



Universitat Politècnica de València
Programa de Doctorado en Biotecnología

**Investigación de la Leucemia Mieloide Aguda
mediante el desarrollo de modelos *in vitro* e
*in vivo***

Elisa González Romero

Dirigida por

Dr. Rafael Pascual Vázquez Manrique
Dr. José Vicente Cervera Zamora
Dr. Miguel Ángel Sanz Alonso

Diciembre 2021



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA



RAFAEL PASCUAL VÁZQUEZ MANRIQUE, Doctor en Ciencias Biológicas por la Universidad de Valencia e Investigador principal del grupo de Biomedicina Molecular, Celular, y Genómica del Instituto de Investigación Sanitaria La Fe, **JOSÉ VICENTE CERVERA ZAMORA**, Doctor en Medicina por la Universidad de Valencia y Jefe de la Unidad de Genética y Diagnóstico Prenatal del Hospital Universitari i Politècnic La Fe, **MIGUEL ÁNGEL SANZ ALONSO**, Catedrático y Profesor Honorario del Departamento de Medicina de la Universidad de Valencia e Investigador Emérito del Instituto de Investigación Sanitaria La Fe,

CERFITICAN:

Que la tesis doctoral titulada “INVESTIGACIÓN DE LA LEUCEMIA MIELOIDE AGUDA MEDIANTE EL DESARROLLO DE MODELOS *IN VITRO* E *IN VIVO*” que presenta Elisa González Romero, licenciada en Biología, para optar al título de Doctora por la Universitat Politècnica de València, ha sido realizada en el grupo de Hematología y Hemoterapia del Instituto de Investigación Sanitaria La Fe bajo su dirección y que reúne las condiciones para ser defendido por su autora.

Valencia, Diciembre de 2021.

Fdo.: Dr. Rafael P. Vázquez Manrique

Fdo.: Dr. José V. Cervera Zamora

Fdo.: Dr. Miguel Ángel Sanz Alonso

A mi familia

AGRADECIMIENTOS

AGRADECIMIENTOS

En primer lugar, me gustaría dar las gracias a mis directores, por haberme guiado hasta aquí.

Rafa, gracias a ti he podido desarrollar esta Tesis. Desde mis inicios en el laboratorio con el trabajo fin de máster me has guiado y apoyado, trabajando conmigo muchas veces codo con codo en el laboratorio. Ha sido una suerte contar con un director tan dedicado y cercano.

José, gracias por haber apostado por mi para formar parte del equipo y para desarrollar este trabajo. Además, tu implicación en el continuo avance de la investigación desde una perspectiva multidisciplinar es un ejemplo a seguir.

Al Dr. Miguel Ángel Sanz, quisiera darle las gracias por su inspiradora trayectoria en el campo de la oncohematología y su pasión por la investigación que transmite a todos los miembros del grupo.

Gracias a todos los que forman y han formado parte del grupo de Investigación en Hematología y Hemoterapia del IIS La Fe. Gracias a Mireia, Cristian, Mireya, Claudia, Alejandra, Cristóbal y Marta por compartir tantas horas de trabajo, congresos y “cenas alternativas”.

Al pequeño duende del laboratorio Cristina quisiera darle las gracias, por además de ser mi amiga, ser mi apoyo incondicional en el trabajo. Ha sido un placer vivir a tu lado tantos buenos momentos, y algunos malos también. A Alessandro quisiera agradecer su enorme implicación en el día a día del laboratorio. Su pasión y tenacidad por la investigación es contagiosa y su apoyo y amistad durante estos años han sido cruciales para el desarrollo de esta Tesis.

Tengo un cariño muy especial a mis compañeros del grupo BMCG. Gracias a Chema, Rafa, Elena, Teresa, Gema, Regina, Ana Pilar, Carla, Lorena, Loli, Cristina, Belén, Cristina, Julia y muchos mas. Gracias por ese año que pasé en vuestro laboratorio donde aprendí tanto y por vuestra ayuda durante todos estos años. Gracias especialmente a Rafa, Ana Pilar, Cristina y Julia por introducirme y guiarme en el mundo gusanero.

También quisiera dar las gracias a todos mis compañeros del IIS La Fe. Gracias a Javi, Inma, Álvaro, Jessica, Laura, Raquel y Roger por todos estos años de amistad. Ir a veros al laboratorio siempre es un chute de energía. Gracias por tantas risas dentro y fuera del trabajo.

A mis padres. Sin ellos no habría podido llegar hasta aquí. Su apoyo incondicional es mi motivación para superarme día tras día. Gracias por todo vuestro cariño, apoyo y comprensión durante todo este proceso, y siempre. Gracias además a toda mi familia, porque en cada comida familiar recargo las pilas para afrontar con mas felicidad el día a día.

Por último, quisiera dar las gracias a mi compañero de vida. Gracia Adrián por haber estado a mi lado y acompañarme y ayudarme en los altibajos de la Tesis. Me has sacado una sonrisa hasta en los días más difíciles. Gracias por hacerme feliz.

RESÚMENES

RESÚMENES

RESUMEN

La leucemia mieloide aguda (LMA) se trata de un grupo heterogéneo de desórdenes hematológicos producidos por alteraciones genéticas en las células precursoras mieloides. Las mutaciones en la enzima Isocitrato deshidrogenasa 2 (IDH2) son unas de estas alteraciones. Las mutaciones más frecuentes en esta proteína afectan a las posiciones R140 y R172, provocando una ganancia de función con la producción del oncometabolito D-2-hidroxisuccinato (2-HG). A pesar de que ambas inducen la producción de 2-HG, la mutación R172 produce mayor cantidad de oncometabolito, presenta menos concurrencias con otras alteraciones genéticas y se asocia a una peor respuesta a la quimioterapia y un mayor riesgo de recaída. Dadas sus implicaciones en el desarrollo de la enfermedad, la mutación *IDH2*^{R172} se ha propuesto como subgrupo independiente de LMA.

Gracias a la investigación con organismos modelo ha sido posible estudiar en profundidad el papel de las mutaciones genéticas en el desarrollo de numerosas enfermedades, como en la LMA. A pesar de ello, aún quedan muchos aspectos por explicar acerca de las implicaciones de las alteraciones genéticas, como las mutaciones en *IDH2*, en el desarrollo de la enfermedad. Por todo ello, en esta Tesis se han desarrollado nuevas estrategias de edición génica mediante el sistema CRISPR/Cas9 con el objetivo de desarrollar nuevos modelos *in vitro* e *in vivo* de mutaciones implicadas en LMA.

En relación al desarrollo del modelo *in vitro*, debido a la baja eficiencia de transfección de los plásmidos CRISPR en las líneas celulares leucémicas, el método más empleado para introducir los elementos CRISPR han sido principalmente vectores lentivirales. Para evitar los inconvenientes de este tipo de vectores, en esta Tesis se ha desarrollado una estrategia alternativa para la introducción de la nucleasa Cas9 y los guías CRISPR. El gen codificante de la Cas9, junto a un promotor y terminador de expresión, se introdujeron en el genoma de células NB4 mediante transducción con lentivirus, generando una línea celular con expresión constitutiva de la nucleasa. Por otro lado, se desarrolló un sistema sencillo de producción de los guías CRISPR mediante

PCR con los elementos esenciales para su expresión y la expresión del reportero *GFP* de forma opcional.

Con el objetivo de optimizar la técnica y probar su eficiencia en distintas dianas se modificaron dos genes implicados en LMA. Estos fueron el gen *IDH2*, en el cuál se buscó introducir la mutación R172, y el gen *MYBL2*, empleado con anterioridad en trabajos del grupo. Finalmente, las eficiencias de edición obtenidas se compararon con el uso de complejos de ribonucleoproteínas CRISPR, muy utilizados por su alta eficiencia. Mientras que los complejos de ribonucleoproteínas presentaron una mayor eficiencia de corte, la eficiencia de edición de la mutación R172 fue similar en ambas estrategias. Mediante secuenciación masiva se confirmó y caracterizó esta edición y se comprobó que la maquinaria de edición no había producido cortes inespecíficos en regiones similares del genoma. De esta forma, la nueva metodología desarrollada permitió editar de forma precisa líneas celulares leucémicas con eficiencias similares a otras técnicas CRISPR más extendidas y sin producir efectos inespecíficos no deseados.

Por otro lado, gracias a la gran conservación evolutiva del gen *IDH2*, los residuos R140 y R172 se encuentran conservados en la proteína *idh-2* de *Caenorhabditis elegans*. Se empleó la estrategia co-CRISPR para desarrollar y seleccionar cepas mutantes con las mutaciones ortólogas a R140 y R172, y una cepa con ambas mutaciones. A pesar de la conservación, no se observó el aumento del oncometabolito 2-HG esperado en las cepas mutantes en comparación con la cepa salvaje control N2. Esto podría deberse a la activación de vías de degradación del 2-HG. Por ello, es necesario profundizar en el estudio de las vías de degradación del 2-HG para desarrollar modelos de investigación que desarrollen las alteraciones moleculares observadas en los pacientes

Para concluir, la estrategia desarrollada de introducción de elementos CRISPR en líneas celulares, junto a los modelos producidos en *C. elegans*, permitirán en futuros estudios investigar en detalle los efectos moleculares de mutaciones detectadas en pacientes de LMA, su implicación en el desarrollo y pronóstico de la LMA y comprender su papel en la estratificación de los pacientes.

ABSTRACT

Acute Myeloid Leukaemia (AML) is a heterogeneous group of haematological disorders caused by genetic alterations in myeloid precursors. Mutations in the Isocitrate dehydrogenase enzyme are among these alterations. The most frequent mutations in this protein affect R140 and R172 positions, leading to a gain of function with the production of the oncometabolite D-2-hydroxyglutarate (2-HG). Although both induce the 2-HG production, the R172 mutation generates greater amount of oncometabolite, has fewer co-occurrences with other genetic alterations and is associated with worse chemotherapy response and higher relapse risk. Due to its implication in the disease progression, *IDH2*^{R172} has been proposed as new subgroup of AML.

Research models have made possible to study the role of genetic mutations in disease development, such as in AML. Despite this progress, the implications of many genetic alterations, as *IDH2* mutations, in disease development remain unexplained. Therefore, in this Thesis we have developed new gene editing strategies using the CRISPR/Cas9 system with the aim of developing new *in vitro* and *in vivo* models of mutations involved in AML.

Regarding *in vitro* model, due to the low transfection efficiency of CRISPR plasmids in leukemic cell lines, the most commonly method used for introducing CRISPR elements have been mainly lentiviral vectors. To avoid the disadvantages of this type of vectors, in this Thesis we have developed an alternative strategy for introducing Cas9 nuclease and CRISPR guides. The gene encoding the Cas9, together with its expression promoter and terminator, were introduced into NB4 genome by lentiviral transduction producing a stable cell line that constitutively express the nuclease. On the other hand, a simple system for the production of CRISPR guides by PCR with essential elements of expression was developed and with *GFP* reporter expression optionally.

In order to optimise the technique and test its efficiency in different targets, two genes involved in AML were modified. These were *IDH2* gene, in which R172 mutation was introduced, and *MYBL2* gene. Finally, editing efficiencies obtained with the new strategy were compared with CRISPR ribonucleoproteins methodology, widely used for its high efficiency. Whereas ribonucleoprotein complexes showed higher cut

efficiencies, the efficiency of edition of R172 mutation efficiency was similar in both strategies. These results were validated and characterized by means of next generation sequencing, and no off-target effects were found. Therefore, the new developed methodology allows precise gene editing in leukemic cell lines with similar efficiencies with other popular CRISPR techniques and without off-target effects.

On the other hand, thanks to the high evolutive conservation of *IDH2* gene R140 and R172 residues are conserved in *Caenorhabditis elegans* *idh-2* protein. The co-CRISPR strategy was used to produce and select mutant strains with ortholog mutations to R140, R172 and one strain with both mutations. Despite the high conservation, the expected increase in oncometabolite 2-HG concentration was not detected in mutant strains compared to the N2 wild type strain. This result could be produced by an activation of 2-HG degradation pathways. So, an exhaustive study of 2-HG degradation pathways is necessary to develop research models with molecular alterations noticed in patients.

In conclusion, the new developed strategy for CRISPR elements introduction in cell lines, together with *C. elegans* models, will allow an in-depth research of molecular effect of mutations detected in patients, its implication in AML progression and prognosis and understand their role in patient stratification.

RESUM

La leucèmia mieloide aguda (LMA) es tracta d'un grup heterogeni de desordres hematològics produïts per alteracions genètiques en les cèl·lules precursors mieloides. Les mutacions en l'enzim Isocitrato deshidrogenasa 2 (IDH2) son d'aquestes alteracions. Les mutacions més freqüents en aquesta proteïna afecten a les posicions R1240 i R172, produint un guany de funció amb la producció de l'oncometabolit D-2-hidroxiglutarat (2-HG). A pesar que ambdues indueixen la producció de 2-HG, la mutació R172 produeix mes quantitat de oncometabolit, presenta menys co-ocurrències con altres alteracions genètiques i s'associa a una pitjor resposta a la quimioteràpia i un major risc de recaiguda. Dades les seues implicacions en el desenvolupament de la malaltia, la mutació *IDH2*^{R172} s'ha proposat com subgrup independent de LMA.

Gràcies a la investigació amb organismes model ha sigut possible estudiar en profunditat el paper de les mutacions genètiques en el desenvolupament de nombroses malalties, com en la LMA. Malgrat això, encara queden molts aspectes per explicar sobre les implicacions de les alteracions genètiques, com les mutacions en *IDH2*, en la LMA. Per tot això, en aquesta Tesis s'han desenvolupat noves estratègies d'edició gènica mitjançant el sistema CRISPR/Cas9 amb l'objectiu de desenvolupar nous models *in vitro* i *in vivo* de les mutacions implicades en la LMA.

En relació al model *in vitro*, degut a la baixa eficiència de transfecció dels plasmids CRISPR en les línies cel·lulars leucèmiques, el mètode més emprat per a introduir els elements CRISPR han sigut principalment vectors lentivirals. Per a evitar els inconvenients d'aquesta mena de vectors, en aquest treball s'ha desenvolupat una estratègia alternativa per a la introducció de la nucleasa Cas9 i els guies CRISPR. El gen codificant de la Cas9, al costat d'un promotor i terminador d'expressió, es van introduir al genoma de cèl·lules NB4 mitjançant transducció amb lentivirus, generant una línia cel·lular amb expressió constitutiva de la nucleasa. D'altra banda, es va desenvolupar un sistema fàcil de producció dels guies CRISPR mitjançant *PCR* amb els elements essencials d'expressió i amb l'expressió del reporter *GFP* de manera opcional.

Amb l'objectiu d'optimitzar la tècnica i provar la seua eficiència en diferents dianes es van modificar dos gens implicats en LMA. Aquests van ser el gen *IDH2*, en el qual es va buscar introduir la mutació R172, i el gen *MYBL2*. Finalment, les eficiències

d'edició obtingudes amb la nova estratègia es van comparar amb l'ús de complexos ribonucleoproteïnes CRISPR, molt utilitzats per la seua alta eficiència. Mentre que els complexos de ribonucleoproteïnes van presentar una major eficiència de tall, l'eficiència d'edició de la mutació R172 va ser similar en les dues estratègies. Mitjançant seqüenciació massiva es va confirmar i caracteritzar aquesta edició i es va comprovar que la maquinària d'edició no havia produït talls inespecífics en regions similars del genoma. D'aquesta manera, la nova metodologia desenvolupada permet editar de manera precisa línies cel·lulars leucèmiques amb eficiències similars a altres tècniques CRISPR més esteses i sense produir efectes inespecífics no desitjats.

D'altra banda, gràcies a la gran conservació evolutiva del gen *IDH2*, els residus R140 i R172 es troben conservats en la proteïna *idh-2* de *Caenorhabditis elegans*. Es va utilitzar l'estratègia co-CRISPR per a desenvolupar i seleccionar ceps mutants amb les mutacions ortòlogues a R140 i R172, i un cep amb dues mutacions. Malgrat l'alta conservació, no es va observar l'augment del oncometabolit 2-HG esperat en els ceps mutants en comparació amb el cep salvatge control N2. Això podria produir-se per l'activació de vies de degradació del 2-HG. En estudis futurs és necessari investigar les vies de degradació del 2-HG per a desenvolupar models d'investigació que presente les alteracions moleculars observades en els pacients

Per a concloure, l'estratègia desenvolupada d'introducció d'elements CRISPR en línies cel·lulars, al costat dels models produïts en *C. elegans* permetran en estudis futurs investigar detalladament els efectes moleculars de mutacions detectades en pacients, la seua implicació en el desenvolupament i prognosi de la LMA i comprendre el seu paper en l'estratificació dels pacients.

ÍNDICE

AGRADECIMIENTOS	vii
RESÚMENES	xi
RESUMEN	xi
ABSTRACT.....	xiii
RESUM	xv
LISTADO DE FIGURAS	xxi
LISTADO DE TABLAS	xxiii
LISTADO DE ABREVIATURAS	xxv
INTRODUCCIÓN.....	3
1. Leucemia.....	4
1.1. Leucemia Mieloide Aguda	5
1.2. Isocitrato deshidrogenasa 2	8
2. Edición génica	15
2.1. Edición génica en la era pre nucleasas.....	15
2.2. Meganucleasas, ZFN y TALEN	16
2.3. Sistema de edición CRISPR/Cas9.....	19
2.4. Otras nucleasas CRISPR	25
2.5. Edición génica en hematología	26
3. <i>Caenorhabditis elegans</i> como modelo de estudio de cáncer.....	29
HIPÓTESIS Y OBJETIVOS	39
MATERIAL Y MÉTODOS	45
Capítulo I. Edición de genes por CRISPR <i>in vitro</i>	45
1. Diseño de las secuencias guías de ARN	45
2. Generación de un vector CRISPR para editar el gen <i>MYBL2</i>	46
3. Generación del vector pEGR-1.....	48
4. Generación de las construcciones por PCR que expresan las guías de ARN.....	50
5. Mantenimiento y cultivos de las líneas celulares empleadas en este trabajo.....	52
6. Edición en la línea celular HEK293	53
7. Optimización de la edición en células HEK293	53
8. Edición del gen <i>IDH2</i> en las células HEK293	54
9. Ensayo de eficiencia de corte mediante Endonucleasa T7 I	55
10. Ensayo de eficiencia de edición mediante <i>RFLP</i>	58
11. Edición del gen <i>IDH2</i> en líneas celulares leucémicas	58
12. Desarrollo de líneas celulares con expresión constitutiva de la nucleasa Cas9.....	59
13. Eficiencia de nucleofección de las líneas leucémicas.....	61
14. Optimización de la edición del gen <i>IDH2</i> en células NB4-Cas9.....	61
15. Edición del gen <i>IDH2</i> en células NB4.....	62
16. Análisis de la eficiencia de edición y los posibles efectos <i>off-targets</i> mediante secuenciación masiva. 62	
Capítulo II. Edición de genes por CRISPR <i>in vivo</i>	64
1. Material.....	64
2. Cepas <i>C. elegans</i>	65
3. Estudio de conservación evolutiva.....	65
4. Generación de gusanos <i>knock-in</i> mediante el sistema CRISPR/Cas9	65

5.	Genotipado de gusanos por <i>PCR</i>	68
6.	Sincronización de gusanos.....	68
7.	Detección y cuantificación del oncometabolito 2-HG.....	69
	Material y Métodos comunes.....	70
	Secuenciación Sanger.....	70
	RESULTADOS.....	77
	Capítulo I. Edición <i>in vitro</i>	77
1.	Desarrollo de un sistema de producción de los elementos CRISPR mediante <i>PCR</i> y optimización en células HEK293.....	79
2.	El uso de dos guías aumenta la eficiencia de corte en el gen <i>IDH2</i>	82
3.	Creación de líneas celulares leucémicas con expresión constitutiva de la nucleasa Cas9...85	
4.	Alta eficiencia de transfección de las construcciones guía desarrolladas en las células NB4-Cas9.....	85
5.	Las construcciones guía producen la misma eficiencia de edición que los complejos <i>RNPs</i> , pero menor porcentaje de <i>indels</i>	87
6.	El análisis de la edición producida con las construcciones CRISPR por secuenciación masiva muestra resultados similares de edición génica sin efectos <i>off targets</i>	90
	Capítulo II. Edición <i>in vivo</i>	96
1.	Establecimiento de posibles modelos de las mutaciones en <i>idh2</i>	96
2.	Estudio de la presencia del oncometabolito 2-HG en las cepas RMV479, RVM480 y RVM481.....	99
	DISCUSIÓN.....	109
	Las construcciones guía generadas por <i>PCR</i> son funcionales.....	109
	Las construcciones guía generadas por <i>PCR</i> son eficientes en la línea celular NB4-Cas9....	111
	La eficiencia de las construcciones CRISPR generadas por <i>PCR</i> es comparable a la de los complejos <i>RNPs</i>	113
	La secuenciación masiva confirma las eficiencias de corte y edición obtenidas.	115
	Mediante secuenciación masiva no se detectan cambios significativos en los <i>off targets</i> analizados.....	116
	Detección del oncometabolito 2-HG en los modelos desarrollados en <i>Caenorhabditis elegans</i>	118
	CONCLUSIONES.....	127
	ANEXO I.....	133
	ANEXO II.....	139
	REFERENCIAS.....	145

LISTADO DE FIGURAS

INTRODUCCIÓN.....	3
Figura 1.1: Hematopoyesis	4
Figura 1.2: Categorías de genes más frecuentemente mutados en la LMA	6
Figura 1.3: Enfermedades asociadas a la desregulación de la actividad de las enzimas IDH..10	
Figura 1.4: Mecanismos de actuación de las mutaciones en <i>IDH1/2</i> y el oncometabolito D-2-HG.....	12
Figura 1.5: Nucleasas pre-CRISPR.....	18
Figura 1.6. Sistema CRISPR/Cas en <i>Streptococcus pyogenes</i>	20
Figura 1.7: Clasificación del sistema CRISPR y de las nucleasas asociadas.....	21
Figura 1.8. Sistema de edición génica CRISPR/Cas9.....	22
Figura 1.9. Aplicaciones del sistema CRISPR/Cas9 en la investigación y desarrollo de terapias en hematología	27
MATERIAL Y MÉTODOS.....	45
Figura 2.1. Localización de las secuencias guías en el exón 4 del gen <i>IDH2</i>	46
Figura 2.2: Vector PX458.....	48
Figura 2.3. Generación del vector pEGR1	49
Figura 2.4. Estrategia de generación de las construcciones guía mediante el vector pEGR1.51	
Figura 2.5. Vector hCas9	52
Figura 2.6. Diseño <i>ssODN</i>	55
Figura 2.7. Ensayo de la enzima Endonucleasa T7 I	57
Figura 2.8. Vector LentiCRISPR V.2.....	59
Figura 2.9: Estrategia de edición génica co-CRISPR. Los complejos de <i>RNPs</i> pre ensamblados se inyectan en la gónada de un hermafrodita adulto.	66
RESULTADOS.....	77
Figura 3.1: Distintas estrategias implementadas para editar las células leucémicas NB4.	78
Figura 3.2: Selección del ratio vector-construcción guía óptimo en células HEK293.....	81
Figura 3.3: Resultados de edición con las construcciones generadas y vectores CRISPR en HEK293.....	82
Figura 3.4: Eficiencias de corte producidas con las guías diseñados para <i>IDH2</i>	83
Figura 3.5: Edición en células HEK293	84
Figura 3.6: Funcionalidad de la nucleasa hSpCas9 insertada en las líneas celulares NB4-Cas9 y HL60-Cas9	86
Figura 3.7. Eficiencia de transfección de las construcciones desarrolladas.	87
Figura 3.8. Optimización edición en NB4-Cas9	88
Figura 3.9: Ensayo <i>RFLP</i> para analizar el porcentaje de edición en células NB4-Cas9.	88

Figura 3.10. Edición mediante complejos <i>RNP</i> en células NB4.....	90
Figura 3.11: Lecturas obtenidas para cada amplicón	92
Figura 3.12: Clasificación de las lecturas obtenidas por el programa CRISPResso2.....	93
Figura 3.13: Gráficas de la distribución de las modificaciones producidas sobre <i>IDH2</i>	94
Figura 3.14. Principales frecuencias alélicas editadas.....	95
Figura 3.16: Árbol evolutivo del gen <i>IDH2</i> . Obtenido con la base de datos Ensembl.	97
Figura 3.17. Homología de la proteína IDH2 humana y la proteína IDH-2 de <i>C. elegans</i>	98
Figura 3.18: Cromatogramas de los enantiómeros L y D y espectros de espectrometría de masas para el D-2-HG)	100
Figura 3.19: Gráfica de las concentraciones medias de D/L-2-HG obtenidas de las distintas cepas estudiadas.	101
ANEXO II.....	139
Figura 3.15: Análisis de los <i>off targets</i> mediante CRISPResso2 <i>Compare</i>	139

LISTADO DE TABLAS

INTRODUCCIÓN.....	3
Tabla 1.1 Clasificación de la LMA y neoplasias relacionadas según la OMS (2016) ¹	7
Tabla 1.2. Estratificación del riesgo en base a la genética por la ENL 2017 ¹	7
Tabla 1.3. Clasificación genética de la LMA propuesta en E. Papaemmanuil <i>et al.</i> , New England Journal of Medicine, 2016 ¹	8
MATERIAL Y MÉTODOS	45
Tabla 2.8: Parámetros empleados en el RPM	70
RESULTADOS.....	77
Tabla 3.1. Recopilación de los porcentajes de <i>indels</i> obtenidos en las diferentes optimizaciones con la construcción pU6-MYBL2 en células HEK293.	81
Tabla 3.2: Recopilación de los porcentajes de <i>indels</i> producidos con los distintos guías diseñados para el gen <i>IDH2</i> y de la eficiencia de la PCR del casete de la nucleasa Cas9.	83
Tabla 3.3: Eficiencias obtenidas de edición en el gen <i>IDH2</i> empleando 5 μ M y 10 μ M de ssODN	85
Tabla 3.4: Eficiencias de producción de <i>indels</i> y de edición mediante el uso de las construcciones desarrolladas y los complejos RNPs.	89
Tabla 3.5. Resultados detección de los enantiómeros D y L del 2-HG mediante LC-MS.	102
Tabla 3.6. Media de las concentraciones de 2HG obtenidas	103
ANEXO I.....	133
Tabla 2.1: Secuencias guía empleadas en los ensayos <i>in vitro</i>	133
Tabla 2.2. Secuencia del ssODN empleado para introducir la mutación R172 en células....	133
Tabla 2.3: Potenciales <i>off targets</i> de los guías <i>IDH2_3</i> e <i>IDH2_8</i>	134
Tabla 2.4. Cepas de <i>C. elegans</i> empleadas	134
Tabla 2.5: Secuencias guía empleadas en <i>C. elegans</i>	134
Tabla 2.6. Secuencias de los ssODNs empleados en <i>C. elegans</i>	135
Tabla 2.7. Potenciales <i>off targets</i> de los guías <i>idh-2_ R123Q</i> e <i>idh-2_ R155K</i>	135
Tabla 2.9. <i>Primers</i> empleados para la amplificación de la secuencia U6-sgRNA en el vector PX458.	135
Tabla 2.10. <i>Primers</i> diseñados para generar las construcciones <i>sgRNA</i> mediante <i>PCRs</i>	136
Tabla 2.11. <i>Primers</i> empleados para amplificar el casete de expresión de Cas9 del vector hCas9.	136
Tabla 2.12. <i>Primers</i> empleados para amplificar la región editada de los genes <i>MYBL2</i> e <i>IDH2</i> en las líneas celulares	136
Tabla 2.13. <i>Primers</i> usados en la preparación de la librería (<i>PCR 1</i>).	137
Tabla 2.14. <i>Primers</i> usados en la preparación de la librería (<i>PCR 2</i>).	137

Tabla 2.15: <i>Primers</i> utilizados para realizar el <i>screening de las mutaciones</i> de <i>idh-2</i> en <i>C. elegans</i>	138
Tabla 2.16: <i>Primers</i> empleados para secuenciación Sanger en <i>C. elegans</i>	138
Tabla 2.17. Cepas bacterianas empleadas en este trabajo.....	138
Tabla 2.18. Vectores empleados en este trabajo.....	138

LISTADO DE ABREVIATURAS

2-HG:	2- hidroxiglutarato
ADN:	Ácido desoxirribonucleico
ADP:	Adenosina 5'-difosfato
ARN:	Ácido ribonucleico
ATP:	Adenosina trifosfato
Cas11:	<i>CRISPR-associated protein 11</i> . Proteína 11 asociada a CRISPR
Cas12a:	<i>CRISPR-associated protein 12a</i> . Proteína 12a asociada a CRISPR
Cas13a:	<i>CRISPR-associated protein 13a</i> . Proteína 13a asociada a CRISPR
Cas14-DETECTR:	<i>DNA Endonuclease Targeted CRISPR Trans Reporter</i> .
Cas7:	<i>CRISPR-associated protein 7</i> . Proteína 7 asociada a CRISPR
Cas9:	<i>CRISPR-associated protein 9</i> . Proteína 9 asociada a CRISPR
CAT:	Ciclo de los ácidos tricarbóxicos
CMV:	Citomegalovirus
CO₂:	Dióxido de carbono
CRISPR:	<i>Clustered Regularly Interspaced Short Palindromic Repeats</i> . Repeticiones palindrómicas cortas agrupadas y regularmente interespaciadas.
CRISPR GUARD	<i>CRISPR Guide RNA Assisted Reduction of Damage</i> . Reducción de daños mediante ARN guía CRISPR.
crRNA:	<i>CRISPR RNA</i> . ARN CRISPR.
dCas9:	<i>dead Cas9</i> . Cas9 inactivada.
DSB:	<i>Double strand break</i> . Corte de doble cadena.
EGFP:	<i>Enhanced green fluorescence protein</i> . Proteína verde fluorescente mejorada
FDA:	<i>Food and Drug Administration</i>
gRNA:	<i>Guide RNA</i> . ARN guía
HDR:	<i>Homologous direct repair</i> . Reparación por homología directa.
HIF1-α:	<i>Hypoxia-inducible factor 1</i> . Factor inducible por hipoxia 1- α
IDH:	Isocitrato deshidrogenasa
iPSC:	<i>Induced pluripotent stem cells</i> . Células humanas pluripotentes inducidas
LMA:	Leucemia mieloide aguda
miARNs:	<i>micro RNAs</i> . micro RNA.
NAD:	Nicotidamina adenina dinucleótido
NADH:	Forma reducida de NAD
NADP:	Cofactor nicotinamida adenina dinucleótido fosfato
NGM:	<i>Nematode growth medium</i> . Medio de crecimiento de nematodos.
NGS:	<i>Next generation sequencing</i> . Secuenciación de última generación.
NHEJ:	<i>Non-homologous end joining</i> . Unión de extremos no homóloga.
NLS:	<i>Nuclear localization signal</i> . Señal de localización nuclear.
OMS:	Organización Mundial de la Salud.
PAM:	<i>Protospacer adjacent motif sequence</i> . Secuencia motivo adyacente a protoespaciador.
PCR:	Polymerase chain reaction. Reacción en cadena de la polimerasa.
RNP:	<i>Ribonucleoprotein</i> . Ribonucleoproteína.

ROS:	<i>Reactive oxygen species</i> . Especies reactivas de oxígeno.
RP:	Retinitis pigmentosa
SHERLOCK:	<i>Specific High-Sensitivity Enzymatic Reporter UnLOCKing</i> . Desbloqueo específico de reportero enzimático de alta sensibilidad
SMD:	Síndrome Mielodisplásico
SNP:	<i>Single nucleotid polymorphism</i> . Polimorfismos de nucleótido único
ssDNA:	<i>single-stranded DNA</i> . ADN de cadena única.
ssODN:	<i>Single stranded oligodeoxyribonucleotide</i> . Oligodesoxiribonucleótido de cadena única
ssRNA:	<i>single-stranded RNA</i> . ARN de cadena única.
T7 EI:	Endonucleasa T7 I
TALE:	<i>Transcription-activator-like effector</i> . Efector similar a activadores de la transcripción.
TALEN:	<i>Transcription-activator-like effector nuclease</i> . Nucleasa de efector similar a activadores de la transcripción.
TMEJ:	<i>Theta-medated end joining</i> . Unión de extremos mediada por Theta.
tracrRNA:	<i>trans-activating crRNA</i> . ARN <i>CRISPR</i> trans activador.
VIH:	Virus de la inmunodeficiencia humana
ZFN:	<i>Zinc fingers nuclease</i> . Nucleasa dedos de Zinc.
α-CG:	α -cetogluturato

INTRODUCCIÓN

INTRODUCCIÓN

Con la llegada de la secuenciación del genoma humano se inició la llamada era post-genómica y con ella un reto formidable para la biomedicina, comprender el impacto de toda la información genómica en la compresión de las enfermedades.

Los estudios de secuenciación del genoma han revelado que en la Leucemia Mieloide Aguda (LMA) están implicados un gran número de genes, dándose complejos patrones mutacionales con mutaciones concurrentes o excluyentes entre sí. Estas alteraciones genéticas se asocian a características clínicas y pronósticas específicas y, en algunos casos, se han podido asociar a subgrupos de LMA concretos. Esto ha permitido una mejor clasificación de los pacientes y el desarrollo de protocolos de actuación y fármacos específicos. Sin embargo aún se desconoce en buena medida la relación directa entre muchas de estas alteraciones con las diversas características clínicas, el pronóstico y la respuesta al tratamiento, así como el modo en que cooperan entre sí en el desarrollo de la enfermedad. Para esclarecer el papel de estas mutaciones es necesario reproducirlas de forma rigurosa en modelos *in vitro* e *in vivo* para poder caracterizarlas funcionalmente.

La tecnología CRISPR/Cas9 ha supuesto una verdadera revolución en la edición génica, siendo merecedora del Premio Nobel de Química 2020 concedido a las científicas Emmanuelle Charpentier y Jennifer Doudna. Con esta tecnología se ha editado el genoma de numerosos organismos como células humanas, primates, ratones, pez cebra, varios invertebrados y microorganismos. En el ámbito de la biomedicina se ha aplicado para desarrollar modelos de enfermedades humanas hereditarias y soluciones terapéuticas basadas en la capacidad para corregir la causa genética de la enfermedad.

En el presente trabajo se han establecido nuevas estrategias para desarrollar modelos de investigación *in vitro* e *in vivo* de mutaciones de LMA mediante el sistema CRISPR/Cas9. Las líneas celulares leucémicas son células con una eficiencia de transfección muy baja, por lo que introducir los vectores de expresión de la nucleasa Cas9 resulta una tarea compleja. Para evitar este obstáculo, se ha desarrollado un protocolo para introducir los elementos guía de la nucleasa Cas9 en forma de pequeñas moléculas de ADN en líneas celulares leucémicas. De esta forma se ha conseguido editar

los genes *IDH2* y *MYBL2* implicados en LMA. Para los estudios *in vivo*, se ha escogido el organismo modelo *Caenorhabditis elegans*, aprovechando la alta conservación evolutiva de algunos genes implicados en la LMA, como *IDH2*. Mediante el empleo de nuevos protocolos de edición génica con el sistema CRISPR/Cas9 diseñados al efecto, ha sido posible generar nuevas cepas de este modelo animal con mutaciones en el gen *IDH2*. Estos nuevos modelos nos ayudarán a comprender el papel de estos genes en la estratificación de la LMA

1. Leucemia

Las leucemias son un grupo de desórdenes hematológicos caracterizados por la proliferación descontrolada de precursores de células hematológicas. Si la estirpe de células comprometidas son los precursores linfoides las leucemias se clasifican como linfocíticas. Estos precursores dan lugar a células T, B y *natural killers*. Por otro lado, se clasifican como mieloides cuando las células afectadas son los precursores mieloides. Las células mieloides incluyen los monocitos, macrófagos, neutrófilos, basófilos, eosinófilos, eritrocitos y megacariocitos. A su vez, de un modo general, cada uno de estos grupos se puede diferenciar en leucemia aguda, cuando las células malignas son menos maduras e invaden rápidamente la médula ósea, afectando a su funcionalidad y desarrollándose la enfermedad rápidamente, o crónica, cuando las células afectadas principalmente son más diferenciadas y funcionales, no se da una invasión tan alta de la médula ósea y los síntomas se desarrollan lentamente (**Figura 1.1**).

