and better clinical management of *PAX6*-associated disorders.

Somatic or gonosomal mosaicism might explain part of the phenotypic variability and/or reduced penetrance in aniridia,

similarly to that reported in other dominant ocular disorders (Faivre et al., 2006; Beryozkin et al., 2016). Post-zygotic variants (PZV), arising as errors in DNA replication at early embryonic stages, have recently been revealed as a novel source of *de novo* and somatic mosaic variants thanks to the advent of more sensitive genotyping technologies (Acuna-Hidalgo et al., 2015, 2016). The developmental stage at which PZVs arise has a major influence on their frequency and distribution in affected tissues and thus on phenotypic expressivity and the recurrence risk in offspring (Acuna-Hidalgo et al., 2016). Interestingly, recent studies have showed that PZVs could account for up to 10% of rare neurodevelopmental disorders such as intellectual disability, epilepsy, and autism (Acuna-Hidalgo et al., 2015; Stosser et al., 2017; Myers et al., 2018); therefore, mosaicism may be underestimated in sporadic cases of other developmental diseases.

In aniridia, up to two-thirds of patients are sporadic (Netland et al., 2011; Lee and Colby, 2013) and thought to carry *de novo* mutations that are generally not detected in parental blood samples. Nevertheless, germline or even low-level somatic mosaicism in one of the parents cannot be ruled out, as multiple tissues are not usually tested during genetic screening. Germline mosaicism has long been postulated in *PAX6*-related disorders to explain the disease co-occurrence in several affected siblings with no family history of eye disorders (Reed and Falls, 1955; Gronskov et al., 1999; Deml et al., 2016; Riera et al., 2017). However, to our knowledge, *PAX6* mosaicism has never been reliably confirmed in a suspected family, mainly due to technical limitations in genetic testing and/or constraints on the availability of germinal or somatic DNA samples other than blood. To date, only a small number of exceptional cases of somatic mosaicism for 11p13 microdeletions have been reported in aniridia (Robinson et al., 2008; Erez et al., 2010; Huynh et al., 2017).

Here, we identify parental mosaicism in three families with *PAX6*-related ODAs in which gonosomal PZVs have been accurately assessed through quantitative droplet-digital PCR (ddPCR). Detection of mosaicism was very relevant to these families in terms of genetic counseling, which allowed for more accurate determination of the risk of recurrence in future offspring. Remarkably, our findings also highlight the inter-generational variable expressivity in these families, which seems to be clearly explained by the presence of parental mosaic *PAX6* mutations as well as by the further transmission of these mutations to the offspring. Thus, our work confirms that mosaicism is an underestimated cause of phenotypic variability and disease recurrence in aniridia and other *PAX6*-related ODAs.

MATERIALS AND METHODS

Patients

We studied a cohort of 247 unrelated Spanish families with ODAs consisting of 78 with congenital aniridia, 33 with anterior segment dysgenesis and 136 with other *PAX6*-related phenotypes, i.e., ocular coloboma, microphthalmia, isolated foveal or optic nerve hypoplasia. Patients and affected and healthy relatives were recruited from two public hospitals from Madrid (Spain), Fundación Jiménez Díaz University Hospital and University Hospital Ramón y Cajal. Genomic DNA was obtained from peripheral blood, saliva, urine, and sperm using standard procedures.

This study was designed in compliance with the tenets of the Helsinki Declaration, and patient enrollment was approved by the ethics committees of both institutions. All participants or their legal guardians provided written informed consent prior to their participation in this study.

Molecular Screening

Probands were screened according to a previously reported genetic algorithm for molecular diagnosis of *PAX6* defects (Blanco-Kelly et al., 2013). Pathogenic *PAX6* variants were screened by Sanger or next-generation sequencing (NGS). 11p13 microdeletions were studied by MLPA and/or custom CGH-arrays, as previously reported (Blanco-Kelly et al., 2017).

The proband of Family 1 was screened using a customized 151-gene panel (unpublished data). Both probands of Families 2 and 3 were studied by means of a custom 260-gene panel, as previously described (Ceroni et al., 2018). Briefly, library capture of all coding and non-coding exons and 20 bp of intronic boundaries was performed using HaloPlex or SureSelect QXT technologies (Agilent Technologies, Santa Clara, CA, United States). Massive sequencing was carried out using Illumina MiSeq or NextSeq 500 platforms running on paired-end mode at a minimum of 450X.

Bioinformatic analysis was performed using standard procedures and custom in-house pipelines for mapping, variant calling, and annotation. Pathogenicity prediction of missense variants was performed using CADD¹, M-CAP², and Alamut software (Interactive Biosoftware, France), which includes SIFT, Polyphen, MutationTaster, and Align GVGD. Population frequencies of the detected variants were assessed using gnomAD³ and CIBERER Spanish Variant Server⁴. Variants were also searched in the literature and in the *PAX6* database⁵. All variants detected by NGS were validated by Sanger sequencing using specifically designed primers (Available On Request). The mutation nomenclature was referred to the canonical RefSeq *PAX6* isoform NM_000280.4.

¹<http://cadd.gs.washington.edu/>

²<http://bejerano.stanford.edu/mcap/>

³<http://gnomad.broadinstitute.org/>

⁴<http://csvs.babelomics.org/>

⁵http://lsdb.hgu.mrc.ac.uk/home.php?select_db=PAX6

SNaPshot Assays

Amplified *PAX6* exons 5, 6, and 10 were analyzed for the presence of c.120C>A, c.178T>C, and c.771G>A variants using the ABI Prism SNaPshot Multiplex kit (Thermo Fisher). Reactions were performed in a 20 μ L mix containing 3 μ L purified PCR amplicons, 3 μ L SNaPshot Ready Multiplex Ready Reaction Mix, and 0.4 μ M of specific primers. Single base extension and further post-extension SAP treatment were performed following the manufacturer's protocol. Labeled products were separated on an ABI3130xl Genetic Analyzer and analyzed with GeneMapper v4 software.

ddPCR

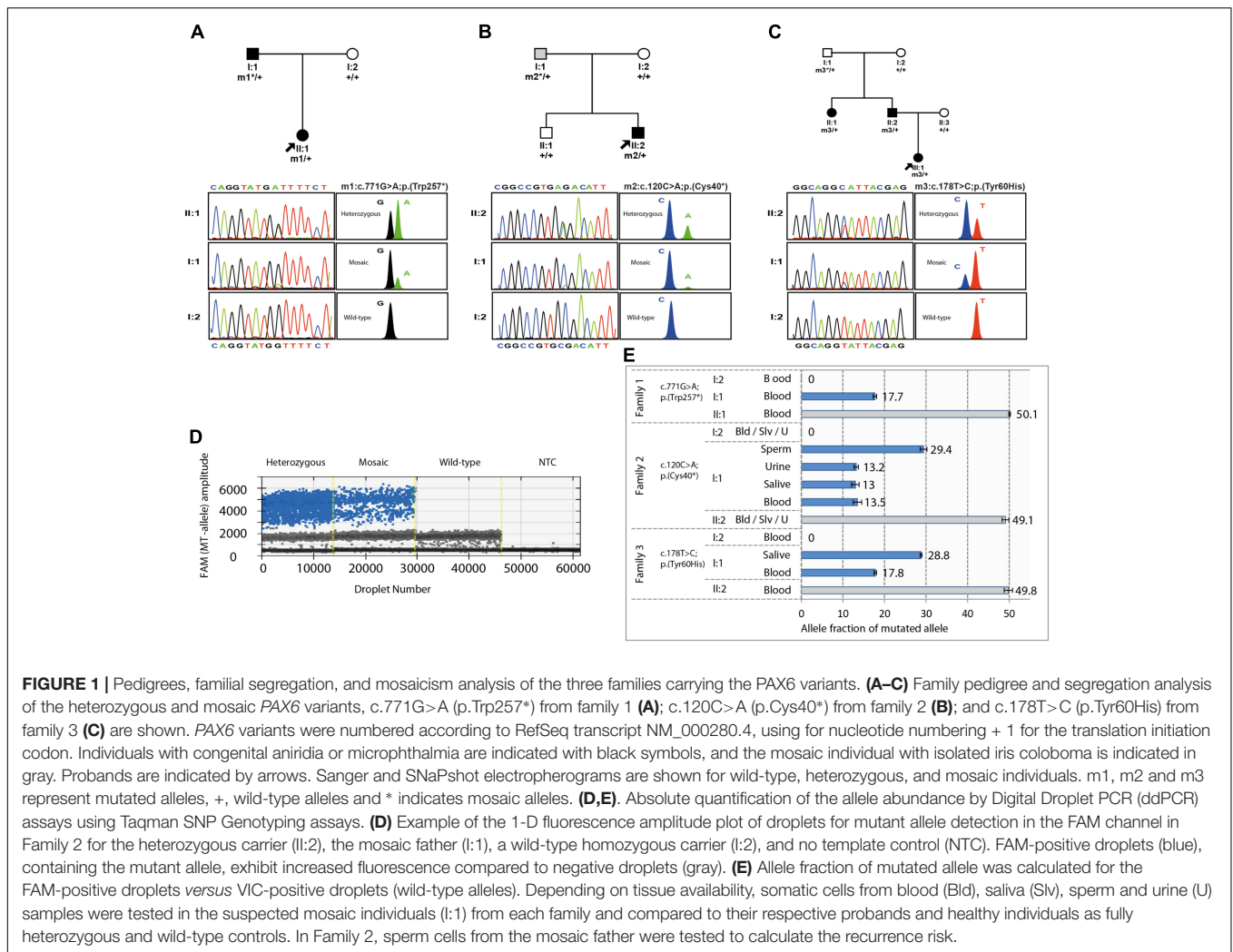
Custom or commercial TaqMan SNP Genotyping Assays (Thermo Fisher) were used for genotyping the *PAX6* variants c.120C>A, c.178T>C, and c.771G>A (rs121907929; commercial Assay ID: C_152371166_10). DNA samples were evaluated by ddPCR using the Droplet Digital PCR QX200 System (Bio-Rad Laboratories, Hercules, United States) in a 20 μ L PCR mix containing 10 μ L 2x ddPCR Supermix for probes (Bio-Rad), 900 nM target-specific PCR primers, 250 nM FAM-labeled (mutant-allele) probe, and 250 nM VIC-labeled (wild-type allele) probe. Droplet emulsion was thermally cycled on a C1000 Touch Thermal Cycler (Biorad) in the following conditions: denaturing at 95°C for 10 min, 40 cycles of PCR at 94°C for 30 s, and a single step of annealing/extension at 56.4°C for 1 min, and a final step to deactivate enzyme at 98°C for 10 min. Each sample was run in quadruplicate. Data were analyzed using Quantasoft v1.7 software (Bio-Rad) using Rare Event Detection settings.

RESULTS

Detection of Parental Mosaicism in *PAX6*

We investigated the presence of parental mosaicism associated with disease-causing *PAX6* variants in 3 families from a large cohort of Spanish patients with aniridia ($n = 78$) and other ODAs ($n = 169$). First, heterozygous *PAX6* variants were identified in probands of these 3 families through targeted NGS using custom approaches for *PAX6* screening. We found the previously reported nonsense variants c.120C>A;p.(Cys40*) and c.771G>A;p.(Trp257*) (Vincent et al., 2003; Graziano et al., 2007) in 2 probands with aniridia, and a novel likely pathogenic missense variant, c.178T>C;p.(Tyr60His), in a proband with non-syndromic microphthalmia. This last variant in the paired domain of *PAX6* affected a highly conserved tyrosine residue (GERP⁺⁺ score 5.4 and PhyloP score 9.3), was predicted to be pathogenic by all *in silico* tools, had not been publicly reported in population and mutation databases, and segregated dominantly in two consecutive generations of affected individuals.

The occurrence of germinal or gonosomal mosaicism was suspected in 2 of the 3 families as the most plausible factor to explain either the unusual recurrence of *PAX6*-related microphthalmia in the offspring of a healthy couple (Family 3, **Figure 1C**), or the remarkable inter-generational phenotypic variability between a proband and her father (Family 2, **Figure 1B**), respectively. In this sense, Sanger sequencing



revealed some traces of the mutant alleles (c.120C>A or c.178T>C) in paternal blood samples, for which the peak heights were substantially lower than those of their heterozygous children. Surprisingly, a third case of potential mosaicism was suspected in the affected father of Family 1. This individual displayed a similar biased sequencing pattern of the mutant allele c.771G>A (**Figure 1A**). Moreover, unequal amplification of the mutant and wild-type alleles was evidenced in all the suspected mosaic individuals when SNaPshot was used (**Figure 1**). Haplotyping analysis confirmed paternity in the three families (data not shown).

Mosaicism was finally evaluated by ddPCR using allele-specific Taqman genotyping, which enabled us to obtain a more accurate assessment of the allele imbalance by quantifying the mutant allele fraction (AF) (**Figure 1D**). We confirmed the presence of somatic mosaic *PAX6* alleles in these 3 individuals who showed variable AFs of 13–29% depending on the tested tissue (**Figure 1E**). As somatic mosaic variants were vertically transmitted to affected offspring in these families, we assumed that mutant alleles should also be present in the mosaic individuals in a relevant fraction of gonadal tissue. In this sense,

we were able to confirm gonosomal mosaicism in the father of Family 2, detecting the mutant allele in gamete-forming cells at a ratio of 29%, as well as in several somatic tissues, such as blood, saliva, and urine cells, with a lower AF of 13%. As a result, we accurately established that the transmission risk of the variant p.(Cys40*) in this family was about one-third. Unfortunately, sperm samples were unavailable in the other two mosaic fathers.

Overall, we have confirmed three cases of mosaicism in a cohort of 247 families with ODAs, representing a rate of 1.2%. Considering only aniridia, the mosaic rate is 3.8% with 2 families in a cohort of 78 families.

Phenotypic Variability Associated With Mosaic *PAX6* Mutations

Table 1 summarizes all the currently available clinical data from mosaic fathers and their respective affected relatives. Fully heterozygous carriers of the nonsense *PAX6* variants presented classical symptoms of congenital aniridia, including marked iris hypoplasia, photophobia, nystagmus, low visual acuity, as well as congenital or early-onset cataracts. Additional progressive

TABLE 1 | Genetic and clinical features of the families carrying postzygotic PAX6 variants.

Family ID	Causing variant	Patient ID	Age	Genotype (mosaic AF)	Iris anomaly	Microphthalmia	Nystagmus	Congenital or early-onset cataracts	Glaucoma	Keratopathy	Other features
1	c.771G>A p.(Trp257*)	I:1 (father)	50	Mosaic (18%, blood)	Complete aniridia (B)		+ (B)	+ (B)	+ (B)		Cataract surgery
		II:1	29	HET	Complete aniridia (B)		+ (B)	+ (B)	+ (B)	Limbal insufficiency (B)	Growth retardation: cataract surgery; amaurosis (RE)
2	c.120C>A p.(Cys40*)	I:1 (father)	59	Mosaic (13%, blood, saliva, and urine; 29%, sperm)	Coloboma (B)			+ (B)			
		II:2	25	HET	Complete aniridia (B)		+ (B)	+ (B)	+ (B)	Corneal opacity with limbal insufficiency (B)	Surgical interventions for limbal transplantation and implantation of Ahmed glaucoma valve (RE); BCVA: FC/0.05.
3	c.178T>C p.(Tyr60His)	I:1 (father)	80	Mosaic (18%, blood; 29%, saliva)							Asymptomatic; no ocular features
		II:1	43	HET		+ (B)	+ (B)	+ (B)			High hyperopia; secondary corectopia after iridectomy.
		II:2	38	HET		+ (B)	+ (B)	+ (B)	+ (B)		High hyperopia; secondary corectopia after iridectomy.
		III:1	10	HET		+ (B)	+ (B)	+ (B)		High hyperopia; secondary corectopia after iridectomy.	

Nucleotide numbering reflects cDNA in the reference sequence NM_000280.4 for PAX6. AF, allele fraction; B, bilateral; BCVA, best corrected visual acuity; FC, finger counting; HET, heterozygous; ID, identification; and RE, right eye.

ocular features such as glaucoma and AAK underlying limbal insufficiency were later developed. The proband of family 1, a 29-year-old female carrying the variant p.Trp257*, also suffered from mild growth retardation in childhood. However, she showed no signs of ataxia or intellectual disability, both systemic features previously reported in a patient with this same mutation (Graziano et al., 2007). Surprisingly, her affected father (I:1) also suffered from congenital aniridia and cataracts, despite carrying the mutant allele in only 18% of his blood cells. Unlike the daughter, he presented an apparently milder disease outcome with no signs of AAK at 50 years of age. By contrast, the mosaic carrier of the variant p.Cys40* (Family 2, I:1), with a slightly lower grade of mosaicism (13%) in several somatic tissues, showed neither congenital aniridia nor any apparent serious visual impairment. He was only diagnosed with iris coloboma and cataracts in both eyes in the fourth decade of life, after the birth of his son, who had aniridia.

The three heterozygous carriers of the missense p.(Tyr60His) variant presented a similar ocular phenotype of bilateral microphthalmia, congenital cataracts, nystagmus, and high hyperopia, without showing any systemic signs. Mosaicism was identified in the asymptomatic father (I:1, Family 3) of two affected siblings. Quantitative data from blood and buccal cells showed a relatively high mutant AF (18 and 29%, respectively). However, he has never complained of visual impairment, and his eyes were found to be normal in a thorough ophthalmic examination carried out in the 8th decade of life.

DISCUSSION

Phenotypic variability is a well-established phenomenon for *PAX6* defects (Hanson et al., 1999; Hingorani et al., 2012) that could be hypothetically attributable to genetic or epigenetic modifiers, pleiotropy, mosaicism, and/or gene dosage (Schedl et al., 1996; Hanson et al., 1999; Vincent et al., 2003; Chou et al., 2015; Cvekl and Callaerts, 2017; Yasue et al., 2017). Here, we present 3 families from our Spanish cohort of ODAs manifesting distinctly inter-generational differences in disease severity and/or ocular phenotype. We confirmed that the variable expressivity observed in all of these families resulted from parental mosaic *PAX6* mutations. Both families showed somatic mosaic alleles with a relatively high AF (> 10%) in healthy or mildly affected individuals and their vertical transmission to affected offspring, which suggests the existence of gonosomal PZVs. In one of them, gonosomal mosaicism was able to be assessed thanks to the availability of both somatic and gonadal samples, resulting in better estimation of the transmission risk of the mutation to future offspring. The estimated genetic risk of about one-third found in this family is quite higher than the risk level expected in a sporadic case carrying a *de novo* variant, which is usually considered to be negligible (Acuna-Hidalgo et al., 2016). In the other two families, the exact risk could not be assessed due to the unavailability of sperm samples. However, both fathers likely carried the pathogenic *PAX6* variant in a proportion of their germ line. Remarkably, the identification of mosaicism in an affected individual with aniridia, as we report here, involves modifying

the initial genetic counseling, since the likelihood of recurrence is obviously lower than in heterozygous carriers. In view of our findings, parental mosaicism should be considered in sporadic cases with *PAX6*-associated diseases, even if parents do not show any ocular abnormalities.

In the two mosaic individuals carrying nonsense *PAX6* mutations, similar levels of somatic mosaicism led to variable clinical outcomes ranging from congenital aniridia to considerably milder manifestations. Interestingly, we are the first to confirm low-level mosaicism in an affected individual clinically diagnosed with congenital aniridia but who manifested some apparently milder secondary ocular features compared to his fully heterozygous daughter. However, the mosaic father of a second family only showed minor iris defects that went unnoticed until adulthood. Therefore, mosaicism levels in blood are not a reliable indicator of the range and severity of ocular features. The variable phenotypic expression observed in these mosaic individuals could be influenced by different factors, including specific mutational effects, gene dosage, and lineage and timing-dependent *PAX6* functionality, not only during ocular development but also in adult tissues (Shaham et al., 2012; Gregory-Evans et al., 2014; Cvekl and Callaerts, 2017). In this sense, previous studies with transgenic and, more recently, genetically edited animal models confirmed that the patterning and differentiation of different ocular structures and cell lineages require different thresholds of *PAX6* activity throughout eye embryogenesis (Davis-Silberman et al., 2005; Yasue et al., 2017). Specifically, somatic *PAX6* mosaicism in CRISPR-edited mouse embryos caused varying degrees of ODAs, as recently published (Yasue et al., 2017). Therefore, the identification of mosaic individuals for *PAX6* mutations both in humans and animal models manifesting variable expressivity strongly reinforces the idea that *PAX6* dosage plays a specific role in both eye development and phenotypic modulation. According to our findings and those reported using CRISPR models, somatic *PAX6* mosaicism might explain some of the approximately 10% of patients with aniridia who apparently lack *PAX6* mutations (Ansari et al., 2016; Bobilev et al., 2016; Vasilyeva et al., 2017) or even mild cases of ODAs (Hanson et al., 1999). The widespread use of more sensitive techniques, such as targeted deep sequencing and/or ddPCR, will allow for the identification of additional cases with somatic or gonosomal PZVs arising at early stages (Acuna-Hidalgo et al., 2016). However, due to limitations on tissue sampling, it will remain challenging to detect PZVs with low AF that could appear later in embryonic development or those restricted to ocular tissues.

It has long been postulated that germline mosaicism in healthy parents could be the most likely explanation for the exceptional disease recurrence in affected siblings with no family history of aniridia (Reed and Falls, 1955; Gronskov et al., 1999) or *PAX6*-related microphthalmia (Deml et al., 2016; Riera et al., 2017). To date, this possibility has not been fully confirmed due to the limited availability of somatic and gonadal tissues for genetic testing. Here, we report an apparent autosomal dominant 3-generation family with microphthalmia in which we first assumed a reduced penetrance due to both disease recurrence in the first

affected siblings and healthy parents. However, NGS analyses only revealed a novel likely pathogenic missense variant in *PAX6*, a highly penetrant gene. Similar missense mutations in the paired-domain of *PAX6* have been occasionally reported to cause microphthalmia (Dansault et al., 2007; Chassaing et al., 2013; Riera et al., 2017; Patel et al., 2018); thus, this variant seemed to be the cause of the phenotype in the family. To delve deeper into the causes of this unusual disease recurrence, we exhaustively investigated the possibility of parental germline mosaicism in the unaffected first generation. In the asymptomatic father of this family, an unexpectedly high mutant AF was detected in several somatic tissues from ectoderm and endoderm lineages indicating that the mosaic PZV had to emerge early in the first embryonic cell divisions. Additionally, the lack of ocular signs in this mosaic individual suggests that the *PAX6* activity had to reach an adequate threshold through optic cup morphogenesis. We could not discern whether this fact was related to a gene dosage effect underlying a lower mutant ratio in embryonic eye or if this variant is hypomorphic, similar to that described for other missense *PAX6* variants (Tang et al., 1997; Vincent et al., 2003). Despite the rarity of *PAX6*-associated microphthalmia, with fewer than 10 variants described to date (Deml et al., 2016), it is worth noting that disease recurrence has putatively been associated with mosaicism in a healthy progenitor in two additional families (Deml et al., 2016; Riera et al., 2017). Thus, the possibility of *PAX6* mosaicism should also be considered when explaining occurrence in two or more affected siblings with microphthalmia in the absence of family history.

Recent evidence has suggested that transmission of parental mosaicism could explain up to 10% of apparently *de novo* mutations in some pathologies, and thus mosaicism rates might be higher than previously expected (Acuna-Hidalgo et al., 2016; Stosser et al., 2017). In our cohort of patients with aniridia, mosaic PZVs were identified in at least 3.8% of families, though this rate is likely underestimated. We cannot rule out that some of our patients without *PAX6* mutations or some parents of *de novo* cases might present a low-grade of mosaicism. Therefore, larger systematic studies should be carried out using high-sensitivity ultradeep sequencing, single cell analysis or ddPCR in several somatic tissues to obtain a better estimation of the actual mosaicism rate as well as to elucidate the actual relevance of PZVs in the pathogenesis of *PAX6*-associated diseases.

In brief, our work confirms the long-held hypothesis of germinal mosaicism in aniridia to explain disease recurrence in the offspring of healthy parents. Additionally, we demonstrate for the first time that mosaic *PAX6* variants cause phenotypic variability from congenital aniridia, mild forms of isolated iris coloboma, and even asymptomatic individuals. Consequently,

our findings have important clinical implications for genetic counseling and ophthalmic management in congenital aniridia and related developmental eye disorders. Finally, somatic and/or gonosomal mosaicism should be taken into consideration as genetic factors to explain not only extraordinary cases of co-occurrence in siblings but also intra-familial variable expressivity, as well as some of apparently *de novo* cases and even uncharacterized *PAX6*-negative cases in aniridia.

AUTHOR CONTRIBUTIONS

MC and MM contributed the conception and design of the study. MT, MM, PR, and CV performed genetic analyses. MT, MM, PR, M-AM-P, and MV participated in the analysis and interpretation of genetic data for the work. MG, FB-K, DR-Z, GR, FM-N, ST-S, BG, and CA participated in the recruitment, clinical evaluation, and interpretation of clinical data. MC wrote the first draft of the manuscript. MT, MM, and MV wrote and/or prepared sections of the manuscript. All authors contributed to manuscript revision, reading and approving the submitted version.

FUNDING

This work was supported by Spanish Institute of Health Carlos III (ISCIII) and Spanish Ministry of Economy and Competitiveness funding partially supported by the following European Regional Development Funds (ERDF): PI17/01164 and SAF2013-46943-R (to MC), PI14/0948 and PI17/01659 (to M-AM-P), the Spanish Center for Biomedical Network Research on Rare Diseases (CIBERER, 06/07/0036 grant, to CA and M-AM-P); by Regional Government of Madrid (CAM, B2017/BMD3721) (to CA and M-AM-P), the Spanish National Organization of the Blind (ONCE) (to MV), the Mutua Madrileña Foundation and the Spanish Foundation of Rare Diseases (FEDER) (to MC), the University Chair UAM-IIS-FJD of Genomic Medicine, and the Ramon Areces Foundation (to CA). MC was supported by the ISCIII Miguel Servet Program (CP12/03256 and CPII17/00006) and MT was supported by a Ph.D. fellowship from the Conchita Rábago Foundation.

ACKNOWLEDGMENTS

The authors thank the families for their participation in this study, the Spanish Association of Aniridia for its helpful support with patient sampling and Oliver Shaw for writing assistance.