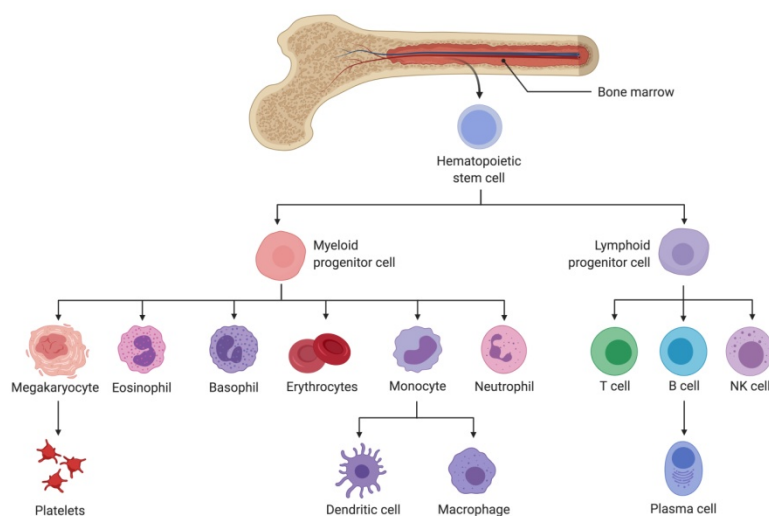


Figura 1.1: Hematopoyesis. Imagen creada con Biorender.com

1.1. Leucemia Mieloide Aguda

La leucemia mieloide aguda (LMA) es un subtipo heterogéneo de leucemias que se caracteriza por englobar numerosos desórdenes hematológicos derivados de una proliferación clonal de precursores mieloides (blastos). Se considera LMA cuando la concentración de blastos es igual o mayor al 20 % de células en sangre o médula ósea. Comprende el 80 % de las leucemias agudas en la edad adulta y el 15-20 % de las leucemias pediátricas, afectando en mayor proporción a varones (60 %) (Brown *et al.*, 2017). A pesar de los avances en el tratamiento de estos trastornos, y exceptuando algunos subgrupos específicos, actualmente tan solo el 35-40 % de los pacientes adultos menores de 60 años y entre el 5-15 % de los mayores de 60 años alcanzan la supervivencia libre de enfermedad prolongada (Döhner *et al.*, 2015). La enfermedad refractaria es común y la recaída representa la principal causa de fallo de tratamiento (Papaemmanuil *et al.*, 2016).

La clasificación de la LMA se ha basado clásicamente en características morfológicas y citogenéticas. Gracias al desarrollo de nuevas tecnologías de secuenciación masiva del genoma ha sido posible la detección de nuevas anomalías genéticas, se ha profundizado en los mecanismos moleculares implicados y ha sido posible la caracterización de la arquitectura clonal de la enfermedad (Arber *et al.*, 2016).

En términos generales, los casos de LMA presentan una media de 13 mutaciones, donde los genes más recurrentemente mutados se pueden clasificar en ocho grupos: metilación del ADN (*DNMT3A*, *TET1*, *TET2*, *IDH1*, *IDH2*), modificación de la cromatina (*EZH2*, *ASXL1*, *KDM6A*), vías de señalización (*FLT3*, *NRAS*, *PTPN11*, *KRAS*), desarrollo hematopoyético (*CEBPA*, *RUNX1*, *GATA2*), supresores de tumores (*TP53*, *WT1*, *PHF6*), *splicing* de ARN (*SRSF2*, *U2AF1*, *SF3B1*), complejos de cohesina (*SMC1A*, *SMC3*, *STAG2*, *RAD21*) y en el gen *NPM1*, codificante de una proteína núcleo-citoplasmática multifuncional (**Figura 1.2**). Estos grupos muestran patrones complejos de concurrencia y exclusividad entre ellos (Shi *et al.*, 2018). Otra de las características que se reveló con estos estudios moleculares es la capacidad de adquisición de nuevas mutaciones somáticas en el curso de la enfermedad, produciendo una alta diversidad clonal con distintas respuestas al tratamiento. Además, mutaciones en determinados genes se clasifican como mutaciones iniciadoras o *drivers*, desencadenantes de la enfermedad,

mientras que otras se consideran mutaciones secundarias, pudiendo jugar un papel relevante en la progresión de la enfermedad o condicionando la respuesta al tratamiento.

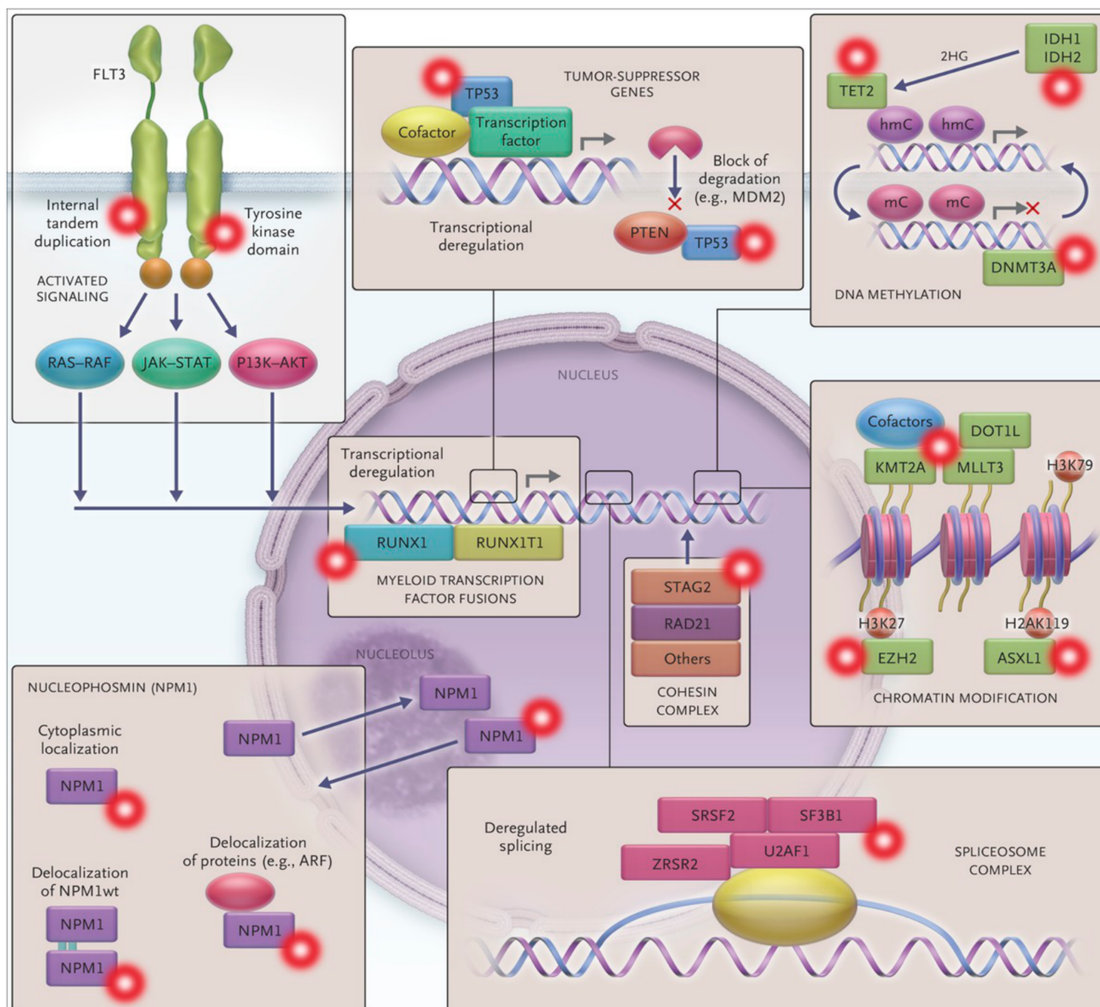


Figura 1.2: Categorías de genes más frecuentemente mutados en la LMA. Las mutaciones en genes de señalización confieren una ventaja proliferativa a través de las vías de señalización RAS-RAF, JAK-STAT y PI3K-AKT. Las mutaciones en factores de transcripción conducen a la desregulación transcripcional y alteración de la diferenciación hematopoyética. En el gen *NPM1* las mutaciones dan como resultado la localización citoplásmica aberrante de la proteína NPM1 y las proteínas que interaccionan con ella. Las mutaciones en genes del complejo *splicing* están involucradas en el procesamiento aberrante del ARN. Las mutaciones en el complejo cohesina pueden afectar a la segregación cromosómica y a la regulación transcripcional. Las mutaciones en genes reguladores de la cromatina conducen a desregulaciones de la epigenética. Mutaciones en genes supresores de tumores conducen a la desregulación transcripcional y la degradación a través de MDM2 y PTEN (Döhner *et al.*, 2015).

Esta nueva información se está aplicando en el desarrollo de una mejor estratificación de los pacientes con LMA. En la última revisión de la guía de clasificación de los tumores Hematológicos de la Organización Mundial de la Salud (OMS) de 2016 se fijaron dos subgrupos propuestos en 2008, LMA con *NPM1* mutado y LMA con

mutaciones bialélicas en *CEBPA*, y se propusieron dos nuevas categorías: LMA con la fusión *BCR-ABL1* y LMA con *RUNX1* mutado (**Tabla 1.1**).

Tabla 1.1 Clasificación de la LMA y neoplasias relacionadas según la OMS (2016)¹

LMA con alteraciones genéticas recurrentes
LMA con t(8;21) (q22;q22.1); <i>RUNX1-RUNX1T1</i>
LMA con inv(16)(p13.1q22) o t(16;16)(p13.1;q22); <i>CBFB-MYH11</i>
Leucemia Promielocítica Aguda (LPA) con <i>PML-RARA</i>
LMA con t(9;11)(p21.3;q23.3); <i>MLL3-KMT2A</i>
LMA con t(6;9)(p23;q34.1); <i>DEK-NUP214</i>
LMA con inv(3)(q21.3q26.2) o t(3;3)(q21.3;q26.2); <i>GATA2, MECOM</i>
LMA con (megacarioblástica) con t(1;22)(p13.3;q13.3); <i>RBM15-MKL1</i>
Entidad provisional: LMA con <i>BCR-ABL1</i>
LMA con <i>NPM1</i> mutado
LMA con mutaciones bialélicas en <i>CEBPA</i>
Entidad provisional: LMA con <i>RUNX1</i> mutado
LMA con cambios asociados a mielodisplasia
Neoplasias relacionadas con el tratamiento
LMA no categorizadas
Sarcoma Mieloide
Proliferación mioelode relacionada con el síndrome de Down

¹Peterson *et al.*, 2016

La importancia de la inclusión de las alteraciones genéticas también se vio reflejada en la actualización de las recomendaciones de la *European LeukemiaNet*. En esta actualización, un panel de 22 expertos internacionales propusieron una nueva estratificación de riesgo de los pacientes de LMA considerando las alteraciones genéticas (Döhner *et al.*, 2017) (**Tabla 1.2**).

Tabla 1.2. Estratificación del riesgo en base a la genética por la ENL 2017¹.

Categoría de riesgo	Alternación genética
Favorable	t(8;21) (q22;q22.1), <i>RUNX1-RUNX1T1</i> inv (16)(p13.1q22) o t(16;16)(p13.1;q22), <i>CBFB-MYH11</i> Mutaciones en <i>NPM1</i> sin <i>FLT3-ITD</i> o con <i>FLT3-ITD</i> ^{low} Mutaciones bialélicas en <i>CEBPA</i>
Intermedio	Mutaciones en <i>NPM1</i> y <i>FLT3 ITD</i> ^{high} <i>NPM1</i> salvaje sin <i>FLT3-ITD</i> o con <i>FLT3-ITD</i> ^{low+} (sin lesiones genéticas de riesgo adverso) t(9;11) (p21.3;q23.3), <i>KMT2A</i> Anormalidades citogenéticas no clasificadas como favorables o adversas
Adverso	t(6;9)(p23;q34.1), <i>DEK-NUP214</i> t(v;11q23.3), reordenamiento de <i>KMT2A</i> t(9;22)(q34.1;q11.2), <i>BCR-ABL1</i> inv(3)(q21.3q26.2) o t(3;3)(q21.3;q26.2), <i>GATA2, MECOM(EVI1)</i> -5 o del(5q), -7; -17/abn(17p) Cariotipo complejo, cariotipo monosomal <i>NPM1</i> salvaje y <i>FLT3-ITD</i> ^{high} <i>RUNX1</i> mutado <i>ASXL1</i> mutado <i>TP53</i> mutado

¹Döhner *et al.*, 2017

Con el objetivo de desarrollar una nueva clasificación basada en las características genéticas, el equipo de E. Papaemmanuil estudió 111 genes relacionados con cáncer en 1540 pacientes de LMA y combinaron estos resultados con información citogenética y datos clínicos. De este análisis surgieron tres nuevos subgrupos genómicos con relevancia clínica no incluidos en la clasificación de la OMS: mutaciones en genes cromatina-*spliceosoma*, aneuploidía de *TP53* y de forma provisional, la mutación *IDH2*^{R172} (**Tabla 1.2**) (Papaemmanuil *et al.*, 2016).

A pesar de que estos estudios pusieron de manifiesto la relevancia de las alteraciones genéticas como factores pronóstico, solamente se han incorporado un número pequeño de alteraciones génicas (*RUNX1*, *ASXL1*, *TP53*, *NPM1*, *FLT3* y *CEBPA*) en las guías actuales de estratificación de riesgo (Döhner *et al.*, 2017). Esto se debe principalmente a la falta de asociaciones fiables entre genes individuales, características clínicas y pronóstico.

Tabla 1.3. Clasificación genética de la LMA propuesta en E. Papaemmanuil *et al.*, New England Journal of Medicine, 2016¹.

LMA con mutación en <i>NPM1</i>
LMA con mutaciones en genes de cromatina, <i>splicing</i> -RNA, o ambos*
LMA con mutaciones en <i>TP53</i> , aneuploidía cromosómica, o ambos**
LMA con inv (16)(p13.1q22) o t(16;16)(p13.1;q22); <i>CBFB-MYH11</i>
LMA con mutaciones bialélicas en <i>CEBPA</i>
LMA con t(15;17)(q22;q12); <i>PML-RARA</i>
LMA con t(8;21)(q22;q22); <i>RUNX1-RUNX1T1</i>
LMA con fusión en el gen <i>KMT2A</i> ; t(x;11) (x;q23)***
LMA con inv (3)(q21q26.2) o t(3;3)(q21;q26.2); <i>GATA2</i> , <i>MECOME(EVII)</i>
LMA con la mutación <i>IDH2</i> ^{R172} y ninguna otra alteración característica de grupo
LMA con t(6;9) (p23;q34); <i>DEK-NUP214</i>
LMA con mutaciones <i>driver</i> , pero incluidas en ninguna característica de grupo
LMA sin mutaciones <i>driver</i> detectadas
LMA que cumple los criterios para 2 subgrupos génicos

* La clasificación en este subgrupo necesita de una o mas mutaciones *drivers* en *RUNX1*, *ASXL1*, *BCOR*, *STAG2*, *EZH2*, *SRSF2*, *SF3B1*, *USAF1*, *ZRSR2* o *MLL*^{PTD}. En presencia de otras mutaciones características de otros subgrupos (como inv (16), t(15;17), t(6;9), genes de fusión *KMT2A*, cariotipo complejo o mutaciones *drivers* en *TP53*, *NPM1* o bialélicas en *CEBPA*) se necesitan dos o mas mutaciones en genes de cromatina-*splicing*.

**La clasificación en este subgrupo necesita de mutaciones en *TP53*, cariotipo complejo, o en ausencia de otras alteraciones características de grupo, uno o más de las siguientes: -7/7q, -5/5q, -4/4q, -9q, -12/12p, -17/-17p, -18/18q, -20/20q, +11/11q, +13, +21 o +22.

***En el grupo se incluyen múltiples patrones de fusión de *MLL*, con implicaciones clínicas dependiendo del patrón específico de fusión.

¹Papaemmanuil *et al.*, 2016

1.2. Isocitrato deshidrogenasa 2

Las enzimas isocitrato deshidrogenasas (IDH) catalizan la carboxilación oxidativa del isocitrato a α -cetoglutarato (α -CG) y dióxido de carbono (CO₂). En las células

eucariotas se encuentran tres enzimas IDH parálogas, IDH1, IDH2 e IDH3. A pesar de catalizar la misma reacción están codificadas por genes distintos y difieren en la localización subcelular, cofactores, regulación alostérica y mecanismos reguladores (Tommasini-Ghelfi *et al.*, 2019a). Las tres son enzimas esenciales del ciclo de los ácidos tricarbóxicos (CAT) o ciclo de Krebs, jugando un papel esencial en el metabolismo energético y balance reducción y oxidación (redox).

Las enzimas IDH1 e IDH2 presentan muchas similitudes. Son dependientes del cofactor nicotinamida adenina dinucleótido fosfato (NADP) y de Mn^{2+} o Mg^{2+} , son homodímeros, formadas por un dominio grande, uno pequeño y un dominio *clasp*, y presentan una homología del 70 % en sus secuencias. Por otro lado, se diferencian en su localización celular, ya que mientras IDH1 se localiza en el citosol y en peroxisomas, IDH2 se localiza en la mitocondria. Por último, la enzima IDH3 es dependiente de nicotinamida adenina dinucleótido (NAD) y está formada por heterodímeros codificados en distintos genes (*IDH3A*, *IDH3B* e *IDH3G*). Su actividad está regulada por las necesidades energéticas de la célula, activándose por adenosina 5'-difosfato (ADP) y NAD e inhibiéndose por la forma reducida de NAD (NADH), adenosina trifosfato (ATP) y α -CG. Al igual que IDH2 se localiza en la mitocondria. Además, mientras las reacciones catalizadas por IDH1 e IDH2 son reversibles, IDH3 cataliza solamente la conversión de isocitrato a α -CG (Dang *et al.*, 2017).

Gracias a los avances en la investigación genómica ha sido posible identificar mutaciones en *IDH1* e *IDH2* en un gran abanico de tumores, como en distintos tipos de cáncer cerebral, desde glioblastomas primarios, secundarios y de grado II/III, astrocitomas y oligodendrogliomas de grado I/II. También se detectaron desórdenes hematológicos como LMA, y con menos frecuencia en síndromes mielodisplásicos y neoplasias mieloproliferativas. Además contribuyen a la patogénesis de desórdenes esqueléticos, varios tipos de condrosarcomas, colonangiocarcinoma intrahepático y carcinoma de tiroides. Por último mutaciones en estas enzimas son características de subtipos de cáncer de pecho (Dang *et al.*, 2017; Tommasini-Ghelfi *et al.*, 2019b). Las mutaciones en IDH1 e IDH2 implicadas en estas enfermedades se dan en heterocigosis.

Por otro lado, las mutaciones en *IDH3* se han asociado a enfermedades hereditarias de retina que se caracterizan por la degradación progresiva de

fotoreceptores y/o de las células del epitelio pigmentario. Dentro de este grupo, la retinitis pigmentaria (RP) es la enfermedad más común. Mutaciones *missense* homocigotas en el codón 304 de *IDH3A* se han asociado con el fenotipo más severo de la enfermedad. Otras variantes de este gen se han asociado a RP autosómica recesiva y a pseudocoloboma macular. Mutaciones en *IDH3B* también se han asociado con RP autosómica recesiva. A pesar de encontrar variantes en *IDH3C*, no está claro aún su papel en la enfermedad (Tommasini-Ghelfi *et al.*, 2019b) (**Figura 1.3**).

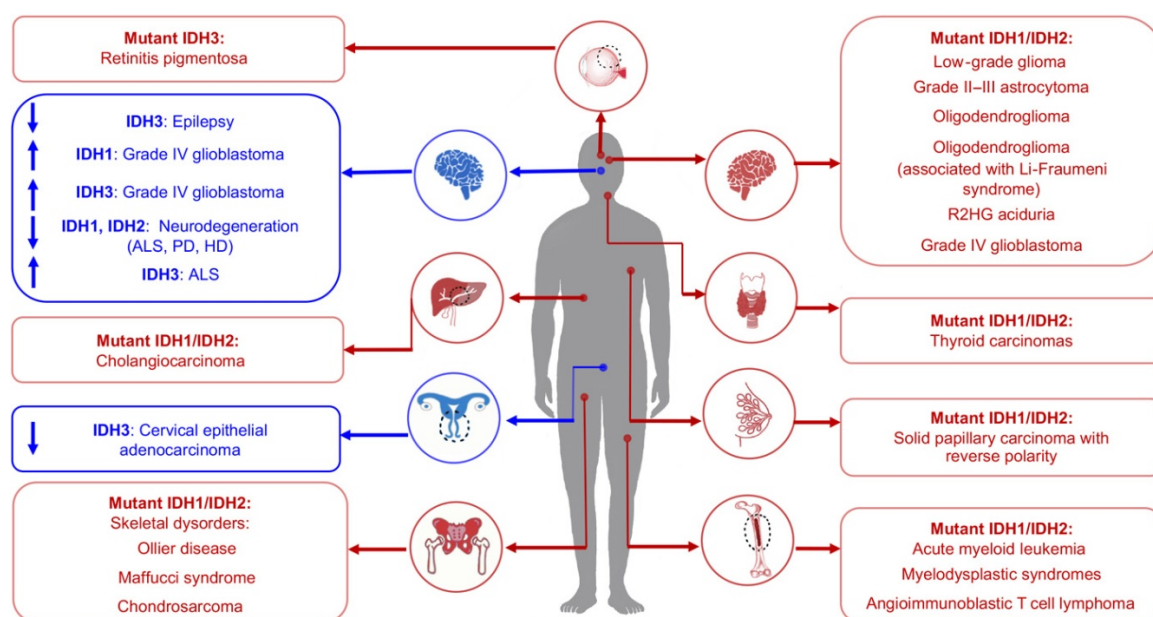


Figura 1.3: Enfermedades asociadas a la desregulación de la actividad de las enzimas IDH. Los recuadros azules indican las enfermedades producidas por alteraciones de la expresión. Los recuadros rojos indican las enfermedades debidas a mutaciones (Tommasini-Ghelfi *et al.*, 2019b).

Centrándonos en la LMA, las mutaciones encontradas en *IDH1* e *IDH2* son de tipo *missense*, desencadenando un cambio de aminoácido en la secuencia de proteína. En *IDH1* la mutación más frecuente afecta al residuo arginina 132, produciendo un cambio a cisteína (R132C), o a histidina (R132H). En *IDH2* los residuos más frecuentemente mutados son el 140, presentando una sustitución de arginina por glutamina (R140Q) y el 172, sustituyendo el residuo de arginina principalmente por lisina (R172K), y en menor frecuencia por triptófano (R172W) (Medeiros *et al.*, 2017). Mientras que las mutaciones *IDH1*^{R132} e *IDH2*^{R172} se encuentran en otros tipos de tumores, la *IDH2*^{R140} es exclusiva de la LMA (Pirozzi *et al.*, 2021).

Estas mutaciones producen una ganancia de función de las enzimas mutadas *IDH1* (m*IDH1*) e *IDH2* (m*IDH2*). Los residuos afectados se localizan en los dominios de

unión al sustrato. Al producirse estas mutaciones se modifica la afinidad al cofactor y la conformación del centro activo de la enzima, produciendo que aumente la afinidad por NADP y disminuyendo la afinidad por el isocitrato. Esto desemboca en una enzima mutada que cataliza la reacción contraria, la reducción del α -CG a D-2-hidroxiglutarato (D-2-HG). El D-2-HG está presente en células sanas a baja concentración, con dos formas enantiómeras, R (D) o S (L) y todavía no se ha encontrado su función bajo estas condiciones. Es estructuralmente muy similar al α -CG, y al aumentar su concentración en presencia de las enzimas mIDH1 y mIDH2 actúa como un inhibidor competitivo de las dioxigenasas dependientes de α -CG. Induce una serie de modificaciones en la célula, como alteraciones epigenéticas, inhibición del factor inducible a hipoxia 1- α (HIF1- α), cambios en el ambiente redox celular y alteraciones en el colágeno, entre otras. De esta forma, el D-2-HG actúa como oncometabolito en numerosas rutas, produciendo la inhibición de la diferenciación celular y una auto renovación celular patológica (Cerchione *et al.*, 2021; Clark *et al.*, 2016) (**Figura 1.4**). Las alteraciones en los patrones de metilación y expresión génica son similares a los presentes en los casos de mutaciones en el gen *TET2*. Unido a que las mutaciones en ambos genes son mutuamente excluyentes puede indicar vías moleculares en común (Pirozzi *et al.*, 2021).

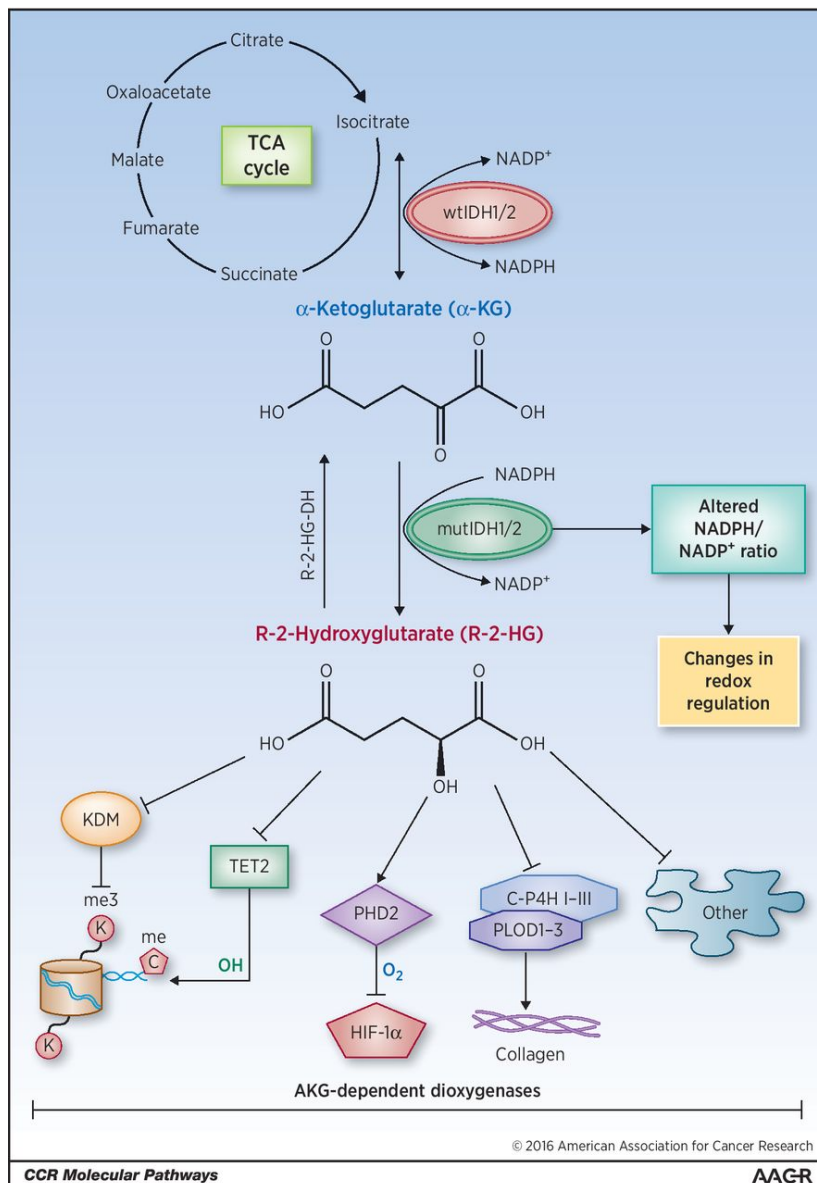


Figura 1.4: Mecanismos de actuación de las mutaciones en *IDH1/2* y el oncometabolito D-2-HG. Las enzimas *IDH1* e *IDH2* del ciclo de Krebs catalizan la conversión de isocitrato a α -CG. Al mutar adquieren una ganancia de función por la cual transforman el α -CG en el oncometabolito D-2-HG. Esta molécula se une a las dianas de α -CG produciendo numerosas alteraciones como: Modificaciones epigenéticas a través de la inhibición directa de enzimas ADN metilasas de la familia *ten-eleven trasloction* (TET) e histona lisina desmetilasas (KDM), alteración de la maduración del colágeno por la inhibición de la prolil-hidroxilación, la activación de la enzima prolil-hidroxilasa 2 (PHD2), inhibición del factor de transcripción inducible por hipoxia 1- α (HIF-1 α). El consumo de NADPH deriva en una reprogramación metabólica que conlleva una desregulación de la expresión génica, inestabilidad genética por la inhibición directa de enzimas de reparación homóloga del ADN alkB dependientes de α -CG, o indirectamente por la alteración e la expresión de genes de reparación del ADN (Clark *et al.*, 2016).

A pesar de que las mutaciones en *mIDH1* y *mIDH2* producen el aumento del oncometabolito D-2-HG no son funcionalmente equivalentes. Se ha descrito la pérdida de heterocigosidad en casos de mutaciones en *mIDH2* por la cual desaparece el alelo salvaje, pero no en *mIDH1*. Esta diferencia puede deberse a que *mIDH1* esté limitado por la producción de α -CG de *IDH1* salvaje, mientras que *mIDH2* tiene más

independencia de un dominio salvaje por el nivel mantenido de α -CG que hay presente en la mitocondria (Pirozzi *et al.*, 2021; Ward *et al.*, 2013). Por otro lado, se ha descrito que la cantidad de D-2-HG producido por mIDH1 es menor que por mIDH2 y, a su vez, la mutación R172 produce mayor cantidad de oncometabolito que la R140 (Medeiros *et al.*, 2017; Pirozzi *et al.*, 2021).

Centrándonos en la mutaciones de IDH2, en la clasificación genómica de la LMA propuesta por E. Papaemmanuil (2016), la mutación $IDH2^{R172}$ se propuso como subgrupo independiente. Esto se debe a que, a pesar de su baja frecuencia en la cohorte estudiada (sólo un 1 %), los pacientes con esta mutación presentan características clínicas y moleculares específicas (Medeiros *et al.*, 2017). Por una parte, R172 presenta patrones de concurrencia con pocas mutaciones, mientras que R140 presenta una fuerte correlación con mutaciones en el gen *NPM1* (con la que la mutación R172 es mutuamente excluyente) y en el gen *DNMT3A* (Papaemmanuil *et al.*, 2016). Además los pacientes con la mutación R140 presentan de forma general, mayores tasas de respuesta completa al tratamiento, mayor supervivencia global y menor riesgo de recaída comparados con pacientes de LMA con la mutación R172 (Pirozzi *et al.*, 2021). Para esclarecer el papel pronóstico de las mutaciones de *IDH2* en LMA se X y colaboradores realizaron un meta análisis con 12747 pacientes. Las mutaciones R172 y R140 se asociaron a una supervivencia global favorable pero con tasas de remisión completas menores en pacientes con R172 (Q. Xu *et al.*, 2017). Por otro lado, otros estudios no han encontrado ninguna asociación entre las mutaciones de *IDH2* con implicaciones pronosticas, pudiendo deberse a otras mutaciones concurrentes en los pacientes que no se tuvieron en cuenta (Pirozzi *et al.*, 2021).

Los modelos de investigación han sido claves en el estudio de los efectos moleculares de las mutaciones en *IDH2*, así como en el testado de fármacos. Los modelos *in vitro* se han desarrollado principalmente mediante lentivirus codificantes del cADN de la proteína mutada bajo el control de un promotor inducible (Kernytzky *et al.*, 2015; Schwartzman *et al.*, 2019), aunque también se ha empleado el sistema de edición génica CRISPR/Cas9 para introducir la mutación puntual R140 en blastos de pacientes empleando lentivirus integrativos (Brabetz *et al.*, 2017). Por otro lado, se han generado modelos de ratón *xenograft* de la mutación R140 mediante la introducción de células de

pacientes con esta mutación en ratones inmunodeprimidos (Cathelin *et al.*, 2018; K. Yen *et al.*, 2017). Exceptuando el caso del modelo desarrollado mediante CRISPR, el resto de modelos producidos no expresan de forma endógena las mutaciones de *IDH2*.

El estudio en estos modelos ha contribuido al desarrollo de fármacos que inhiben las mutaciones de *IDH2*. Uno de ellos, enasidenib o AG-221 (Idhifa; Celgene Corp) (NCT03683433), es un inhibidor específico de las mutaciones R140Q y R172K (K. Yen *et al.*, 2017). Actúa uniéndose al sitio alostérico de la enzima mutada, inactivándola e induciendo la disminución de la producción del oncometabolito y por tanto, induciendo la diferenciación de los blastos. Actualmente se encuentra en fase de estudio I/II/III aprobados por la *U.S. Food and Drug Administration* (FDA) para tratamiento de pacientes con LMA refractaria o en recaída. En el primer ensayo clínico fase I/II llevado a cabo se estudió el efecto del tratamiento con enasidenib en 345 pacientes con enfermedades hematológicas con mutaciones en *IDH2*. En este estudio la supervivencia media de los pacientes con LMA refractaria y/o recaída tratados fue de 8,8 meses (Stein *et al.*, 2019; Pirozzi *et al.*, 2021). Una gran mejora en comparación con la supervivencia media de 3,3 meses obtenida con otras terapias (Roboz *et al.*, 2014). Estos prometedores resultados han sentado la base para el empleo de enasidenib como tratamiento para extender la remisión y posteriormente realizar un trasplante de médula ósea. Como inconveniente se han reportado efectos secundarios derivados de este tratamiento. Se han descrito casos de síndromes de diferenciación que produce una gran liberación de citocinas (6 %), hiperbilirubinemia (10 %) y trombocitopenia (7 %) (Fathi *et al.*, 2018; Pirozzi *et al.*, 2021). En paralelo, se están desarrollando ensayos clínicos para evaluar el efecto de enasidenib como terapia de mantenimiento post trasplante (NCT03515512 y NCT03728335) y la combinación con azacitidina en casos nuevos diagnosticados con mutaciones en *IDH2* (NCT02677922). A pesar del riesgo de desarrollar el síndrome de diferenciación, el beneficio probado en los pacientes ha servido para conseguir la aprobación de la Agencia Europea del Medicamento para iniciar su uso en Europa (Ahmadmehrabi *et al.*, 2021).