REFERENCES

- Acuna-Hidalgo, R., Bo, T., Kwint, M. P., van de Vorst, M., Pinelli, M., Veltman, J. A., et al. (2015). Post-zygotic point mutations are an underrecognized source of *de novo* genomic variation. *Am. J. Hum. Genet.* 97, 67–74. doi: 10.1016/j.ajhg.2015.05.008
- Acuna-Hidalgo, R., Veltman, J. A., and Hoischen, A. (2016). New insights into the generation and role of *de novo* mutations in health and disease. *Genome Biol.* 17:241. doi: 10.1186/s13059-016-1110-110.1186/s13059-016-1110-1
- Ansari, M., Rainger, J., Hanson, I. M., Williamson, K. A., Sharkey, F., Harewood, L., et al. (2016). Genetic analysis of 'PAX6-Negative' individuals with Aniridia or Gillespie syndrome. *PLoS One* 11:e0153757. doi: 10.1371/journal.pone.0153757
- Asuma, N., Yamaguchi, Y., Handa, H., Tadokoro, K., Asaka, A., Kawase, E., et al. (2003). Mutations of the *PAX6* gene detected in patients with a variety of

- optic-nerve malformations. *Am. J. Hum. Genet.* 72, 1565–1570. doi: 10.1086/375555
- Beryozkin, A., Levy, G., Blumenfeld, A., Meyer, S., Namburi, P., Morad, Y., et al. (2016). Genetic analysis of the Rhodopsin gene identifies a mosaic dominant retinitis Pigmentosa mutation in a healthy individual. *Invest. Ophthalmol. Vis. Sci.* 57, 940–947. doi: 10.1167/iovs.15-18702
- Blanco-Kelly, F., Palomares, M., Vallespin, E., Villaverde, C., Martin-Arenas, R., Velez-Monsalve, C., et al. (2017). Improving molecular diagnosis of aniridia and WAGR syndrome using customized targeted array-based CGH. *PLoS One* 12:e0172363. doi: 10.1371/journal.pone.0172363
- Blanco-Kelly, F., Villaverde-Montero, C., Lorda-Sanchez, I., Millan, J. M., Trujillo-Tiebas, M. J., and Ayuso, C. (2013). Guidelines for genetic study of aniridia. *Arch. Soc. Esp. Ophthalmol.* 88, 145–152. doi: 10.1016/j.oftal.2012.07.006S0365-6691(12)00384-X
- Bobilev, A. M., McDougal, M. E., Taylor, W. L., Geisert, E. E., Netland, P. A., and Lauderdale, J. D. (2016). Assessment of PAX6 alleles in 66 families with aniridia. *Clin. Genet.* 89, 669–677. doi: 10.1111/cge.12708
- Ceroni, F., Aguilera-Garcia, D., Chassaing, N., Bax, D. A., Blanco-Kelly, F., Ramos, P., et al. (2018). New GJA8 variants and phenotypes highlight its critical role in a broad spectrum of eye anomalies. *Hum. Genet.* doi: 10.1007/s00439-018-1875-2 [Epub ahead of print].
- Chassaing, N., Causse, A., Vigouroux, A., Delahaye, A., Alessandri, J. L., Boespflug-Tanguy, O., et al. (2013). Molecular findings and clinical data in a cohort of 150 patients with anophthalmia/microphthalmia. *Clin. Genet.* 86, 326–334. doi: 10.1111/cge.12275
- Chou, C. M., Nelson, C., Tarle, S. A., Pribila, J. T., Bardakjian, T., Woods, S., et al. (2015). Biochemical basis for dominant inheritance, variable penetrance, and maternal effects in RBP4 congenital eye disease. *Cell* 161, 634–646. doi: 10.1016/j.cell.2015.03.006
- Cvekl, A., and Callaerts, P. (2017). PAX6: 25th anniversary and more to learn. *Exp. Eye Res.* 156, 10–21. doi: 10.1016/j.exer.2016.04.017
- Dansault, A., David, G., Schwartz, C., Jaliffa, C., Vieira, V., de la Houssaye, G., et al. (2007). Three new PAX6 mutations including one causing an unusual ophthalmic phenotype associated with neurodevelopmental abnormalities. *Mol. Vis.* 13, 511–523.
- Davis-Silberman, N., Kalich, T., Oron-Karni, V., Marquardt, T., Kroeber, M., Tamm, E. R., et al. (2005). Genetic dissection of Pax6 dosage requirements in the developing mouse eye. *Hum. Mol. Genet.* 14, 2265–2276. doi: 10.1093/hmg/ddi231
- Deml, B., Reis, L. M., Lemyre, E., Clark, R. D., Kariminejad, A., and Semina, E. V. (2016). Novel mutations in PAX6, OTX2 and NDP in anophthalmia, microphthalmia and coloboma. *Eur. J. Hum. Genet.* 24, 535–541. doi: 10.1038/ejhg.2015.155ejhg2015155
- Dubey, S. K., Mahalaxmi, N., Vijayalakshmi, P., and Sundaresan, P. (2015). Mutational analysis and genotype-phenotype correlations in southern Indian patients with sporadic and familial aniridia. *Mol. Vis.* 21, 88–97.
- Erez, A., Li, J., Geraghty, M. T., Ben-Shachar, S., Cooper, M. L., Mensing, D. E., et al. (2010). Mosaic deletion 11p13 in a child with dopamine beta-hydroxylase deficiency—case report and review of the literature. *Am. J. Med. Genet. A* 152A, 732–736. doi: 10.1002/ajmg.a.33269
- Faivre, L., Williamson, K. A., Faber, V., Laurent, N., Grimaldi, M., Thauvin-Robinet, C., et al. (2006). Recurrence of SOX2 anophthalmia syndrome with gonosomal mosaicism in a phenotypically normal mother. *Am. J. Med. Genet. A* 140, 636–639. doi: 10.1002/ajmg.a.31114
- Graziano, C., D'Elia, A. V., Mazzanti, L., Moscano, F., Guidelli Guidi, S., Scarano, E., et al. (2007). A de novo nonsense mutation of PAX6 gene in a patient with aniridia, ataxia, and mental retardation. *Am. J. Med. Genet. A* 143A, 1802–1805. doi: 10.1002/ajmg.a.31808
- Gregory-Evans, C. Y., Wang, X., Wasan, K. M., Zhao, J., Metcalfe, A. L., and Gregory-Evans, K. (2014). Postnatal manipulation of Pax6 dosage reverses congenital tissue malformation defects. *J. Clin. Invest.* 124, 111–116. doi: 10.1172/JCI7046270462
- Gronskov, K., Rosenberg, T., Sand, A., and Brondum-Nielsen, K. (1999). Mutational analysis of PAX6: 16 novel mutations including 5 missense mutations with a mild Aniridia phenotype. *Eur. J. Hum. Genet.* 7, 274–286. doi: 10.1038/sj.ejhg.5200308
- Gupta, S. K., De Becker, I., Tremblay, F., Guernsey, D. L., and Neumann, P. E. (1998). Genotype/phenotype correlations in Aniridia. *Am. J. Ophthalmol.* 126, 203–210. doi: 10.1016/S0002-9394(98)00191-3
- Hanson, I., Churchill, A., Love, J., Axton, R., Moore, T., Clarke, M., et al. (1999). Missense mutations in the most ancient residues of the PAX6 paired domain underlie a spectrum of human congenital eye malformations. *Hum. Mol. Genet.* 8, 165–172. doi: 10.1093/hmg/8.2.165
- Hingorani, M., Hanson, I., and van Heyningen, V. (2012). Aniridia. *Eur. J. Hum. Genet.* 20, 1011–1017. doi: 10.1038/ejhg.2012.100ejhg2012100
- Huynh, M. T., Boudry-Labis, E., Duban, B., Andrieux, J., Tran, C. T., Tampere, H., et al. (2017). WAGR syndrome and congenital hypothyroidism in a child with a Mosaic 11p13 deletion. *Am. J. Med. Genet. A* 173, 1690–1693. doi: 10.1002/ajmg.a.38206
- Jia, X., Guo, X., Xiao, X., Li, S., and Zhang, Q. (2010). A novel mutation of PAX6 in Chinese patients with new clinical features of Peters' anomaly. *Mol. Vis.* 16, 676–681.
- Lee, H. J., and Colby, K. A. (2013). A review of the clinical and genetic aspects of Aniridia. *Semin. Ophthalmol.* 28, 306–312. doi: 10.3109/08820538.2013.825293
- Myers, C. T., Hollingsworth, G., Muir, A. M., Schneider, A. L., Thuesmun, Z., Knupp, A., et al. (2018). Parental mosaicism in “De Novo” epileptic encephalopathies. *N. Engl. J. Med.* 378, 1646–1648. doi: 10.1056/NEJMc1714579
- Nallathambi, J., Neethirajan, G., Shashikant, S., Vijayalakshmi, P., and Sundaresan, P. (2006). PAX6 missense mutations associated in patients with optic nerve malformation. *Mol. Vis.* 12, 236–242.
- Netland, P. A., Scott, M. L., Boyle, J. W. T., and Lauderdale, J. D. (2011). Ocular and systemic findings in a survey of aniridia subjects. *J. AAPOS* 15, 562–566. doi: 10.1016/j.jaapos.2011.07.009
- Patel, N., Khan, A. O., Alsahli, S., Abdel-Salam, G., Nowilaty, S. R., Mansour, A. M., et al. (2018). Genetic investigation of 93 families with microphthalmia or posterior microphthalmos. *Clin. Genet.* 93, 1210–1222. doi: 10.1111/cge.13239
- Reed, T. E., and Falls, H. F. (1955). A pedigree of aniridia with a discussion of germinal mosaicism in man. *Am. J. Hum. Genet.* 7, 28–38.
- Riera, M., Wert, A., Nieto, I., and Pomares, E. (2017). Panel-based whole exome sequencing identifies novel mutations in microphthalmia and anophthalmia patients showing complex Mendelian inheritance patterns. *Mol. Genet. Genomic Med.* 5, 709–719. doi: 10.1002/mgg3.329
- Robinson, D. O., Howarth, R. J., Williamson, K. A., van Heyningen, V., Beal, S. J., and Crolla, J. A. (2008). Genetic analysis of chromosome 11p13 and the PAX6 gene in a series of 125 cases referred with aniridia. *Am. J. Med. Genet. A* 146A, 558–569. doi: 10.1002/ajmg.a.32209
- Sannan, N. S., Gregory-Evans, C. Y., Lyons, C. J., Lehman, A. M., Langlois, S., Warner, S. J., et al. (2017). Correlation of novel PAX6 gene abnormalities in aniridia and clinical presentation. *Can. J. Ophthalmol.* 52, 570–577. doi: 10.1016/j.cjco.2017.04.006
- Schedl, A., Ross, A., Lee, M., Engelkamp, D., Rashbass, P., van Heyningen, V., et al. (1996). Influence of PAX6 gene dosage on development: overexpression causes severe eye abnormalities. *Cell* 86, 71–82. doi: 10.1016/S0092-8674(00)8078-1
- Shaham, O., Menuchin, Y., Farhy, C., and Ashery-Padan, R. (2012). Pax6: a multi-level regulator of ocular development. *Prog. Retin. Eye Res.* 31, 351–376. doi: 10.1016/j.preteyeres.2012.04.002
- Skeens, H. M., Brooks, B. P., and Holland, E. J. (2011). Congenital Aniridia variant: minimally abnormal irides with severe limbal stem cell deficiency. *Ophthalmology* 118, 1260–1264. doi: 10.1016/j.ophtha.2010.11.021
- Stosser, M. B., Lindy, A. S., Butler, E., Retterer, K., Piccirillo-Stosser, C. M., Richard, G., et al. (2017). High frequency of mosaic pathogenic variants in genes causing epilepsy-related neurodevelopmental disorders. *Genet. Med.* 20, 403–410. doi: 10.1038/gim.2017.114

- Tang, H. K., Chao, L. Y., and Saunders, G. F. (1997). Functional analysis of paired box missense mutations in the PAX6 gene. *Hum. Mol. Genet.* 6, 381–386. doi: 10.1093/hmg/6.3.381
- Thomas, S., Thomas, M. G., Andrews, C., Chan, W. M., Proudlock, F. A., McLean, R. J., et al. (2014). Autosomal-dominant nystagmus, foveal hypoplasia and presenile cataract associated with a novel PAX6 mutation. *Eur. J. Hum. Genet.* 22, 344–349. doi: 10.1038/ejhg.2013.162
- van Heyningen, V., and Williamson, K. A. (2002). PAX6 in sensory development. *Hum. Mol. Genet.* 11, 1161–1167. doi: 10.1093/hmg/11.10.1161
- Vasilyeva, T. A., Voskresenskaya, A. A., Kasmann-Kellner, B., Khlebnikova, O. V., Pozdeyeva, N. A., Bayazutdinova, G. M., et al. (2017). Molecular analysis of patients with Aniridia in Russian federation broadens the spectrum of PAX6 mutations. *Clin. Genet.* 92, 639–644. doi: 10.1111/cge.13019
- Vincent, M. C., Pujo, A. L., Olivier, D., and Calvas, P. (2003). Screening for PAX6 gene mutations is consistent with haploinsufficiency as the main mechanism leading to various ocular defects. *Eur. J. Hum. Genet.* 11, 163–169. doi: 10.1038/sj.ejhg.52009405200940
- Yasue, A., Kono, H., Habuta, M., Bando, T., Sato, K., Inoue, J., et al. (2017). Relationship between somatic mosaicism of Pax6 mutation and variable developmental eye abnormalities—an analysis of CRISPR genome-edited mouse embryos. *Sci. Rep.* 7:53.

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2018 Tarilonte, Morín, Ramos, Galdós, Blanco-Kelly, Villaverde, Rey-Zamora, Rebolleda, Muñoz-Negrete, Tahsin-Swafiri, Gener, Moreno-Pelayo, Ayuso, Villamar and Corton. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

TITLE: Activation of cryptic donor splice sites by non-coding PAX6 variants contributes to congenital aniridia

SHORT RUN TITTLE: Congenital aniridia caused by silent splicing site variants in PAX6.

AUTHORS: María Tarilonte¹, MSc; Patricia Ramos¹, Jennifer Moya¹, Guillermo Sanz², MD; Fiona Blanco-Kelly¹, MD, PhD; Saoud T Swafiri¹, MD; Cristina Villaverde¹; Raquel Romero¹, PhD; Blanca Gener³, MD; Patrick Calvas^{4,5}, MD, PhD; Carmen Ayuso¹, MD, PhD; Marta Corton¹, PhD.

¹ Department of Genetics & Genomics, Instituto de Investigación Sanitaria-Fundación Jiménez Díaz University Hospital- Universidad Autónoma de Madrid (IIS-FJD, UAM), Centre for Biomedical Network Research on Rare Diseases (CIBERER), Madrid, Spain.

² Department of Ophthalmology, Fundación Jiménez Díaz University Hospital (FJD), Madrid, Spain.

³ Department of Genetics, Cruces University Hospital, BioCruces Health Research Institute, Bilbao, Spain.

⁴ Service de Génétique Médicale, Hôpital Purpan, CHU Toulouse, Toulouse, France

⁵ INSERM U1056, Université Toulouse III, Toulouse, France

CORRESPONDING AUTHOR:

Marta Corton (MC), PhD:

Department of Genetics

Fundación Jiménez Díaz University Hospital

Av. Reyes Católicos nº 2.

28040 Madrid, Spain

mcorton@fjd.es

Phone: +34915504872

ABSTRACT

PURPOSE: The homeodomain transcription factor *PAX6* cause a wide spectrum of ocular developmental anomalies, such as congenital aniridia and microphthalmia. Here, we aimed to functionally assess the implication on splicing of seven non-coding, missense and non-canonical splicing variants in exon 6, which represents the main hotspot region for loss-of-function *PAX6* variants.

METHODS: By locus-specific analysis of *PAX6* using classical and/or targeted next generation sequencing, we screened a Spanish cohort of 106 patients with *PAX6*-related diseases. Functional splicing assays were performed *in vitro* by a minigene approach and, when available, in patients-derived lymphocytes cell lines.

RESULTS: Five out seven variants, including 3 synonymous, one small exonic deletion and a non-canonical splice variant, showed anomalous splicing patterns yielding partial skipping or elongation of exon 6.

CONCLUSION: We described new spliceogenic mechanisms for *PAX6* variants mediated by creating or strengthening 5 different cryptic donors on exon 6. Our work revealed that the activation of cryptic *PAX6* splicing sites seems to be a recurrent and underestimated cause of aniridia. Our findings demonstrate the importance of the extensive screening of non-coding *PAX6* regions and functional assessing of silent variants to unravel the missing heritability and improve the variant interpretation for genetic counseling in aniridia.

KEYWORDS

PAX6, aniridia, non-coding variants, mosaicism, cryptic splice site

INTRODUCTION

Congenital aniridia (MIM #106210) is a panocular rare disease with a worldwide incidence of 1:50,000-100,000 births¹, and is characterized by the misdevelopment of iris and fovea. Aniridia is caused by the haploinsufficiency of the *Paired box gene 6 (PAX6)*, a highly conserved homeodomain-containing transcription factor that plays a pivotal role in morphogenesis and maintenance of ocular structures². To date, about 500 dominant pathogenic SNVs and copy-number changes have been reported in *PAX6* (LOVD *PAX6* database³), leading to a spectrum of different ocular developmental anomalies (ODA)^{1,4}. The *PAX6*-associated phenotypes might be explained by a genotype-phenotype correlation model, in which mutation extent and gene dosage determine the residual activity of the mutant *PAX6* and therefore, the clinical outcome^{1,5}. Most of aniridia cases are caused by variants introducing a premature truncating codon (PTC). In contrast, missense variants are most associated with atypical phenotypes such as Peter's anomaly, microphthalmia and isolated foveal hypoplasia⁶. Although *PAX6* haploinsufficiency is usually highly penetrant, phenotypic variability is usually observed⁷, even in individuals carrying the same variant. Iris anomalies are variable in patients with aniridia ranging from the complete lack of the iris, partial hypoplasia, coloboma, to full presentation showing only minor structural defects⁸. Some cases of *PAX6*-associated intra-familial variability might be explained by postzygotic mosaicism, as we recently reported⁹. However, the causes of the variable expressivity are not well known yet and it could involve different genetic and non-genetic mechanisms.

With the advent of NGS and CGH arrays, pathogenic *PAX6* variants are found in 90% of patients with aniridia^{10,11}. A proportion of the missing heritability could be potentially explained by variants located in non-coding *PAX6* regions or affecting cis-regulatory sequences^{10,12}, that are not usually assessed in routine genetic diagnosis. Similarly, other variants may also be dismissed such as those affecting non-canonical splice sites (NCSS) as well as other silent changes distant to the ultraconserved splice sites. However, they might potentially interfere with exon/intron recognition either by affecting auxiliary splicing regulatory elements (SRE) and/or by creating or strengthening cryptic splice sites¹³. The impact on splicing of these variants of unknown significance (VUS), remains largely unknown mainly due to the underlying difficulties to confirm their pathogenicity in a clinical setting. To date, only few of the potentially spliceogenic *PAX6* variants have been functionally characterized using RNA assays in patient-derived lymphocyte-cell lines (LCLs)¹⁴⁻¹⁶ or alternatively, by means of *in vitro* minigene approaches¹⁷⁻²⁰. In this way, it has been described that disruption of the canonical donor GT dinucleotide of *PAX6* exon 6 lead to an unexpected partially skipped exon by activating indirectly a cryptic deep-exonic donor site^{14,15}. The predicted protein consequence is an in-frame deletion affecting 36 residues of the DNA-binding *PAX6* paired-domain instead of a PTC. Thus, functional

assessment of potentially splicing variants is essential to provide an accurate diagnosis and to improve genotype-phenotype associations in *PAX6*-associated diseases.

Herein, we identified novel spliceogenic mechanisms in *PAX6* exon 6 for 5 silent exonic and NCSS variants that were identified in familial cases with remarkable variable iris phenotype. These variants lead to a direct or indirect activation of at least 5 different cryptic donor splicing sites. Our work reveals new insights into the splicing regulation of this highly conserved exon and also into the potential involvement of these splicing-associated mechanisms with the phenotypic expression of different iris outcome.

MATERIALS AND METHODS

Cohort description

A cohort of 106 patients with ODA, including congenital aniridia, iris coloboma, Peter's anomaly, foveal hypoplasia and/or microphthalmia that had been referred for the genetic study of *PAX6* to University Hospital Fundación Jiménez Díaz (Madrid, Spain). Genomic DNA was isolated from peripheral whole blood or saliva by standard methods. All participants or their legal guardians provided written informed consent prior to their participation and ethical approval was obtained from the Ethics Research Committee of our institution according to the tenets of the Declaration of Helsinki and further review.

***PAX6* screening**

SNVs and microdeletions affecting *PAX6* or its 3' regulatory enhancers were screened by Sanger, NGS, MLPA and CGH-arrays, as previously reported^{10,11,21}. Targeted NGS was performed by means of: i) ANIRIDIA custom panels for the entire *PAX6* gene and 16 known enhancers¹², and up to 260 genes for other eye developmental diseases, that were developed using HaloPlex or SureSelect QXT technologies (Agilent Technologies, Santa Clara, CA), as previously described^{10,22} or ii) commercial clinical exome approaches (*TruSightOne Sequencing Panel* kit, Illumina, San Diego, California, US or the *Clinical Exome Solution* kit, Sophia Genetics, Lausanne, Switzerland) in order to prioritize potentially causal variants in *PAX6* and a subpanel of 142 ODA-associated genes. Bioinformatic analysis was performed for mapping, variant calling and annotation using standard procedures and custom in-house pipelines. Population frequencies of the detected variants were assessed using gnomAD and CIBERER Spanish Variant Server (<http://csvs.babelomics.org/>).

Identified variants were sequentially assessed for pathogenicity¹⁰ and classified following the ACMG recommendations²³. All likely pathogenic variants and VUS were validated by Sanger sequencing. Canonical RefSeq *PAX6* isoform NM_000280 was used for variant nomenclature. Pathogenic predictions were assessed using different *in silico* tools, including CADD and DANN^{24,25}. *In silico*

analysis of missense and in-frame variants was also conducted using SIFT, PolyPhen2, Align GV-GS and Mutation Taster tools, implemented in Alamut Visual v.2.12 (Interactive Biosoftware, Sophia Genetics, Lausanne, Switzerland). 3D structure analysis of the effect of non-synonymous changes were carried out by PyMol software (<https://pymol.org>) using the crystallographic model for *PAX6* PBD ID 6PAX²⁶.

***In silico* splicing analysis**

To identify potential spliceogenic variants on exon 6, several *in silico* tools, including Human Splicing Finder (HSF), MaxEntScan, NNSPLICE and Splicing Sequence Finder from Alamut, were used to analyze mutant as well as wild-type sequences in order to predict the impact of the variants on splicing, the strengths of the canonical and potential cryptic splicing sites. HSF, ESEfinder, Rescue-ESE and Ex-Skip were used to predict putative SRE-associated sites.

Minigene constructions

Wild-type minigene for exon 6 was generated by amplifying genomic DNA encompassing this exon and 200bp flanking intronic sequences, which include the alternative exon 5a of the *PAX6* isoform NM_000164, using specific primers containing BamHI and NotI tails (Figure S1). Amplicons were further cloned directly on the expression pSPL3 vector²⁷. Seven mutant minigenes for potential spliceogenic variants were further obtained by site-directed mutagenesis with QuikChange Lightning kit (Agilent). Primer details are described in Table S1.

RNA splicing assays

For *in vivo* *PAX6* transcriptional analysis, LCLs were established by Epstein Barr virus (EBV)-mediated transformation of blood lymphocytes from a carrier of the variant c.174C>T (ANI-0062; III:3) and two healthy controls. For minigene assays, HEK-293 and ARPE-19 line cells were transiently transfected with wild-type and mutant constructs using Jet-PEI reagent (Polyplus Transfection, New York, USA) in 2 independent replicates, as previously described¹⁰. Total RNA was isolated using TRIzol reagent (ThermoFisher Scientific, Waltham, USA), and cDNA was synthesized from 2.5 µg of RNA with Superscript IV First-Strand Synthesis System (ThermoFisher). Splicing analysis was performed by RT-PCR using vector and/or *PAX6*-specific primers (Table S1). *GAPDH* was used as internal control. PCR-amplified products were analyzed by electrophoresis using the Fragment Analyzer (Advanced Analytical, Agilent) and on a 2% agarose gel. Bands were purified with NucleoSpin Gel & PCR Clean-up kit (Macherey–Nagel, Bethehem, Pensilvania, USA) and further sequenced. For sequencing of complex band patterns, amplicons were first cloned using TOPO-TA Cloning kit (ThermoFisher).

To quantify the relative abundance of each transcript, semi-quantitative fluorescent PCRs were performed in duplicate using a HEX-labeled oligo (Table S1) and analyzed with LIZ-500 Size Standard on the ABI3130xl Genetic Analyzer and the GeneMapper software (ThermoFisher).

Droplet Digital PCR assays

Mosaicism was evaluated in family ANI-0062 by allele-specific genotyping. A custom TaqMan SNP Genotyping Assay (ThermoFisher) was specifically designed for the *PAX6* variant c.174C>T (VIC and FAM-labeled wild-type and mutant allele, respectively). Droplet Digital PCR (ddPCR) analysis was carried out in quadruplicate in DNA from blood and saliva samples using the ddPCR QX200 System (Bio-Rad, Hercules, USA), as previously described⁹.

RESULTS

Identification of exonic and intronic variants that potentially disrupts splicing

PAX6 screening in our cohort of ODA patients revealed 6 variants affecting evolutionary well-conserved nucleotides on exon 6 for which potential spliceogenic effects have been predicted (Figure 1A, Tables 1 and S2). These included: (i) variants affecting exonic or intronic non-canonical nucleotides of the natural donor site: c.356G>T and c.357+6T>G; (ii) synonymous variants c.174C>T;p.(Gly58=) and c.249T>G;p.(Val83=) that alter two predicted exonic cryptic 5'SS; and (iii) variants located at two additional predicted exonic cryptic donors (c.178T>C and c.326_331dup) that might affect potential exonic SRE binding sites (Figure 1A). Other potentially pathogenic variants were excluded in the whole *PAX6* locus or in other ODA-related genes. All of these variants seemed to segregate dominantly with disease in at least 2 generations or were found as *de novo* events (Figure S1). Germinal mosaicism was suspected in 3 families given the disease recurrence in 2 siblings of the first affected generation. The asymptomatic mother of family ANI-062 carried the synonymous variant c.174C>T in a variable degree of somatic mosaicism (11% and 17% in blood and saliva, respectively), as confirmed by ddPCR-based allele-specific quantification (Figure 2A-B). Similarly, gonosomal mosaicism in family ANF-060 was also confirmed for the variant c.178T>C as we previously reported⁹. Unfortunately, mosaicism analysis could not be carried out in family ANF-0048 since DNA samples were unavailable.

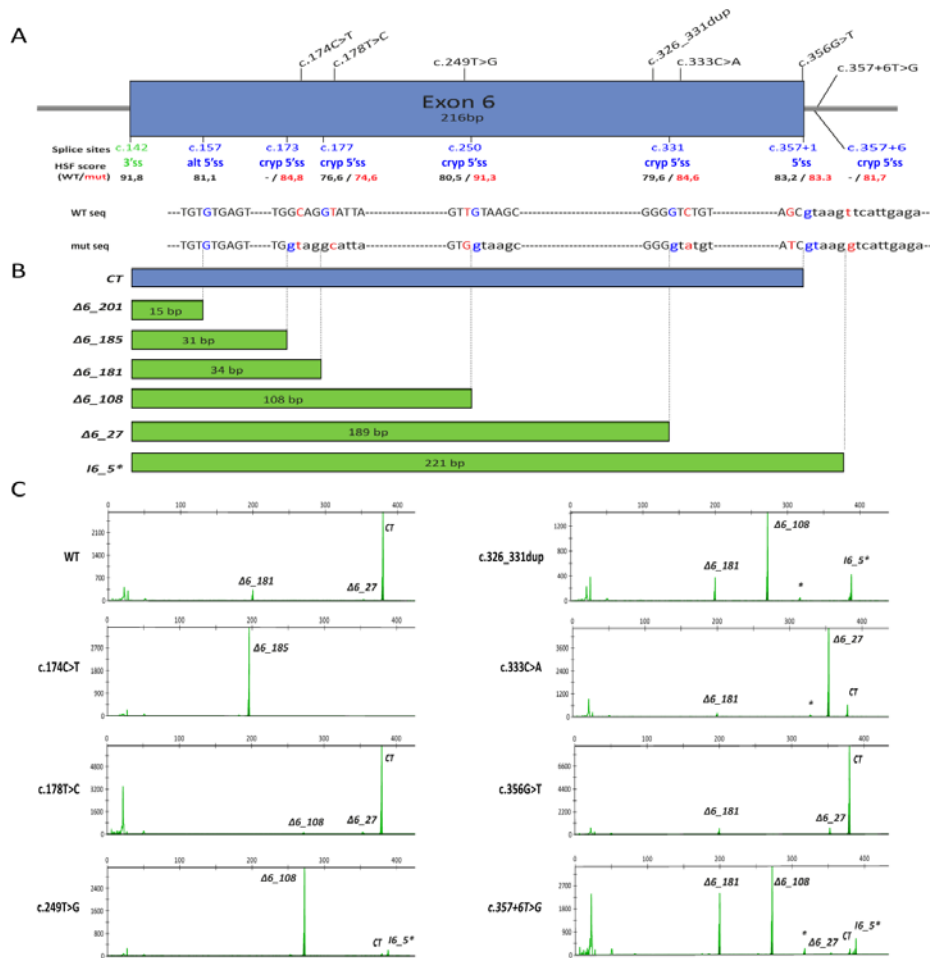


Figure 1. Splicing outcomes for 7 putative splice variants located in PAX6 exon 6. **A).** Schematic representation of exon 6 showing the localization of all canonical and predicted alternative 5'donor (blue) and 3' acceptor (green) splicing sites (SS) and the analyzed variants. The splice prediction scores obtained from Human Splicing Finder tool, wild-type and mutated sequences (in red), exon-intron boundaries and the G nucleotides (in blue) located at the position +1 of the natural (c.357+1), alternative (c.157) and cryptic donor splicing sites (c.173, c.177, c.250, c.331 and c.357+6), are represented. Exonic and intronic nucleotides are shown in upper and lower case, respectively. **B).** Representation of all the mRNA transcripts detected in wild type and mutant RNA assays. These included the full length canonical transcript (CT) of 216 bp and 6 alternative isoforms, represented in blue and in green, respectively. Alternative isoforms include 5 partially skipped mRNAs ($\Delta 6_{201}$, $\Delta 6_{185}$, $\Delta 6_{181}$, $\Delta 6_{108}$ and $\Delta 6_{27}$) that correspond with partial deletions of exon 6 sizing 201, 185, 181, 108 and 27 bp, respectively and an elongated isoform of 221 bp that retained 5 bp of intron 6 ($I6_5^*$). **C).** Semi-quantitative capillary electropherograms of transcripts yielded for the different minigene assays. RNA from transfected ARPE-19 cells was amplified by RT-PCR using a gene-specific forward primer (HEX-labeled c.148_exon6) and a vector-specific reverse primer and further analyzed by capillary electrophoresis. Fragment sizes (bp) and relative fluorescent units are indicated on the x- and y-axes, respectively. Wild-type minigene shows a predominant expected peak of 381 bp corresponding to the correct splice amplicon. At least 5 aberrant splice amplicons were obtained sizing 196, 199, 273, 354, 381 and 385 bp that corresponded $\Delta 6_{185}$, $\Delta 6_{181}$, $\Delta 6_{108}$, $\Delta 6_{27}$ and $I6_5^*$ isoforms, respectively. All bands were confirmed by sequencing except minor isoforms of 317 and 385 bp. (*).