A pesar de los grandes avances realizados en el desarrollo de nuevos fármacos, no se ha alcanzado la curación total de estos pacientes y en algunos casos se han registrado resistencias primarias y adquiridas al tratamiento. Las resistencias primarias

al tratamiento se han asociado con la concurrencia con mutaciones en los genes *NRAS*, *CEBPA* y *RUNX1* (Wang *et al.*, 2021). Por otro lado, en las resistencias adquiridas se han descrito casos en los que se ha adquirido una segunda mutación en *trans* en el gen *IDH2* en las posiciones Q316E e I319M, inhibiendo la unión del fármaco a la proteína mutante (Intlekofer *et al.*, 2018). La adquisición de mutaciones en los genes *IDH1*, *KRAS*, *NRAS*, *BCOR* y *TET2* también se ha asociado con la generación de resistencias (Wang *et al.*, 2021).

Los importantes efectos de las mutaciones en *IDH2* junto con su alta prevalencia en varios tipos de cáncer hacen que sea un gen ideal para el desarrollo de nuevos modelos de investigación que expresen de forma endógena las mutaciones más frecuentes para ayudar a una mejor comprensión de su implicación en el desarrollo de la enfermedad.

2. Edición génica

Desde el descubrimiento de la doble hélice de ADN los científicos han codiciado realizar cambios específicos en el genoma de células y organismos. De esta creciente necesidad surgió la edición génica, entendida como la introducción de modificaciones específicas en el genoma de organismos vivos. La capacidad de modificar la información genética es una herramienta fundamental en la investigación biológica y se ha aplicado en numerosos campos, como la mejora de especies de cultivo para aumentar su resistencia a entornos hostiles o su composición nutricional, la mejor comprensión de los procesos biológicos de numerosas especies y el desarrollo de modelos de enfermedades humanas y nuevos fármacos con prometedores resultados.

2.1. Edición génica en la era pre nucleasas

La primera estrategia empleada para modificar el genoma fue el uso de transposones y sustancias químicas que producen alteraciones espontáneas e impredecibles. Estas estrategias resultaban ser muy laboriosas y no permiten la introducción de cambios específicos ni dirigidos. En la década de 1980 comenzó el desarrollo de diferentes estrategias de edición génica que han evolucionado con el avance de la investigación y la tecnología. La primera estrategia de edición génica directa fue la modificación dirigida sin emplear rotura del ADN (Capecchi, 2005). Esta técnica se

basa en la introducción de una molécula de ADN homóloga a una región de interés con una mutación específica. Mediante la conversión por recombinación de los brazos homólogos de la molécula con la diana es posible introducir la modificación de interés para producir cambios funcionales en este gen (*knock-in*) o para silenciarlo (*knock-out*). La adición de genes de selección, como por ejemplo genes de resistencia a antibióticos, facilitó la selección de clones editados. A pesar de ser un gran avance, esta estrategia depende en gran medida del factor azaroso de interacción entre la molécula de ADN y la región diana, lo que la hace muy ineficiente (en el rango de 10^{-7} a 10^{-6}) (J.-S. Kim, 2016).

2.2. Meganucleasas, ZFN y TALEN

Un punto de inflexión en la evolución de las técnicas de edición génica fue la descripción de un gran aumento de la eficiencia de edición tras generar un corte de doble cadena (*DSB*, de las siglas en inglés *Double-Strand Break*) en la hebra de ADN (Rouet *et al.*, 1994; Rudin *et al.*, 1988). Los *DSBs* son eventos que inducen inestabilidad en la célula, por lo que se activa la maquinaria de reparación celular para proteger la funcionalidad del genoma. La reparación del ADN se realiza mediante dos mecanismos: la reparación precisa por homología directa (*HDR*, siglas del inglés *Homologous Direct Repair*) o la reparación no homóloga (*NHEJ*, *Non-homologous end joining*), que se caracteriza por una elevada tasa de errores. En los primeros estudios se introdujeron cortes mediante las nucleasas de restricción de levadura HO y I-SceI, obteniéndose un aumento de la eficiencia de introducción del ADN molde en más de dos órdenes de magnitud respecto al uso de la recombinación homóloga sin cortes inducidos. Además, se describió la reparación mediante mecanismos no homólogos, que resultaban en deleciones y/o inserciones de nucleótidos (*indels*) y recombinaciones (Rouet *et al.*, 1994; Rudin *et al.*, 1988). Estas nucleasas (o meganucleasas por su tamaño) producen el *DSB* tras el reconocimiento de una secuencia grande de entre 12 a 40 pares de bases (**Figura 1.5A**). A pesar del gran salto producido al asociar el aumento de eficiencia con los *DSBs*, las meganucleasas presentan muchas desventajas, entre ellas la necesidad de una gran secuencia homóloga presente en el *loci* a editar, ineficiencia para producir cambios bialélicos y que su uso se restringe a unos pocos tipos celulares. A pesar de ello, marcaron el inicio de la era de la edición génica mediante nucleasas.

Unos años más tarde, en 1996 el investigador Chandrasegara y su equipo desarrollaron las nucleasas unidas a dedos de Zinc (*ZFNs*, siglas del inglés *Zinc Fingers Nucleases*) con la fusión del dominio nucleasa de la enzima FokI con dominios de dedos de Zinc de unión al ADN (Y. G. Kim *et al.*, 1996). Cada dominio dedo de Zinc es capaz de interactuar con una secuencia específica de 3 nucleótidos (**Figura 1.5B**). De esta forma los *ZFNs* reconocen la secuencia diana de forma modular, con lo que se pueden disponer de varios *ZFNs* en un orden determinado para reconocer un sitio muy específico. Con esta tecnología se consiguió modificar de forma dirigida el genoma de *Drosophila melanogaster* (Bibikova *et al.*, 2002), *Caenorhabditis elegans* (Morton *et al.*, 2006), *Zebrafish* (Doyon *et al.*, 2008), plantas (Lloyd *et al.*, 2005; Townsend *et al.*, 2009) y células humanas (Porteus, 2003). Pese a presentar un diseño más sencillo que las meganucleasas, son necesarios conocimientos específicos de ingeniería de proteínas para adaptar los dominios dedos de zinc a cada secuencia diana, lo que restringe su aplicabilidad. Además, se describió un efecto citotóxico significativo derivado de modificaciones producidas en otras secuencias similares del genoma (*off targets*) que inducían muerte celular (H. J. Kim *et al.*, 2009). Sin embargo, el problema de la citotoxicidad se resolvió con el descubrimiento de las proteínas *TALE* (siglas en inglés de *transcription-activator-like effector*) del género *Xanthomonas*. Las proteínas *TALE* son un tipo de proteínas de unión al ADN con capacidad de asociarse a modificadores transcripcionales o endonucleasas (Boch *et al.*, 2009). Por ello, presentan un gran potencial para la edición génica. El dominio de unión al ADN de las *TALE* está formado por 33-35 aminoácidos que forman repeticiones en tándem. Cada dominio puede diferir en las posiciones 12 y 13, que determinan la unión a un nucleótido específico. Así es posible encontrar principalmente cuatro tipos de dominios de unión al ADN: NN, NG, HD y NI, que se unen preferentemente a los nucleótidos G/A, T, C y A respectivamente (Bhardwaj *et al.*, 2021). La unión de estos dominios con el dominio FokI generó una nueva clase de nucleasas programables, las *TALENs* (Miller *et al.*, 2011) (**Figura 1.5C**). Distintas combinaciones de los dominios de unión al ADN dirigen a las *TALEN* a localizaciones específicas. La estrategia más popular es emplear dos unidades *TALEN* que al dimerizar producen un *DSB* (Bhardwaj *et al.*, 2021). Otras variantes presentan distintas enzimas acopladas a las *TALE* en lugar de la enzima FokI, como las endonucleasas PvuII, I-TevI, I-Anil, I-Onul y MutH (Becker *et al.*, 2021). Todas ellas

presentan el inconveniente de que son específicas de secuencia. A pesar de su dificultad de diseño, las *TALEN* se siguen empleando en el desarrollo de modelos de investigación en numerosas especies, como hongos (Maruyama, 2021), plantas (Matres *et al.*, 2021), animales (Vats *et al.*, 2021) y células humanas (Kwon *et al.*, 2021). A pesar de los inconvenientes de estas tres nucleasas, su tecnología se ha mejorado y empleado en el desarrollo de terapias de diversas enfermedades, actualmente en fase de estudio en ensayos clínicos: Meganucleasas (NCT04030195, NCT04171843, NCT04649112 y NCT03666000); *ZFN* (NCT03041324, NCT02702115); y *TALEN* (NCT03190278, NCT02746952, NCT02808442, NCT04150497, NCT03226470, NCT04416984, ALLO-605, NCT04093596, NCT04142619, NCT03939026, NCT04696731).

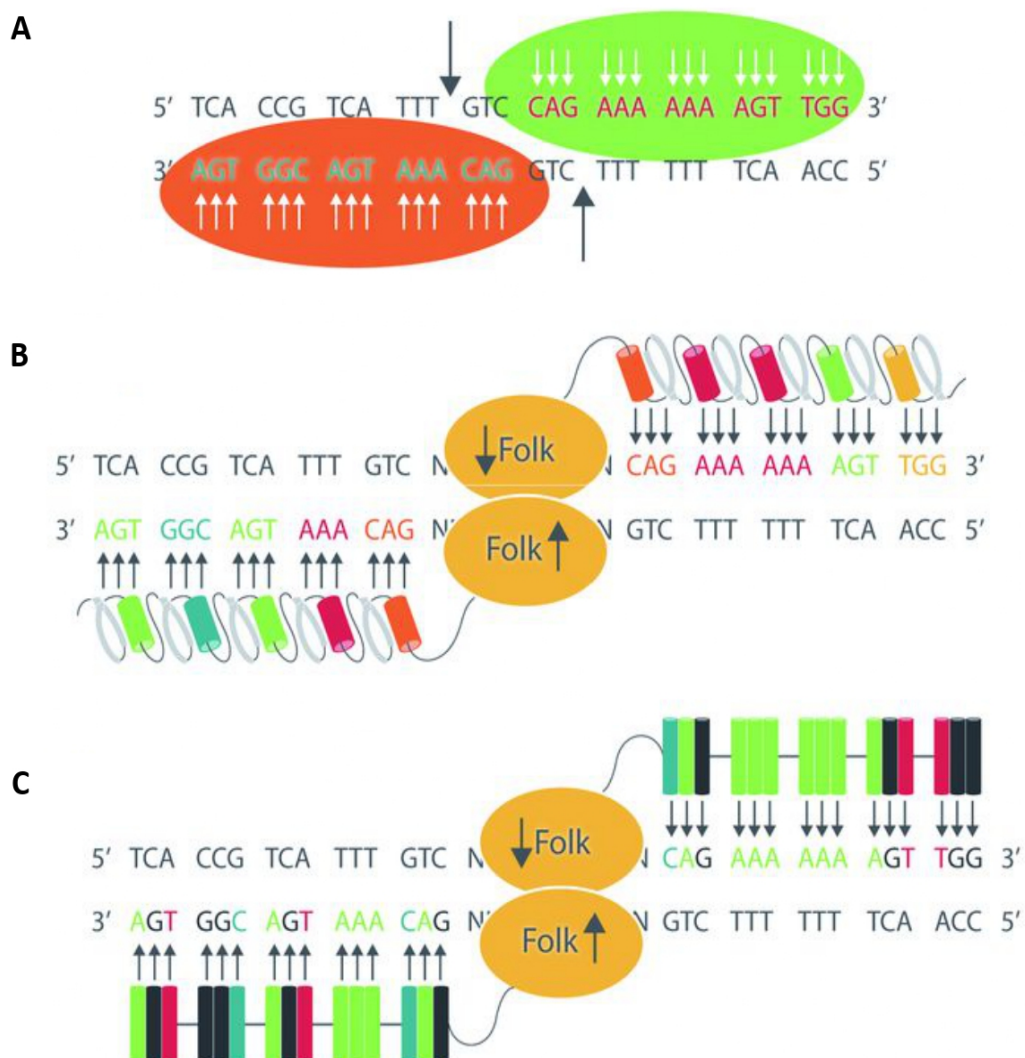


Figura 1.5: Nucleasas pre-CRISPR. Las flechas blancas indican las uniones de hidrógeno entre los nucleótidos y proteínas; B. Nucleasas Dedos de Zinc; C. *TALENs*. Las flechas negras pequeñas indican las uniones de hidrógeno entre los nucleótidos y proteínas. Las flechas negras grandes indican el corte de las nucleasas en el ADN (González-Romero *et al.*, 2019).

2.3. Sistema de edición CRISPR/Cas9

El sistema que ha revolucionado las estrategias de edición génica ha sido el sistema CRISPR/Cas9, de las siglas en inglés de *Clustered Regularly Interspaced Short Palindromic Repeats/CRISPR-associated protein 9*. Mientras que las anteriores estrategias se basan en la interacción ADN-proteína, el sistema CRISPR/Cas9 está basado en la interacción ADN-ARN mediante enlaces Watson-Crick. Aquí radica la diferencia fundamental entre los dos grupos de herramientas de edición génica, siendo el sistema CRISPR/Cas9 mucho más sencillo de diseñar y sintetizar.

Las secuencias CRISPR fueron detectadas por primera vez en la eubacteria *Escherichia coli* (Ishino *et al.*, 1987) y en la arqueobacteria *Haloferax mediterranei* (Mojica *et al.*, 1995). En estos organismos el sistema CRISPR está formado por un conjunto de repeticiones cortas de nucleótidos (protoespaciadores) alternadas por otras secuencias, denominadas espaciadores, que derivan de secuencias de bacteriófagos y plásmidos invasores. En estudios llevados a cabo por el grupo de F. Mojica se probó que los organismos que presentaban estos espaciadores son inmunes frente a los patógenos con los que presentan homología en esa secuencia, y se propuso que actuaban como un sistema inmunitario primario adaptativo (F.J.M. Mojica *et al.*, 2005; F. J. M. Mojica *et al.*, 2009)(**Figura 1.6**). En estudios posteriores se han descrito un gran número de microorganismos con distintos tipos de sistemas CRISPR (**Figura 1.7**).

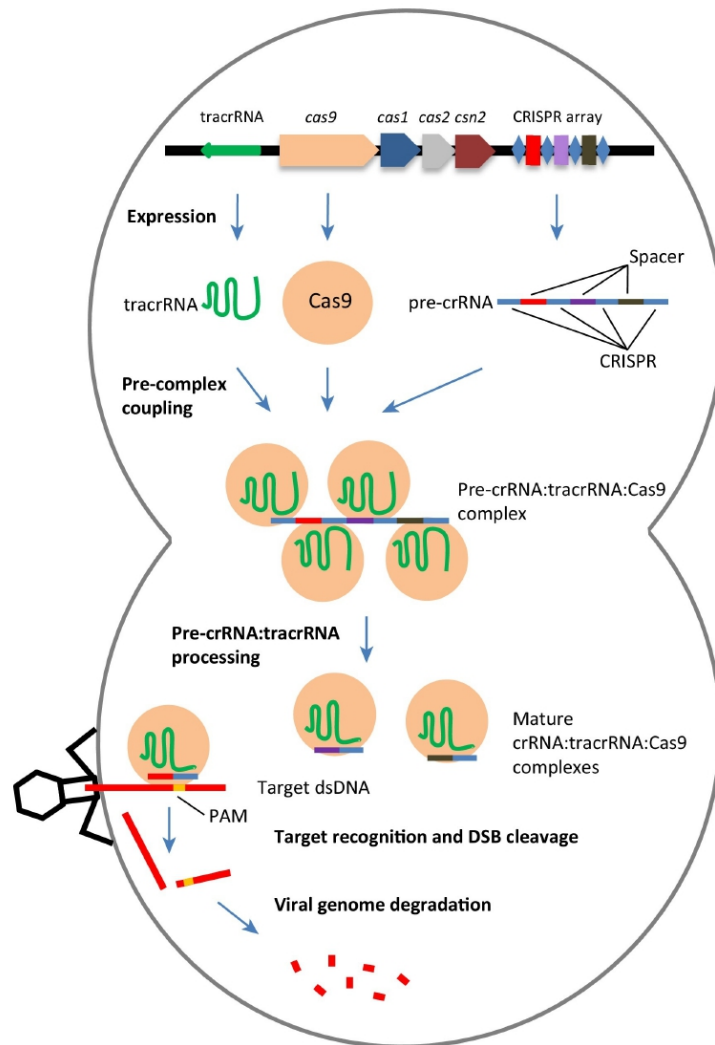


Figura 1.6. Sistema CRISPR/Cas en *Streptococcus pyogenes*. El locus CRISPR está compuesto por cuatro genes Cas, una región codificante de *tracrRNA* y un *array* de espaciadores CRISPR. Tras la transcripción de estos elementos, el *tracrRNA*, proteínas Cas9 y el precursor-*crRNA* (*pre-crRNA*), se ensamblan formando un pre-complejo. Una ARN III endonucleasa es capaz de cortar los *pre-crRNA*, dando lugar a complejos activos de *crRNA:tracrRNA:Cas9*. Cuando estos complejos se asocian a un ADN invasor con una secuencia PAM apropiada y homología a la secuencia espaciadora del *crRNA*, la Cas9 cataliza el *DSB*, produciendo la degradación de ese ADN (Mojica *et al.*, 2016).

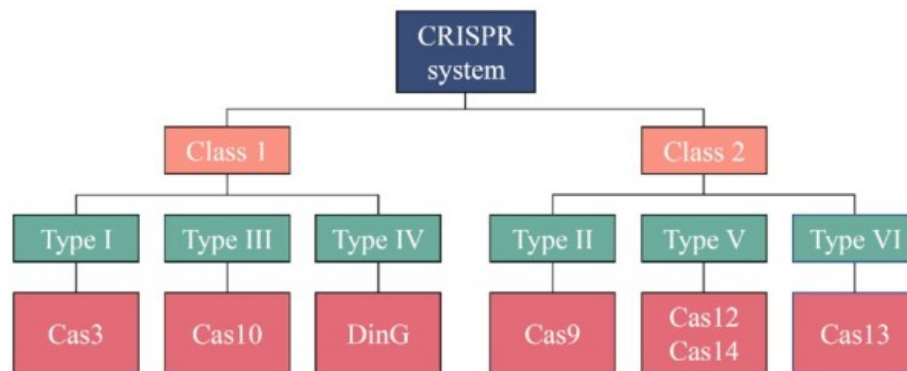


Figura 1.7: Clasificación del sistema CRISPR y de las nucleasas asociadas. Se trata de una clasificación funcional y estructural. El sistema se ha clasificado en clase 1 y 2 y estas a su vez en seis tipos. La principal diferencia entre las clases, radica en que la clase 2 emplea solo una proteína Cas, mientras la clase 1 presenta complejos multiproteicos. El sistema CRISPR clase I, incluye los tipos I, III y IV, que a su vez engloban las nucleasas Cas3, Cas10 y DinG. Por otro lado, la clase 2 presenta los tipos II, V y VI, donde se encuentran las proteínas Cas más aplicadas en la edición génica, la Cas9, Cas12, Cas14 y Cas13 (W. Liu *et al.*, 2021).

La expresión ectópica de los elementos del sistema CRISPR ha permitido llevar a cabo edición génica, tanto por *NHEJ* como *HDR*. En 2012 se desarrolló la primera aplicación del sistema CRISPR/Cas9 para la introducción de *DSB* en una localización específica en el ADN (Jinek *et al.*, 2012). A raíz de este descubrimiento, el sistema CRISPR/Cas9 ha producido una revolución en la investigación biológica. Las investigadoras Jennifer Doudna y Emmanuelle Charpentier fueron galardonadas en 2020 con el Premio Nobel de Química por su contribución en la adaptación del sistema natural CRISPR a la edición génica. Al contrario que los sistemas de edición anteriores, el sistema CRISPR/Cas9 se caracteriza por su sencillez. Los componentes esenciales que guían a la nucleasa Cas9 a su diana son dos elementos de ARN. El CRISPR *RNA* (*crRNA*) de 20 pares de bases con homología a la secuencia de interés y el *trans-activating crRNA* (*tracrRNA*). Estas dos moléculas se hibridan entre sí para formar el *guide RNA* (*gRNA*). Este último se hibrida a la nucleasa Cas9 y juntos migran a la localización genómica que presenta homología al *crRNA*. La región diana, además de presentar homología al *crRNA*, debe presentar una secuencia de nucleótidos específica denominada *protospacer adjacent motif sequence* (*PAM*) (Figura 1.8A). Esta secuencia es específica de cada nucleasa Cas, siendo la secuencia 5' NGG 3' la específica para la Cas9 aislada de *Streptococcus pyogenes*. Si la secuencia PAM se localiza aguas arriba de la secuencia homóloga, la Cas9 es capaz de producir un *DSB* gracias a sus dos dominios nucleasa (HNH y RuvC). El *DSB* producido con esta estrategia se ha empleado para producir la

inactivación de genes o la introducción de *indels* mediante la reparación por *NHEJ*, o la introducción de modificaciones específicas mediante *HDR* (Mojica *et al.*, 2016) (Figura 1.8B).

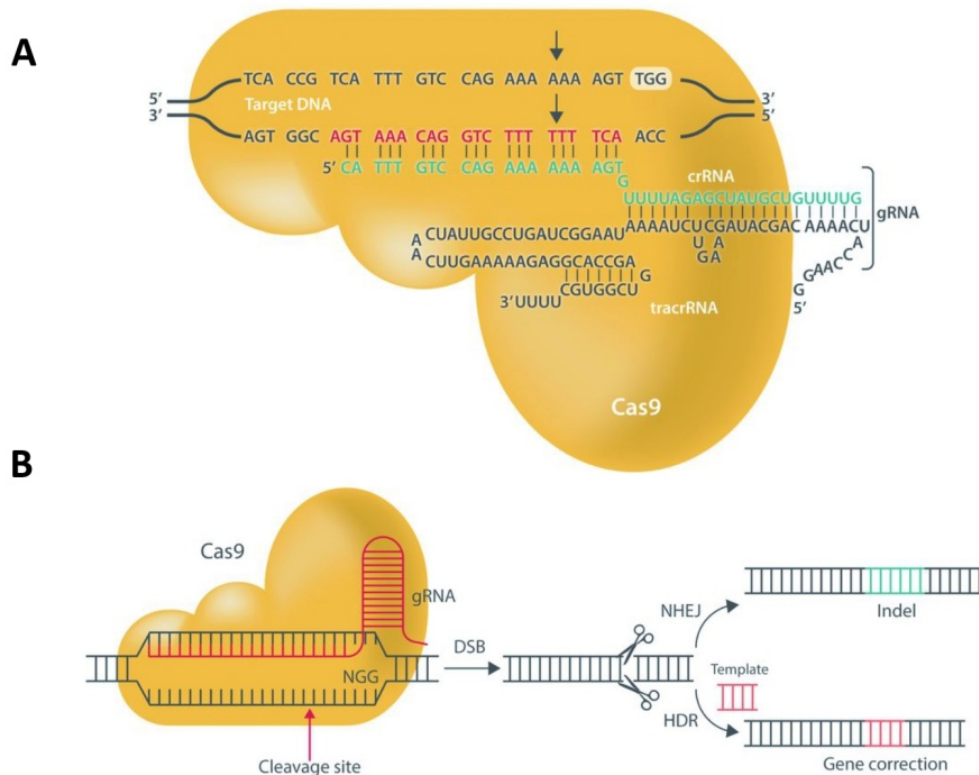


Figura 1.8. Sistema de edición génica CRISPR/Cas9. A Esquema de los elementos del sistema CRISPR/Cas9. Los elementos de ARN *crRNA* (nucleótidos verdes) y *tracrRNA* se hibridan gracias a una región homóloga formando el *gRNA*. La nucleasa Cas9 es capaz de unirse al *gRNA* a través del *tracrRNA* y juntos se dirigen a la diana en el ADN (nucleótidos rojos). La nucleasa separa las dos hélices del ADN y en presencia de la secuencia *PAM* 5'-NGG-3' (en este esquema representada por TGG) produce el *DSB*. Las flechas negras señalan los puntos aproximados de corte. B. Tras el corte de la Cas9, el *DSB* puede ser reparado por el sistema *NHEJ* o mediante *HDR*. El *NHEJ* es un sistema propenso a errores, que puede generar *indels*, modificando la pauta de lectura o introduciendo codones de *stop* prematuros e inactivando el gen. Por otro lado, en presencia de una molécula de ADN molde, el *DSB* puede ser reparado mediante la vía *HDR*, permitiendo introducir modificaciones específicas (González-Romero *et al.*, 2019).

Mediante modificaciones de la estructura de la nucleasa Cas9 se han derivado otros sistemas de edición. Mediante la inactivación de uno o los dos dominios nucleasa de la enzima surgieron la enzima nicasa y la enzima deadCas9 (dCas9) respectivamente. Mientras la primera enzima produce un solo corte en una de las hebras de ADN, la segunda solamente tiene la capacidad de unión al ADN. La Cas9 nicasa está indicada para introducir cambios puntuales (Trevino *et al.*, 2014) con mayor especificidad ya que es posible emplearlas en parejas, donde una nucleasa presenta homología con la cadena

con sentido y otra enfrentada en la cadena opuesta generarían un *DSB*. Esta estrategia emplea un reconocimiento de 20 más 20 pb lo que hace muy improbable los eventos de *off targeting*. Por otro lado, la dCas9 puede co-expresarse con activadores (CRISPRa) o represores (CRISPRi) de la expresión, siendo una herramienta muy útil en estudios de expresión génica. También se puede emplear para producir modificaciones epigenéticas acoplando la dCas9 a un modificador epigenético. Por otro lado, también se ha empleado para el marcaje de localizaciones específicas de la cromatina, acoplando la expresión de una proteína fluorescente, para inducir la interacción de regiones distantes de la cromatina e introducir modificaciones puntuales sin necesidad de producir un *DSB* (*base editing*) (Brezgin *et al.*, 2019).

Gracias a la simplicidad de la técnica, solamente son necesarios conocimientos básicos de biología molecular para diseñar un sistema CRISPR dirigido a la diana de interés. De esta forma el sistema CRISPR/Cas9 ya se ha empleado más allá del ámbito de la investigación biológica, dando lugar a numerosos fármacos para el tratamiento de diversas enfermedades, como infecciones bacterianas, cáncer, desórdenes hematológicos y enfermedades oculares (Hirakawa *et al.*, 2020). Por otro lado, esta estrategia presenta una serie de inconvenientes que se deben considerar con cautela, especialmente la mayor probabilidad que las anteriores nucleasas de producir *off targets*. La nucleasa Cas9 tolera ciertas diferencias entre el *crRNA* y el ADN, lo que hace más probable que se hibride en regiones genómicas similares y produzca *DSBs* indeseados. Estos cortes pueden afectar a genes esenciales, como *TP53*, cuya inhibición puede acarrear un aumento de la predisposición a transformación tumoral.

No es posible evitar totalmente los efectos *off targets*, pero la selección de una determinada secuencia diana y del *gRNA* puede disminuir el número de *off targets* y aumentar la eficiencia de edición. Para esta tarea se han desarrollado un gran número de herramientas *online* de predicción de los potenciales *off targets*. Estas herramientas permiten descartar secuencias *gRNA* con un número alto de secuencias similares y seleccionar aquellas con menos *off-targets*. Si a pesar de ello es necesario disminuir el número de *off targets*, se ha probado que *gRNAs* de secuencia más corta o truncados aumentan la especificidad del complejo *gRNA*:Cas9. Otras modificaciones como la adición de dos nucleótidos guanina o de una estructura en horquilla en el extremo 5' de

la secuencia homóloga también aumenta la especificidad, reduciendo la actividad *off-targets* (Vicente *et al.*, 2021). Con el mismo objetivo se han desarrollado distintas variantes de la nucleasa Cas9 más específicas conocidas como *Cas9 High Fidelity*. Presentan modificaciones que reducen la afinidad de la Cas9 con secuencias de ADN no totalmente homólogas (Kleinstiver *et al.*, 2016; Vakulskas *et al.*, 2018). Recientemente se han desarrollado estrategias para evitar la unión de la nucleasa a regiones inespecíficas. Coelho y colaboradores han desarrollado un sistema de reducción de *off-targets* denominado *CRISPR Guide RNA Assisted Reduction of Damage (CRISPR GUARD)*. Esta estrategia consiste en bloquear la unión del *sgRNA* con zonas inespecíficas a través de la competición con complejos inactivos de Cas9. El complejo inactivo se genera empleando secuencias guía complementarias al sitio *off-target* de menos de 16 nucleótidos. Esta longitud le permite unirse a la Cas9 pero no activa la actividad nucleasa (Coelho *et al.*, 2020). Por otro lado, también se está implementando el uso de proteínas anti-CRISPR derivadas de fagos con capacidad de evasión del sistema CRISPR en bacterias. Estas proteínas actúan inhibiendo los procesos de unión al ADN, unión del *crRNA* con la Cas9 e inhibiendo la producción de *DSB* (Aquino-Jarquín, 2021).

A pesar de las mejoras desarrolladas para garantizar la seguridad de la edición con CRISPR, las implicaciones éticas de sus aplicaciones en células somáticas son serias, pero aún son mayores en gametos y embriones. Aunque muchos países mantienen una política estricta de no realizar modificaciones en la línea germinal humana por estos motivos, al menos por ahora es posible que se haya llevado a término la gestación de dos niñas modificadas genéticamente en China (Cyranoski, 2019). Por ello, las principales figuras implicadas en el desarrollo de esta técnica de edición han firmado una moratoria global para retrasar la edición génica en la línea germinal humana (Lander *et al.*, 2019).

Otra de las consideraciones a tener en cuenta para llevar esta técnica a la práctica clínica es la eficiencia. A pesar del aumento de eficiencia en comparación con otras técnicas de edición, la edición mediante *HDR* sigue siendo subóptima. Actualmente se están desarrollando diferentes estrategias para su mejora *in vitro*, que deben ser testados antes de ser trasladados a la clínica. Por otro lado, se ha observado la presencia de anticuerpos y linfocitos T en humanos adultos y neonatos contra la

nucleasa Cas9. Estos anticuerpos son principalmente contra las nucleasas Cas9 aisladas de *Staphylococcus aureus* y *Streptococcus pyogenes* (Charlesworth *et al.*, 2019). Se cree que una infección previa de estas bacterias habría generado la respuesta inmune, y esto ha despertado la alarma sobre una posible reacción inmunogénica tras la administración de los reactivos de edición génica, en un hipotético intento de terapia con CRISPR. Entre las estrategias para minimizar o evitar esta reacción puede ser el uso de otras nucleasas CRISPR, el desarrollo de predictores *in silico* de predisposición inmunogénica, la comprensión de la respuesta inmune innata contra Cas9 con el objetivo de mejorar el diseño de los vectores empleados, identificar las regiones antigénicas de la nucleasa para desarrollar una desinmunización y, por último, aplicar inmunosupresores para reducir ese rechazo (González-Romero *et al.*, 2019).

2.4. Otras nucleasas CRISPR

Por otro lado, a pesar de que la Cas9 aislada de *S. pyogenes* es la más empleada, también se han aislado otras nucleasas que se han incluido en la rutina de la edición génica. La nucleasa Cas12a o Cpf1 se aisló del género *Acidaminococcus* y de la especie *Lachnospiraceae bacterium*. En esta nucleasa, el *crRNA* no necesita de la molécula *tracrRNA* para unirse a Cpf1, reconoce una secuencia PAM rica en timina que varía dependiendo del organismo, y corta en puntos distintos de las dos hebras de ADN formando extremos romos con 4 o 5 nucleótidos en el extremo 5' (Zetsche *et al.*, 2015). La nucleasa Cas13a o C2c2, al contrario que las anteriores, presenta actividad ARNasa y solamente necesita del elemento *crRNA* para unirse a la secuencia diana y algunos de esta nucleasa no necesitan de una secuencia PAM para producir el corte (Seohyun Kim *et al.*, 2021). Una vez unido a su *single-stranded RNA (ssRNA)* homólogo, la Cas13a localiza la diana de ARN y produce el corte. Tras este evento se produce una activación de corte en *trans* de *ssRNA* no específicos, o lo que es lo mismo, es capaz de cortar las moléculas de ARN que estén a su alrededor de forma inespecífica (Granados-Riveron *et al.*, 2018). Recientemente se ha descubierto la nucleasa Cas14a en el superfilum de arqueas extremófilas DAPNN (Harrington *et al.*, 2018). Presenta una serie de características que la diferencian de las anteriores nucleasas al ser la nucleasa CRISPR más pequeña descrita hasta la fecha y no necesita de una secuencia PAM para realizar el corte. Actúa sobre *single-stranded DNA (ssDNA)* y tras esto adquiere actividad de

corte en *trans* de *ssDNA* no específico (Aquino-Jarquin, 2019). Tanto la Cas13a como la Cas14a se han empleado en el desarrollo de plataformas de detección llamadas SHERLOCK (*Specific High-Sensitivity Enzymatic Reporter UnLOCKing*)(Gootenberg *et al.*, 2017) y Cas14-DETECTR (*DNA Endonuclease Targeted CRISPR Trans Reporter*) (Harrington *et al.*, 2018). Estas estrategias incluyen un primer paso de reconocimiento y corte de la secuencia diana y, posteriormente, un corte colateral inespecífico de *ssRNA* o *ssDNA* unidos a marcadores fluorescentes, produciendo señales fluorescentes cuantificables.

Por último, en septiembre de 2021 se ha descrito una nueva nucleasa, la Cas7-11. El equipo de Gootenberg y Abudayyeh descubrieron esta nueva nucleasa tras estudiar casi 11 billones de proteínas bacterianas *in silico*. Esta nucleasa está formada por la fusión del dominio putativo de Cas11 con múltiples subunidades Cas7. Los autores han programado la nucleasa Cas7-11 de *Desulfonema ishimotonii* (DiCas7-11) para inducir el corte de ARN en *Escherichia coli* y células de mamífero, observando una gran eficiencia sin detección de *off targets*. Esto es una ventaja frente a la nucleasa Cas13a, que en las mismas condiciones, produce una alta toxicidad en las células (Özcan *et al.*, 2021).

2.5. Edición génica en hematología

En el campo hematológico, el sistema de edición CRISPR/Cas9 se ha aplicado tanto en investigación básica como en estudios clínicos traslacionales (**Figura 1.9**). La propia naturaleza de las células hematopoyéticas permite abordar fácilmente estudios de edición génica con distintos objetivos. Su fácil extracción a partir de muestras de pacientes, su crecimiento en suspensión y fácil expansión han permitido la aplicación del sistema CRISPR/Cas9 en el estudio funcional de alteraciones concretas en líneas celulares, el desarrollo de modelos de investigación y el descubrimiento de nuevas dianas terapéuticas y fármacos.

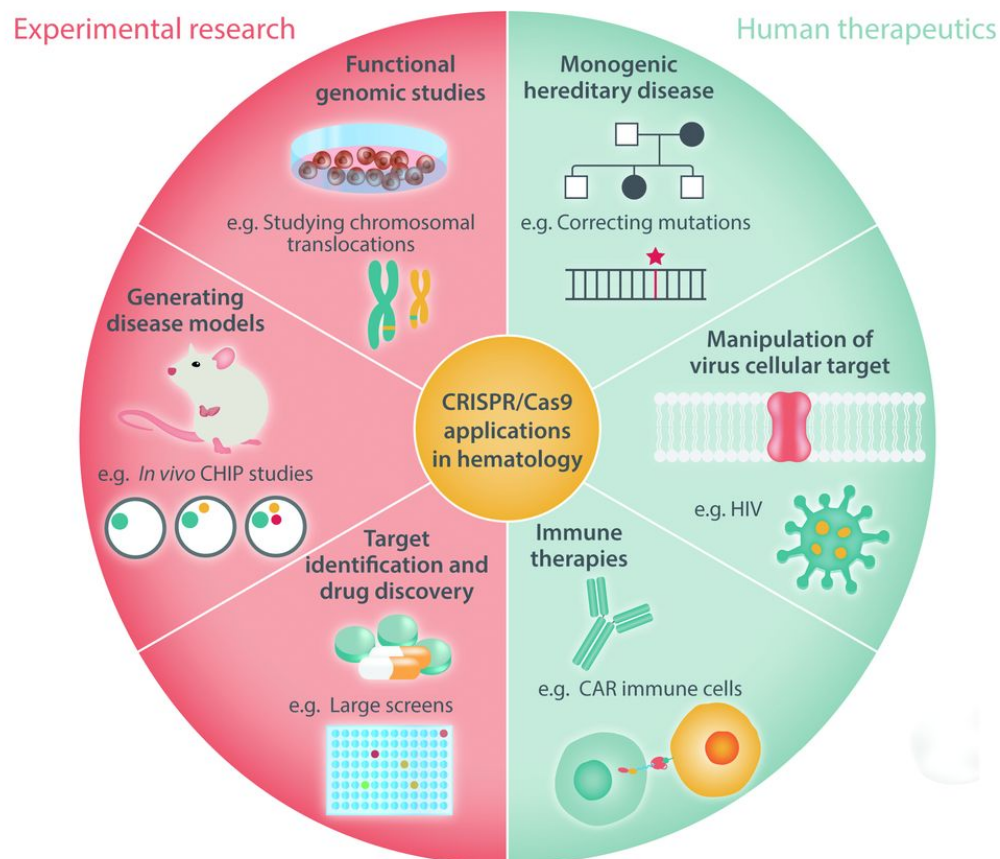


Figura 1.9. Aplicaciones del sistema CRISPR/Cas9 en la investigación y desarrollo de terapias en hematología. Esta técnica de edición génica ha permitido el desarrollo de numerosas investigaciones básicas, que han permitido esclarecer los efectos funcionales de determinadas alteraciones, y generar modelos de investigación e identificar nuevos fármacos. Por otro lado, este sistema de edición se está aplicando directamente en la corrección de mutaciones causales de enfermedades hereditarias monogénicas, alteración de las dianas celulares de virus para evitar la propagación de infecciones y el desarrollo de terapias inmunológicas (González-Romero *et al.*, 2019).

Muchos estudios han probado la capacidad del sistema CRISPR/Cas9 para introducir alteraciones concretas en las células hematopoyéticas, ayudando al estudio de su papel en el desarrollo de la enfermedad. Por ejemplo, la inactivación del gen *ASXL1* mediante el mecanismo *NHEJ* en células U937 sirvió para asociar esta modificación con defectos en la diferenciación de monocitos y macrófagos y la alteración de expresión de genes implicados en varias vías celulares asociadas a Síndrome Mielodisplásico (SMD) (Z. J. Wu *et al.*, 2018). Por otro lado, mediante el uso de un molde de ADN con la secuencia correcta y el mecanismo de *HDR* se han generado modelos de investigación con las mutaciones encontradas en los pacientes, como la mutación P95H en el gen *SRSF2* (J. Zhang *et al.*, 2015). Por último, realizando varios cortes en el ADN ha sido posible reproducir algunos reordenamientos, como t(8;21)/*RUNX1T1* y t(4,11)/*KMT2A-*

AFF1/AFF1-KMT2A en líneas celulares y células primarias (Castaño *et al.*, 2016; Torres *et al.*, 2014). Estos ejemplos muestran como este sistema de edición permite desarrollar de forma precisa modelos celulares que ayudan a relacionar fenotipos de la enfermedad con mutaciones genéticas concretas.

Mediante el sistema de *HDR* y el uso de un ADN molde también se han corregido mutaciones asociadas a enfermedad. Por ejemplo, se empleó esta técnica para corregir una mutación de pérdida de función en el gen *ASXL1* en una línea celular de leucemia mieloide crónica (Valletta *et al.*, 2015). Además de líneas celulares, muchos investigadores han aplicado la edición génica en células primarias ya que resultan una aproximación más directa a la terapia. Las células humanas pluripotentes inducidas (*iPSC*) son un claro ejemplo, ya que es posible aislar células hematopoyéticas de pacientes, reprogramarlas a *iPSC*, corregirlas mediante edición génica, diferenciadas a células hematológicas inmaduras y reintroducidas en el mismo individuo. Algunos ejemplos de esta aproximación sería la corrección del gen *HAX1* en células *iPSC* aisladas de un paciente con neutropenia congénita severa (Pittermann *et al.*, 2017). A pesar de los grandes avances producidos gracias a la tecnología CRISPR/Cas9 existen limitantes para poder desarrollar de forma óptima modelos de investigación en LMA. En el ámbito de las líneas celulares, la transfección de los elementos CRISPR es muy ineficiente debido a las características de las células, por lo que normalmente se emplean vectores víricos. Con el objetivo de mejorar la seguridad y la eficacia de la técnica es necesario desarrollar nuevos formatos de introducción de los componentes CRISPR en este tipo de células.

El sistema CRISPR también se ha empleado *in vivo* para desarrollar modelos animales de desórdenes hematológicos. En ratón por ejemplo, se han desarrollado modelos germinales de mutaciones en el factor de *splicing ZRSR1* (Horiuchi *et al.*, 2018) y en los factores F8 y F9 de la hemofilia A y B (C.-T. Yen *et al.*, 2016), entre otros. Además ha sido posible un gran número de genes a la vez (Rathe *et al.*, 2015). Estos experimentos realizados tanto *in vitro* como *in vivo* dan gran cantidad de información acerca de la evolución clonal, la complejidad genómica, patrones de concurrencia y nuevas resistencias a tratamientos (Zheng *et al.*, 2021).

Por último se están desarrollando numerosos ensayos clínicos basados en la técnica CRISPR/Cas9 para diversas enfermedades hematológicas, como β -talasemia

(NCT03655678, NCT03728322), anemia de células falciformes (NCT03745287, NCT04443907, NCT04474536, NCT04819841, NCT04853576), infección por el virus de la inmunodeficiencia humana (VIH) (NCT03164135), LMA (NCT04037566, NCT04614636, NCT04849910), enfermedades de células B (NCT03398967, NCT03166878, NCT04035434), leucemia linfocítica aguda (NCT04557436, NCT04154709) y linfoma Non-Hodgking (NCT04637763). Es de especial interés la mejora de las terapias basadas en células T con expresión de receptores quiméricos de antígenos específicos de virus y de células tumorales. Por citar algunos ejemplos, se ha probado que la inactivación del gen codificante de la proteína 1 de muerte celular programada (*PD1*) aumenta la eficiencia del tratamiento (Xia *et al.*, 2019). Por otro lado, se ha probado como la eliminación del gen codificante del receptor de glucocorticoides en células T específicas de virus aumenta la eficiencia de esta terapia en pacientes inmunodeprimidos que en paralelo son tratados con esteroides (Kaeuferle *et al.*, 2020).

3. *Caenorhabditis elegans* como modelo de estudio de cáncer

El gusano *Caenorhabditis elegans* se ha empleado ampliamente como organismo modelo para estudiar la enfermedad humana (Apfeld *et al.*, 2018; Culetto, 2000; Markaki *et al.*, 2010). Este animal invertebrado presenta una serie de características que han permitido su uso en estudios genéticos sobre el cáncer. *C. elegans* es un pequeño nematodo que de adulto puede alcanzar 1 mm de longitud. Su mantenimiento en el laboratorio es muy sencillo, siendo posible cultivar cientos de gusanos en placas Petri con la bacteria *Escherichia coli* como fuente de alimento. Su ciclo de vida es corto ya que los huevos pueden eclosionar y desarrollar gusanos adultos en solamente tres días. Presenta dos sexos, machos y hermafroditas, lo que es una gran ventaja en la manipulación genética ya que permite tanto realizar auto fecundación como fecundación cruzada. Además, fue el primer organismo multicelular con todo su genoma secuenciado. Los principales descubrimientos realizados en este modelo han sido el ARN de interferencia, la vía genética de la apoptosis y el uso de la proteína verde fluorescente como marcador biológico. Gracias a ello, el gusano es el organismo multicelular mejor caracterizado, empleándose principalmente en investigación sobre la biología del desarrollo y neurobiología.

La aplicación de *C. elegans* como organismo modelo ha aportado conocimientos importantes sobre procesos que se presentan alterados en cáncer, como progresión del ciclo celular, invasión y migración celular, señales de factores de crecimiento, apoptosis e inestabilidad del genoma, entre otros (Shahzad *et al.*, 2021). La contribución del gusano en la investigación del cáncer se basa en tres pilares. Primero, el 45-75 % de los genes humanos asociados a patologías presentan ortólogos en *C. elegans* (Silverman *et al.*, 2009). Por ello, es posible caracterizar genes oncogénicos y supresores de tumores humanos mediante el estudio de sus ortólogos y sus vías de actuación en el gusano. Por ejemplo, el estudio de *C. elegans* permitió el descubrimiento de los micro *RNAs* (Grad *et al.*, 2003; Lim *et al.*, 2003), y en concreto, la identificación de la familia de *let-7* (Reinhart *et al.*, 2000). En el gusano, su expresión regula la migración celular en el adulto y es la familia de miARNs más altamente conservada, hasta en humanos, donde actúa como supresores de tumores inhibiendo el crecimiento del tumor y la metástasis (Ma *et al.*, 2021). Otro caso similar es *lin-4*, el primer miARNs descrito en *C. elegans* (Lee *et al.*, 1993). Su ortólogo en humanos, miR-125b, es de los miARNs más importantes por su regulación de varias vías del desarrollo y progresión de varios tipos de cáncer (Peng *et al.*, 2021). La alta conservación de los factores de *splicing* en la evolución también ha hecho posible desarrollar modelos de mutaciones implicadas en SMD y melanoma uveal. Al contrario que en modelos de ratón, en *C. elegans* fue viable la introducción en homocigosis de mutaciones *missense* en el ortólogo del gen *SF3B1*, *sftb-1* (Serrat *et al.*, 2019). Además, fue posible recrear patrones de concurrencia con otros factores de *splicing* mutados que normalmente son mutuamente excluyentes, permitiendo esclarecer las interacciones entre mutaciones que en otros modelos no son viables (Serrat *et al.*, 2019).

En segundo lugar, a pesar de que no se ha podido observar el desarrollo de tumores en *C. elegans*, es posible observar variaciones de procesos bien caracterizados, como alteraciones del desarrollo y diferenciación celular. Esto puede ayudar a la identificación de nuevas vías y genes involucrados en el desarrollo del cáncer. Esto se ha aplicado en el campo de los tumores cerebrales (Shahzad *et al.*, 2021). En este punto, los modelos murinos han sido un instrumento base para esclarecer la biología de los tumores cerebrales humanos, pero presentan algunas limitaciones, como la

dependencia de trabajo con ratones inmunodeprimidos, complicados implantes de tumores humanos o líneas celulares en el ratón, además de los altos costes de mantenimiento de las colonias, protocolos experimentales tediosos y largos procesos para obtener los certificados éticos. Por el contrario, *C. elegans* es un organismo más simple y genéticamente más fácil de modificar y de mantener. A pesar de que el gusano no presenta tumores cerebrales, ha sido posible estudiar las vías moleculares alteradas. En glioblastoma, en un 40-50 % de los casos se da la inactivación del gen *PTEN*, y en gliomas se activa la vía de receptor tirosina quinasa/Ras/PI3K en un 88 %. En gusano, el aumento de la señal de Ras produce una alteración de la formación de la vulva, generando un fenotipo multivulva hiperproliferativo (Muv) (Shahzad *et al.*, 2021). Gracias a la conservación evolutiva se ha generado un modelo de glioma con la activación del gen ortólogo a *KRAS* (*let-60*) y la inactivación del ortólogo de *PTEN* (*daf-18*), que presenta el fenotipo Muv. Este modelo ofrece una alternativa para el estudio de las sinergias entre las vías de señalización comúnmente alteradas en gliomas (Shahzad *et al.*, 2021).

La meiosis es otra de las vías que está muy caracterizada en *C. elegans*. La homología de los genes supresores de tumores *BRCA1* y *BARD1* con los ortólogos en gusano *brc-1* y *brd-1* ha permitido el estudio de su implicación en la reparación del ADN durante la meiosis (Li *et al.*, 2021). Gusanos con la inactivación de estos genes mostraron alta sensibilidad a la irradiación y mayor número de machos en la descendencia tras auto fecundación. Se concluyó que mientras el gen *brc-1* es prescindible para la inactivación del cromosoma sexual durante la espermatogénesis, *brd-1* juega un papel importante en la reparación del ADN entre secuencias heterólogas en la oogénesis y estabiliza RAD-51. En estudios recientes se relacionaron las mutaciones de estos genes con un aumento de *indels* tras múltiples generaciones. En estos casos la vía de reparación más empleada resultó ser la *theta-mediated end joining* (*TMEJ*) en lugar de *NHEJ*, produciendo una reparación ineficiente. Estos perfiles mutacionales resultaron ser similares a los encontrados en células tumorales con deficiencia de *BRCA1*, sugiriendo que la reparación por *TMEJ* puede contribuir a la carcinogénesis (Li *et al.*, 2021).

Por último, el gusano es una potencial plataforma de testado de fármacos. Se trata de un sistema económico y rápido de descubrimiento de fármacos ya que permite

realizar pruebas de validación sobre gusanos vivos y visualizar cambios en la morfología y/o la locomoción. Por ejemplo, se ha empleado para investigar la acción de nuevos inhibidores de la síntesis de ácidos grasos. La sobreexpresión de la sintasa de ácidos grasos se ha asociado a varios tipos de cáncer, como próstata, ovario, pecho, colorrectal y de pulmón. En *C. elegans* se confirmó el efecto del fármaco PFI09 en la inhibición de la síntesis de partículas lipídicas (Qu *et al.*, 2021). Estos son algunos de los numerosos ejemplos de la aplicabilidad de *C. elegans* en el estudio de las vías moleculares alteradas en cáncer. Las tecnologías de cribado masivo de fármacos se han adaptado al gusano, permitiendo monitorizar el efecto de diferentes compuestos en placas individuales, de 98, 384 o 1536 pocillos. De esta forma se pudieron testar 14,000 compuestos, de los que 300 indujeron cambios de fenotipo (Silverman *et al.*, 2009).

Por otro lado, la entrada en escena del sistema CRISPR/Cas9 ha supuesto un gran avance en el desarrollo de modelos en *C. elegans*. Los elementos esenciales de este sistema (*gRNA*, *tracrRNA* y Cas9) pueden introducirse en el gusano en forma de plásmidos de ADN, ARN mensajero o complejos ribonucleoproteicos (*RNPs*). La forma seleccionada es inyectada en la gónada de un gusano hermafrodita, produciendo los cambios en la descendencia. Se han ideado distintas metodologías para seleccionar los animales editados de forma rápida y fácil o de obtener una población enriquecida en gusanos editados, con el objetivo de agilizar la selección de animales modificados. Un marcador muy empleado es el plásmido pRF4, que expresa el alelo dominante *rol-6(su1006)* (Peixoto *et al.*, 1998), causando que los animales se desplacen en movimientos circulares similares a un sacacorchos que se puede visualizar fácilmente con un microscopio de disección. En los casos en los que se emplea un molde para introducir una mutación puntual, la estrategia co-CRISPR presenta grandes ventajas. Esta radica en editar simultáneamente el gen de interés y un gen marcador, obteniendo un fenotipo visible en los gusanos editados y un marcador de la actividad nucleasa (H. Kim *et al.*, 2014). La aplicación del sistema CRISPR/Cas9 en *C. elegans* ha permitido reducir enormemente los tiempos invertidos en la generación de modelos.

Por todo lo expuesto anteriormente, *C. elegans* presenta un gran potencial para desarrollar estudios sobre los procesos moleculares involucrados en la LMA. Por medio de la tecnología CRISPR/Cas9 es posible desarrollar modelos con las mutaciones

encontradas en los pacientes y recrear patrones de concurrencia que arrojen información sobre las causas de las diferencias clínicas y de respuesta a tratamientos observada en pacientes. Esto es aplicable a las mutaciones más frecuentes en el gen *IDH2*. El CAT es una vía altamente conservada debido a su papel esencial en el metabolismo energético celular y con ello, también sus enzimas intermediarias, como IDH2. Por ello, las mutaciones de este gen encontradas en pacientes pueden tener efectos moleculares similares en *C. elegans*. Siguiendo esta estrategia, es posible estudiar las diferencias moleculares entre las mutaciones R140 y R172 que expliquen su diferente estratificación en los subgrupos de LMA.

HIPÓTESIS Y OBJETIVOS

HIPÓTESIS Y OBJETIVOS

HIPÓTESIS

Gracias al desarrollo de las nuevas tecnologías de secuenciación masiva, las mutaciones del gen *IDH2* se han relacionado con características clínicas y de pronóstico específicas en pacientes con leucemia mieloide aguda. Sin embargo, se desconoce por qué las mutaciones más frecuentes en este gen (R140 y R172) se asocian a subclases distintas de LMA, a pesar de producir el mismo efecto en la proteína. Para ello, la aplicación del novedoso sistema de edición CRISPR/Cas9, que ya se ha empleado para modificar el genoma de numerosos organismos, permitiría generar modelos específicos de estas mutaciones. Respecto a los modelos celulares, es sabido que las líneas celulares leucémicas presentan grandes dificultades de transfección y, por tanto, es necesario idear nuevas metodologías para editar de forma eficiente estas células. En cuanto a los modelos animales, el organismo modelo *Caenorhabditis elegans* presenta numerosas ventajas de manipulación y técnicas específicas para producir nuevas cepas, portadoras de transgenes y/o mutaciones generadas a voluntad. Todo ello, sumado a la alta conservación evolutiva entre los genes *IDH2* humano y su ortólogo en *C. elegans*, hace que sea un potencial modelo de estudio de las mutaciones en *IDH2*.

La hipótesis de este trabajo es que es posible desarrollar nuevas estrategias de edición génica que permitan producir de forma rápida y eficiente modelos *in vitro* e *in vivo* de mutaciones implicadas en la LMA. Con este fin nos planteamos, por un lado, aumentar la eficiencia de transfección y edición en las células leucémicas empleando los elementos del sistema CRISPR en su mínima expresión, como pequeñas construcciones de ADN y, por otro lado, generar cepas mutantes de *C. elegans* en las que poder reproducir los efectos moleculares observados en pacientes.

OBJETIVOS

El objetivo global de esa Tesis Doctoral es la aplicación del sistema CRISPR/Cas9 para desarrollar modelos de investigación *in vitro* e *in vivo* de las mutaciones más frecuentes del gen *IDH2* asociadas a leucemia mieloide aguda. Para ello, los objetivos específicos son los siguientes:

- i. Desarrollo de una metodología alternativa de introducción de los elementos CRISPR en células leucémicas.
- ii. Comparación de los resultados obtenidos mediante la nueva metodología con otras estrategias de introducción de los elementos CRISPR.
- iii. Desarrollo de cepas de *Caenorhabditis elegans* con las mutaciones más frecuentes del gen *IDH2*.
- iv. Estudio de los efectos funcionales de las mutaciones en *IDH2* en las nuevas cepas generadas.

MATERIAL Y MÉTODOS

MATERIAL Y MÉTODOS

Capítulo I. Edición de genes por CRISPR *in vitro*

1. Diseño de las secuencias guías de ARN

El objetivo de editar el gen *MYBL2* fue la optimización del corte de la nucleasa Cas9. Para ello, se escogió una región del intrón 3 del gen para diseñar la secuencia diana. Se empleó la secuencia disponible del gen en la base de datos del *National Center for Biotechnology Information* (NCBI) y se diseñó la secuencia guía manualmente, con una longitud de 18 nucleótidos.

Para realizar el diseño de las secuencias guías de *IDH2* se estudió mediante secuenciación Sanger la presencia de polimorfismos de nucleótido único (*SNP*) en las líneas celulares NB4, HL60, MOLM-13 y Katsumi-1. Esto se hizo porque cualquier cambio presente, que no esté descrito en la base del genoma humano, podría inhibir la hibridación del molde empleado para introducir la mutación R172. Para ello se amplificó la región de interés del exón 4 del gen *IDH2* mediante la enzima Taq ADN polimerasa *PCR Master Mix 2x* (Thermo Fisher Scientific, Waltham, Massachusetts, U.S, K0172) y *primers* específicos. Una vez confirmada la ausencia de *SNPs*, se emplearon distintos predictores web para el diseño de las secuencias guía <http://chopchop.cbu.uib.no/> (Labun *et al.*, 2019) y <http://www.e-crisp.org/E-CRISP/> (Heigwer *et al.*, 2014). Cada herramienta utiliza algoritmos propios para escoger los guías en función de su eficiencia y el número de posibles *off targets*. Se seleccionaron los ocho guías comunes en las dos listas producidas (**Figura 2.1**) (**Anexo I, Tabla 2.1**).

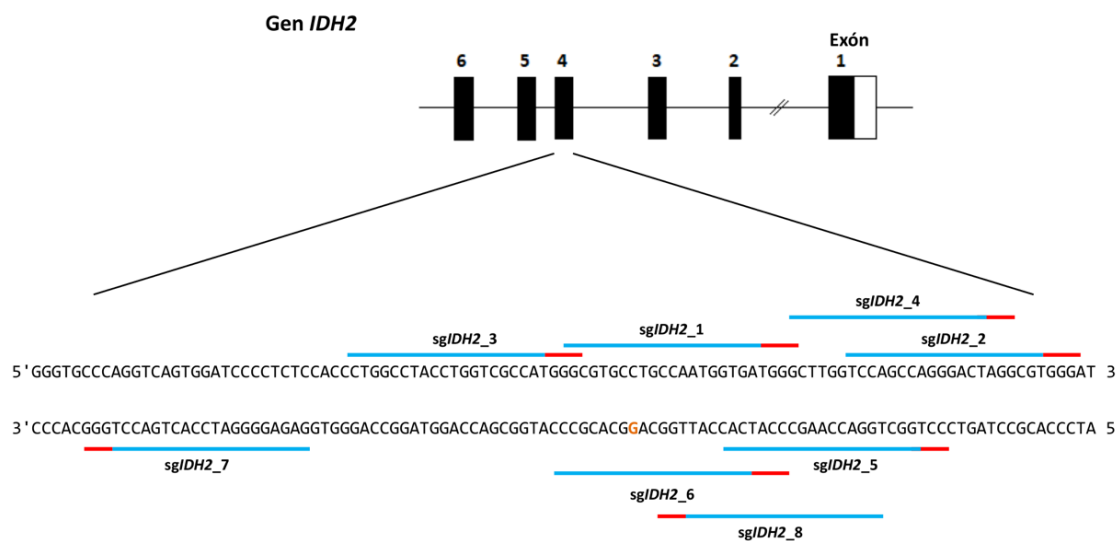


Figura 2.1. Localización de las secuencias guías en el exón 4 del gen *IDH2*. Cada línea señala la secuencia diana, donde la región azul indica la secuencia que se utilizó como guía y en rojo la secuencia PAM. En naranja se resalta la posición del nucleótido que sufre la mutación G>A. Se consideraron aquellos guías lo más próximos posible a la posición de la mutación.

2. Generación de un vector CRISPR para editar el gen *MYBL2*

Para realizar esto empleamos el vector pX458 (Addgene, Watertown, Massachusetts, USA, #48138) (Ran et al., 2013). Este vector presenta el gen de la Nucleasa Cas9 del organismo *Streptococcus piogenes* con sus codones humanizados (hSpCas9), sitio de clonación para una secuencia guía y el gen codificante de la proteína verde fluorescente (*EGFP*) (**Figura 2.2**). pX458 es un vector ampliamente utilizado en los experimentos *in vitro* (Ran et al., 2013). La expresión de la hSpCas9 se encuentra controlada por el *enhancer* del citomegalovirus humano (CMV) y el promotor β -actina de pollo. Ambas secuencias promueven una fuerte expresión de la nucleasa de forma constitutiva en una gran variedad de líneas celulares de mamífero (Hitoshi et al., 1991). La secuencia hSpCas9 se encuentra flanqueada por secuencias de localización nuclear (NLS). Estas secuencias son reconocidas por proteínas carioferinas, que permiten su transporte al núcleo tras la traducción. También presenta tres secuencias de epítomos Flag que permiten su detección mediante anticuerpos específicos. El gen codificante de hSpCas9 y el de la proteína *GFP* se encuentran acoplados transcripcionalmente. Entre ambos se encuentra la secuencia del péptido T2A, que permite tras la traducción, que la nucleasa Cas9 y *GFP* sean dos proteínas independientes. Por otro lado, la expresión de *crRNA* y *tracrRNA* se encuentran bajo el control del promotor U6 (pU6). Se trata de un promotor específico de la ARN polimerasa III y es activo en cualquier tipo celular.

Permite la transcripción eficiente de secuencias de ARN pequeñas y muy acotada a la secuencia requerida. Como requisito de este promotor, para que se produzca la transcripción la secuencia guía debe presentar una guanina en su primera posición. (Orioli *et al.*, 2012). Por último, este vector presenta el gen de resistencia al antibiótico ampicilina.

Para clonar el guía de *MYBL2* en este vector, la secuencia guía fue sintetizada como *primers* antiparalelos (Integrate DNA Technologies, Coralville, Iowa, United States), con la secuencia 5'-CACC-3' en el extremo 5' del *primer* directo y 5'-AAAC-3' al extremo del *primer* reverso, correspondientes a los extremos cohesivos producidos por la enzima de restricción BbsI (Thermo Fisher Scientific, Ref. ER1011). *Primer* directo y reverso se hibridaron para formar una doble hebra mezclando 4,5 μ L de cada *primer* a 100 μ M y 1 μ L de Tampón Duplex 10X (acetato potásico 1M y Hepes 300mM pH 7,5). La hibridación se lleva a cabo mediante una rampa de temperatura, de 95 $^{\circ}$ C hasta 25 $^{\circ}$ C, disminuyendo 1 grado por minuto.

Mediante la enzima BbsI se digirió el vector, produciendo extremos cohesivos complementarios a las secuencias añadidas en los *primers*. La ligación de vector y *primers* hibridados se realizó mediante la ligasa DNA T4 (Thermo Fisher Scientific, cat no EL0014). El producto de la ligación fue transformado en bacterias electrocompetentes *Escherichia coli* TOP10 mediante el equipo MicroPulser™ Electroporator (BIO-RAD, Hercules, California, United States). Estas células se sembraron en placas de cultivo LB con el antibiótico de selección del vector, ampicilina. Tras incubar durante 16 h a 37 $^{\circ}$ C se analizaron las colonias mediante *PCR* colonia con un *primer* específico de vector y otro específico del inserto introducido. Esa *PCR* colonia se realizó mediante *PCR* Master Mix 2x (Thermo Fisher Scientific K0172). Las colonias positivas fueron amplificadas en medio LB líquido con ampicilina y el vector se purificó mediante el kit QIAprep Spin Miniprep (QIAGEN, cat no 27104). El inserto se secuenció mediante método *Sanger* para confirmar la inserción correcta de la secuencia guía.

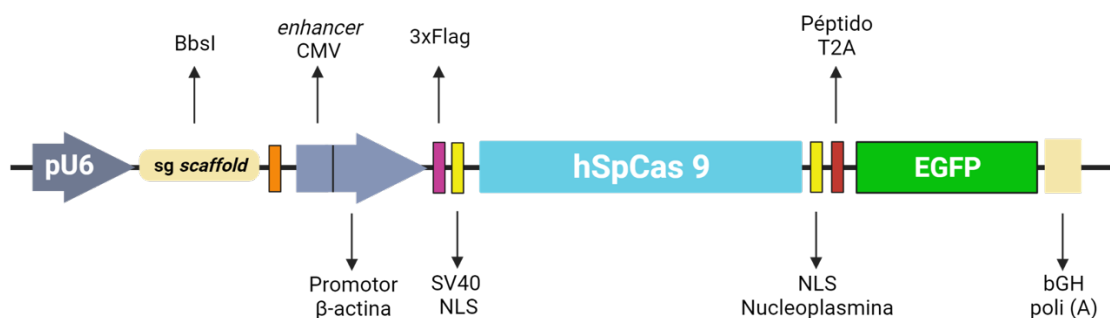


Figura 2.2: Vector PX458. Este vector presenta codificados los elementos necesarios del sistema CRISPR. El promotor pU6 controla la expresión de la secuencia guía en células de mamífero. Mediante la enzima BbsI es posible clonar una secuencia guía de interés en la región *sg scaffold*. Esta, junto al *tracrRNA* (naranja), constituyen el *sgRNA*. La nucleasa hSpCas9 (azul) se expresa bajo el control del *enhancer* CMV y el promotor de β -actina. Además, está flanqueada por secuencias Flag (rosa) y secuencias de localización nuclear (NLS) (amarillo). El gen codificante de la proteína GFP (verde) se expresa conjuntamente con la nucleasa. Gracias al péptido T2A (rojo) la proteína GFP se traduce de forma independiente. Imagen creada con Biorender.com.

3. Generación del vector pEGR-1

Para desarrollar un sistema sencillo y fácil de producción de guías CRISPR con un reportero necesitábamos un vector que presentara tanto el sitio de clonación de una secuencia guía y sus elementos de expresión, como el casete de expresión de GFP. Los vectores comerciales disponibles no presentan estos elementos de forma contigua, si no que se encuentran separados por otros elementos, como reguladores de la expresión bacteriana o marcadores de selección como resistencia a antibióticos. Por ello se desarrolló un vector propio donde estos elementos se encuentran próximos entre si, el vector pEGR1.

Para la generación del vector pEGR-1 se partió de los vectores pX458 y pEGFP-N1. El vector pEGFP-N1 (Clontech Laboratories, Mountain View, California, United States) expresa el gen codificante de la proteína GFP bajo el control del promotor/*enhancer* CMV que permite su expresión en células de mamífero. También presenta un sitio de reconocimiento de la enzima de restricción AflIII al final de la secuencia terminadora y de poliadenilación SV40 poly(A) y resistencia al antibiótico kanamicina. Sobre el vector pX458 se amplificó por PCR la región del sitio de clonación de la secuencia guía, promotor y terminador de expresión. Para ello se empleó la enzima ADN Polimerasa Phusion High Fidelity (Thermo Fisher Scientific, cat no F530L) y *primers* específicos con la secuencia diana de la enzima de restricción AflIII (Thermo Fisher

Scientific, cat no. ER0831) (5'-CTTAAG-3') en su extremo 5'. Estos *primers* hibridan a 36 pb del comienzo del promotor y a 91 pb tras el terminador de la transcripción. Este producto de *PCR* y el vector pEGFP-N1 se digirieron con la enzima AflII, generando extremos cohesivos. Al linealizar el vector pEGFP-N1 con esta enzima es posible clonar el casete de expresión de secuencia guía aguas abajo del terminador de *GFP*. El vector de destino fue desfosforilado tras la digestión mediante Fosfatasa Alcalina (New England Biolab, Ipswich, Massachusetts, United States, cat no M02905). Inserto y vector se purificaron mediante el kit QIAquick PCR Purification (QIAGEN, Hilden, Germany, cat no 28104) y ligados mediante la ligasa de ADN T4 (Thermo Fisher Scientific, cat no EL0014) (**Figura 2.3**).

El producto de la ligación fue electroporado en células *Escherichia coli* electrocompetentes Top 10 mediante el equipo MicroPulser™ Electroporator (BIO-RAD) y sembradas en placas LB con el antibiótico de selección kanamicina. Las colonias con el vector fueron seleccionadas mediante *PCR* colonia con *primers* específicos del inserto y amplificadas en cultivo líquido LB con el antibiótico de selección. El vector fue purificado mediante el kit QIAprep Spin Miniprep Kit (QIAGEN, cat no 27104) y el inserto verificado mediante secuenciación *Sanger*. De esta forma construimos el vector final de trabajo pEGR-1.

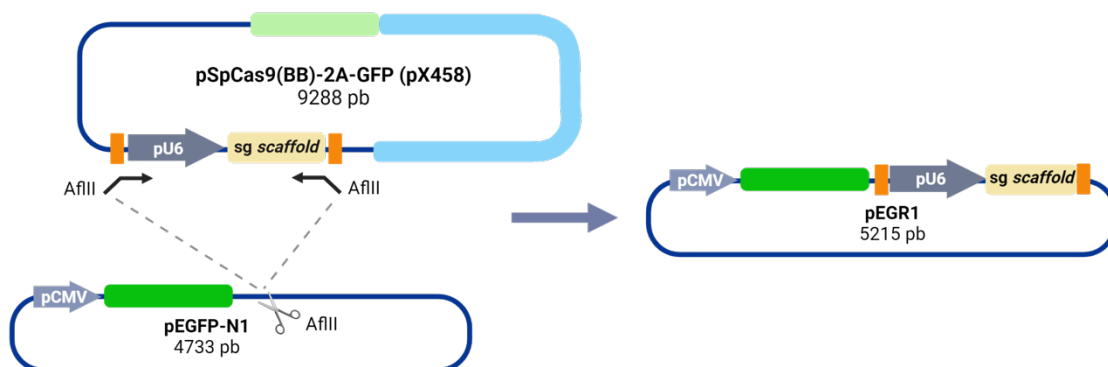


Figura 2.3. Generación del vector pEGR1. El vector pX458 presenta el gen de la Nucleasa hSpCas9 (azul), sitio de clonación y expresión de una secuencia guía (promotor pU6 y *sg scaffold*) y el gen codificante de *GFP* (verde). El vector de destino pEGFP-N1 expresa el gen codificante de la proteína *GFP* (verde) bajo el control del promotor/*enhancer* CMV y de la señal de terminación y poliadenilación SV40 poly(A), junto a la cuál presenta un sitio de reconocimiento de la enzima de restricción AflII. La amplificación del casete de expresión de la secuencia guía de PX458 con *primers* con la diana de la enzima AflII permite su clonación el vector pEGFP-N1, previamente digerido con la misma enzima. La combinación de ambos elementos genera el vector pEGR1. Imagen creada con Biorender.com.

4. Generación de las construcciones por *PCR* que expresan las guías de ARN

Se creó el plásmido pEGR-1 desde el cual amplificar las construcciones guía, de forma que contuviesen los promotores, secuencias de interés y terminadores de la transcripción, además de un gen reportero (*EGFP*), que podría ser añadido a la construcción a voluntad. A partir del vector pEGR1 se generaron las construcciones que codifican las secuencias guías de *IDH2* y *MYBL2* por medio de una serie de *PCR* de fusión. Para realizar cada una de las construcciones solamente es necesario modificar los *primers* específicos con las secuencias guía. Las secuencias se incorporan al extremo 5' de los *primers* con homología al *sg scaffold*. Uno de ellos presenta la secuencia directa del guía y otro con la secuencia reversa (**Figura 2.4**). La flecha azul representa el *primer* directo, FPCR1_F, y la flecha roja el *primer* reverso, FPCR2_R. Las colas naranjas representan las secuencias guía. Partiendo de estos dos *primers*, se generan los 3 fragmentos principales que servirán para generar las construcciones.