To assess the functional effect on splicing, RNA assays were performed using patient-derived LCLs or alternatively *in vitro* minigene constructs. In addition to the above 6 described variants, we also assessed the previously reported synonymous variant c.333C>A²⁸ that also was predicted to alter an additional putative cryptic donor (Tables 1 and S2). A complex tissue-independent splicing pattern was found in all mutant cells from different lineage, including LCLs (Figure 2C), retinal ARPE-19 (Figure 1C) and renal HEK-293 lines (data not shown). Overall, five out of the 7 potentially spliceogenic variants showed an aberrant mRNA splicing pattern in which the canonical transcript (CT) was absent or markedly decreased (Figure 1C and Table S3). Consequently, they lead to a total or partial silencing of the canonical 5' splicing site (5'SS), therefore impairing the correct exon 6 inclusion (Figure 1A). By contrast, the 2 missense variants assessed showed similar correctly spliced mRNAs than wild-type; thus, they were considered as neutral in terms of pre-mRNA processing. However, both changes, which affect high-conserved residues, are thought to have a pathogenic effect by altering the 3D structure of the paired-domain (Figure S3).

Nucleotide change	Residue change	Family	Inheritance	Reference	ACMG classification	<i>In silico</i> splicing prediction	Minigene outcome	Mayor protein consequence
c.174C>T	p.(G58=)	ANI-062	AD*	Filatova, 2019	3	Creation of a new cryptic 5'SS in c.173	Aberrant Δ6_185 isoform. Complete loss of CT.	p.(Arg59Serfs*13)
c.178T>C	p.(Y60H)	ANF-060	AD*	Tarilonte, 2018	3	Loss of a cryptic 5'SS in c.177. Altering SF2 ESE sites.	No effect on splicing	p.(Tyr60His)
c.249T>G	p.(V83=)	ANI-048	AD**	This study	3	Strengthening of cryptic 5'SS in c.250	Aberrant Δ6_108 isoform. Loss of CT.	p.(Lys86_Ser121del)
c.326_331dup	p.(E109_G110dup)	ANI-053	<i>de novo</i>	This study	4	Disruption of a cryptic 5'SS in c.331 Creation of a new ESS site (hnRNPA1)	Three aberrant isoforms: Δ6_108, Δ6_181 and I6_5. Complete loss of CT.	p.(Lys86_Ser121del)
c.333C>A	p.(V111=)	N.D. [#]	N.D. [#]	Murton, 2000	3	Creation or strengthening of a cryptic 5'SS in c.331	Aberrant Δ6_27.isoform. Loss of CT.	p.(Val111Cysfs*4)
c.356G>T	p.(S119I)	ANF-037	Suspected AD	This study	4	No effect on canonical 5'SS. Altering SRp40/SRp55 ESE sites.	No effect on splicing	p.(Ser119Ile)
c.357+6T>G	p.(?)	ANI-001	AD	This study	3	Creation new 5'SS in c.357+6. Secondary weakening of natural donor c.357+1	4 aberrant isoforms: Δ6_108, Δ6_181, I6_5 and Δ6_27. Loss of CT.	p.(Lys86_Ser121del) and p.(Arg59Serfs*13)

Table 1. Overview of the potential spliceogenic variants in exon 6 of PAX6 assessed in this study. Splicing predictions and *in vitro* outcomes of the potential spliceogenic variants are shown. The pre-mRNA isoforms detected included the canonical transcript (CT), 4 skipped isoforms (Δ6_185, Δ6_181, Δ6_108 and Δ6_27) with partial deletions of exon 6 sizing 185, 181, 108 and 27 nucleotides, respectively and an elongated isoform retaining 5 nucleotides of intron 6 (I6_5).

[#] Variant reported in the PAX6 LOVD database and not found in our cohort. * Gonosomal mosaicism was confirmed in an unaffected progenitor using ddPCR-mediated quantification. ** Suspected germinal mosaicism in the 1st familial generation with unaffected progenitors. 5'SS: donor splicing site; ESE: exonic splicing enhancer, ESS: exonic splicing silencer

Con formato

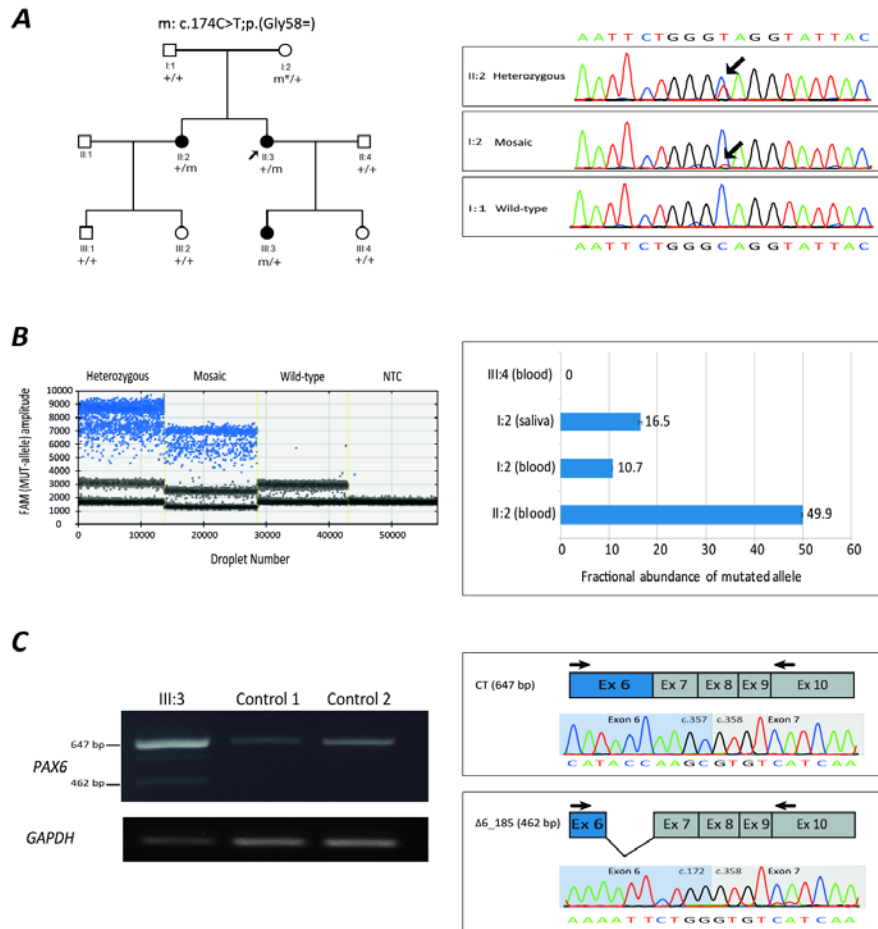


Figure 2. Genetic findings in family ANI-062. A) Family pedigree and segregation analysis of the heterozygous and mosaic variant c.174C>T in *PAX6*, are shown. Individuals with congenital aniridia are indicated with black symbols. Probands are indicated by arrows. Sanger electropherograms are shown for wild-type, heterozygous, and mosaic individuals. m, +, and * represents mutated, wild-type and postzygotic mosaic alleles. *PAX6* variant was numbered according to RefSeq transcript NM_000280.4, using for nucleotide numbering +1 for the translation initiation codon. **B)** Absolute quantification of the allele abundance by digital droplet PCR (ddPCR) assay was carried out using a custom Taqman SNP Genotyping assay in somatic cells from blood and/or saliva. The 1-D fluorescence amplitude plot of FAM genotyping for mutant allele amplification for a heterozygous carrier (II:2), the suspected mosaic mother (II:1), a wild-type homozygous carrier (III:4), and no template control (NTC). FAM-positive droplets (blue) exhibit increased fluorescence compared to negative droplets (grey). Allele ratio of mutated vs wild-type alleles was calculated for the FAM-positive versus VIC-positive droplets. **C)** RT-PCR analysis was carried out in RNA from EBV-transformed blood lymphocytes of an affected heterozygous carrier (III:3) and two unaffected controls with wild-type alleles (+/+), using primers (arrowheads) targeting *PAX6* exons 6 and 10. RT-PCR analysis of *GAPDH* was performed as control. Agarose gel electrophoresis of RT-PCR amplicons showed a single expected 647-bp fragment corresponding the correct spliced transcript (CT) in all samples. Patient III:3 carrying the c.174C>T variant show an additional fragment of 462 nucleotides (isoform Δ6_185). Further Sanger sequence analysis of the aberrant transcript confirmed a partial skipping of 185-bp downstream the position c.172.

Three synonymous variants create exonic cryptic 5' splice sites.

Ex vivo and/or in vitro RNA assays for the 3 studied synonymous variants confirmed the *in silico* splicing predictions consisting of the creation or strengthening of 3 different strong exonic cryptic 5'SS located 185, 107 and 27-nt upstream of the canonical GT dinucleotide, respectively (Figure 1A).

The variant c.174C>T lead to an aberrant transcript $\Delta 6_{185}$ due to the partial skipping of the last 185-nt of exon 6 both in a mutant patient-derived LCL (Figure 2C) and in a minigen assay (Figures 1 and S2). This variant creates a strong cryptic 5'SS on exonic position c.173 that is alternatively used instead of the natural donor (Figure 1A).

Minigene assay for the novel variant c.249T>G revealed a prominent defective transcript caused by partial skipping of 108-nt ($\Delta 6_{108}$) that produce an in-frame deletion p.(Lys86_Ser121del) (Figures 1 and S2, Table S2). This change is predicted to potentiate a previously described cryptic 5'SS located at the exonic position c.250, 107-bp upstream of the natural donor site⁹⁻¹¹.

The previously reported c.333C>A also produced a defective splicing pattern in a minigene assay. An imbalance of the wild-type isoform towards a major partially skipped transcript $\Delta 6_{27}$ was observed (Figure 1 and Table S2). This aberrant isoform, which lacks the last 27-nt of exon 6, could potentially cause an in-frame deletion of 9 residues (p.(Cys112_Val120)). This variant potentiate an *in silico* predicted weak exonic cryptic 5'SS at c.331 (Figure 1).

c.357+6T>G creates a new intronic cryptic 5' splice site.

The variant c.357+6T>G, which affects one of the non-canonical nucleotides of the natural 5'SS, was predicted to have a dual effect on splicing through slightly weakening of the canonical donor but also creating a potential novel GT donor at the intronic position +6 (Tables 1 and S2). Minigene assay confirmed the defective use of this intronic 5'SS that cause an elongated isoform by retaining the first 5-nt of intron 6. Strikingly, this aberrant isoform did not seemed to be the major splicing outcome mediated by this variant (Table S3). Surprisingly, a plethora of at least other 3 different aberrant mRNA isoforms were also detected corresponding to different aberrant partially skipped isoforms ($\Delta 6_{108}$, $\Delta 6_{27}$ and $\Delta 6_{181}$). Thus, this intronic variant seems to lead to an unpredicted activation of 3 distant exonic cryptic 5'SSs, the above mentioned located at positions c.250 and c.331 and a novel alternative donor at c.177 (Figure 1).

An exonic in-frame variant cause aberrant splicing defects through activation of an ESS site.

We assessed if a *de novo* duplication c.326_331dup could affect the splicing process modulating the above described cryptic 5'SS site at c.331. Although no changes in the strength of this potential alternative donor were envisaged, this variant is predicted to potentially alter the SRE-mediated

splicing regulation by creating a new putative hnRNP A1 site (Tables 1 and S2). A complex splicing pattern of different aberrant transcripts was yielded in which the skipped $\Delta 6_{108}$ and the elongated 16_5 isoforms were overrepresented (Figure 1 and Table S2). Interestingly, this variant seems to indirectly silence the natural donor and consequently, promotes an alternative use of other cryptic 5'Ss, mainly those located at c.250 and c.357+6 (Figure 1).