Mediante la pareja de *primers* FPCR_4F y FPCR_2R se genera el fragmento que presenta el casete de expresión de *GFP*, el promotor y la secuencia guía de interés (**Figura 2.4.a**). La pareja FPCR_1F y FPCR_1R sirve para producir el fragmento con la secuencia guía y el terminador SV40 poli(A) (**Figura 2.4.b**). Por último, la combinación de los *primers* FPCR_2F y FPCR_2R genera un fragmento que presenta el promotor pU6 y la secuencia guía introducida (**Figura 2.4.c**). Gracias a las regiones solapantes entre estas *PCRs* es posible generar la construcción pU6-guía empleando los *primers* FPCR_3F y FPCR_3R fusionando el fragmento 2.4.a y 2.4.b (**Figura 2.4.d**), y la construcción *GFP*-pU6-guía mediante FPCR_5F y FPCR_3R, fusionando los fragmentos 2.4.b y 2.4.c (**Figura 2.4.e**). Todas las construcciones generadas se realizaron con la enzima ADN Polimerasa Phusion High Fidelity (Thermo Fisher Scientific, cat no F530L) y fueron purificadas mediante columna con el kit MinElute PCR *Purification Kit* (QIAGEN, cat no 28004).

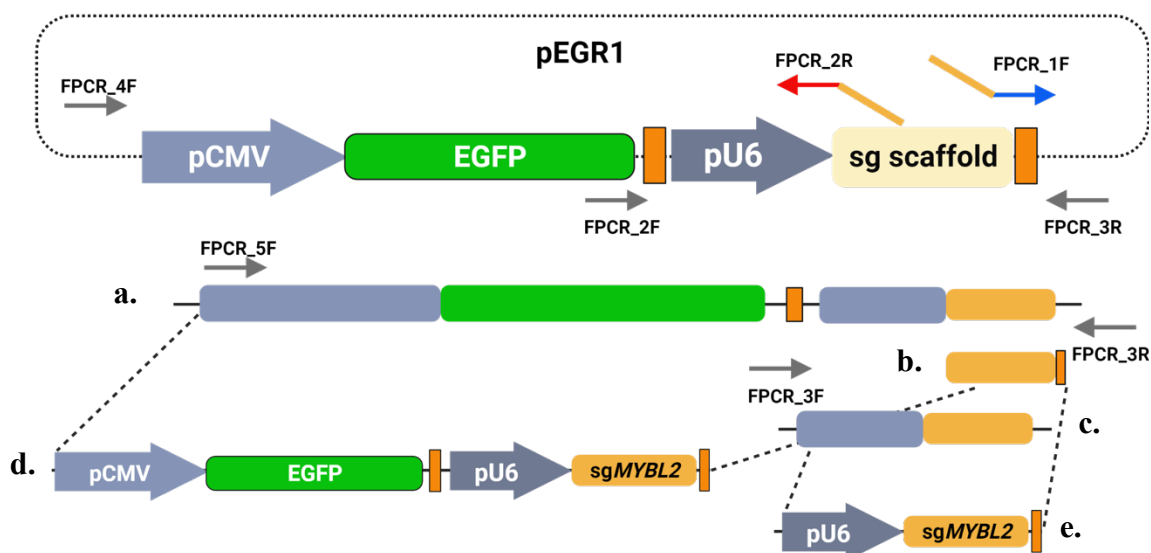


Figura 2.4. Estrategia de generación de las construcciones guía mediante el vector pEGR1. Este sistema permite generar fragmentos de expresión de secuencias guía con o sin el reportero EGFP mediante una combinación de sencillas PCR. Imagen creada con Biorender.com.

La generación de la construcción pU6-MYBL2 con extremos fosforilados (pU6-MYBL2-P), se produjo fosforilando el extremo 5' de los primers FPCR_3F y FPCR_3R mediante la Polinucleotido Quinasa T4 (New England BioLaBS, cat no M0201S). Tras la fosforilación, los *primers* fueron purificados con el kit *MinElute PCR Purification Kit* (QIAGEN, cat no 28004) y utilizados en la PCR final para generar la construcción fosforilada.

La construcción del complejo de expresión de la hSpCas9 se generó mediante amplificación por PCR sobre el vector hCas9 (Addgene, #41815) (Mali *et al.*, 2013) (Figura 2.5). Este vector expresa la nucleasa hSpCas9 bajo el control del promotor/*enhancer* CMV. Para la amplificación por PCR se empleó la enzima ADN Polimerasa Phusion High Fidelity (Thermo Fisher Scientific, cat no F530L) y *primers* específicos. El *primer* directo, hCas9_F hibrida a 186 pb aguas arriba del promotor/*enhancer* CMV, y el *primer* reverso hCas9_R, que hibrida a 326 pb aguas abajo del gen codificante de hSpCas9, amplificando la zona terminadora y la señal poli(A). El producto de PCR fue purificado mediante columna con el kit de Purificación de PCR QIAGEN (cat no 28104).

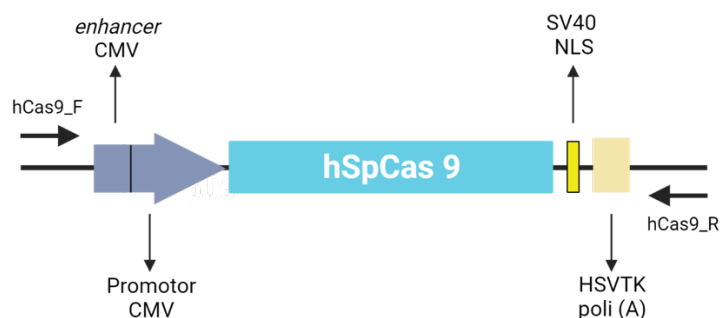


Figura 2.5. Vector hCas9. En detalle se muestra la región de interés para generar la construcción con el complejo de expresión de hSpCas9. Este vector presenta el gen codificante de la nucleasa hSpCas9 (azul). La expresión de este gen está bajo control del promotor y *enhancer* CMV (gris). También presenta codificada una señal de localización nuclear (SV40 NLS) y una señal de terminación de la transcripción (HSVTK poli (A)). En detalle se muestra la localización de los *primers* empleados para amplificar esta región. Imagen creada con Biorender.com

5. Mantenimiento y cultivos de las líneas celulares empleadas en este trabajo.

Para testar si las construcciones guías generadas son capaces de guiar a la nucleasa Cas9 en el núcleo celular hasta la secuencia homóloga del genoma se empleó la línea celular de fácil manejo HEK293 (Graham *et al.*, 1977). Se trata de células humanas embrionarias de riñón de rápido crecimiento, fáciles de manipular y transfectar (Stepanenکو *et al.*, 2015). Las células HEK293 se cultivaron en medio *Dulbecco's Modified Eagle Medium* (DMEM 1X) (Gibco, Thermo Fisher Scientific, cat no 41966-029), suplementado con 0,5 % de Penicilina/Estreptomicina 100X Solution (Biowest, Nuailé, France, cat no L0022-020) y 10 % de suero bovino fetal (Gibco, cat no 10270-106). Las células se incubaron a 37 °C y 5 % CO₂.

Las líneas celulares leucémicas escogidas fueron: NB4, establecida a partir de un paciente con leucemia promielocítica aguda (Lanotte *et al.*, 1991), HL60 (Collins *et al.*, 1977) y Katsumi-1 (Asou *et al.*, 2003) establecidas a partir de pacientes con leucemia mieloide aguda, y MOLM-13 (Matsuo *et al.*, 1997), establecida a partir de un paciente con leucemia monocítica aguda. Se cultivaron en medio RPMI 1640 (1X) (Gibco, cat no 21875-034), suplementado con 1 % Penicilina/Estreptomicina 100X Solution (Biowest, cat no L0022-020) y 10 % suero bovino fetal (Gibco, cat no 10270-106) a 37 °C y 5 % CO₂.

6. Edición en la línea celular HEK293

Para testar la efectividad de las construcciones generadas por *PCR* que codifican las guías CRISPR, se transfectaron en células HEK293 mediante el reactivo Lipofectamina 3000 (Invitrogen, Carlsbad, California, United States, cat no 15282465) (Hirko et al., 2003). El procedimiento para las distintas pruebas realizadas fue el siguiente. Veinte y cuatro horas antes de la transfección se sembraron 9×10^4 células HEK293 por pocillo en placas de 24 pocillos con medio de cultivo completo. Uno de los pocillos fue transfectado con el vector comercial pMAXGFP (Lonza, Basel, Switzerland), con expresión de la proteína GFP, sirviendo como control positivo de la transfección. Para probar la estrategia *cold shock* descrita por el grupo de D. Zho y J.N Feder, en la cuál registraon un aumento de la eficiencia de edición cultivando las células a temperaturas inferiores a 37 °C (Guo et al., 2018), las células transfectadas se cultivaron a 32 °C. Para todas las condiciones de transfección y cultivos testados, cuarenta y ocho horas post transfección se extrajo el ADN de las células mediante el kit comercial QIAamp DNA Mini Kit (QIAGEN, cat no 51304). Tras esto se prosiguió con la comprobación de la eficiencia de corte de la nucleasa mediante el ensayo con la enzima endonucleasa T7 EI.

7. Optimización de la edición en células HEK293

Distintas cantidades de construcción pU6-guía se testaron para obtener la máxima eficiencia de corte. Para ello se emplearon distintas cantidades de la construcción pU6-MYBL2 y del vector hCas9. Se estudiaron los siguientes ratios molares de vector hCas9-construcción pU6-MYBL2: 500 ng hCas9 – 23,3 ng pU6-MYBL2 (Ratio 1:1); 250 ng hCas9 – 12 ng pU6-MYBL2 (Ratio 1:1); 250 ng hCas9 – 35 ng pU6-MYBL2 (Ratio 1:3); 250 ng hCas9 – 60 ng pU6-MYBL2 (Ratio 1:5); y 150 ng hCas9 – 7 ng pU6-MYBL2 (Ratio 1:1).

Para comprobar si la fosforilación de los extremos 5' de las construcciones producía un aumento de la eficiencia de corte se transfectó la construcción pU6-MYBL2-P junto al vector hCas9. En paralelo, se transfectaron también la construcción no fosforilada y el vector pX458, ambos con la misma secuencia guía. Para transfectar las construcciones se emplearon 250 ng del vector hCas9 y 35 ng de construcción. En el caso del vector PX458, se transfectaron 500 ng.

Por último, el casete de expresión de hSpCas9 en forma de *PCR* se transfectó junto a la construcción pU6-*MYBL2*. Se testaron dos cantidades, 140 ng y 250 ng, manteniendo la cantidad de construcción guía a 35 ng.

8. Edición del gen *IDH2* en las células HEK293

Partiendo de las mejores condiciones obtenidas con la construcción pU6-*MYBL2* se pasó a seleccionar el guía más eficiente para editar el gen *IDH2*. Se generaron ocho construcciones guías contra el gen *IDH2* partiendo de las ocho secuencias guías diseñadas. Cada guía fue transfectada, junto al vector hCas9 en las células HEK293. Se emplearon 35 ng de construcción guía y 250 ng de vector hCas9. Los guías pU6-*IDH2*_3 y pU6-*IDH2*_8 se transfectaron también de forma conjunta debido a que se localizan muy próximas y flanquean la posición de la mutación R172. Una vez seleccionada el guía más eficiente, se empleó el casete de expresión de la nucleasa hSpCas9 en forma de *PCR* en lugar del vector hCas9 para comparar las eficiencias de corte.

Para la introducción de la mutación *IDH2*^{R172} se empleó la combinación de los guías pU6-*IDH2*_3 y pU6-*IDH2*_8, el vector hCas9 y una molécula oligonucleotídica de cadena única (*ssODN*) (*Integrate DNA Technologies*) homóloga a la región de *IDH2* a modificar como donador (**Figura 2.6**) (**Anexo I, Tabla 2.2**). La secuencia *ssODN* presenta el cambio que produce la mutación R172, el cambio del codón 172 AGG por AAG. Se testaron dos concentraciones del *ssODN*, 5 μ M y 10 μ M. También se probó la eficiencia de edición empleando un *ssODN* con la secuencia anti sentido. Cuarenta y ocho horas post transfección se extrajo el ADN de las células y se realizó el estudio de la eficiencia de edición.

Secuencia diana *IDH2*

5' - CCT - GGC - TGG - ACC - AAG - CCC - ATC - ACC - AT**T** - **GGC** - AGG - CAC - **GCC** - **CAT** - GGC - GAC - CAG - GTA - GGC - CAG - GGT - GGA - 3'

IDH2_8 *IDH2_3*

Secuencia *ssODN*

5' - CCT - GGC - TGG - ACT - AAA - CCG - ATA - ACG - AT**C** - **GGA** - AAG - CAC - **GCG** - **CAT** - GGC - GAC - CAG - GTA - GGC - CAG - GGT - GGA - 3'

IDH2_8 *IDH2_3*

Figura 2.6. Diseño *ssODN*. Esquema del diseño *ssODN* centrado en la zona con modificaciones. La primera secuencia representa la secuencia salvaje del gen *IDH2*. La segunda secuencia representa el *ssODN*. Ambas secuencias se representan divididas por tripletes. En negrita se resaltan los nucleótidos que han sido modificados en el *ssODN*. Además de la mutación (nucleótido naranja), es necesario modificar las secuencias *PAM* de los guías *IDH2_3* e *IDH2_8* para evitar que la nucleasa Cas9 corte el gen tras la edición (nucleótidos rojos). Se modificaron mediante cambios sinónimos que no alteran la secuencia proteica. El cambio sinónimo de la secuencia *PAM* del guía *IDH2_3* introduce un sitio de reconocimiento de la enzima de restricción HhaI 5'-GCGC-3' (New England BioLabs, cat no R0139S). Esto permite evaluar la eficiencia de edición mediante ensayos de *RFLP*. Al no ser posible modificar la parte esencial de la secuencia *PAM* de *IDH2_8* sin alterar la secuencia de la proteína se introdujeron 7 cambios silenciosos a lo largo de la secuencia diana del guía para evitar el corte por la nucleasa (nucleótidos en azul). La molécula *ssODN* se completa con dos brazos homólogos de 35pb, presentado un tamaño final de 98 nucleótidos. El *ssODN* fue sintetizado como *Ultramer Oligonucleotides* (Integrated DNA Technologies).

9. Ensayo de eficiencia de corte mediante Endonucleasa T7 I

La enzima Endonucleasa T7-I (T7 EI) se emplea para determinar el porcentaje de *indels* producidos en la región genómica donde se ha guiado a la nucleasa y, en consecuencia, la eficiencia de los sistemas CRISPR/Cas9 (**Figura 2.7**). Tras 48 h post transfección se extrajo el ADN de las células, y se amplificó la región modificada mediante la ADN polimerasa Phusion High-Fidelity y *primers* específicos. Las células transfectadas con el vector pMAXGFP se emplearon como control negativo de edición. Los productos de las *PCRs* se comprobaron y purificaron mediante electroforesis en gel de agarosa al 0,8 %. El fragmento de *PCR* es purificado del gel mediante purificación por columna con el kit comercial E.Z.N.A Gel Extraction (OMEGA bio-tek, Norcross, Georgia, United States, cat no D2500-02). Una vez purificado el producto de *PCR* se siguieron las instrucciones comerciales de la T7-EI (New England Biolabs, cat no M0302L). El producto amplificado por *PCR* presenta un conjunto heterogéneo de secuencias, donde algunas presentan los *indels* producidos tras el corte por la nucleasa, y otras no presentan cambios. Este conjunto fué sometido a una rampa de temperaturas para producir la desnaturalización de las hebras de ADN y una hibridación rápida, de forma que se pueden hibridar secuencias con *indels* con secuencias salvajes. En este caso, se crea un *heteroduplex* u horquilla en la zona de no homología entre secuencias. Este

heteroduplex es reconocido por la enzima T7-EI y produce un corte de doble cadena. Así, tras el corte por la T7-EI, se producen fragmentos de un tamaño más pequeño, y las secuencias que se han hibridado con sus homólogos permanecen con su tamaño original (Reyon et al., 2012).

Los productos obtenidos fueron separados mediante electroforesis en un gel de poliacrilamida al 10 % y se tiñó con SYBR Safe DNA Gel Stain (Invitrogen, cat no S33102). Como resultado se espera observar una banda de mayor tamaño, correspondiente a las secuencias sin digerir y dos bandas de tamaño menor generadas por el corte de la enzima T7 EI. El experimento se realizó por duplicado, donde a una de las muestras no se le adiciona la T7 EI. El duplicado sirvió para asegurar que las bandas observadas se deben al efecto de la enzima y no a bandas inespecíficas de la *PCR*.

Se empleó el programa ImageJ para estimar el porcentaje de *indels* producidos. Se midió la intensidad integrada de las bandas correspondientes a los productos digeridos y los no digeridos. Mediante la siguiente fórmula se calcularon los porcentajes de *indels*:

$$Indel (\%) = 100 \times (1 - \sqrt{1 - f_{cut}})$$

f_{cut} corresponde a la fórmula $(b+c)/(a+b+c)$, donde b y c corresponden a la intensidad de banda de los productos digeridos y a la intensidad de la banda no digerida.

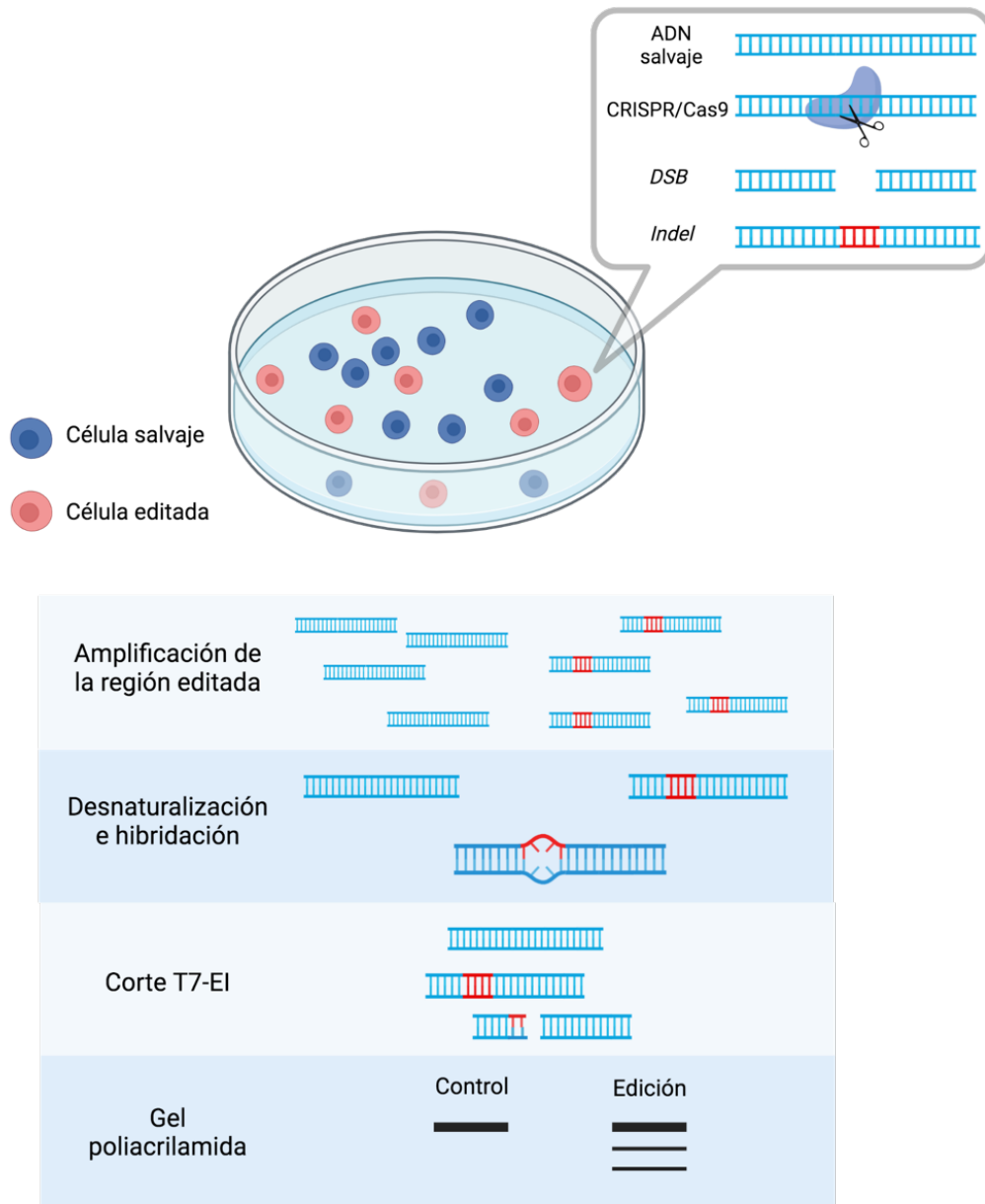


Figura 2.7. Ensayo de la enzima Endonucleasa T7 I. A las 48 h post transfección se extrae el ADN genómico de todas las células y se amplifica la región genómica donde se ha guiado la nucleasa Cas9. Los *primers* empleados se encuentran a distancias dispares del punto de corte esperado de la Cas9. Esto hace que los posibles cambios introducidos en el proceso de reparación no se encuentren en la parte central del amplificado. De esta forma se obtiene un conjunto formado por amplificados con la secuencia salvaje sin editar y otros editados (rojo). Este conjunto se somete a una rampa de temperaturas que fuerza la desnaturalización de la doble hebra del ADN y la hibridación aleatoria. En este paso se puede dar la hibridación de una hebra editada con una salvaje, formándose una horquilla en la zona de no homología entre ambas hebras. Esta estructura es reconocida por la enzima T7-EI, produciendo un corte de doble cadena y generando dos fragmentos de ADN más cortos. Esta diferencia de tamaños se observa en un gel de poliacrilamida y se emplea para cuantificar el porcentaje de *indels* producidos. Imagen creada con Biorender.com

10. Ensayo de eficiencia de edición mediante *RFLP*

La determinación de la eficiencia de edición de la mutación R172 se llevó a cabo mediante ensayo de *RFLP*. La diana de HhAI introducida en el ADN genómico por recombinación homóloga, ya que el *ssODN* contiene este cambio, permite detectar y cuantificar el porcentaje de edición producido.

Una vez extraído el ADN de las células a las 48 h post transfección, la región del gen *IDH2* a analizar se amplificó usando los mismos *primers* empleados en el ensayo de T7-IE. El producto de *PCR* se comprueba y purifica siguiendo el mismo protocolo que para el ensayo de T7-EI. El producto fue digerido mediante la enzima HhAI, siguiendo las recomendaciones del protocolo comercial. Tras la digestión, el producto se carga en un gel de poliacrilamida al 10 %. Al igual que en el experimento de T7-I, el gel es teñido con SYBR Safe DNA Gel Stain. Si la enzima detecta su diana produce un corte en la doble hebra generando dos fragmentos de tamaños diferentes. De esta forma se espera, al igual que en el ensayo de la T7-EI, tres bandas tras la digestión: una de mayor tamaño, correspondiente a las secuencias sin editar (homodúplex), y dos más pequeñas, resultado de las secuencias digeridas (heterodúplex). Al igual que en el apartado anterior, cada ensayo se lleva a cabo por duplicado, donde una de las réplicas es procesada con el protocolo de digestión, pero sin adicionar enzima y el ADN de las células transfectadas con el vector pMAX se utiliza como control negativo de edición.

La imagen del gel es procesada mediante el programa ImageJ y se miden las intensidades de banda relativas. Para calcular la eficiencia de edición se emplea la fórmula presentada en el apartado anterior, $(b+c)/(a+b+c)$, donde b y c corresponden a la intensidad de banda de los productos digeridos y a la intensidad de la banda no digerida.

11. Edición del gen *IDH2* en líneas celulares leucémicas

Tras la optimización y prueba de efectividad de las construcciones guías desarrolladas por *PCR* en las células HEK293 se pasó a emplearlas sobre las líneas celulares leucémicas. Estas células presentan eficiencias de transfección muy bajas (Zhou et al., 2016). Por ello, se desarrollaron líneas celulares leucémicas con expresión constitutiva de la nucleasa hSpCas9 mediante transducción con lentivirus. Una vez

establecidas estas líneas celulares se emplearon para probar la eficiencia de edición de las construcciones guía desarrolladas.

12. Desarrollo de líneas celulares con expresión constitutiva de la nucleasa Cas9

Para desarrollar líneas con la integración del gen codificante de la nucleasa y sus reguladores de expresión primero se generaron lentivirus con el vector de transferencia lentiCRISPR V.2 (Addgene, #52961) (Sanjana et al., 2014) (**Figura 2.8**).

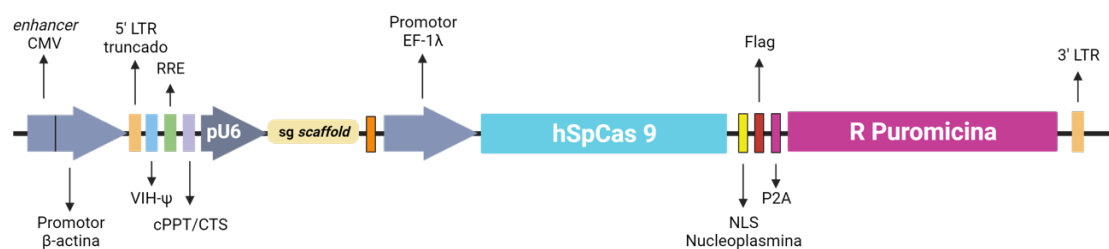


Figura 2.8. Vector LentiCRISPR V.2. Diagrama que describe la estructura del vector LentiCRISPR v.2. Este vector viral. presenta el gen codificante de la hSpCas9 bajo el control del promotor EF-1 α . Acoplado transcripcionalmente se encuentra el gen de resistencia al antibiótico puomicina. Separando ambos genes se encuentra la secuencia del péptido 2A. Tras la traducción, este péptido es capaz de auto cortarse, lo cual permite que ambas proteínas se expresen de forma independiente. Por último, la hSpCas9 presenta una secuencia NLS que permite su transporte al núcleo tras la traducción. Estas secuencias se encuentran flanqueadas por una serie de elementos génicos que permiten el empaquetamiento del vector en las partículas víricas. Estos elementos son la señal de empaquetamiento VIH-1 Ψ ; la secuencia RRE y las secuencias cPPT-CTS. Además, presentan secuencias repetidas terminales (LTR) en los extremos 5' y 3' que permiten la integración de la hSpCas9, el gen de resistencia a puomicina y sus elementos reguladores en el genoma de la célula huésped. Estas secuencias LTR se encuentran modificadas por cuestiones de bioseguridad (Durand et al., 2011). Imagen creada con Biorender.com

Los lentivirus se desarrollaron mediante un sistema de segunda generación, empleando el vector de empaquetamiento psPAX2 (Addgene, #12260;) (Trono lab, no publicado), y el vector de envoltura fue pMD2.G (Addgene, #12259) (Trono lab, no publicado).

La producción de lentivirus se realizó en células HEK293T en cultivo. Se mantienen en medio de cultivo *Dulbecco's Modified Eagle Medium* (DMEM 1X) (Gibco, cat no 41966-029), suplementado con 0.5 % de Penicilina/Estreptomina 100X Solution (Biowest, cat no L0022-020) y 10 % de suero bovino fetal (Gibco, cat no 10270-106). Las células se incuban a 37 °C y 5 % CO₂. Un día antes a la transfección, se sembraron 3x10⁶ células HEK293T en una placa Petri de 10 cm. Al día siguiente las células fueron transfectadas mediante el método de Cl₂Ca con el vector de transferencia lentiCRISPR

V.2 y los vectores psPAX2 y pMD2.G. A las 48 h post transfección, se recoge el sobrenadante donde han sido liberados los virus generados. Se filtra mediante un filtro de nitrocelulosa de 0,45 μm y almacenados a $-80\text{ }^{\circ}\text{C}$ hasta su uso.

Se infectaron 5×10^5 células NB4, HL60, MOLM-13 o Katsumi-1 con 100 μL del sobrenadante con virus y Polibrene (Merck, Kenilworth, New Jersey, United States, cat no TR-1003-G) a una concentración final de 4 $\mu\text{g}/\mu\text{L}$. Tras 24 h, las células se centrifugaron y se resuspendieron en medio RPMI completo con el antibiótico de selección puromicina (0,6 $\mu\text{g}/\text{mL}$) (InvivoGen, San Diego, California, United States, cat no ant-pr-1). Tras mantenerlas siete días en cultivo, se aisló el ADN de las células transducidas mediante el kit comercial QIAamp DNA Mini Kit (QIAGEN, cat no 51304). Se confirmó la presencia de la nucleasa hSpCas9 mediante *PCR* empleando la ADN Taq polimerasa (PCR Master Mix 2x, Thermo Fisher Scientific K0172) con *primers* específicos de la nucleasa. De las distintas líneas transducidas las líneas celulares positivas fueron NB4, HL60 y MOLM13, que se denominaron NB4-Cas9, HL60-Cas9 y MOLM-13-Cas9 para distinguirlas de las líneas celulares de origen. Las líneas celulares positivas se mantienen en medio suplementado con 0,2 $\mu\text{g}/\text{mL}$ de puromicina.

Se testó la funcionalidad de la hSpCas9 insertada empleando la construcción pU6-*IDH2*_1. Para introducir las construcciones que codifican las secuencias guías, en estas células se utilizó la técnica de nucleofección de Amaxa (Lonza). Esta tecnología se basa en el desarrollo de poros en la membrana celular mediante pulsos eléctricos. Estos poros permiten que el ADN exógeno entre al citoplasma nuclear y sea translocado al núcleo. Para ello también es necesario el uso de tampones de nucleofección específicos de la línea celular con la que se esté trabajando. Amaxa ha desarrollado protocolos y buffers de nucleofección específicos para las líneas HL60 y NB4, pero en el momento de desarrollo de los experimentos no ofrecían protocolos específicos para la línea celular MOLM-13. Por ello, las líneas MOLM-13 y MOLM-13-Cas9 fueron descartadas para continuar con los experimentos.

Para adaptar la cantidad de construcción guía necesaria para editar las células leucémicas se partió de las concentraciones recomendadas por el protocolo comercial. El protocolo comercial de Amaxa recomienda emplear hasta 2 μg de plásmido de ADN para las líneas celulares seleccionadas. Se consideró una relación equimolar para

calcular la cantidad de guía a emplear. Considerando 2 μg de plásmido de ADN de 3 kb se empleó 266 ng de construcción pU6-*IDH2_1*. Además de esta cantidad se testaron también 400 ng y 800 ng de la construcción pU6-*IDH2_1*. Siguiendo esas condiciones se nucleofectaron 2×10^5 células. Se empleó el aparato 4D-Nucleofector System (Lonza, Basilea, Switzerland) y los programas específicos para cada línea celular, CZ-100 para las células NB4-Cas9 y EN-138 en el caso de las HL60-Cas9. Tras la nucleofección las células se cultivan en placas de 24 pocillos con medio RPMI completo y se incuban a 37 °C y 5 % de CO₂. Se utilizó el vector pMAXGFP como control positivo de transfección, comprobando la expresión de GFP en ese pocillo a las 24 h. A las 48 h post transfección se realiza la extracción de ADN de cada muestra, seguido del protocolo del ensayo de eficiencia de corte por la enzima T7-EI.

13. Eficiencia de nucleofección de las líneas leucémicas

La eficiencia de nucleofección se testó en las células HL60-Cas9 y NB4-Cas9. Se nucleofectó la construcción GFP-pU6-*MYBL2* (2033 pb), el vector PX458-*sgMYBL2*, de mucho mayor tamaño que la construcción guía (9288 pb) y el vector comercial pMAXGFP de tamaño medio (2030 pb). Se nucleofectaron 2×10^5 células con 800 ng de vector PX458, 800 ng de construcción GFP-pU6-*MYBL2* o 500 ng de vector pMAXGFP. Los programas de nucleofección empleados son los mismos que se han detallado en el apartado anterior. A las 24 h post transfección se analizó el ratio de células GFP positivas en cada condición mediante el citómetro FACSDiva 8.0.1 (BD, Franklin Lakes, New Jersey, United States). Estos experimentos se repitieron por triplicado.

14. Optimización de la edición del gen *IDH2* en células NB4-Cas9

Empleando la combinación de guías más eficiente para editar el gen *IDH2* en las células HEK293 (los guías *IDH2_3* e *IDH2_8*) primero se optimizó la cantidad óptima de construcción para editar el gen *IDH2* en las células NB4-Cas9. Se compararon las diferencias de edición empleando dos cantidades, 800 ng y 1,5 μg totales entre ambas construcciones. Estos experimentos se realizaron por triplicado y se calculó su eficiencia de corte mediante el ensayo T7-EI a las 48 h post nucleofección.

También se testó la cantidad de *ssODN* más eficiente para introducir la mutación R172 en estas células. El *ssODN* testado en HEK293 se transfectó a 10, 100, 200, 300 y

500 μ M, junto a 1,5 μ g de las construcciones guías. Las eficiencias fueron calculadas mediante ensayo por *RFLP*. Todos los experimentos se repitieron por triplicado.

15. Edición del gen *IDH2* en células NB4

Las células NB4 sin expresión constitutiva de la nucleasa hSpCas9 se editaron mediante complejos *RNPs* y la construcción de expresión de hSpCas9.

Para desarrollar los complejos de *RNPs* se adquirieron los guías en formato RNA como AltR™ CRISPR-Cas9 *crRNA* (Integrate DNA Technologies) (*crIDH2_3* y *crIDH2_8*); la molécula *tracrRNA* en formato Alt-R™ CRISPR-Cas9 *tracrRNA* (Integrated DNA Technologies, cat no 1072533) y la nucleasa hSpCas9 recombinante como Alt-R™ S.p. Cas9 Nuclease 3NLS (Integrated DNA Technologies, cat no 1074182).

Los AltR™ CRISPR-Cas9 *crRNA* presentan una secuencia complementaria al Alt-R™ CRISPR-Cas9 *tracrRNA* por la cual pueden hibridarse. Cada *crRNA* se mezcla a condiciones equimolares con el *tracrRNA* para formar dos dúplex de 20,3 μ g (*sgIDH2_3* y *sgIDH2_8*). La hibridación se lleva a cabo a 95 °C durante 5 min siguiendo las especificaciones comerciales. 17 μ g de Alt-R™ S.p. Cas9 Nuclease 3NLS se añade a cada dúplex y se incubaron a temperatura ambiente durante 10 min para formar los dos complejos *RNPs*. Doscientas mil células NB4 fueron nucleofectadas con cantidades finales de 40,6 μ g de *sgRNAs* y 34 μ g de nucleasa Cas9. Para los experimentos de edición se añadió 100 μ M de *ssODN* a la nucleofección. Como en los apartados anteriores, un pocillo fue nucleofectado con el vector pMAXGFP, y se comprobó la presencia de células GFP positivas a las 24 h como control de la nucleofección. A las 48 h se extrajo el ADN de las distintas muestras y se llevaron a cabo los correspondientes ensayos de eficiencia de corte o de edición. Estos experimentos se realizaron por triplicado.

Por último, a pesar de la baja eficiencia de corte obtenida en las células HEK293 con la hSpCas9 en forma de PCR, se probó a nucleofectar 800 ng del casete de expresión la hSpCas9 en forma de PCR, junto 1,5 μ g de los guías de *IDH2*.

16. Análisis de la eficiencia de edición y los posibles efectos *off-targets* mediante secuenciación masiva.

Mediante la herramienta web Cas-OFFinder (<http://www.rgenome.net/cas-offinder/>) (Bae et al., 2014) se seleccionaron los potenciales *off-targets* para los guías

IDH2_3 e *IDH2_8*. Se obtuvieron cuatro posibles *off targets* para el guía *IDH2_3* y doce para el guía *IDH2_8* (**Anexo I, Tabla 2.3**).

El estudio se realizó con el ADN de las células NB4 editadas mediante *RNPs* (NB4 *RNP*); células NB4 editadas mediante casete de hSpCas9 en forma de PCR y las construcciones guías (NB4 *PCR*); y células NB4-Cas9 editadas mediante construcciones guías (NB4-Cas9 Edición). Además, se incluyó el ADN células NB4 y NB4-Cas9 sin modificar como controles.

Los amplicones se generaron mediante dos tandas de *PCR*. En la primera reacción, se emplearon *primers* específicos para cada diana que, además de la secuencia homóloga, presentan una secuencia adaptadora. Tras llevar a cabo la primera reacción, las *PCRs* fueron purificadas mediante el método Ampure XP (Beckman Coulter, Brea, California, United States, cat no A63881). En la segunda reacción, las *PCRs* fueron reamplificadas con un *primer* directo complementario a la secuencia adaptadora introducida en la primera reacción y con un *primer* reverso, que además de ser complementario al adaptador presenta una secuencia identificadora. Gracias a esta secuencia es posible diferenciar de que muestra procede cada secuencia. La segunda tanda de *PCRs* también fue purificada mediante Ampure XP. La librería se generó mezclando las *PCR* finales a concentración equimolar siguiendo el protocolo comercial. Para realizar la secuenciación, la mezcla se cargó en un cartucho Micro MiSeq Reagent Kit V2 (500 ciclos) (Illumina, San Diego, CA, USA) y secuenciado mediante la plataforma MiSeq (Illumina).

Los archivos *fastq.gz* generados tras la secuenciación se analizaron mediante el programa informático CRISPResso2 (Clement et al., 2019). Este programa, con versión web (<https://crispresso.pinellolab.partners.org/submission>) y por línea de comandos, permite la cuantificación y visualización de los resultados de la edición y la evaluación de los efectos *off targets*. De forma automática, este programa realiza primero un filtrado de las lecturas de baja calidad y elimina los adaptadores, si fuera necesario. Continúa con el alineamiento de las secuencias con el amplicón de referencia y cuantifica la proporción de *indels* introducidos por el mecanismo de *NHEJ* y la edición de la mutación R172 mediante *HDR*. Tras el análisis, genera varios gráficos resumen que

permiten evaluar la edición producida y archivos de texto con los resultados para realizar análisis posteriormente (Pinello et al., 2016).

Para el estudio de las posibles modificaciones en los *off targets* seleccionados se empleó la herramienta CRISPResso2 Compare. Esta variante del mismo *software* permite comparar análisis previos realizados con CRISPResso2 sobre un mismo amplicón entre distintas muestras (muestra control y muestra editada), y a su vez compararlas con la secuencia de referencia.

Capítulo II. Edición de genes por CRISPR *in vivo*.

1. Material

En este apartado se especifican los materiales y soluciones relevantes para el trabajo con el gusano nematodo *C. elegans*.

1.1. Medio NGM

El medio *Nematode Growth Medium* (NGM) es el medio de crecimiento de *C. elegans*. Está compuesto por NaCl (3 g/L), Agar (20 g/L), Peptona (2,5 g/L), MgSO₄ 1M (1 mL/L), colesterol 5 mg/mL (1 mL/L), Nistatina 12,5 mg/mL (1 mL/L), CaCl₂ 1M (1 mL/L), KPO₄ 1M (25 mL/L) y agua Milli-Q. Tras autoclavar los compuestos se añaden el MgSO₄, colesterol, nistatina, CaCl₂ y KPO₄ al medio. Se plaquea en placas Petri en condiciones de esterilidad y se deja solidificar.

1.2. Solución M9

La solución M9, que es un suero fisiológico, se emplea para recoger los gusanos y se usa también para cultivarlos, junto con otros componentes. Además, para sincronizar los cultivos, es capaz de mantener vivos los animales, y permite su crecimiento en ausencia de comida, desde el estadio de embrión al de larva de estadio L1. Se prepara disolviendo Na₂HPO₄ (6 g/L), KH₂PO₄ (3 g/L), NaCl (5 g/L) y MgSO₄ 1M (1 mL/L) en agua Milli-Q y posterior autoclave.

1.3. Solución de lisis

Esta solución se emplea para lisar los gusanos y obtener el ADN genómico. Sus componentes son *Buffer PCR 10X Taq DNA Polimerase* (Merck, Ref 11146173001) (1 mL/ 5 mL), *MgCl PCR 10X Taq DNA Polimerase* (150 µL/ 5 mL), β- Mercaptoetanol (Thermo Fisher, Ref 10367100) (50 µL/ 5 mL), *Tween 20* (Sigma, Ref P9416-50ML) (22,5 µL/ 5 mL).

Se mantiene a 4 °C y antes de su uso se le añade Proteinasa K (Merck, Ref 3115801001) (5 µL/ 100 µL solución de lisis) para que pueda romper la cutícula del gusano.

1.4. Solución de lejía

Esta solución es empleada para la sincronización de un gran número de gusanos mediante la lisis de gusanos adultos hermafroditas, que contienen muchos embriones en el útero, y la obtención de la descendencia en estadio de embrión. Está compuesta por NaClO comercial (1 mL), NaOH 1M (2,5 mL), Solución M9 (4 mL) y ddH₂O (500 µL).

2. Cepas *C. elegans*

Los cultivos de *C. elegans* se han desarrollado como se ha descrito previamente (Brenner, 1974). Brevemente, los gusanos se cultivan en placas de Petri (55 mm x 16 mm) con medio NGM con una capa de bacteria OP50. Todas las cepas se mantienen a 20 °C. Se empleó la cepa salvaje estándar Bristol N2, o comúnmente denominada N2 (Brenner, 1974) obtenida del *Caenorhabditis Genetic Center* (CGC, Minneapolis, MN, USA). Las cepas generadas fueron establecidas mediante microinyección en la cepa N2 (**Anexo I, Tabla 2.4**).

3. Estudio de conservación evolutiva

Se emplearon distintas bases de datos públicas para estudiar la homología entre los genes ortólogos *IDH2* e *idh-2* en *C. elegans*. Mediante la base de datos Ensembl (Howe et al., 2021) se estudió la conservación filogenética del gen *IDH2* y su similitud con *idh-2*. Con esta base de datos también se compararon los dominios funcionales predichos de los transcritos codificados. Por último, se empleó la base de datos DIOT (Hu et al., 2011) para estudiar la similitud entre las secuencias de aminoácidos codificadas.

4. Generación de gusanos *knock-in* mediante el sistema CRISPR/Cas9

Gracias a la gran rapidez con la que avanza la tecnología de edición CRISPR/Cas9 nos planteamos emplear la técnica co-CRISPR (H. Kim et al., 2014; Paix et al., 2017) para desarrollar los nuevos modelos de investigación dada su alta eficiencia y rapidez. Empleamos una adaptación realizada por el grupo de investigación del Dr. Julián Cerón (Instituto IDIBELL, Barcelona, España). Esta estrategia radica en introducir una mutación puntual en el gen *dpy-10*, a la vez que se modifica el gen de interés. La modificación del

gen *dpy-10* produce una alteración de la cutícula del gusano, haciendo que solo pueda desplazarse mediante movimientos en espiral (movimiento *rol*) (Arribere et al., 2014), si ha habido una edición en *dpy-10*. Este co-marcador nos indica que la maquinaria CRISPR ha funcionado correctamente en el animal inyectado y que, con cierta probabilidad, también podría haberse modificado el gen de interés (un 5-50 % de la descendencia *Roller* en la F1 presentan la edición en el gen de interés, dependiendo de los guías empleados) (Paix et al., 2017). Si, por el contrario, se produce el corte por la nucleasa pero no la edición por recombinación homóloga en *dpy-10*, los gusanos desarrollarían un tamaño anormalmente corto del cuerpo o fenotipo *dumpy*. (Figura 2.9). En ambos casos sabemos que CRISPR ha funcionado, y que es posible que nuestro gen de interés pueda estar editado.

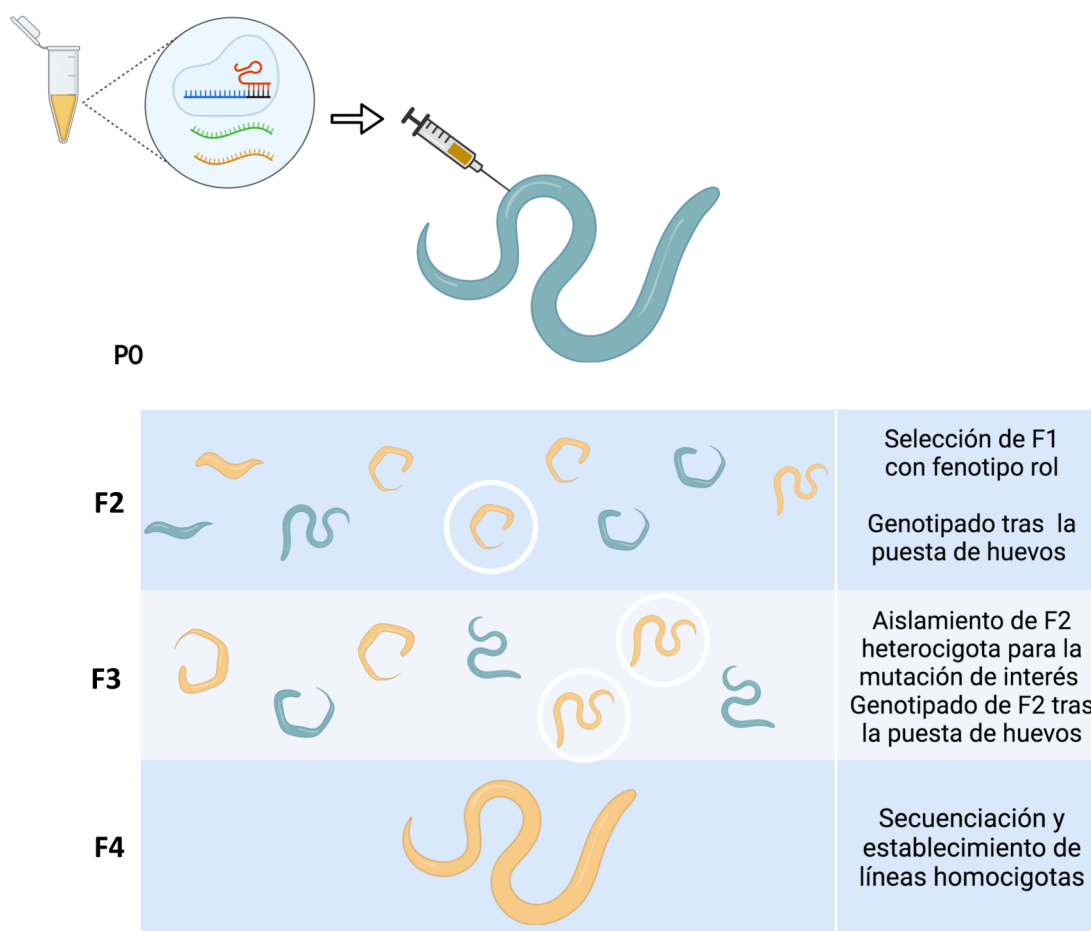


Figura 2.9: Estrategia de edición genética co-CRISPR. Los complejos de RNPs pre ensamblados se inyectan en la gónada de un hermafrodita adulto. Los animales F1 que presentan el fenotipo rol son aislados, permitiendo que se reproduzcan por autofecundación. Tras la puesta de huevos, la F1 es genotipada para detectar los gusanos heterocigotos. Los gusanos F2 de las placas heterocigotas son aislados en placas individuales para su autofecundación. De nuevo, tras la puesta de los huevos se genotipan para identificar a los homocigotas. Una vez establecidas las líneas homocigotas se procede a la confirmación de la edición mediante secuenciación Sanger. Imagen creada con Biorender.com.

Las secuencias guías se diseñaron con el predictor *web* <http://crispr.cos.uni-heidelberg.de> (Stemmer et al., 2015). Como criterios de selección se consideraron: alto contenido en nucleótidos guanina y citosina (50-75 %); presencia de la secuencia (N18)GGNGG (Farboud et al., 2015); evitar el nucleótido citosina en la posición 20 (N19)CNGG (Doench et al., 2014); eficiencia y especificidad. Con el objetivo de editar el gen *idh-2* se valoraron los guías cuyo punto de corte estimado estuviera más cercano a la posición de la mutación (**Anexo I, Tabla 2.5**).

Los *crRNA* correspondientes fueron sintetizados por *Integrate DNA Technologies*, así como el *crRNA* para el gen *dpy-10* (Arribere et al., 2014; S. Kim et al., 2014). Al igual que los empleados para editar líneas celulares, estos *crRNA* presentan una secuencia que permite la hibridación del *tracrRNA*. La edición se realizó mediante *ssODNs*. Las secuencias *ssODN* además de la mutación de interés se diseñaron con una serie de cambios silenciosos (**Tabla 2.6**). Esto permite la detección del alelo mutante en los gusanos mediante *PCR* con un *primer* específico y diferenciar gusanos homocigotos y heterocigotos para la mutación.

Para llevar a cabo la co conversión, se mezcló la nucleasa recombinante Alt-R™ S.p. Cas9 Nuclease 3NLS con Alt-R™ CRISPR-Cas9 *tracrRNA*, los AltR™ CRISPR-Cas9 *crRNA* y los *ssODN*. Las concentraciones finales fueron: 4,5 μM Nucleasa Cas9, 32 μM *tracrRNA*; 10 μM *crdpy-10*, 35 μM del *crRNA* de interés, 1,22 μM *ssODN dpy-10 (cn64)* y 175 μg/μL *ssODN* de interés. La mezcla se centrifuga (10 min a 13000 rpm) y se incubó a 37 °C durante 15 min para la formación de los complejos *RNPs*. La mezcla se inyectó en la gónada de gusanos hermafroditas N2. La descendencia F2 con fenotipo *rol* es separada en placas individuales para confirmar el genotipo mediante *PCR*. Una vez confirmada la presencia de la mutación en heterocigosis, se aíslan 20 gusanos de la descendencia F3 para identificar los homocigotos. Mediante secuenciación Sanger se confirma la presencia de la mutación en la F4.

Por último, se estudiaron los posibles efectos inespecíficos del sistema CRISPR/Cas9. La herramienta crispr.cos.uni-heidelberg.de empleada para el diseño también predice los potenciales *off targets* de los guías seleccionados. De esta forma se obtuvo un *off-target* para el guía *idh-2_R123Q* y tres para *idh-2_R155K* (**Tabla 2.7**).

Tras extraer el ADN genómico (consultar apartado de Genotipado), se realizó una *PCR* con *primers* específicos para estudiar estas regiones. Mediante secuenciación *Sanger* se comprobó la presencia de mutaciones. Tras comprobar la ausencia de mutaciones, las cepas se retrocruzaron con la cepa salvaje N2 para asegurar que no haya ninguna modificación no deseada.

5. Genotipado de gusanos por *PCR*

Para el genotipado se recogieron 12 gusanos en un tubo eppendorf de 0,5 mL con 10 μ L de solución de lisis (consultar apartado Materiales específicos). Los tubos se incubaron 1 h a 65 °C y 15 min a 95 °C para lisar los gusanos y obtener el ADN genómico. Una vez finalizada la incubación se emplearon 2 μ L de la lisis directamente en la *PCR*. Las *PCRs* de genotipado se realizaron con la Taq ADN polimerasa PCR Master Mix 2x (Thermo Fisher Scientific K0172).

Para la detección de las mutaciones en *idh-2* mediante *PCR* se empleó un *primer* reverso homólogo a la mutación y los cambios silenciosos introducidos por el *ssODN*. El *primer* reverso específico del alelo salvaje solo se hibrida en la secuencia salvaje. Para ambos se emplea el mismo *primer* directo. Los resultados de la *PCR* se visualizaron en gel de agarosa al 1 %. Finalmente, mediante secuenciación *Sanger* se confirmó la presencia de las mutaciones.

6. Sincronización de gusanos

Se emplearon tres placas Petri con gusanos adultos con huevos para sincronizar la población mediante la solución de lejía. Los gusanos se recogieron con solución M9 y se traspasaron a un tubo eppendorf de 15 mL mediante una pipeta de vidrio. El tubo se centrifugó (1 min, 1500 rpm) para depositar los gusanos. En el sobrenadante quedó la solución M9 con restos de bacteria OP50. El pellet de gusanos se resuspendió en 4 mL de solución de lejía (consultar apartado Materiales específicos) y se agitó durante 5 min. Una vez lisados los gusanos se centrifugó y se lavó el pellet de huevos obtenido con solución M9, eliminando los restos de lejía, que puedan afectar al crecimiento de los huevos. Tras un segundo lavado los huevos se mantuvieron en 2 mL de solución M9 y se incubaron a 20 °C en agitación durante la noche para que eclosionaran. Al día siguiente

se realizó el conteo del número de larvas obtenido. Este procedimiento se realizó por duplicado para obtener un número suficiente de gusanos.

7. Detección y cuantificación del oncometabolito 2-HG

Veinte mil gusanos L1 sincronizados se cultivaron en placas Petri de 10 cm con OP50 (10000 gusanos/placa). Se cultivaron durante 48 h o 72 h a 20 °C hasta que la población alcanzó el estadio de larva L4 o adulto. Tras alcanzar el estadio, los 200000 gusanos se recogieron con solución M9 y se traspasaron a un tubo eppendorf de 15 mL. Los tubos se centrifugaron para eliminar los restos de bacteria un par de veces en solución M9, y un último paso en agua Milli-Q. En el último lavado se eliminó todo el sobrenadante, dejando el pellet de gusanos sin líquido y se congeló a -80 °C.

La lisis de los gusanos y la detección y cuantificación de los dos enantiómeros de 2-HG se llevó a cabo en la Unidad Analítica (UA) del IIS La Fe. Para ello la Dr. Marta Roca desarrolló un protocolo específico. La lisis de los gusanos se realizó empleando 25 µL del pellet de gusanos con 250 µL MeOH: H₂O (50:50). Se empleó el homogeneizador Precellys (Bertin Instruments), siguiendo un programa de 2 ciclos de 30 s, con una pausa de 10 s, a 6500 rpm. Se recogió el sobrenadante y se centrifugó 5 min a 13000 rpm, a 4 °C. De este extracto se emplearon 150 µL para concentrar y continuar el análisis. El resto del volumen se empleó para cuantificar la proteína total en las muestras mediante el método BCA (Thermo Fisher, Ref 23227) siguiendo el protocolo comercial. Los 150 µL concentrados se resuspendieron en 50 µL 5:95 H₂O: MeOH, ácido acético 0,3 % e hidróxido de amonio 0,1 %. Por último, la muestra se centrifugó (5 min a 13000 rpm, a 4 °C) y se pasó a la placa de inyección para su análisis mediante cromatografía líquida de alta resolución acoplada a espectrometría de masas. El análisis se realizó en un equipo UPLC-HRMS Q Exactive Plus con detector *Orbitrap*. La separación cromatográfica se realizó mediante una columna quiral para la separación de los enantiómeros, sin necesidad de derivatizar. Para el método de cromatografía líquida se empleó una columna Astec Chirobiotic R, 15 cm*2,1, 5µm; T^a columna 25; fase móvil (isocrático) 5:95 H₂O: MeOH, ácido acético 0,3 % e hidróxido de amonio 0,1 %; flujo 0,8 mL/ min; y volumen de inyección 10 µL. Para el *Orbitrap*, las condiciones de *Parallel Reaction Monitoring (RPM)* se detallan en la **Tabla 2.8**. Las transiciones de masas, tanto para L-

2HG como D-2HG fueron 147.0298 > 129 .0193 (cuantitativa); 147.0298 > 103.0401 > 85.0295 (cualitativa).

Tabla 2.8: Parámetros empleados en el RPM

General	
<i>Run time</i>	0 a 5 min
<i>Polarity</i>	Negativa
<i>In-source CID</i>	0.0 eV
<i>Dynamic RT</i>	<i>off</i>
<i>Default charge state</i>	2
<i>Inclusion</i>	<i>on</i>
MS ²	
<i>Microscans</i>	1
<i>Resolution</i>	70,000
<i>AGC target</i>	1e6
<i>Maximum IT</i>	200 ms
<i>Loop count</i>	1
<i>MSX count</i>	1
<i>MSX isochronous ITs</i>	<i>on</i>
<i>Isolation window</i>	1.5 m/z
<i>Isolation offset</i>	0.0 m/z
<i>Fixed first mass</i>	-
<i>(N)CE/ Stepped NCE</i>	10, 20, 40
<i>Spectrum data type</i>	Centroid

Material y Métodos comunes

Secuenciación Sanger

Las secuenciaciones Sanger realizadas en este trabajo se realizaron en tres servicios distintos, según la disponibilidad en cada momento: Servicio de Genética del Hospital La Fe, Plataforma de Genómica del Instituto de Investigación Sanitaria La Fe y el servicio de secuenciación de la empresa Stab Vida.

Las secuenciaciones realizadas en las Unidades del Hospital y del IIS La Fe se realizaron mediante el siguiente procedimiento. Los productos de PCR fueron purificados mediante la enzima *ExoProStar* (Sigma, Ref US77705V) siguiendo el protocolo comercial. La secuenciación de vectores se realizó empleando 100 ng de estos. Tanto la PCR purificada como los vectores se someten a la reacción de secuenciación con *BigDye™ Terminator v1.1 Cycle Sequencing Kit* (Applied Biosystems, Ref 4337450) siguiendo las indicaciones del protocolo. Se emplea 1 µL del *primer* de interés (3,2 µM), 2 µL del tampón específico, 1 µL de Big Dye y la PCR purificada o los 100 ng del vector. Una vez finalizada la reacción, los productos se cargan en el secuenciador ABIPRISM 3130 DNA Analyzer (Applied Biosystems) para su secuenciación

y la obtención de las secuencias de formato FASTA. Las secuencias producidas se visualizan mediante el programa Chromas.

Las secuenciaciones realizadas en el servicio de Stab Vida se realizaron según especifica el protocolo de la empresa. Tras purificar la *PCR* con el mismo protocolo comentado en el párrafo anterior, se envían al servicio de secuenciación junto a 3 μL del *primer* (10 μM). Stab Vida se encarga de la recepción, procesamiento y secuenciación de la muestra. Al finalizar el proceso facilitan las secuencias formato FASTA para su análisis mediante el programa Chromas.

RESULTADOS

RESULTADOS

Capítulo I. Edición *in vitro*

El sistema de edición CRISPR/Cas9 se ha empleado para desarrollar numerosos modelos celulares de mutaciones relacionadas con la LMA. La estrategia más establecida para introducir los elementos CRISPR en las células leucémicas es mediante transducción con partículas víricas, principalmente lentivirus, debido a la baja eficiencia de transfección de este tipo celular. El desarrollo de lentivirus específicos para la edición de diferentes dianas se trata de una estrategia laboriosa, muy costosa y que presenta numerosos riesgos de bioseguridad.

Con el objetivo de desarrollar una estrategia más sencilla que permita editar de forma eficiente el genoma de líneas celulares leucémicas se ha puesto a punto un sistema de producción de pequeñas construcciones de ADN con el casete de expresión de secuencias guía. Partiendo del vector pEGR1 generado en este trabajo (Material y Métodos, apartado 3) y mediante un sistema de *PCRs* y *primers* específicos se han desarrollado nueve construcciones guía que presentan el promotor de expresión en células de mamífero pU6, una secuencia guía específica, la secuencia codificante del *tracrRNA* y un terminador de la transcripción. Estos elementos permiten la expresión de una molécula guía funcional que dirige a la nucleasa Cas9 a su diana. De forma adicional, estas construcciones pueden presentar acoplada la expresión de la proteína verde fluorescente *GFP*, que permite monitorizar la eficiencia de transfección. Para producir estas construcciones mediante *PCRs* se emplea el plásmido pEGR1 desarrollado en este trabajo como molde.

Para probar la eficiencia de estas construcciones primero se realizó una puesta a punto en las células de fácil manejo HEK293, donde se consiguió editar el gen *MYBL2* e *IDH2*. El gen *MYBL2* codifica un factor de transcripción involucrado en ciclo celular, supervivencia y regulación de la diferenciación celular. Desregulaciones de su expresión se han asociado con un gran abanico de tipos de cáncer, entre ellos la LMA (Fuster *et al.*, 2013; Musa *et al.*, 2017). En trabajos previos de nuestro grupo de investigación se optimizaron guías CRISPR para este gen, por lo que se empleó para la optimización del corte de la nucleasa Cas9. Por otro lado, la optimización de la introducción de una mutación concreta se realizó empleando el gen *IDH2*. Este gen codifica una enzima del

ciclo de los ácidos tricarboxílicos. Las mutaciones más recurrentes en este gen (R140 y R172) producen una ganancia de función en la proteína y cataliza la producción del oncometabolito D-2-HG. Este metabolito desemboca en alteraciones epigenéticas y de expresión génica y participa en el desarrollo de algunos tipos de LMA. Se escogió introducir la mutación R172 por su relevancia en la estratificación de los pacientes y a que está menos estudiada que la R140.

Tras la optimización en HEK293, se trasladó esta estrategia a células leucémicas con expresión constitutiva de la nucleasa hSpCas9. De forma paralela, se aplicaron varias estrategias para aumentar la eficiencia de edición con estas construcciones. Durante el transcurso de este trabajo se produjo la explosión del uso de las ribonucleoproteínas CRISPR y una mayor disponibilidad comercial de proteínas Cas9 recombinantes. Las *RNPs* se aplicaron en la edición del gen *IDH2* y se comparó con las construcciones desarrolladas. Para finalizar, mediante secuenciación masiva de última generación se validó y caracterizó la edición producida en el gen *IDH2* y los posibles cortes inespecíficos producidos por la nucleasa con las distintas estrategias empleadas (**Figura 3.1**).

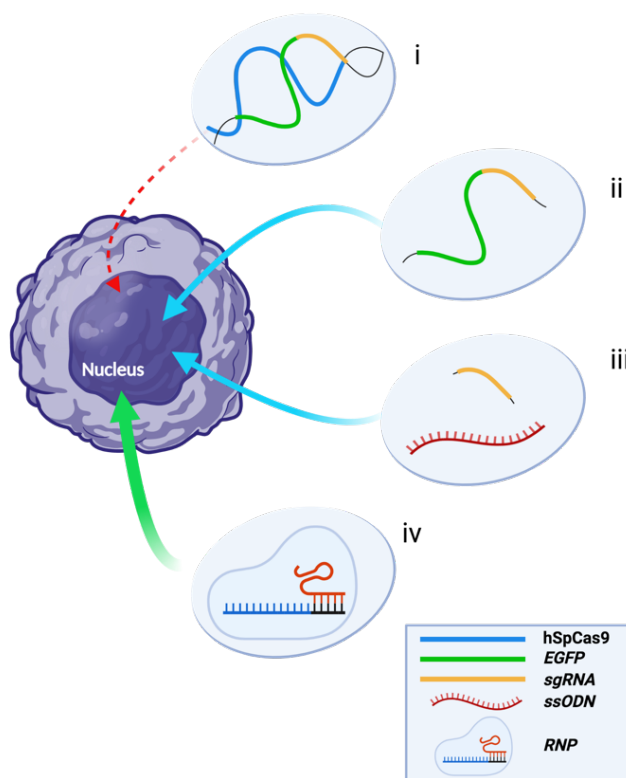


Figura 3.1: Distintas estrategias implementadas para editar las células leucémicas NB4. i. Plásmidos codificantes de la nucleasa Cas9, *sgRNA* y el marcador *EGFP*; ii. Construcciones codificantes de *sgRNA* y *EGFP*; iii. Construcciones codificantes del *sgRNA* junto a *ssODN* como ADN molde de edición; iv. Ribonucleoproteínas. Las flechas indican las eficiencias de transfección y edición. Ordenadas de menor a mayor eficiencia serían la flecha roja discontinua, azul y verde. Figura creada con Biorender.com

1. Desarrollo de un sistema de producción de los elementos CRISPR mediante PCR y optimización en células HEK293

Nuestra estrategia comienza con la creación de un vector que permita producir secuencias cortas codificantes del casete de expresión de un *sgRNA* concreto junto al casete de expresión de *EGFP*. Para ello se generó el vector pEGR1. Este vector presenta los elementos esenciales para producir las construcciones que inducen la expresión de los *sgRNA* y del reportero *EGFP*. Los vectores comerciales disponibles no presentan estos elementos de forma contigua, si no que se encuentran separados por otros elementos, como reguladores de la expresión en bacterias o marcadores de selección como resistencia a antibióticos. Por ello se desarrolló un vector propio donde estos elementos se encuentran próximos entre si y es posible amplificarlos de forma conjunta mediante *PCRs*.

Los elementos de expresión del *sgRNA* en el vector pEGR1 presentan todos los componentes necesarios para su expresión. Primero, el promotor pU6, un promotor de expresión constitutiva en células de mamífero que permite la expresión de ARN pequeños por la RNA polimerasa tipo III, produciendo transcritos cortos y de secuencia muy concreta (Orioli et al., 2012). Por otro lado, presenta la secuencia conservada del *sgRNA*, formada por el sitio de clonación de la secuencia guía y la secuencia del *tracrRNA*, elemento necesario para la unión a la nucleasa Cas9. Finalmente, presenta la señal de terminación poli (A) que evita la generación de colas de poliadenina. De forma adicional, es posible generar estas construcciones con el reportero *GFP*. La expresión de este marcador está controlada por el promotor/*enhancer* CMV y la secuencia terminadora y de poliadenilación SV40 poli(A). El promotor CMV es un promotor dependiente de la ARN polimerasa II de fuerte expresión en cualquier tipo celular y ampliamente utilizado para la expresión ectópica de genes exógenos (Z. L. Xu et al., 2001).

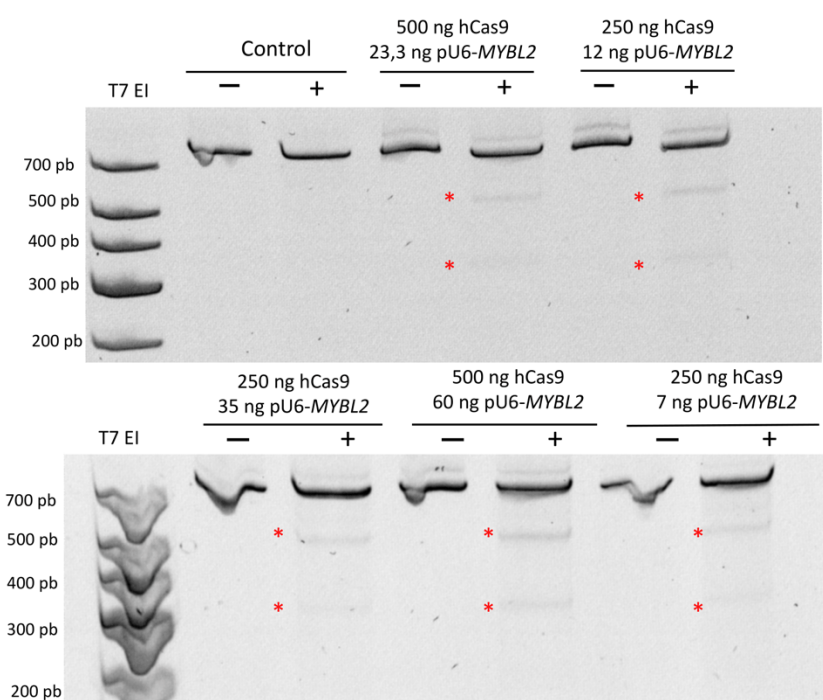
Empleando este vector como molde se desarrolló un protocolo de *PCRs* que permite producir construcciones con un *sgRNA* específico con un tamaño de 465 pb y construcciones con la guía y el marcador *EGFP* de 2033 pb. Para ello, primero se generan distintas *PCRs* con los componentes básicos. En una *PCR* se amplifica la región de expresión de *EGFP* junto al promotor pU6 y una secuencia guía específica. Por otro lado,

mediante otra combinación de *primers* se amplifica solamente el promotor pU6 junto a la secuencia guía. La última *PCR* necesaria presenta la secuencia guía y la secuencia de terminación poli (A). Finalmente, gracias a la homología en la región de las secuencias guía es posible generar las construcciones pU6-*sgRNA* o *EGFP*-pU6-*sgRNA* incorporando las secuencias terminadoras.

Siguiendo esta estrategia, mediante dos rondas de *PCR* se generó una construcción con expresión de un guía para el gen *MYBL2* (pU6-*MYBL2*). La construcción pU6-*MYBL2* se empleó para optimizar las condiciones de edición en las células HEK293. Primero se estudió la cantidad de construcción necesaria para producir corte en la diana genómica co-transfectando la construcción guía y el vector hCas9, codificante de la nucleasa hSpCas9. Para ello se estudiaron distintos ratios construcción/vector: 1:1, 1:3 y 1:5, y distintas cantidades de vector. El ADN genómico de las células transfectadas se extrajo a las 48h, y mediante ensayo de la enzima T7 El se calcularon las eficiencias de corte. En la **Tabla 3.1** se detallan los resultados obtenidos en cada réplica. Se observa cómo se correlacionan el aumento de la cantidad de vector hCas9 y construcción guía con el aumento de la eficiencia de corte. Sin embargo, a partir de 250 ng de vector hCas9 y 35 ng de construcción guía la eficiencia disminuye. Por ello, la condición más eficiente es el ratio 1:3 con 250 ng de vector hCas9 y 35 ng de construcción, con una media de corte de 10,3 %, seguida del ratio 1:5 con 250 ng hCas9 (8 %). El ratio 1:1 con 500 ng de construcción guía presenta una eficiencia de corte similar (7,66 %). Por último, las condiciones menos eficientes fueron los ratios 1:1, con 250 ng de hCas9 (4,31 %) y 150 ng de hCas9 (3,43 %) (**Figura 3.2**). Por ello, se seleccionó la condición de 250 ng de hCas9 y 35 ng de construcción guía para los siguientes experimentos.

Tabla 3.1. Recopilación de los porcentajes de *indels* obtenidos en las diferentes optimizaciones con la construcción pU6-MYBL2 en células HEK293.

Condiciones	500 ng hCas9- 23.3 ng pU6-MYBL2 (Ratio 1:1)	250 ng hCas9- 12 ng pU6-MYBL2 (Ratio 1:1)	250 ng hCas9-35 ng pU6-MYBL2 (Ratio 1:3)	250 ng hCas9- 60 ng pU6-MYBL2 (Ratio 1:5)	150 ng hCas9-7 ng pU6-MYBL2 (Ratio 1:1)	250 ng hCas9- 35 ng pU6-MYBL2-P	500 ng pX458-MYBL2	140 ng PCR Cas9-35 ng pU6-MYBL2	250 ng PCR Cas9- 35 ng pU6-MYBL2
Réplica 1	5,20	3,45	5,24	6	4,5	14	4,3	1,83	3
Réplica 2	3	5,8	8	6	3	8	11	1,44	7
Réplica 3	14	4	13	14,7	3,7	6	3	0,2	2,4
Réplica 4	8,43	4	14,79	5,3	2,5	13,9	13		
Media <i>Indels</i> (%±ES)	7,66±2,3	4,31±0,5	10,26±2,2	8±2,2	3,43±0,4	10,49±2	7,83±1,2	1,16±0,5	4,2±1,5

**Figura 3.2: Selección del ratio vector-construcción guía óptimo en células HEK293.** Resultados del ensayo de la enzima T7 EI en gel de poliacrilamida al 10%. Las bandas producidas por la enzima se indican con asteriscos rojos.