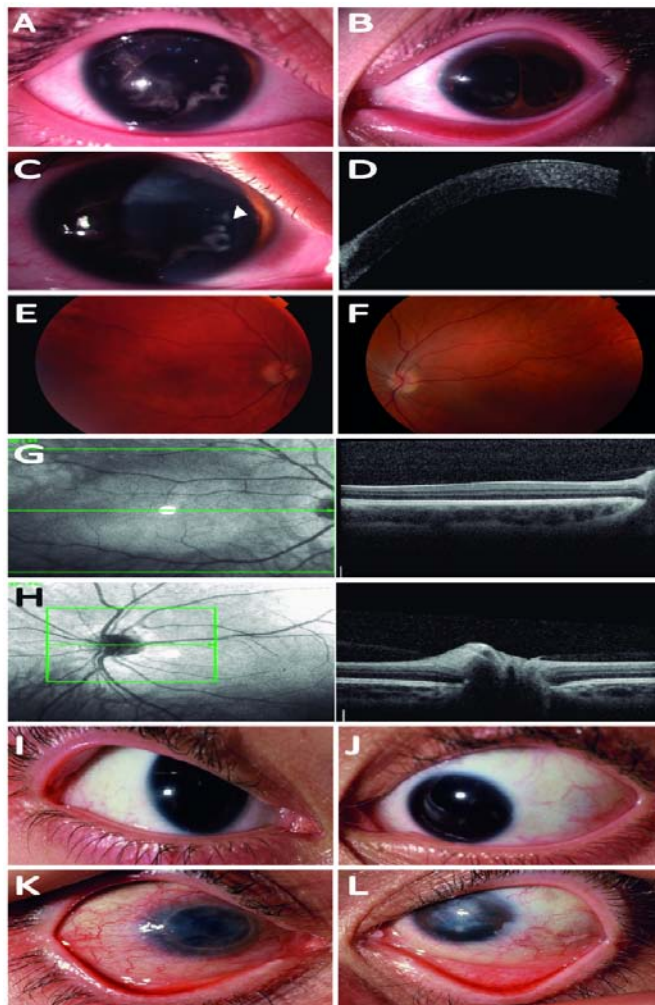
Classification of variants according to the splice defects and phenotypic correlations

The 5 spliceogenic variants showed a variable range of correctly and abnormally spliced *PAX6* transcripts as calculated by semi-quantitative capillary electrophoresis (Table S3). The c.174C>T, which yielded only aberrant transcripts causing a PTC (p.(Arg59Valfs*12), seemed to be the most severe variant. The variants c.326_331dup and c.357+6T>G that did show none or negligible levels of the full-length isoform, yield a mixture of different in-frame and frameshift protein consequences, and therefore were classified as moderately severe. Finally, the variants c.249T>G and c.333C>A were predicted as less severe giving they yielded some residual levels of correct spliced RNA and also mostly lead to in-frame deletions.

Despite this attempt of classification, no obvious genotype–phenotype correlations were identified. All the heterozygous carriers of these spliceogenic *PAX6* variants presented nystagmus, low visual acuity, early-onset cataracts and corneal opacity occurring as result of limbal stem cell deficiency. However, the iris phenotype varied widely between individuals of the same family and even fellow eyes (Table 2). Figure 3A-J shows the ophthalmic features of a patient of family ANI-0062 (III:3) carrying the presumed most severe variant c.174C>T. She presented a markedly asymmetric iris phenotype with near-total absence of iris in right eye, but the presence of a rudimentary iris and polycoria in the fellow eye. A similar presentation of partial aniridia and iris coloboma was presented in a carrier of the c.249T>G (ANI-0048, III:3). Interestingly, the variable expressivity of this variant seemed to be even more extreme in her father (II:2) who had full irises with only minor structural defects (Table 2). Intra-familial phenotypic expressivity was also observed in two patients carrying the variant c.357+6T>G and manifesting different iris and corneal defects (Figure 3I-L). The proband of family ANI-001 presented complete aniridia but her father suffered from iris atrophy and coloboma. This individual was clinically diagnosed of bilateral Peter’s anomaly given the presence of congenital central corneal leukoma.

Family_Patient	Variant	Age at last examination	Iris defects	Foveal hypoplasia	Nystagmus	Early-onset cataracts	Keratopathy	Other ocular features	Es Con formato
ANI-001_II:1	c.357+6T>G	38	Atrophic iris and iridocorneal synechiae (B) Inferior coloboma (LE)	NA	(+) B	(+) B	(+) B Congenital central leukoma		
ANI-001_III:2	c.357+6T>G	14	Complete aniridia	NA	(+) B	(+) B	(+) B	High myopia and strabismus	Skeletal congenital anomalies of right low limb and seizures.
ANI-048_II:2	c.249T>G	68	Full iris with abnormal structure (B)	NA	(+)	(+) B	(+) B	Corectopia and endotropia (OD); glaucoma	Diabetes (38 y)
ANI-048_III:3	c.249T>G	28	Nasal coloboma (RE) and partial aniridia (LE)	(+) B	(+)	(+) LE	(+) B	High myopia and endotropia	Obesity
ANI-053_II:2	c.326_331dup	1	Complete aniridia	(+) B	(+)	(-)	(-)		
ANI-62_III:3	c.174C>T	23	Partial aniridia (B; more severe in RE)	(+) B	(+)	(+) B	(+) B	Optic nerve hypoplasia (B) and high myopia	
ANF-037_III:2	c.356G>T p.(Ser119Ile)	10	(-)	(+) B	(+)	(+)	(-)	Microphthalmia (B), posterior embryotoxon, pupillary membrane, corectopia, pale papilla and high myopia.	Mild mental retardation, seizures, overweight and hyperinsulinemia.
ANF-060_II:1	c.178T>C p.(Tyr60His)	43	(-)	(-)	(+)	(+)	(-)	Microphthalmia (B) and high hyperopia. Secondary corectopia after iridectomy.	

Table 3. Clinical features of the affected individuals. B: bilateral; LE: left eye; RE: right eye; HET: heterozygous; NA: not available.



Con formato

Figure 3. Ophthalmological findings from affected individuals from families ANI-062 and ANI-001. **(A-H)** Clinical data of patient III:3 from family ANI-0062 carrying the variant c.174C>T at 23 years of age. **(A-C)** Slit lamp photographs of the right eye showed a nearly complete absence of iris, with a small remnant of the nasal iris, cortical cataract and moderate limbal stem cell deficiency with subepithelial fibrosis and *pannus*. **(B)** Slit lamp photograph of the left eye showing partial aniridia with a rudimentary iris, anterior polar cataract and mild limbal stem cell deficiency with slight peripheral corneal opacity. **(D)** Anterior Segment- optical coherence tomography (OCT) image of right eye showing epithelial hyperreflectivity, which clinically corresponds to the area of corneal fibrosis. **(E-F)** Fundus color photographs of both eyes showing small optic discs, a weak tessellated fundus and an abnormal foveal reflex. **(G)** Spectral Domain-OCT horizontal cross-sections of the right eye macula showed a grade IV foveal hypoplasia, characterized by absence of foveal pit, extrusion of inner retinal layers, outer segment lengthening and outer nuclear layer widening. **(H)** OCT of the optic nerve head in the left eye showed a small optic disc with almost no cupping. **(I-L)** Clinical data of patients II:1 and III:2 from family ANI-001 that carry the variant c.357+6 T>G. **(I-J)** Slit lamp photograph of both eyes from patient III:2 (proband) at 14 years-old showed a complete absence of iris, mild limbal stem cell deficiency with slight peripheral corneal opacity. **(K)** Slit lamp photograph of the right eye from patient II:1 at 38 years-old showed the outcome of a corneal graft. Penetrating keratoplasty was performed at 36 years-old. **(L)** Slit lamp photograph of the left eye from patient II:1 showed highly vascularized central leukoma, inferior iris coloboma, polycoria, central and peripheral corneal opacity.

DISCUSSION

Current clinical databases report around 100 splice site *PAX6* variants mainly in ultraconserved GT/AG splice sites but also NCSS and deep-intronic VUS have been described in some patients^{10,20}. Splicing effects can enormously vary depending on the degree of similarity to canonical sites and the sequence surroundings²⁹; thus, splicing consequences cannot always be accurately predicted by *in silico* tools. Therefore, functional characterization by transcript analysis is required to improve their clinical interpretation. In the present study, we carried out a splicing analysis for novel exonic and NCSS variants on *PAX6* exon 6. Given that RNA analysis of the illegitimate *PAX6* expression in EBV-transformed LCLs could only be performed in one family, we developed an alternative minigene assay in retinal ARPE-19 cells. Five out the 7 variants promoted anomalous splicing patterns causing alternative exon/intron recognition that predominate over or totally eliminate the canonical transcript. All these splicing variants non-constitutive yielded partially skipped transcripts; and 4 of them lead to co-existing shortened and elongated *PAX6* isoforms.

The main consequence of most variants was the direct activation of an alternative GT dinucleotides that could putatively be used preferentially or in parallel to the natural donor site. In addition to this site, 6 additional potential cryptic 5'SSs along exon 6 were herein identified in the present and previous studies^{14,15,20}. This is the case of a donor at c.157 resulting in a constitutive alternative skipped isoform ($\Delta 6_{201}$) that has been detected not only in control LCLs¹⁶, but also in our minigene assay when vector-specific primers were used. Other predicted cryptic donor sites along exon 6 could potentially alter the natural splicing when are enhanced by different nucleotide changes. As previously observed in *ex vivo* LCLs^{14,15}, variants affecting the natural GT dinucleotide (i.e. c.357+1G>A and c.357+2insT) activate indirectly a pre-existing 5'SS at position c.250. Other changes outside the canonical 5'SS showed exactly this same missplicing in our *in vitro* assays but this was mediated by other novel mechanisms causing the indirect silencing of the canonical 5'SS. These includes the strengthening of this alternative donor c.250 direct or indirectly by the variant c.249T>G or by SRE-mediated deregulation, respectively. In this last mechanism, a small duplication of c.326_331 creates a new potential silencer hnRNPA1 site in an ESE-enriched region 27-nt upstream the exon/intron boundary, and accordingly abolishes completely the constitutive 5'SS splicing. Our findings highlight a likely important role for this SRE-enriched region in regulating the apparently weak balance between the canonical and alternative c.250 donors.

Our minigene assays have been pinpointed other 3 novel cryptic donor sites at exonic and intronic positions that are activated by different variants. The silent variant c.333C>A clearly potentiates a weak GT donor site at c.331 leading to a 27-bp skipped transcript. The NCSS c.357+6 seems to have

multiple splice effects that includes not only those mediated by the weakening the natural 5'SS, i.e. partially skipped transcripts, but also an elongated isoform due to the creation of a second GT dinucleotide in the exon/intron boundary. Intriguingly, this variant leads to a remarkable activation of a distant weak cryptic 5'SS at c.177. This missplicing has been also residually observed in the control minigene; however, the change c.178T>C did not promote this alternative skipping. Follow-up studies should be performed to confirm if this new cryptic 5'SS could participate in alternative *PAX6* splicing. Here, we also demonstrated that a recently described cryptic donor site at c.173 by *in vitro* assays²⁰, is alternatively used in patient-derived LCLs carrying the synonymous c.174 variant; thus confirming its implication in splicing defects.

Our work highlights that the activation exonic cryptic donors in *PAX6* exon 6 could be a underestimated and recurrent pathogenic mechanism in aniridia. Despite the plethora of alternative splicing events and cryptic 5'SSs here identified, missplicing of exon 6 is likely due to disruption of the weak balance between different nearby 5'SSs with apparently similar strength and other potentially modulating SREs motifs. Similarly, a disease-causing *PAX6* variant affecting a SRE site was described to alter the splicing ratio for the nearby exon 5a between alternative and canonical transcripts and then, disrupting the fine-tune splice regulation for the inclusion of this alternative exon³⁰. In exon 6, the shifting towards misspliced isoforms might be due to an apparently weakness of the canonical 5'SS, as is reflected by a low strength of *in silico* predictors, its suboptimal usage in the control minigene and mainly the presence of constitutive alternative splicing mediated by the cryptic donor c.157¹⁶. We did not detect any putative spliceogenic variant affecting the acceptor site of exon 6 in our cohort of patients. Although 3'SS sites have been described as less disruptive in other pathologies^{29,31}, Filatova et al. have recently characterized *in vitro* 2 NCS variants affecting the natural acceptor by activating intronic nearby cryptic A/T dinucleotides²⁰. Therefore, a more complex scenario could be envisaged for the splicing regulation of this exon at both canonical 3'/5'SSs. This might imply a fine-tune regulation for the correct inclusion of this functionally important exon. Further studies should be addressed to deepen on into the regulation of pre-mRNA processing in *PAX6*.

Our work confirms also that synonymous variants on *PAX6*, which are usually considered as benign or VUS by standard classification criteria, may not be neutral and, therefore missplicing should be considered. Given the mutational overrepresentation in exon 6 of about 100 disease-causing variants, additional silent as well as other non-coding variants may have been dismissed or even undetected in some unsolved cases of aniridia. We recently reported 2 *de novo* deep-intronic variants in *PAX6* intron 6¹⁰. Unfortunately, our minigene approach did not cover the full intronic

region; consequently, their functional assessment has not yet been carried out. Here, we also report a novel case of gonosomal mosaicism in *PAX6* that explain the aniridia recurrence in two affected siblings from asymptomatic parents. Thus, this confirms our previous observation that postzygotic *PAX6* events could be an underestimated mutational mechanism in *PAX6*-associated diseases⁹.

Most of the aberrant isoforms here identified are thought to cause large in-frame deletions of the paired DNA binding domain; thus it would cause a loss of *PAX6* functionality. Severity of clinical phenotypes might correlate to the splicing ratio of normal and aberrant splicing, as described for *CEP290*-associated retinal dystrophies³². Although a classification of severity was attempted based on missplicing outcomes, it was difficult to establish clear genotype-phenotype correlations. Except for the 2 microphthalmia-associated missense variants in which we excluded a spliceogenic role, most of the carriers of splicing variants showed classical manifestations of aniridia. However, variable iris phenotype from normal-sized irises to the complete absence of the iris were observed between individuals carrying the same variant or even between the fellow eyes of the same patient with markedly asymmetric iris phenotypes.

Here, we expanded the spectrum of pathogenic variants in aniridia, identifying novel exonic and NCSS variants that alter natural splicing of exon 6 of *PAX6*; thus confirming any type change at any position might potentially disrupt pre-mRNA processing. Our work also reveals that missplicing mediated by cryptic donor sites could represent recurrent pathogenic mechanisms in aniridia. Therefore, a proportion of the missing heritability in aniridia would be potentially explained by other non-coding variants involving similar splicing defects. Our study herein highlights the importance of exploring the mechanism of cryptic splice sites as the basis for emerging treatments that aim to suppress defective splicing³³⁻³⁵. Functional assessment of putative splicing variants will help for pathogenic classification and thereby allow obtaining a more accurate genetic diagnosis and assessment.