Manteniendo ese ratio, en los siguientes experimentos se estudió el efecto de la fosforilación de los extremos 5' de la construcción pU6-MYBL2. En estudios previos otros autores registraron un aumento de la expresión de productos de *PCR* con los extremos 5' fosforilados (Castanotto *et al.*, 2004). Al co-transfectar la construcción fosforilada con el vector hCas9 no se observó un aumento de la eficiencia de corte. Por otro lado, ambas construcciones presentan una mayor eficiencia de corte que el vector pX458 con la misma secuencia guía para el gen *MYBL2* (7,8 %) (Tabla 3.1) (Figura 3.3).

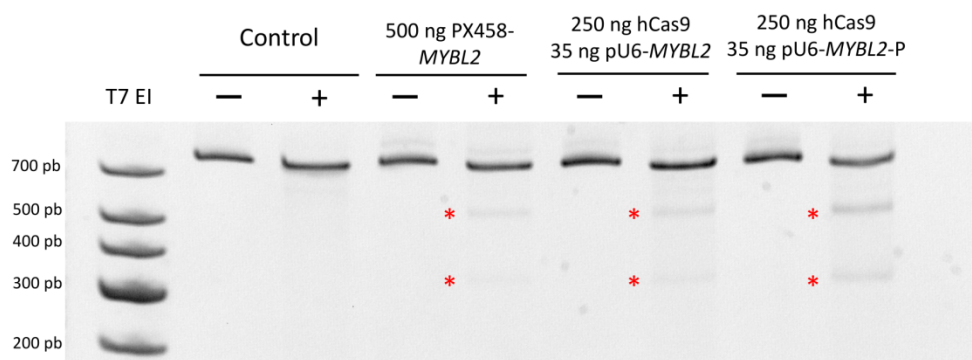


Figura 3.3: Resultados de edición con las construcciones generadas y vectores CRISPR en HEK293. Comparación de la eficiencia de corte entre el vector PX458, la construcción pU6-MYBL2 y la construcción pU6-MYBL2-P. Las bandas producidas por la enzima se indican con asteriscos rojos.

Por último, se comprobó si el uso de una construcción con los elementos esenciales para la expresión de hSpCas9 producida mediante *PCR* producía mayor eficiencia. Esta construcción, al amplificar solo los elementos necesarios para la expresión de la nucleasa, presenta un tamaño mucho más pequeño que el vector hCas9 (5346 pb frente a 9553 pb), ya que no se incluyen los elementos necesarios para la replicación en bacterias y las resistencias a antibióticos

Para ello se amplificó el casete de expresión de hSpCas9 del vector hCas9 (promotor CMV, gen Cas9 y señal poli A). Se co-transfectó 140 ng y 250 ng del casete de hSpCas9 y la cantidad de construcción pU6-sgMYBL2 se ajustó para mantener el ratio 1:3. Los resultados con la enzima T7 EI muestran eficiencias de corte más bajas que con el vector hCas9, con una media de 1,16 % empleando 140 ng de PCR Cas9 y 4,2 % con 250 ng de PCR Cas9 (**Tabla 3.1**).

2. El uso de dos guías aumenta la eficiencia de corte en el gen *IDH2*

Partiendo de las optimizaciones realizadas en los apartados anteriores se prosiguió con la introducción de la mutación *IDH2*^{R172}. Tras confirmar la ausencia de *SNPs* en la región a editar que podrían inhibir la identificación de la diana por el complejo CRISPR en *IDH2* se emplearon los predictores web <http://chopchop.cbu.uib.no/> (Labun *et al.*, 2019) y <http://www.e-crisp.org/E-CRISP/> (Heigwer *et al.*, 2014) para escoger las secuencias guía más eficientes. Partiendo de 8 secuencias guía se generaron 8 construcciones mediante el sistema de *PCRs* sobre el vector pEGR-1.

Cada construcción se co-transfectó con el vector hCas9 en las células HEK293. Además, se testó la co-transfección de las dos construcciones pU6-*IDH2*_3 y pU6-

IDH2_8 juntas dado que ambas flanquean la posición de la mutación a introducir. A las 48 h post transfección, se aisló el ADN de las células y se prosiguió con el análisis de las eficiencias de edición obtenidas. Los resultados obtenidos tras realizar el ensayo de la enzima T7 EI se muestran en la **Tabla 3.2**. Las construcciones 4, 5 y 7 presentan las eficiencias más bajas (2,66, 2,34 y 4,62 % respectivamente), mientras que la construcción 8 registra la media de eficiencias más alta (10,75 %). Por último, la transfección combinada de los guías 3 y 8 resultó ser la más eficiente, consiguiendo duplicar la eficiencia de la construcción 8 con un 21,14 % de eficiencia media de corte (**Figura 3.4.**)

Tabla 3.2: Recopilación de los porcentajes de *indels* producidos con los distintos guías diseñados para el gen *IDH2* y de la eficiencia de la PCR del casete de la nucleasa Cas9.

Condiciones	<i>IDH2_1</i> *	<i>IDH2_2</i> *	<i>IDH2_3</i> *	<i>IDH2_4</i> *	<i>IDH2_5</i> *	<i>IDH2_6</i> *	<i>IDH2_7</i> *	<i>IDH2_8</i> *	<i>IDH2_3 + IDH2_8</i> **	250 ng PCR Cas9- <i>IDH2_3 + IDH2_8</i> **
Réplica 1	7,54	9,28	7,13	5,27	2,34	8,61	4,62	19	24,244	4,45
Réplica 2	8	13	15	2,66	1,17	3,4	3,15	11	28,80	7
Réplica 3	4	5,4	5,4	1,9	2,43	5,34	5,82	9	11,80	2,4
Réplica 4	5,95	5,5	5,5					4	19,70	
Media de <i>Indels</i> (%±ES)	6,37±0,91	8,3±1,81	8,26±2,45	2,66±1	2,34±0,4	5,34±1,52	4,62±0,77	10,75±3,1	21,14±3,6	4,2±1,5

*Se empleó el ratio vector hCas9-construcción guía más eficiente de la tabla 3.1.

**Las cantidades de construcción guía fueron 17,5 ng de cada una, para un total de 35 ng.

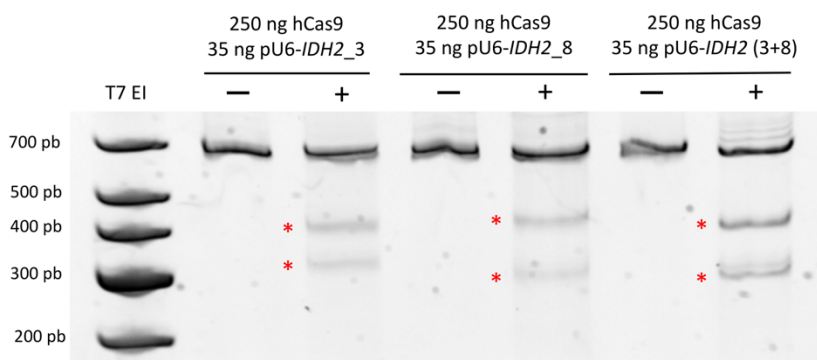


Figura 3.4: Eficiencias de corte producidas con las guías diseñados para *IDH2*. Se muestran los resultados de las guías más eficientes. Las bandas producidas por la enzima se indican con asteriscos rojos. Los resultados de las guías *IDH2_1*, *IDH2_2*, *IDH2_4*, *IDH2_5*, *IDH2_6* e *IDH2_7* se detallan en la Tabla 3.2.

A continuación, se testó la eficiencia de corte sustituyendo el vector hCas9 por el casete de expresión de hSpCas9 amplificado mediante *PCR*. A pesar de haber obtenido una eficiencia baja al emplearlo sobre el gen *MYBL2*, se decidió probarlo en *IDH2* para descartar que esta eficiencia fuera debido a la diana escogida. Tras co-transfectar el casete de expresión de hSpCas9 junto a las dos construcciones guía se obtuvo una media

de eficiencia menor que empleando el vector hCas9 (4,2 %), al igual que en la diana *MYBL2* (Tabla 3.2).

Por último, se llevó a cabo la optimización de la concentración necesaria de *ssODN* para introducir la mutación *IDH2*^{R172}. Para ello se testaron 2 cantidades de *ssODN* (5 μ M y 10 μ M), junto a 250 ng de vector hCas9 y 35 ng totales de construcciones guía. Ambas concentraciones de *ssODN* producen eficiencias de edición similares (1,28 y 1 % respectivamente) (Figura 3.5) (Tabla 3.3).

Para aumentar las eficiencias de edición se probó la estrategia *cold shock*. En el trabajo de Guo *et al.* se describe como la incubación de las células a 32 °C aumenta la eficiencia de edición (Guo *et al.*, 2018). Siguiendo este trabajo, tras transfectar las células HEK293 con vector, construcciones guía y las dos concentraciones de *ssODN*, se incubaron a 32 °C durante 48 h. Pasados dos días se observó que se había producido mucha muerte celular y muchas células se habían levantado de la placa. Debido a las pocas células viables no fue posible realizar una extracción de ADN óptima para el ensayo de la enzima T7 E1.

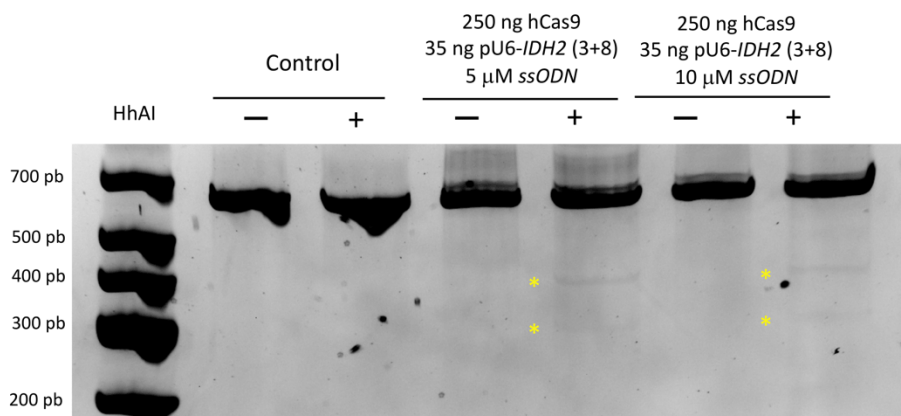


Figura 3.5: Edición en células HEK293. Resultados en gel de poliacrilamida al 10%. Los asteriscos amarillos indican las bandas generadas por la enzima de restricción al detectar la región editada.

Tabla 3.3: Eficiencias obtenidas de edición en el gen *IDH2* empleando 5 μ M y 10 μ M de *ssODN*.

Condiciones	pU6- <i>IDH2</i> _3 + pU6- <i>IDH2</i> _8 + 5 μ M <i>ssODN</i> *	pU6- <i>IDH2</i> _3 + pU6- <i>IDH2</i> _8 + 10 μ M <i>ssODN</i> *
Réplica 1	1	0,5
Réplica 2	1,5	0,86
Réplica 3	1,3	1,17
Media de <i>Indels</i> (%\pmES)	1,28 \pm 0,13	1,02 \pm 0,12

*Se emplean las concentraciones optimizadas en los puntos anteriores. 250 ng del vector hCas9 y 35 ng totales de las construcciones pU6-*IDH2*_3 y pU6-*IDH2*_8.

3. Creación de líneas celulares leucémicas con expresión constitutiva de la nucleasa Cas9

Una vez probada la funcionalidad de las construcciones desarrolladas en las células HEK293 se pasó a trabajar con las líneas celulares leucémicas. Debido a la baja eficiencia de transfección de vectores plasmídicos de gran tamaño en estas células, se generaron líneas celulares con expresión constitutiva de la nucleasa hSpCas9. Para ello se emplearon lentivirus portadores del vector de transferencia LentiCRISPR V.2. Este vector codifica la nucleasa hSpCas9 y el gen de resistencia al antibiótico puromicina y secuencias específicas que permiten su integración en el genoma de las células huésped.

Las líneas celulares leucémicas disponibles en el laboratorio NB4, HL60, MOLM-13 y Katsumi-1 se infectaron con este lentivirus. Tras la infección, las células se cultivaron en presencia del antibiótico puromicina para seleccionar aquellas donde se había insertado el vector lentiviral. Posteriormente, se extrajo el ADN de estas células para detectar por *PCR* la presencia del gen codificante de Cas9. Mediante esta prueba se pudo detectar la inserción en las células NB4, HL60 y MOLM-13 infectadas, que pasaron a nombrarse NB4-Cas9, HL60-Cas9 y MOLM-13-Cas9. En las células Katsumi-1 no se detectó la presencia de la nucleasa.

4. Alta eficiencia de transfección de las construcciones guía desarrolladas en las células NB4-Cas9

Tras generar las líneas NB4-Cas9, HL60-Cas9 y MOLM-13-Cas9 se estudió la funcionalidad de la nucleasa Cas9 integrada. Para ello se comprobó la eficiencia de corte en el gen *IDH2* mediante la nucleofección de la construcción pU6-*IDH2*_1, codificante del guía 1 diseñado para *IDH2*. Para ello se empleó la tecnología de nucleofección de Amaxa.

En el momento de la realización de estos experimentos, Amaxa no ofrecía un kit de nucleofección específico para la línea MOLM-13. Por ello, se descartó esta línea para los siguientes experimentos. En las otras dos líneas se nucleofectaron distintas cantidades de la construcción guía. Para ello se adaptó la cantidad de construcción guía necesaria partiendo del protocolo comercial, correspondiendo a 266 ng de construcción de 447 pb. También se testaron 400 y 800 ng. Tras los análisis con la enzima T7 EI solamente se visualizó corte en la condición de células NB4-Cas9 nucleofectadas con 800 ng de construcción (**Figura 3.6**). Dado que no se observó corte en las pruebas realizadas en HL60-Cas9 se descartó esta línea y se continuó únicamente con las células NB4-Cas9.

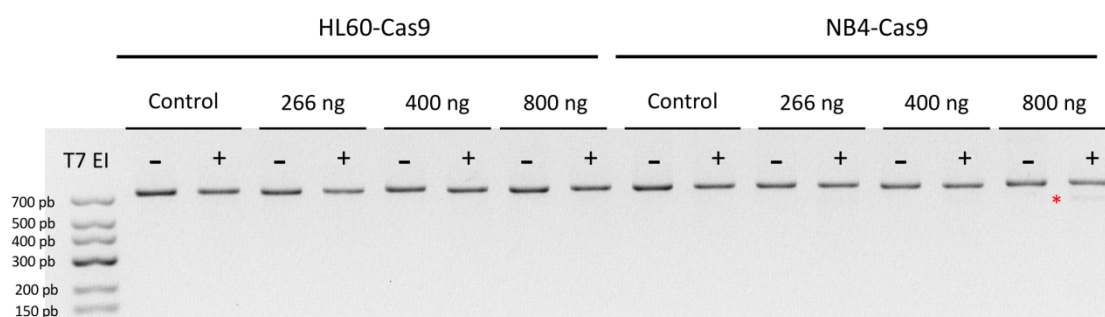


Figura 3.6: Funcionalidad de la nucleasa hSpCas9 insertada en las líneas celulares NB4-Cas9 y HL60-Cas9. Detección de la producción de *indels* trans nucleofectar distintas cantidades de la construcción pU6-*IDH2_1*. Las bandas producidas por la enzima se indican con asteriscos rojos.

A continuación, se estudió la eficiencia de nucleofección empleando la construcción *GFP*-pU6-*MYBL2* producida mediante el sistema de *PCR*. Esta construcción codifica el reportero *GFP* además del guía CRISPR contra el gen *MYBL2*. Con el objetivo de comparar la eficiencia de nucleofección de la construcción con otros vectores CRISPR, se nucleofectaron en paralelo el vector pX458 y el vector comercial pMAXGFP. El plásmido pX458 es un vector comercial empleado en numerosos estudios de edición génica *in vitro*. Expresa la nucleasa hSpCas9 y el reportero fluorescente *EGFP* en células de mamífero. El plásmido pMAXGFP también codifica el reportero *EGFP* y es suministrado en los kits de nucleofección de Amaxa como control positivo de nucleofección. A las 24 h post nucleofección se analizó el ratio las células *GFP* positivas en cada condición mediante citometría de flujo, comprobando así la capacidad de la construcción y de los plásmidos de introducirse en el interior celular. Se observó que el ratio era mayor en las células nucleofectadas con la construcción *GFP*-pU6-*MYBL2* que la obtenida con el vector comercial pX458 (una media de 39,1 % \pm 7, frente a la media

de 2,36 % \pm 0,78 del vector). El ratio de células *GFP* positivas del vector pMAX*GFP* fue de 25,7 % \pm 0,88 (Figura 3.7).

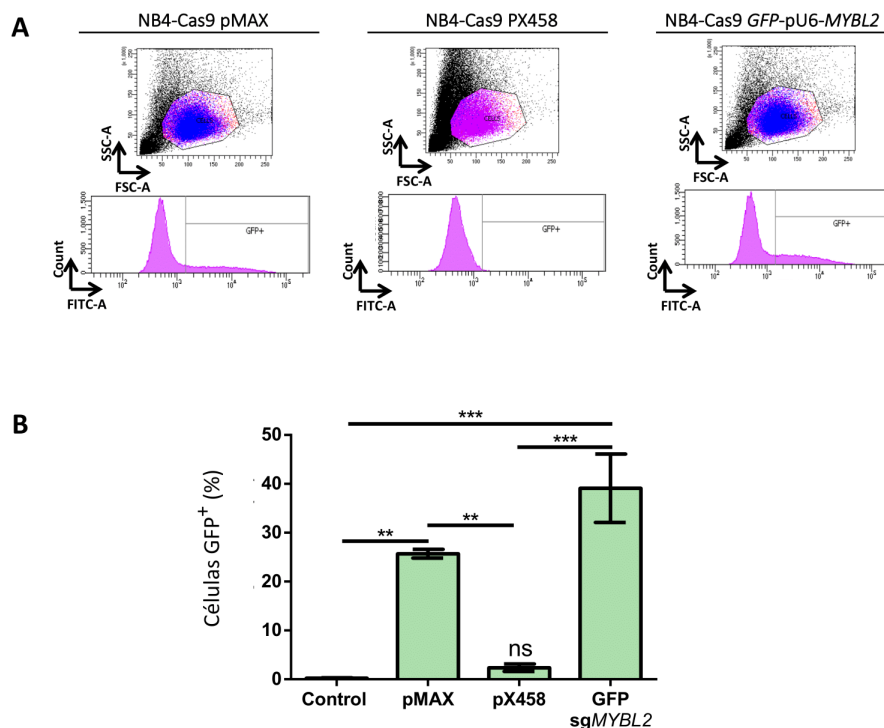


Figura 3.7. Eficiencia de transfección de las construcciones desarrolladas. A) Detección mediante citometría de flujo de células NB4-Cas9 *GFP* positivas nucleofectadas con el vector pMAX, el vector PX458 o la construcción GFP-pU6-sg*MYBL2*. En azul se muestran los ratios de células *GFP* positivas obtenidas en cada experimento. B) Las eficiencias obtenidas con la construcción son significativamente diferentes a la fluorescencia de las células control sin transfectar ($p < 0,001$) y transfectadas con el vector pX458 ($p < 0,001$). La eficiencia del vector pMAX*GFP* también es significativamente diferente al control ($p < 0,01$) y a pX458 ($p < 0,01$). A pesar de observarse una mayor eficiencia al nucleofectar la construcción *GFP*-pU6-sg*MYBL2* y el vector pMAX*GFP*, no existe una diferencia significativa (Test ANOVA unidireccional, con Test Tukey).

Por último, para esclarecer si la ausencia de edición en las células HL60-Cas9 se debía a una nucleofección ineficiente se nucleofectó la misma construcción en estas células. A las 24 h se estudió la eficiencia de transfección mediante citometría de flujo obteniéndose una media alta de células *GFP* positivas, 54,25 % \pm 4,9.

5. Las construcciones guía producen la misma eficiencia de edición que los complejos *RNPs*, pero menor porcentaje de *indels*.

Tras la comprobación de que las construcciones guía producidas mediante *PCRs* se nucleofectan con mayor eficiencia en las células NB4-Cas9 que los vectores CRISPR clásicos, se prosiguió con la edición del gen *IDH2*. Para ello se nucleofectaron las dos guías más eficientes de *IDH2* (pU6-*IDH2*_3 y pU6-*IDH2*_8) en las células NB4-Cas9.

Para optimizar las condiciones, se testaron las concentraciones, 800 ng y 1,5 μg totales de ambas construcciones. Ambas condiciones produjeron eficiencias de corte similares tras los análisis con la enzima T7 EI (13,26 % y 14,42 % respectivamente) (**Figura 3.8**). Se seleccionó la cantidad de 1,5 μg para continuar la optimización de la edición. En cuanto al *ssODN*, se probaron distintas concentraciones del *ssODN* (10, 100, 200, 300 y 500 μM) y del *ssODN* anti sentido (100 y 500 μM). De estas condiciones, el *ssODN* a 100 μM fue la única en dar resultados en el ensayo de *RFLP*, con una media de eficiencia de edición del 2,2 % (**Figura 3.9**) (**Tabla 3.4**). Además de este resultado, a altas concentraciones del *ssODN* (200, 300 y 500 μM) se obtuvo un aumento de la muerte celular. Las concentraciones testadas del *ssODN* anti sentido no produjeron muerte celular, pero tampoco edición.

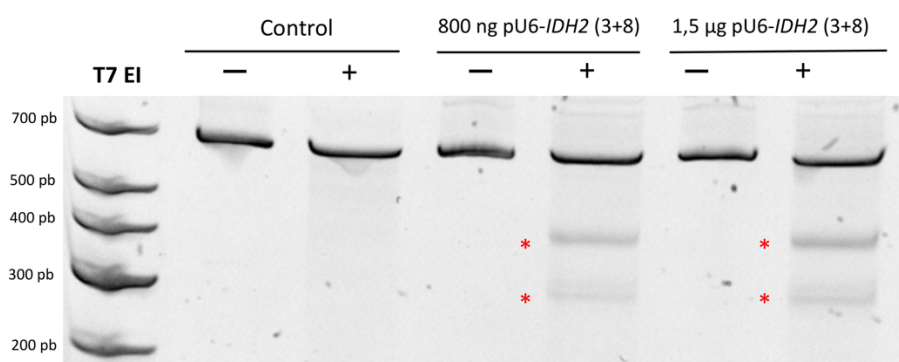


Figura 3.8. Optimización edición en NB4-Cas9. Se muestran los resultados obtenidos con dos cantidades de construcción guía para editar el gen *IDH2* en las células NB4-Cas9. Las bandas producidas por la enzima se indican con asteriscos rojos

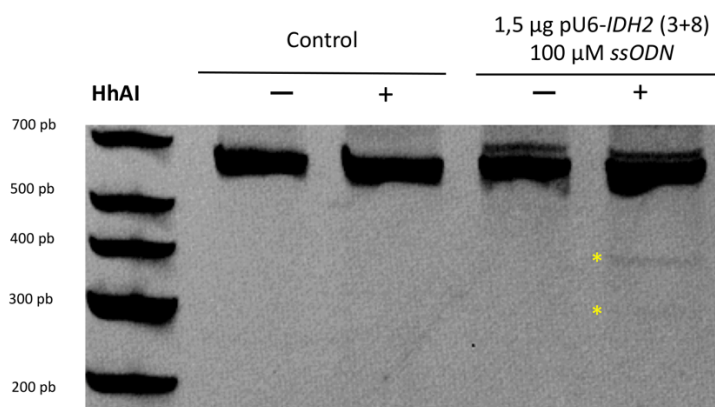


Figura 3.9: Ensayo *RFLP* para analizar el porcentaje de edición en células NB4-Cas9. Las bandas producidas por la enzima se indican con asteriscos amarillos. Se ha modificado la saturación de la imagen para poder visualizar mejor las bandas generadas.

Tabla 3.4: Eficiencias de producción de *indels* y de edición mediante el uso de las construcciones desarrolladas y los complejos *RNPs*.

Condiciones	800 ng (pU6- <i>IDH2_3</i> + pU6- <i>IDH2_8</i>)	1,5 µg (pU6- <i>IDH2_3</i> + pU6- <i>IDH2_8</i>)	1,5 µg (pU6- <i>IDH2_3</i> + pU6- <i>IDH2_8</i>) + 100 µM <i>ssODN</i>	<i>RNP</i> (<i>sgIDH2_3</i> + <i>sgIDH2_8</i>)	<i>RNP</i> (<i>sgIDH2_3</i> + <i>sgIDH2_8</i>) + 100 µM <i>ssODN</i>
Réplica 1	14	21,56	2,8	28,7	2,5
Réplica 2	11,8	7	2,28	34,16	1,6
Réplica 3	14	15	1,5	26,7	2
Media de <i>Indels</i>/Edición (%±ES)	13,26±1,27	14,52±7,29	2,2±0,37	29,85±2,23	2±0,26

Tras probar la eficiencia de edición de la metodología desarrollada nos propusimos compararla con la eficiencia de los complejos de *RNPs*. Estos complejos se emplean en numerosos estudios de edición génica debido a su facilidad de transfección, altas eficiencias de edición y bajo porcentaje de cortes inespecíficos u *off-targets* (Liang *et al.*, 2015). Estos complejos están formados por la nucleasa hSpCas9 recombinante y una molécula guía de ARN (*sgRNA*) formada por *crRNA* y *tracrRNA* sintetizados *in vitro*. Para ello se generaron *sgRNA* para cada uno de los guías empleados, *sgIDH2_3* y *sgIDH2_8*.

Se nucleofectaron 40,6 µg de *sgRNAs* entre ambos guías y 34 µg de nucleasa Cas9 en células NB4 sin expresión constitutiva de Cas9. Para los estudios de edición se empleó 100 µM del *ssODN*. Mediante los ensayos de la enzima T7 EI y *RFLP* se concluyó que los complejos de *RNPs* producen una mayor eficiencia de corte en *IDH2* (29,85 %), pero igual eficiencia de edición de la mutación R172 (2 %) (**Figura 3.10**) (**Tabla 3.4**).

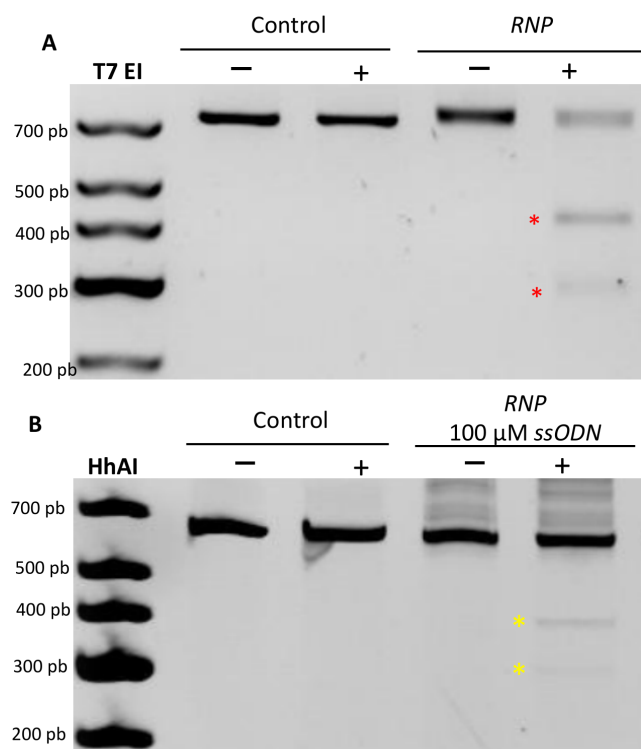


Figura 3.10. Edición mediante complejos RNP en células NB4. A) Ensayo de la enzima T7 EI de células NB4 transfectadas con los complejos RNPs y los guías *sgIDH2_3* and *sgIDH2_8*. Las bandas producidas por la enzima se indican con asteriscos rojos. B) Ensayo RFLP mediante la enzima HhAI de células editadas mediante complejos RNPs y *ssODN*. Las bandas producidas por la enzima se indican con asteriscos amarillos.

Por último, se probó a nucleofectar el amplificado por *PCR* del casete de expresión de la nucleasa hSpCas9 y las construcciones guía en la línea NB4. Tras el ensayo de T7 EI no se obtuvo ningún producto de corte en los genes *MYBL2* e *IDH2*.

6. El análisis de la edición producida con las construcciones CRISPR por secuenciación masiva muestra resultados similares de edición génica sin efectos *off targets*

Con el objetivo de validar y caracterizar la edición producida y comprobar los posibles cortes inespecíficos se empleó la secuenciación masiva de última generación. Mediante el predictor *web* Cas-OFFinder (<http://www.rgenome.net/cas-offinder/>) se obtuvieron los potenciales *off targets* de los guías *IDH2_3* e *IDH2_8*. Como criterio de selección de los potenciales *off targets* se consideró una diferencia máxima de 3 nucleótidos con la secuencia de los guías de *IDH2*. De esta forma se obtuvieron 4 potenciales *off targets* para el guía *IDH2_3* y 12 para el guía *IDH2_8*. De estos, 7 se encuentran en regiones intrónicas, 6 en regiones intragénicas y 3 en regiones intergénicas.

El *pool* de células NB4-Cas9 editadas mediante construcciones guía (NB4-Cas9 Edición), NB4 editadas mediante *RNPs* (NB4 *RNP*) y NB4 editadas con el casete de expresión de hSpCas9 en forma de *PCR* y construcciones guía (NB4 *PCR*) se emplearon para generar una librería de secuenciación por amplicones de los 16 *off targets* y el *on target IDH2*. El ADN de células NB4 y NB4-Cas9 sin editar se emplearon como controles. En esta estrategia, a cada amplicón generado se le añadió un adaptador que presenta una etiqueta que su origen. De esta forma es posible identificar a qué condición pertenece cada lectura producida.

Los archivos *fastq.gz* obtenidos del secuenciador se analizaron mediante el programa CRISPResso2 (Clement *et al.*, 2019) que realiza una cuantificación precisa de los eventos de edición de forma sencilla para investigadores sin experiencia en bioinformática. El programa CRISPResso2 realiza de forma automática varios procesos: i. Filtrado de las lecturas de baja calidad, ii. Eliminación de los adaptadores o *trimming* si fuera necesario, iii. Alineamiento de las secuencias obtenidas con el amplicón de referencia y iv. Cuantificación de la proporción de edición producida por *NHEJ* y *HDR*. Para la cuantificación el programa emplea una ventana de cuantificación estrecha centrada en el sitio teórico de corte de la nucleasa Cas9, unas 3 pb aguas arriba de la secuencia PAM. El empleo de esta ventana permite diferenciar los cambios debido a la edición de errores introducidos durante la secuenciación.

Las coberturas obtenidas en las muestras NB4-Cas9, NB4-Cas9 Edición y NB4 *PCR* fueron altas, con más de 90.961, 98.705 y 94.661 lecturas alineadas por amplicón, respectivamente. En la muestra NB4 *RNP*, todos los amplicones obtuvieron una baja cobertura, siendo el *off target* 8.4 el amplicón menos cubierto, con 132 lecturas alineadas, y el *off target* 8.11 el más cubierto, con 510 lecturas. La muestra NB4 se secuenció con muy baja cobertura. Al no alcanzar los requisitos mínimos del programa no fue posible analizarla (**Figura 3.11**). Por otro lado, en la secuenciación se generó una gran cantidad de lecturas indeterminadas. Es probable que estas lecturas correspondan a las muestras NB4 y NB4 *RNP*, y que por un error en la detección de los *index* no se clasificaran correctamente. Debido a que el secuenciador realiza el *trimming* o la eliminación de la secuencia de los adaptadores no fue posible identificar la procedencia de las lecturas.

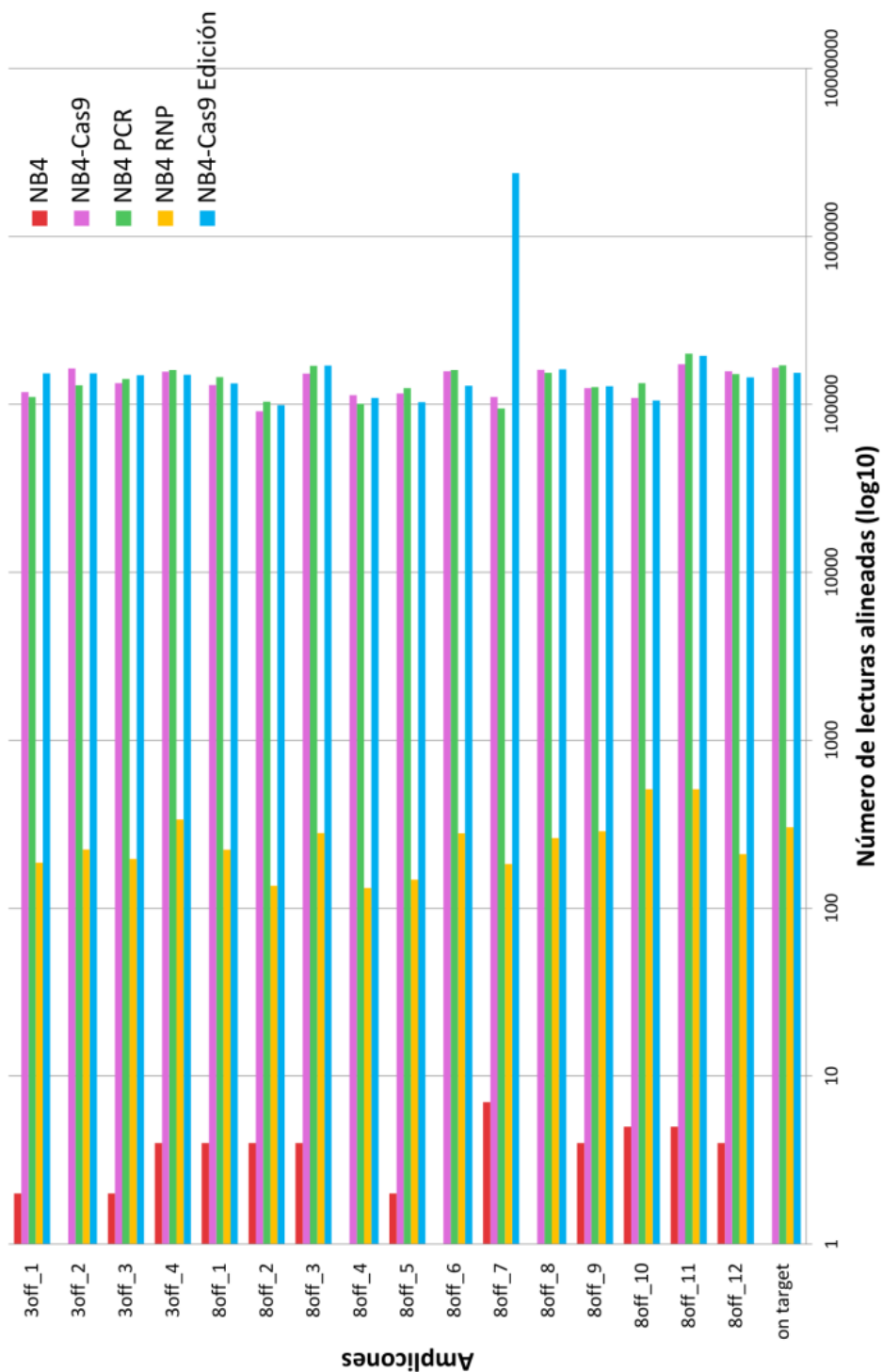


Figura 3.11: Lecturas obtenidas para cada amplicón. Los datos se muestran en base logarítmica de 10.