Acknowledgements: We acknowledge generous support from the families published in this article and Aniridia España Association. We thank Salvador Marti and Virginia Corrochano from CIBERER Biobank (Valencia, Spain) for their help in the generation of LCLs. We also thank Dr. L. Desviat for the kindly gift of the pSPL3 vector. This work was supported Spanish Institute of Health Carlos III (ISCIII)/European Regional Development Fund (ERDF) (PI17_01164 and CP12/03256), Spanish Ministry of Economy and Competitiveness/ ERDF (MINECO, SAF2013-46943-R), Spanish Federation of Rare Diseases (FEDER) and Mutua Madrileña Foundation, funding awarded to Marta Cortón, CIBERER (06/07/0036), the University Chair UAMIIS-FJD of Genomic Medicine, the Ramon Areces Foundation and Regional Government of Madrid (CAM, B2017/BMD3721), funding awarded to

Carmen Ayuso. Marta Corton was sponsored by the Miguel Servet Program (CP12/03256 and CPII17_00006) from ISCIII, Maria Tarilonte received a Conchita Rabago PhD fellowship and Jennifer Moya was sponsored by Regional Government of Madrid/European Regional Development Fund (ERDF) (PEJ-2017-TL/BMD-7421).

REFERENCES

1. Hingorani M, Hanson I, van Heyningen V. Aniridia. *Eur J Hum Genet.* 2012;20(10):1011-1017.
2. Simpson TI, Price DJ. Pax6; a pleiotropic player in development. *Bioessays.* 2002;24(11):1041-1051.
3. Brown A, McKie M, van Heyningen V, Prosser J. The Human PAX6 Mutation Database. *Nucleic Acids Res.* 1998;26(1):259-264.
4. Glaser T, Jepeal L, Edwards JG, Young SR, Favor J, Maas RL. PAX6 gene dosage effect in a family with congenital cataracts, aniridia, anophthalmia and central nervous system defects. *Nat Genet.* 1994;7(4):463-471.
5. van Heyningen V, Williamson KA. PAX6 in sensory development. *Hum Mol Genet.* 2002;11(10):1161-1167.
6. Tzoulaki I, White IM, Hanson IM. PAX6 mutations: genotype-phenotype correlations. *BMC Genet.* 2005;6:27.
7. Lee HJ, Colby KA. A review of the clinical and genetic aspects of aniridia. *Semin Ophthalmol.* 2013;28(5-6):306-312.
8. Skeens HM, Brooks BP, Holland EJ. Congenital aniridia variant: minimally abnormal irides with severe limbal stem cell deficiency. *Ophthalmology.* 2011;118(7):1260-1264.
9. Tarilonte M, Morin M, Ramos P, et al. Parental Mosaicism in PAX6 Causes Intra-Familial Variability: Implications for Genetic Counseling of Congenital Aniridia and Microphthalmia. *Front Genet.* 2018;9:479.
10. Plaisancie J, Tarilonte M, Ramos P, et al. Implication of non-coding PAX6 mutations in aniridia. *Hum Genet.* 2018;137(10):831-846.
11. Blanco-Kelly F, Palomares M, Vallespin E, et al. Improving molecular diagnosis of aniridia and WAGR syndrome using customized targeted array-based CGH. *PLoS One.* 2017;12(2):e0172363.
12. Bhatia S, Bengani H, Fish M, et al. Disruption of autoregulatory feedback by a mutation in a remote, ultraconserved PAX6 enhancer causes aniridia. *Am J Hum Genet.* 2013;93(6):1126-1134.
13. Manning KS, Cooper TA. The roles of RNA processing in translating genotype to phenotype. *Nat Rev Mol Cell Biol.* 2017;18(2):102-114.
14. Hanson IM, Seawright A, Hardman K, et al. PAX6 mutations in aniridia. *Hum Mol Genet.* 1993;2(7):915-920.
15. Axton RA, Hanson IM, Love J, Seawright A, Prosser J, van Heyningen V. Combined SSCP/heteroduplex analysis in the screening for PAX6 mutations. *Mol Cell Probes.* 1997;11(4):287-292.
16. Gronskov K, Rosenberg T, Sand A, Brondum-Nielsen K. Mutational analysis of PAX6: 16 novel mutations including 5 missense mutations with a mild aniridia phenotype. *Eur J Hum Genet.* 1999;7(3):274-286.
17. Weisschuh N, Wissinger B, Gramer E. A splice site mutation in the PAX6 gene which induces exon skipping causes autosomal dominant inherited aniridia. *Mol Vis.* 2012;18:751-757.
18. Miao Q, Ping X, Tang X, et al. Experimental assessment of novel PAX6 splicing mutations in two Chinese families with aniridia. *Gene.* 2017;630:44-48.

19. Zhang R, Linpeng S, Wei X, et al. Novel variants in PAX6 gene caused congenital aniridia in two Chinese families. *Eye (Lond)*. 2017;31(6):956-961.
20. Filatova AY, Vasilyeva TA, Marakhonov AV, Voskresenskaya AA, Zinchenko RA, Skoblov MY. Functional reassessment of PAX6 single nucleotide variants by in vitro splicing assay. *Eur J Hum Genet*. 2019;27(3):488-493.
21. Blanco-Kelly F, Villaverde-Montero C, Lorda-Sanchez I, Millan JM, Trujillo-Tiebas MJ, Ayuso C. Guidelines for genetic study of aniridia. *Arch Soc Esp Ophthalmol*. 2013;88(4):145-152.
22. Ceroni F, Aguilera-Garcia D, Chassaing N, et al. New GJA8 variants and phenotypes highlight its critical role in a broad spectrum of eye anomalies. *Hum Genet*. 2018.
23. Richards S, Aziz N, Bale S, et al. Standards and guidelines for the interpretation of sequence variants: a joint consensus recommendation of the American College of Medical Genetics and Genomics and the Association for Molecular Pathology. *Genet Med*. 2015;17(5):405-424.
24. Kircher M, Witten DM, Jain P, O'Roak BJ, Cooper GM, Shendure J. A general framework for estimating the relative pathogenicity of human genetic variants. *Nat Genet*. 2014;46(3):310-315.
25. Quang D, Chen Y, Xie X. DANN: a deep learning approach for annotating the pathogenicity of genetic variants. *Bioinformatics*. 2015;31(5):761-763.
26. Xu HE, Rould MA, Xu W, Epstein JA, Maas RL, Pabo CO. Crystal structure of the human Pax6 paired domain-DNA complex reveals specific roles for the linker region and carboxy-terminal subdomain in DNA binding. *Genes Dev*. 1999;13(10):1263-1275.
27. Desviat LR, Perez B, Ugarte M. Minigenes to confirm exon skipping mutations. *Methods Mol Biol*. 2012;867:37-47.
28. Murton NJ, Rehman I, Black GC, Inglehearn CF, Churchill AJ. A novel deletion (IVS11+3del4) identified in the human PAX6 gene in a patient with aniridia. *Hum Mutat*. 2000;15(6):582.
29. Sangermano R, Khan M, Cornelis SS, et al. ABCA4 midgenes reveal the full splice spectrum of all reported noncanonical splice site variants in Stargardt disease. *Genome Res*. 2017;28(1):100-110.
30. Epstein JA, Glaser T, Cai J, Jepeal L, Walton DS, Maas RL. Two independent and interactive DNA-binding subdomains of the Pax6 paired domain are regulated by alternative splicing. *Genes Dev*. 1994;8(17):2022-2034.
31. Buratti E, Chivers M, Hwang G, Vorechovsky I. DBASS3 and DBASS5: databases of aberrant 3'- and 5'-splice sites. *Nucleic Acids Res*. 2011;39(Database issue):D86-91.
32. Drivas TG, Wojno AP, Tucker BA, Stone EM, Bennett J. Basal exon skipping and genetic pleiotropy: A predictive model of disease pathogenesis. *Sci Transl Med*. 2015;7(291):291ra297.
33. Albert S, Garanto A, Sangermano R, et al. Identification and Rescue of Splice Defects Caused by Two Neighboring Deep-Intronic ABCA4 Mutations Underlying Stargardt Disease. *Am J Hum Genet*. 2018;102(4):517-527.
34. Cideciyan AV, Jacobson SG, Drack AV, et al. Effect of an intravitreal antisense oligonucleotide on vision in Leber congenital amaurosis due to a photoreceptor cilium defect. *Nat Med*. 2019;25(2):225-228.
35. Levin AA. Treating Disease at the RNA Level with Oligonucleotides. *N Engl J Med*. 2019;380(1):57-70.

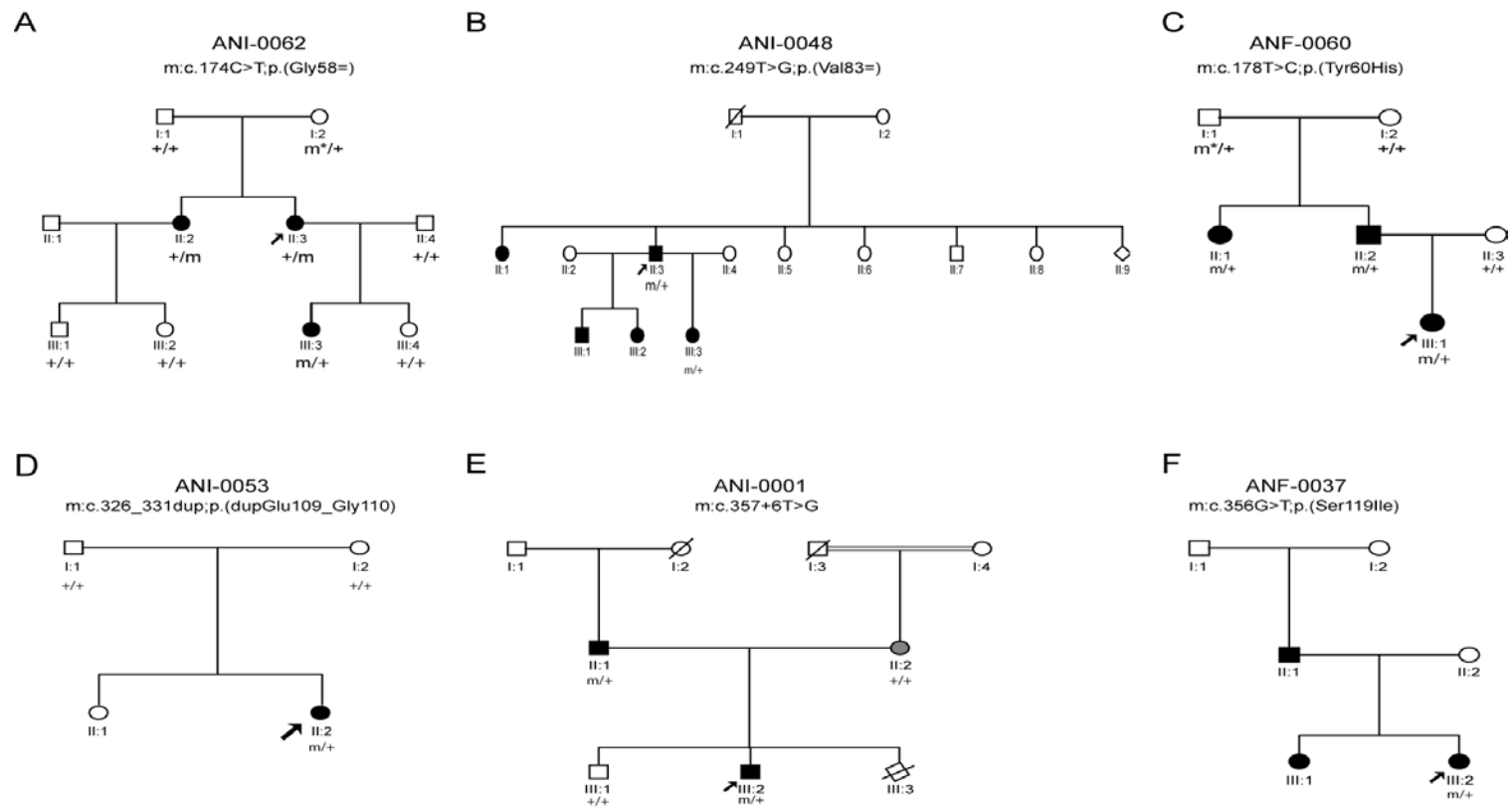


Figure S1. Pedigrees of the five families carrying the PAX6 exon 6 variants reported by the first time in this study. Arrow indicates the proband; +: wild-type allele; m: mutant allele; *: mosaic allele confirmed by allele-specific genotyping using droplet-digital PCR.

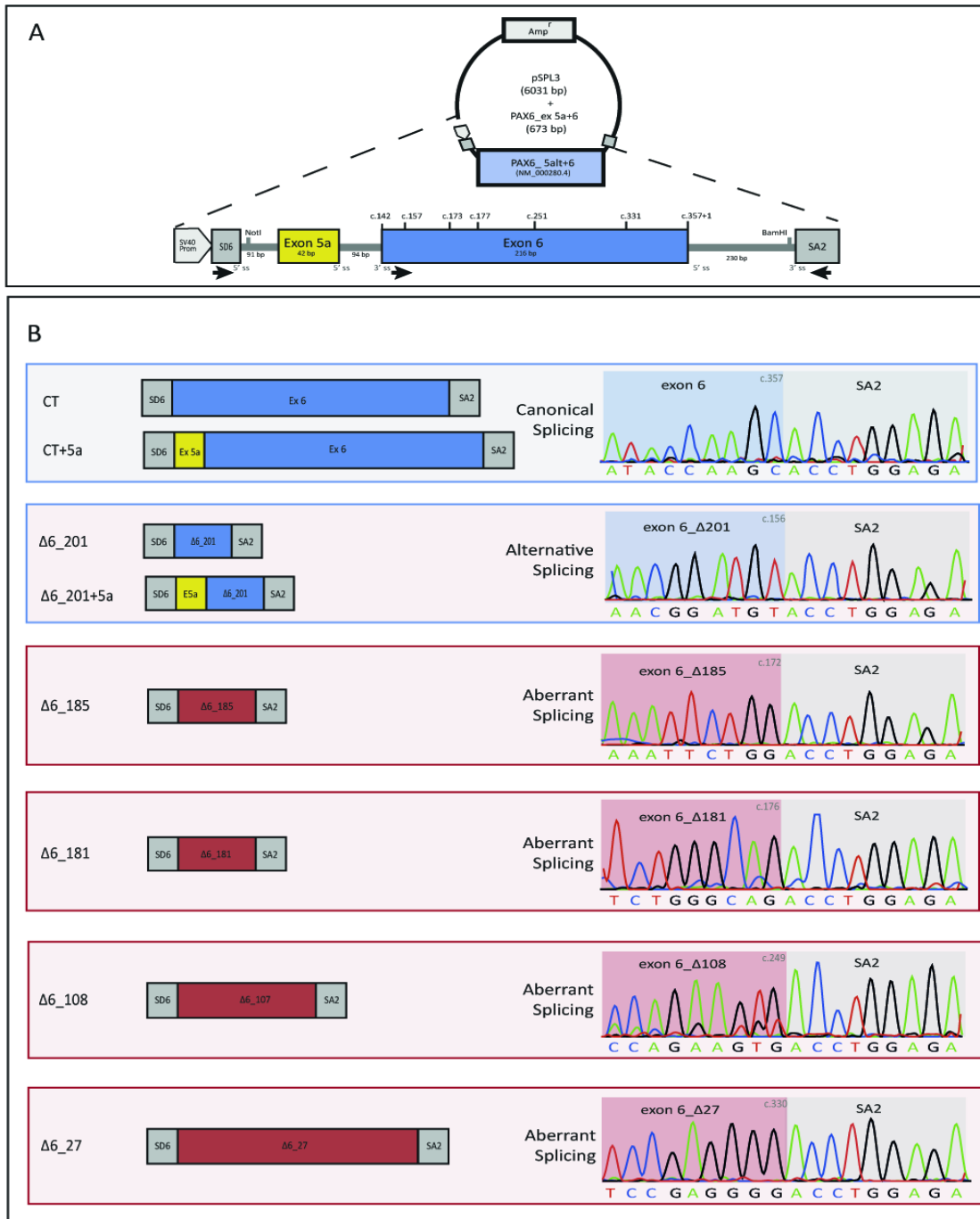
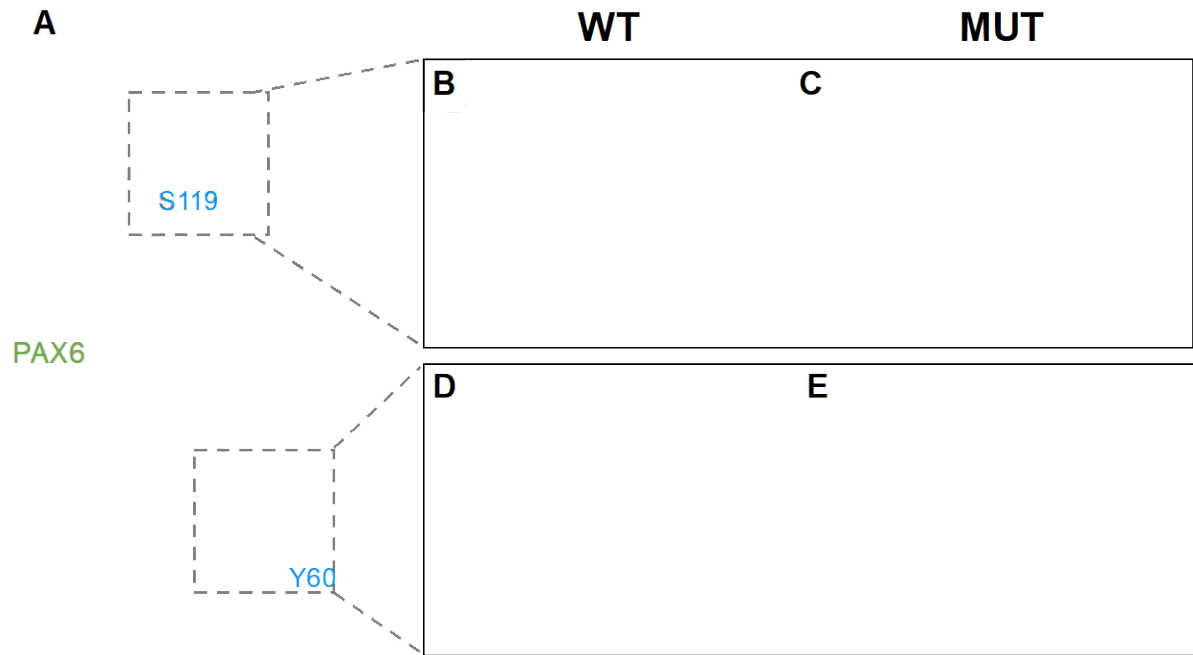


Figure S2. Sanger sequencing of the different pre-mRNA isoforms obtained in the *in vitro* minigene assays. **A.** Schematic representation of the minigene generated from an exon trapping pSLP3 vector (grey boxes) containing exon 6 (blue box) and intronic 200-bp surroundings, which include an alternative exon 5a (yellow box), located only 94-bp upstream. Localizations of analyzed variants as well as specific primers for targeting vector (SD6 and SA2) or exon 6 sequences (c.148) (arrowheads) used in RT-PCR and Sanger analysis are represented. **B.** Schematic representation and Sanger electropherograms of the all splicing events detected in wild type and 7 mutant minigenes assays. These included the canonical transcript (CT), 5 alternative partially skipped mRNA ($\Delta 6_{201}$, $\Delta 6_{201+5a}$, $\Delta 6_{185}$, $\Delta 6_{181}$, $\Delta 6_{108}$ and $\Delta 6_{27}$) with partial deletions of exon 6 sizing 201, 185, 181, 108 and 27 nucleotides, respectively and an elongated isoform that retained 5 bp of intron 6 ($\Delta 6_{5}$). $\Delta 6$: indicates deletion of exon 6 (a.i. $\Delta 6_{185}$ indicates the partial skipping of 185 bp).



DNA

Figure S3. The PAX6-DNA crystal structure. **A.** Overview of PAX6 Paired Domain bound to the DNA. PAX6 is shown as green ribbon and the DNA strand in orange. The inset shows a zoom into the mutation sites. **B-C.** Ribbon diagram of PAX6 in complex with DNA of Wild type and mutated protein, respectively. PAX6 is shown in green and the DNA strand in orange. Residues of interest are shown as cyan sticks. Oxygen atoms are shown in red, nitrogen in blue. Potential hydrogen bonds are indicated by the dashed lines. (B) Residue S119 is shown as cyan balls and sticks. (C) Mutated residue I119 is shown in pale pink. S119 has been reported to be a key residue that contacts the DNA (REF.1). The change of S119 to Isoleucine would result in a loss of a critical contact with the DNA. Two of the three potential hydrogen bonds that form S119 with the DNA are lost after the substitution. The mutation to isoleucine would not only place a hydrophobic group near the phosphate group of the DNA, but also create steric clashes as it is a bulkier residue. **D-E.** Space-filling model of PAX6 showing surface electrostatic charge of (D) the wild type protein and (E) the mutant (blue, positive; red, negative). Residues of interest are shown as cyan sticks. (E) Mutated residue H60 is shown in pale pink. The mutation of Y60 to Histidine causes a change in the protein surface polarity, the mutant showing a more positive charge. Owing to its smaller ring, the change to histidine additionally causes the deformation of the protein surface. Y60 lies in the proximity of the residue G64. Mutations in position 64 that causes the deformation of the protein surface have been previously reported to be very deleterious, indicating a functionally important role of this part of the protein (Kumar, 2015).

Name	Sequence	Tm
Sanger validation		
PAX6_6F	GTGGTTTTCTGTCCACTCC	57.3
PAX6_6R	AGGAGAGAGCATTGGGCTTA	57.3
Minigene construction		
PAX6_int5_F	AATGCGGCCGCTGTAGTTTTTAGGGCTACAAATGTAAT	75.6
PAX6_int6_R	ATAGGATCCTTGACAGCATGCAAATGAAGT	69.9
PAX6_c.174C>T_F	GCCAGTCTCGTAATACCTACCCAGAATTTTACTCACA	71.7
PAX6_c.174C>T_R	TGTGAGTAAAATTCTGGGTAGGTATTACGAGACTGGC	71.7
PAX6_c.178T>C_F	GGAGCCAGTCTCGTAATGCCTGCCAGAAATTTTAC	76.8
PAX6_c.178T>C_R	GTAAAATTCTGGGCAGGTATTACGAGACTGGCTCC	73.6
PAX6_c.249T>G_F	TGGGCTATTTTGCTTACCATTCTGGAGTCGCTAC	74.7
PAX6_c.249T>G_R	GTAGCGACTCCAGAAAGTGGTAAGCAAAATAGCCCA	74.7
PAX6_c.326_331dup_F	GTTATCGTTGGTACAGACCCCCCTCCCCCTCGGACAGTAAT	80.0
PAX6_c.326_331dup_R	ATTACTGTCCGAGGGGGAGGGGGTCTGTACCAACGATAAC	80.0
PAX6_c.333C>A_F	GTTATCGTTGGTACATACCCCCCTCGGACAGTAAT	71.8
PAX6_c.333C>A_R	ATTACTGTCCGAGGGGGTATGTACCAACGATAAC	71.8
PAX6_c.356G>T_F	GTTCTCAATGAACTTACGATTGGTATGTTATCGTTGGTACA GAC	73.8
PAX6_c.356G>T_R	GTCTGTACCAACGATAACATACCAATCGTAAGTTCATTGAG AAC	73.8
PAX6_c.357+6T>G_F	GGGCAGATGTTCTCAATGCACTTACGCTTGGTATGTT	76.8
PAX6_c.357+6T>G_R	AACATACCAAGCGTAAGGTCATTGAGAACATCTGCC	76.1
RNA analysis		
pSPL3_SD6_F	TCTGAGTCACCTTGGACAACC	60.1
pSPL3_SA2_R	ATCTCAGTGGTATTTGTGAGC	54.4
PAX6_c.148-171_F	AACGGATGTGTGAGTAAAATTCTG	59.5
PAX6_c.357+8_SA2_R	TCCAGGTGAACTTACGCTTGG	62.5
PAX6_10R	CATTTGGCCCTTCGATTAGA	60.0
GAPDH_F	CGACCACTTTGTCAAGCTCA	60.0
GAPDH_R	AGGGGAGATTCAAGTGTGGTG	60.0

Table S1. List of primers used in this work.

Nucleotide change	HSF	MaxEntScan	NNSPLICE	SSF	DANN	CADD	GERP	PhyloP100	SIFT	Mutation Taster	Align GV-GD
c.174C>T	— ⇒ 84.8 (+100%)	— ⇒ 6.9 (+100%)	— ⇒ 0.98 (+100%)	74.1 ⇒ 76.8 (+3.7%)	0.630	12.7	5.35	-----	-----	-----	-----
c.178T>C	76.5 ⇒ — (-100%)	7.3 ⇒ — (-100%)	0.74 ⇒ — (-100%)	72.1 ⇒ — (-100%)	0.995	27.7	5.35	6.023	D	D	Class C65
c.249T>G	80.4 ⇒ 91.3 (+13.5%)	4.4 ⇒ 8.1 (+86.7%)	0.51 ⇒ 1.00 (+94.6%)	70.1 ⇒ 82.7 (+18.0%)	0.663	6.33	5.35	-----	-----	-----	-----
c.326_331dup	79.5 ⇒ — (-100%)	0.78 ⇒ — (-100%)	-----	-----	-----	-----	5.35	-----	-----	-----	-----
c.333C>A*	79.5 ⇒ 84.6 (+6.3%)	0.78 ⇒ 7.6 (+100%)	— ⇒ 0.98 (+100%)	— ⇒ 77.8 (+100%)	0.73	13.2	5.35	-----	-----	-----	-----
c.356G>T	83.2 ⇒ 83.3 (+0.1%)	10.1 ⇒ 10.4 (+3.7%)	0.99 ⇒ 0.98 (-1.5%)	77.4 ⇒ 77.9 (+0.6%)	0.995	34	5.35	6.023	D	D	Class C65
c.357+6T>G	— ⇒ 81.7 (+100%)	— ⇒ 4.3 (+100%)	-----	— ⇒ 77.4 (+100%)	0.947	20.1	5.35	5.35	5.35	5.35	5.35

Table S2: *In silico* predictions of the potential spliceogenic variants in exon 6 of PAX6. The splice prediction scores for canonical donor (c.357+1) or cryptic 5'SS obtained from different *in silico* splicing tools are shown. * Variant reported in the PAX6 LOVD database and not found in our cohort. 5'SS: donor splicing site; ESE: exonic splicing enhancer, ESS: exonic splicing silencer, D: deletereous; HSF: Human Splicing Finder tool.

	Observed isoforms							Expected frameshift (%)	Expected in-frame (%)	Severity Ranking	Iris phenotype
	CT (%)	$\Delta 6_{185}$ (%)	$\Delta 6_{181}$ (%)	$\Delta 6_{108}$ (%)	$\Delta 6_{27}$ (%)	$I6_5$ (%)	Other				
Minigene constructs											
<i>WT</i>	86		12		2			12	2		
<i>c.174C>T</i>	0	100						100	0	High (1)	Asymmetric iris
<i>c.178T>C</i>	93			1	6			0	7		Corectopia ^{&}
<i>c.249T>G</i>	3			91				6	91	Less (4)	Intra-individual variability
<i>c.326_331dup</i>	0		14*	66			2	33	66	Moderate (2)	Complete aniridia
<i>c.333C>A</i>	13		3		84			3	84	Less (5)	N.D.
<i>c.356G>T</i>	88		4		8			4	8		Pupillary membrane and corectopia ^{&}
<i>c.357+6T>G</i>	3		33	49	1*	10	4	43	51	Moderate (3)	Interindividual variability
Peak size (bp)	381	196	199	273	354	385	317				
Donor position	c.357+1	c.173	c.177	c.250	c.331	c.357+6					
Protein consequence	---	p.(R59Vfs*12)	p.(R59Sfs*13)	p.(K86_S121del)	p.(C112_V120del)	p.(S121Rfs*5)					

Table S3: Summary of minigene outcomes obtained after semi-quantification by capillary electrophoresis. Distribution, in percentage, of isoforms found for each variant. Nucleotide position reflects the GT dinucleotide of the donor splicing site. Protein consequence is also detailed. In the wild-type minigene, 86% of transcripts corresponded with the canonical transcript (CT). Other 2 alternative residual isoforms ($\Delta 6_{27}$ and $\Delta 6_{181}$) were also observed due to the partial skipping of 27 and 181 nucleotides, respectively. Therefore, the splice defect was only considered when the amount of CT decreased compared to that of the wild-type minigene. * similar levels to WT. $\Delta 6$: indicates deletion of exon 6 (a.i. $\Delta 6_{185}$ indicates the partial skipping of 185 bp). $I6_5$ indicates insertion of the first 5 nucleotides of intron 6. [&] Microphthalmia.