En la **Figura 3.12** se muestra la clasificación de las lecturas realizada por el programa CRISPResso2 para cada muestra. Al analizar las lecturas, la muestra NB4-Cas9 Edición presentó un 10,95 % de lecturas del *on target IDH2* reparadas por el mecanismo *NHEJ* tras el corte de la nucleasa Cas9. La principal modificación producida con este tipo de reparación son deleciones (10,47 % de las lecturas totales), donde el

rango principal de deleción es de 20-29 pb (5,80 % de las lecturas totales), pero alcanzando en algunas lecturas los 112 pb. Con porcentajes más bajos también se encuentran lecturas con sustituciones de bases nucleotídicas (0,44 % de las lecturas totales) e inserciones de bases (0,38 % de las lecturas totales) (**Figura 3.13.B**). La muestra NB4 *RNP* presenta un 22,37 % de lecturas donde se observa una reparación por *NHEJ*. Al igual que se observa en NB4-Cas9 Edición, el principal motivo de reparación en este grupo ha sido la deleción (20,72 % del total), donde el rango principal abarca de 20-29 pb (9,54 % de las lecturas totales). En esta muestra las inserciones y las sustituciones también se dan en menor porcentaje (1,97 % y 0,33 % respectivamente) (**Figura 3.13.C**). Al analizar los resultados de la edición mediada por *HDR* tras el corte de Cas9, obtuvimos en la muestra NB4-Cas9 Edición un 2,44 % de lecturas editadas. Un 0,42 % de estas presentaron todos los cambios del *ssODN*, menos la modificación de la secuencia PAM del *sgIDH2_3*. En la muestra de células NB4 editada con *RNP* se detectó un 1,64 % de lecturas editadas con todos los cambios del *ssODN*.

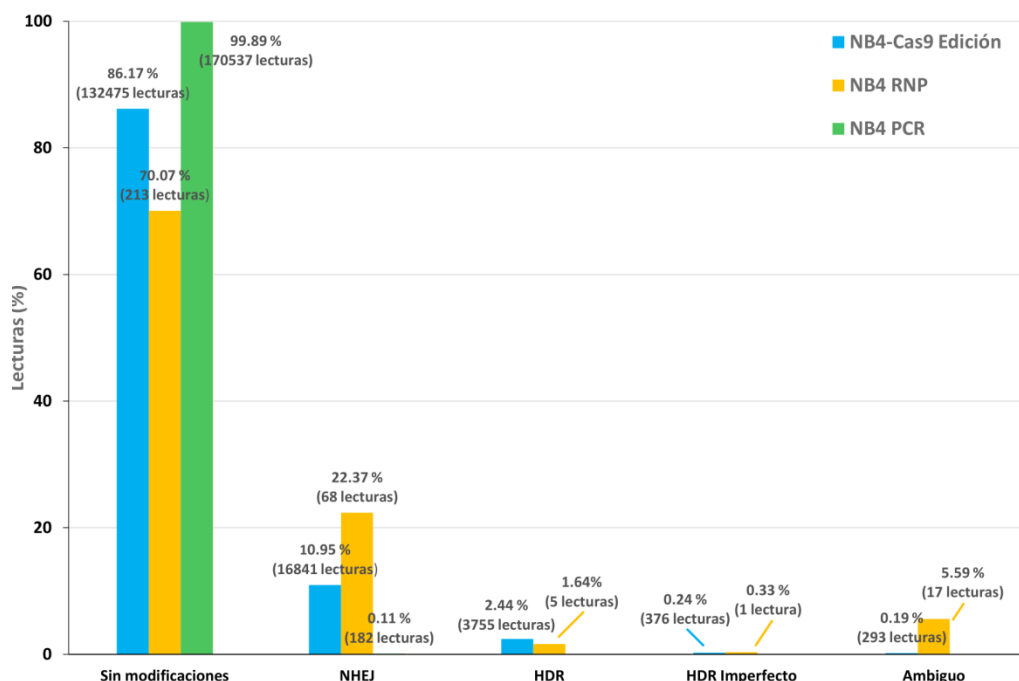


Figura 3.12: Clasificación de las lecturas obtenidas por el programa CRISPResso2. Se clasifican como lecturas sin modificaciones, lecturas editadas con reparaciones mediadas por el mecanismo de *NHEJ*, lecturas editadas mediante el mecanismo *HDR*, lecturas editadas mediante el mecanismo *HDR* pero con más modificaciones y lecturas ambiguas.

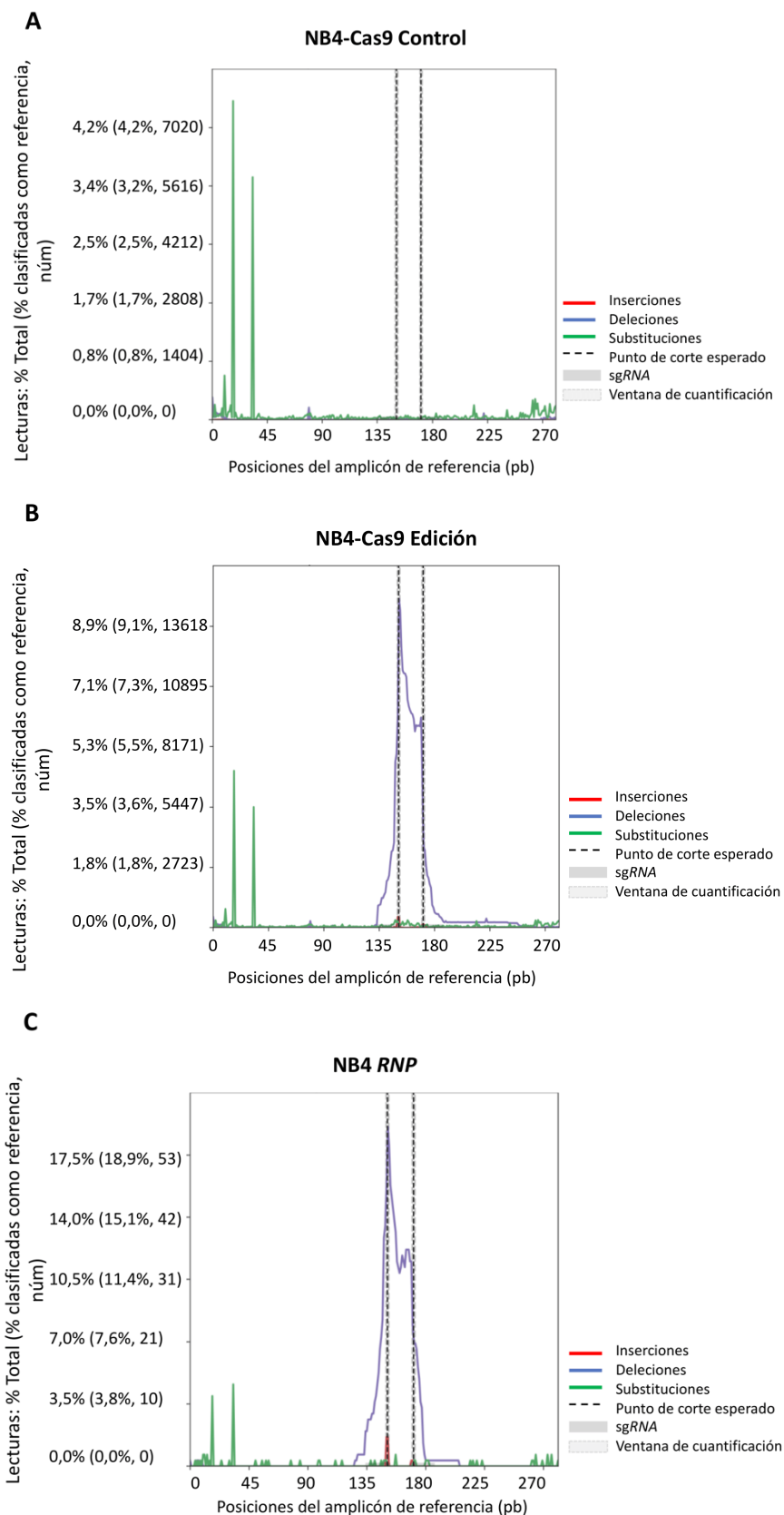


Figura 3.13: Gráficas de la distribución de las modificaciones producidas sobre *IDH2*. Se muestran los resultados de las muestras NB4-Cas9 Control (A), NB4-Cas9 Edición (B) y NB4 RNP (C). Se representa el porcentaje de lecturas con cambios según la posición en el amplicón de referencia. Con distintos colores se identifica si el cambio es una inserción (rojo), deleción (azul) o sustitución (verde). Se observa como el tipo de cambio más frecuente son las deleciones. Las posiciones que más cambios

presentan son aquellas que se encuentran dentro de la ventana de cuantificación (delimitada por los dos puntos de corte predichos de las secuencias guía), aunque también se observan cambios aguas arriba del guía *sgIDH2_8* y aguas abajo de *sgIDH2_3*.

En ambas muestras el programa CRISPResso2 clasificó un pequeño porcentaje de lecturas en la categoría de *HDR* imperfecto (0,24 % en NB4-Cas9 Edición y 0,33 % en NB4 *RNP*) (**Figura 3.12**). Estas lecturas presentan otras modificaciones, como deleciones y/o inserciones, además de los cambios del *ssODN*.

En la **Figura 3.14** se detallan las principales lecturas con cambios obtenidas en estas dos muestras. En la muestra NB4 *PCR* no se detectó edición ni corte en *IDH2*.

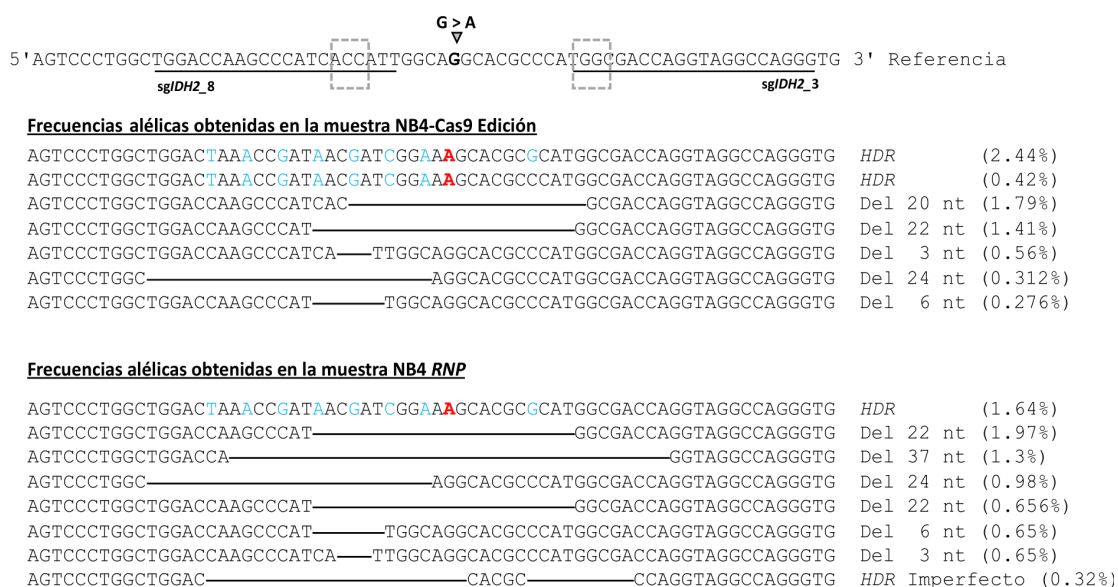


Figura 3.14. Principales frecuencias alélicas editadas. En la primera línea se muestra la secuencia de referencia. Subrayadas se señalan las regiones de los dos *sgRNA*. Los cuadrados con guiones señalan las ventanas de cuantificación empleadas por el programa. La flecha indica la posición de la mutación a introducir. Los nucleótidos azules son los cambios introducidos al emplear el *ssODN* como molde en la reparación por *HDR*. Los nucleótidos rojos señalan la introducción de la mutación R172. Las deleciones producidas durante la reparación por *NHEJ* se indican con líneas.

Por último, el programa empleado clasificó un 0,19 % de las lecturas de NB4-Cas9 Edición y un 5,59 % de NB4 *RNP* en la categoría de secuencias ambiguas (**Figura 3.12**). Estas secuencias ambiguas presentan deleciones del rango de 22-99 pb, que se salen del rango marcado por la ventana de cuantificación del programa. Si las modificaciones no se encuentran en esta región el programa las considera ambiguas. Todas las lecturas clasificadas como ambiguas en la muestra NB4-Cas9 Edición presentaron deleciones que afectan al punto de corte teórico para *sgIDH2_8*, además de su secuencia diana. En

cuanto a la región de *sgIDH2_3*, el 75,42 % de las lecturas ambiguas habían perdido el punto de corte para este guía pero mantienen la secuencia diana. El 24,57 % restante también había perdido esta última región. En cuanto a la muestra NB4 *RNP*, todas las lecturas ambiguas habían perdido las regiones de corte esperado y secuencia diana para ambas guías.

Para realizar un estudio más completo de los posibles cambios en las secuencias *off-targets*, los resultados obtenidos para las muestras NB4-Cas9 control y NB4-Cas9 Edición se compararon entre sí y con la secuencia de referencia mediante el programa CRISPResso2 *Compare*. En la mayoría de los *off-targets* no se obtuvieron variaciones significativas entre ambas muestras ni respecto a la referencia (**Anexo II, Figura 3.15**). Los *off-targets* 3.2, 3.3, 8.2, 8.4, 8.7, 8.9, 8.10 y 8.11 presentaron lecturas que diferían respecto a la secuencia de referencia pero se encontraban en porcentajes similares en la muestra NB4-Cas9 control y en NB4-Cas9 Edición, indicando que se trataba de variaciones características de la línea celular (**Anexo II, Figura 3.15**). La comparación entre NB4 control y NB4 *RNP* no se pudo realizar debido a la baja cobertura de la muestra control. Los *off-targets* de la muestra NB4 *RNP* se estudiaron directamente con el programa CRISPResso2. No se obtuvieron variaciones significativas en los 16 *off-targets*.

Capítulo II. Edición *in vivo*

1. Establecimiento de posibles modelos de las mutaciones en *idh2*

Debido a su papel esencial en el ciclo de Krebs, el gen *IDH2* mantiene una alta conservación entre especies, presentando 264 genes ortólogos (**Figura 3.16**). El gen *idh-2* de *C. elegans* presenta un porcentaje de similitud del 67,82 % con su ortólogo humano según la base de datos Ensembl y la proteína codificada un 78 % según la base de datos DIOPT (**Figura 3.17**). Ambas proteínas presentan dominios similares con las mismas funciones, siendo dominios de unión a substrato, NADP, magnesio y un dominio catalítico.

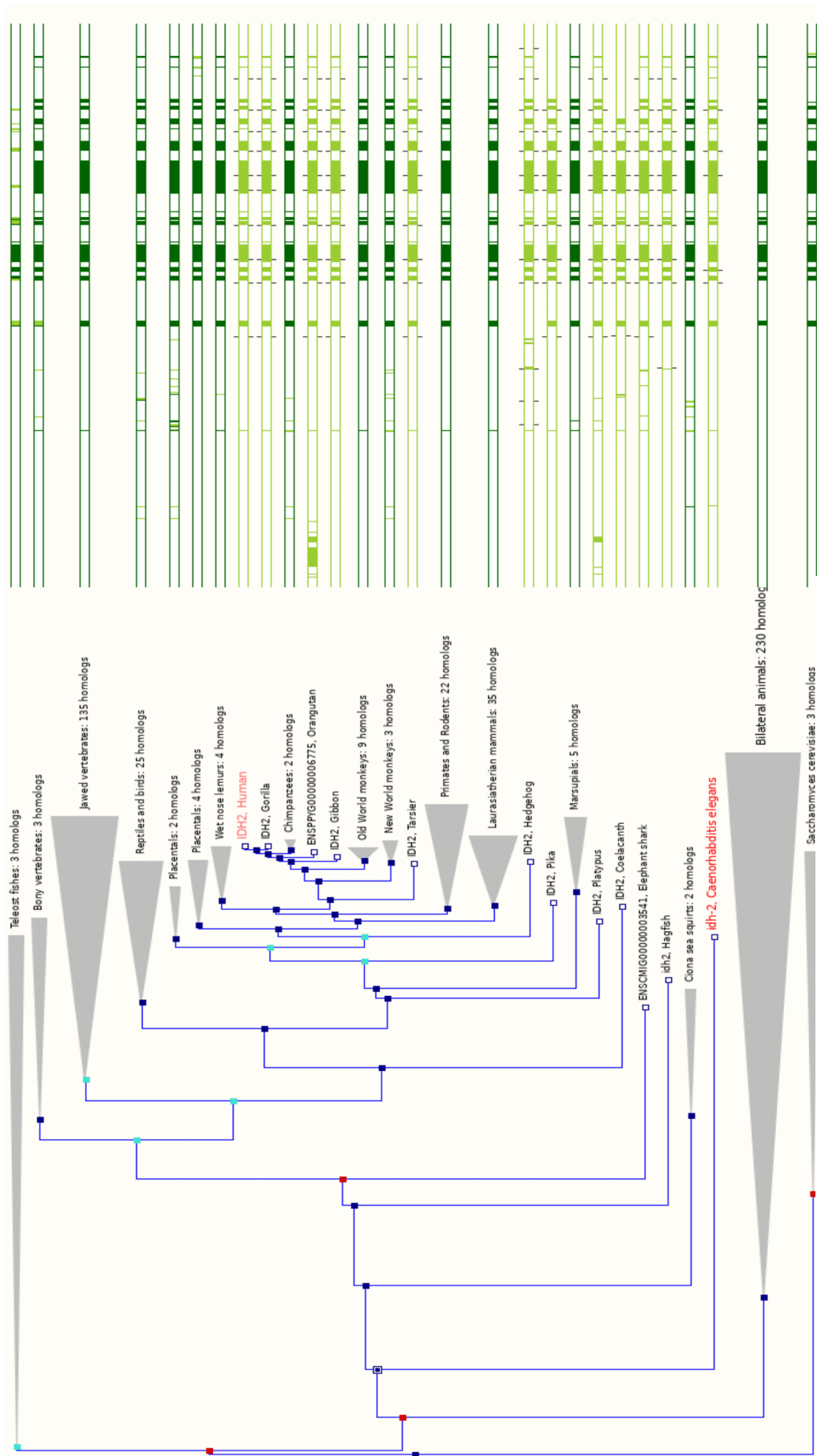


Figura 3.16: Árbol evolutivo del gen *IDH2*. Obtenido con la base de datos Ensembl.

en el gen *idh-2*, un *knock-in* en el gen *dpy-10*. La modificación de este gen produce una alteración del fenotipo muy evidente. A pesar de que ambos eventos de edición no están ligados, el cambio fenotípico debido a la rotura y/o edición en *dpy-10* permite seleccionar aquellos gusanos donde la inyección de los componentes CRISPR ha sido correcta y donde la maquinaria de reparación celular ha resuelto ese corte producido por la nucleasa Cas9.

Mediante esta estrategia se generó la cepa RVM479 *idh-2* (vlt24), con la mutación *idh2*^{R123Q} y la cepa RVM480 *idh-2* (vlt25), con la mutación *idh2*^{R155K}. La cepa RVM481 *idh-2* (vlt26), con ambas mutaciones, se generó cruzando las cepas anteriores. Mediante genotipado por *PCR* se seleccionaron aquellos gusanos homocigotos para las mutaciones y se confirmaron mediante secuenciación *Sanger*. No se detectó ninguna modificación de los potenciales *off targets* analizados mediante secuenciación *Sanger*.

2. Estudio de la presencia del oncometabolito 2-HG en las cepas RVM479, RVM480 y RVM481

Tras desarrollar las dos cepas de *C. elegans* con las mutaciones homólogas a R140Q y R172K (RVM479 y RVM480), y la cepa RVM481 con ambas, se procedió a estudiar si se producía el oncometabolito 2-HG en estos animales.

Tras sincronizar las cepas, se cultivaron 200000 gusanos en placas con medio NGM y bacteria OP50. Estos gusanos se recogieron en los estadios de larva 4 (L4) o adulto y se congelaron a -80 °C. En total se recogieron cinco réplicas de cada cepa, incluida la cepa control silvestre (N2), en estadio L4 y tres réplicas de cada cepa en estadio adulto. La lisis de los animales se realizó en la Unidad Analítica del IIS La Fe mediante la tecnología Precellys. El extracto se dividió en dos partes para realizar la cuantificación de proteína mediante BCA, para posteriormente normalizar los resultados, y la cuantificación del oncometabolito. A pesar que el enantiómero D-2-HG es el que se ha asociado con pacientes de LMA, se incluyó también el enantiómero L-2-HG para caracterizar en mayor medida el efecto de las mutaciones.

La detección y cuantificación de los dos enantiómeros del 2-HG se realizó en la Unidad Analítica mediante cromatografía líquida de alta resolución acoplada a espectrometría de masas siguiendo un protocolo puesto a punto para este análisis (**Figura 3.18**). Las concentraciones detectadas se normalizaron con los valores de la

cuantificación de proteína para obtener μg del enantiómero por mL. En la **Tabla 3.5** se especifican los valores obtenidos para cada enantiómero en cada muestra analizada, y en la **Tabla 3.6** se muestran las concentraciones medias. Al analizar los resultados se observa que no hay diferencias en las concentraciones de los enantiómeros entre las cepas de un mismo estadio (**Figura 3.19**). En el estadio L4 las cepas estudiadas presentan valores que van de 0,0232 a 0,0196 μg de D-2HG/mL y 0,0149 a 0,0151 μg de L-2HG/mL. En el caso de los gusanos analizados en estadio adulto se observa que los valores son inferiores, en el rango de 0,00683-0,00819 μg de D-2HG/mL, y 0,00348 – 0,00448 μg de L-2HG/mL.

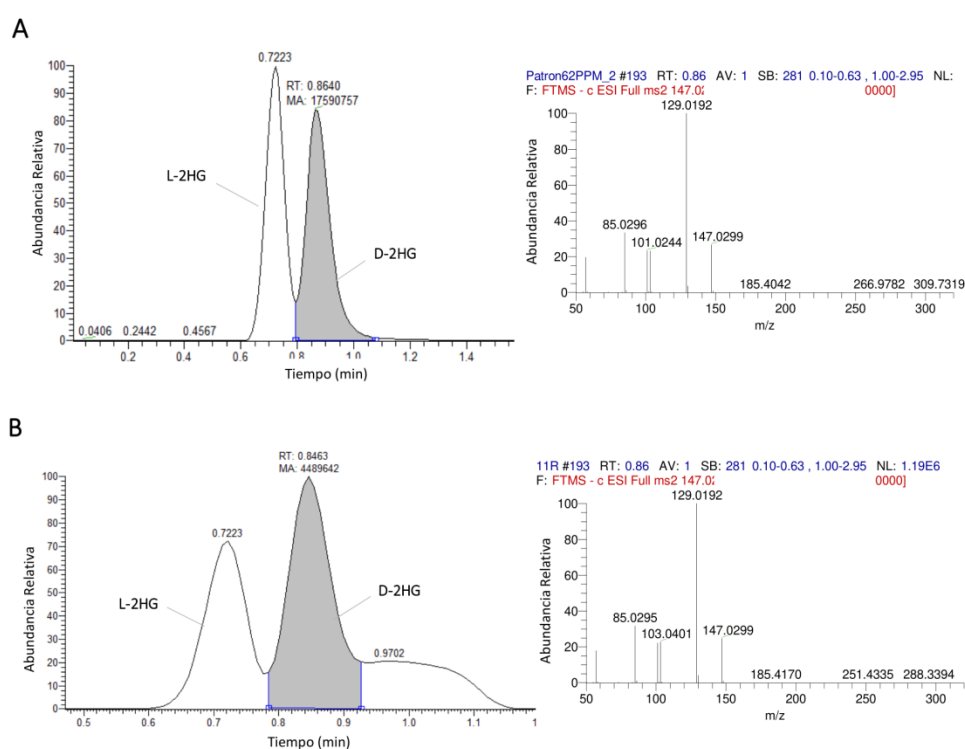


Figura 3.18: Cromatogramas de los enantiómeros L y D y espectros de espectrometría de masas para el D-2-HG. Realizados sobre el patrón de referencia (A) y la muestra a estudiar (B)

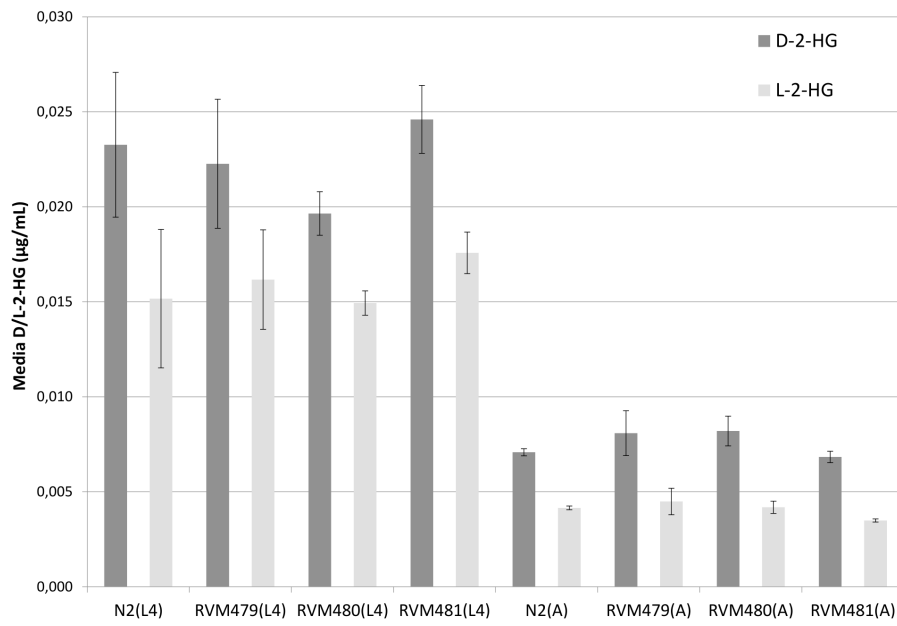


Figura 3.19: Gráfica de las concentraciones medias de D/L-2-HG obtenidas de las distintas cepas estudiadas.

Tabla 3.5. Resultados detección de los enantiómeros D y L del 2-HG mediante LC-MS.

Estadio	Réplica	Cepa	Proteína total (µg/mL)	Concentración D-2-HG (µg/mL)	Concentración L-2-HG (µg/mL)	Concentración D-2-HG corregida (µg/mL)	Concentración L-2-HG corregida (µg/mL)
L4	1	N2	4184,545	158,923	158923,231	0,038	0,027
		RVM479	7370,909	179,649	179648,840	0,024	0,018
		RVM480	5566,364	110,313	110313,328	0,020	0,016
		RVM481	5161,818	162,302	162302,131	0,031	0,021
	2	N2	2998,182	51,630	51630,285	0,017	0,005
		RVM479	4993,636	93,540	93540,216	0,019	0,015
		RVM480	4875,455	87,567	87566,794	0,018	0,014
		RVM481	5052,727	107,903	107902,830	0,021	0,015
	3	N2	5707,273	126,907	126906,575	0,022	0,016
		RVM479	2743,636	44,424	44423,560	0,016	0,011
		RVM480	4670,909	104,819	104818,835	0,022	0,016
		RVM481	4125,455	91,823	91822,576	0,022	0,017
	4	N2	4461,818	95,079	95079,212	0,021	0,016
		RVM479	2789,091	96,591	96591,407	0,035	0,026
		RVM480	5266,364	114,393	114393,209	0,022	0,016
		RVM481	5516,364	133,992	133992,339	0,024	0,017
	5	N2	6952,727	122,157	122156,863	0,018	0,012
		RVM479	5470,909	95,122	95122,109	0,017	0,012
		RVM480	4598,182	74,923	74923,396	0,016	0,013
		RVM481	5389,091	127,488	127488,349	0,024	0,017
Adulto	1	N2	5434,545	40,490	40490,071	0,007	0,004
		RVM479	3684,545	38,395	38394,804	0,010	0,006
		RVM480	4070,909	39,299	39298,563	0,010	0,005
		RVM481	5680,000	35,421	35421,358	0,006	0,004
	2	N2	5757,273	39,756	39755,848	0,007	0,004
		RVM479	5875,455	38,885	38884,899	0,007	0,004
		RVM480	5480,000	38,236	38235,742	0,007	0,004
		RVM481	5984,545	43,073	43072,736	0,007	0,003
	3	N2	5725,455	39,375	39374,635	0,007	0,004
		RVM479	4984,545	35,967	35966,699	0,007	0,004
		RVM480	7193,636	57,256	57255,966	0,008	0,004
		RVM481	6439,091	45,462	45462,083	0,007	0,003

Tabla 3.6. Media de las concentraciones de 2HG obtenidas

	Muestra	Media ($\mu\text{g}/\text{mL}$) \pm ES
D-2HG	N2(L4)	0,0232 \pm 0,00339
	R140(L4)	0,0222 \pm 0,00339
	R172(L4)	0,0196 \pm 0,00114
	R140;R172(L4)	0,0246 \pm 0,00178
	N2(A)	0,00707 \pm 0,000186
	R140(A)	0,00808 \pm 0,00118
	R172(A)	0,00819 \pm 0,000782
	R140;R172(A)	0,00683 \pm 0,000300
L-2HG	N2(L4)	0,0151 \pm 0,00364
	R140(L4)	0,0161 \pm 0,00261
	R172(L4)	0,0149 \pm 0,000641
	R140;R172(L4)	0,0175 \pm 0,00109
	N2(A)	0,004150 \pm 0,00009
	R140(A)	0,00448 \pm 0,000697
	R172(A)	0,00418 \pm 0,000322
	R140;R172(A)	0,00348 \pm 0,000082

DISCUSIÓN

DISCUSIÓN

En el presente proyecto de Tesis Doctoral se han desarrollado estrategias para la generación de modelos *in vitro* e *in vivo* de mutaciones genéticas implicadas en la LMA. Debido a la dificultad de transfección de las líneas celulares hematológicas con plásmidos convencionales, se ha desarrollado una metodología para introducir los distintos elementos del sistema de edición génica CRISPR/Cas9 en forma de pequeñas construcciones generadas por *PCR*, más fáciles de transfectar. Este trabajo ha demostrado que con esta metodología es posible obtener porcentajes de transfección y edición en la línea celular NB4 similares a otras líneas celulares ampliamente utilizadas y a otras estrategias de edición. En esta Tesis también se han introducido las mutaciones de *IDH2* en el organismo modelo *Caenorhabditis elegans*, generando 3 nuevas cepas. Finalmente, se ha llevado a cabo un estudio funcional en estas cepas para comprobar los efectos moleculares de estas mutaciones.

Por todo ello, los resultados obtenidos en esta Tesis abren nuevas vías de desarrollo de modelos de investigación en el ámbito de la LMA. Entre otras aplicaciones, podrían emplearse en la investigación de los mecanismos moleculares implicados en el proceso leucemogénico, el estudio de las interacciones entre mutaciones concurrentes así como los procesos moleculares comunes a otras enfermedades.

Las construcciones guía generadas por *PCR* son funcionales

En este trabajo probamos como las construcciones generadas por *PCR* a partir del vector pEGR1 presentan una eficiencia de edición al menos similar al vector CRISPR pX458 en las células HEK293. La metodología desarrollada se distingue de otras por ser sencilla, rápida y muy económica. Partiendo del vector pEGR1 y de secuencias guías empleadas como *primers*, es posible obtener en un corto periodo de tiempo las construcciones guías, con el gen codificante del marcador *GFP* de forma opcional.

Las eficiencias de edición reportadas en la bibliografía presentan un rango de valores muy amplio, desde trabajos con valores similares a los obtenidos en esta Tesis, hasta trabajos donde han alcanzado valores mucho mayores, como el trabajo de Di Stazio, en el cuál consiguen un 39 % de eficiencia de edición por *HDR* (Di Stazio *et al.*, 2021). Estas notables diferencias pueden deberse a factores que no se conocen aún en

profundidad, debido a la gran velocidad a la que progresa el campo de la edición génica. En el trabajo de Miyaoka *et al.* (2016) se analizó sistemáticamente varias condiciones de edición génica en tres líneas celulares distintas. Este grupo concluyó que las eficiencias de reparación por *NHEJ* y *HDR* están influenciados en gran medida por el contexto, según el tipo celular y la diana escogida. Para el locus *RBM20* fueron capaces de editar mediante *HDR* células HEK293, pero no fueron capaces de reproducirlo de forma eficiente en las células HeLa y en *iPSCs*. Distintos patrones de expresión de los componentes celulares encargados de la reparación del ADN y modificaciones epigenéticas específicas pueden ser las causas de las diferencias de edición entre distintas líneas celulares (Miyaoka *et al.*, 2016). La condensación de la cromatina eucariota es uno de los factores que tiene un gran peso en la eficiencia de corte de la nucleasa Cas9 (Chakrabarti *et al.*, 2019; Horlbeck *et al.*, 2016; Isaac *et al.*, 2016; Yarrington *et al.*, 2018). Issac y colaboradores describieron que las secuencias de ADN localizadas en la región central de los nucleosomas son menos accesibles a la Cas9, produciéndose una gran disminución de la eficiencia de corte (alrededor de 1000 veces respecto a secuencias de ADN libres) debido al bloqueo del reconocimiento de la secuencia PAM (Isaac *et al.*, 2016). Estos resultados están apoyados por los estudios *in vivo* e *in vitro* de Horlbeck (Horlbeck *et al.*, 2016). Por otro lado, se ha comprobado que la hiperacetilación de histonas y la disminución de la marca H3K27me también afectan a la eficiencia de edición, aumentándola o disminuyéndola respectivamente (Chakrabarti *et al.*, 2019). Por ello, la condensación del ADN en las regiones diana de *IDH2* y *MYBL2*, así como marcas epigenéticas características de la línea HEK293 pueden explicar los rangos de eficiencia de edición obtenidos en este trabajo. En este sentido, la caracterización del estado de la cromatina y de las marcas epigenéticas de un gran número de genes en las líneas celulares más empleadas sin duda permitiría optimizar los experimentos de edición mejorando los criterios de diseño de las secuencias guía.

Por otro lado, se probó la eficiencia de edición al emplear un amplificado del casete de expresión de Cas9 mediante *PCR*. De esta forma se introdujo en las células los elementos mínimos de expresión de la nucleasa, eliminando los elementos de replicación en bacterias. Esto hace que las construcciones sean mucho más pequeñas y, en principio, más fáciles de transfectar. Con estas construcciones se obtuvieron

eficiencias de *NHEJ* y *HDR* menores que empleando el vector hCas9. Esta disminución podría deberse a una menor estabilidad de las construcciones lineales, cuyos extremos libres están más expuestos a la acción de nucleasas. Para compensar este efecto se podría explorar una religación de los extremos libres para proteger el ADN. La disminución de la eficiencia también podría deberse a una baja transfección. En este sentido, se ha descrito que el reactivo lipofectamina forma estructuras compactas con el ADN circular que induce la captación celular, pero en presencia de secuencias largas de ADN lineal puede formar estructuras similares a collares de perlas, afectando a la captación celular (Lehner *et al.*, 2013; vonGroll *et al.*, 2006). Sería necesario implementar nuevos métodos de transfección para introducir estos amplificadores, con la condición de mantener una alta viabilidad celular. Los métodos de transfección mediados por polietiliminina (PEI) (Dyson, 2016), electroporación (Teah *et al.*, 2017) y estructuras inorgánicas, como nanopartículas de oro y nanotubos de carbón (Y. Wu *et al.*, 2021) pueden ser opciones alternativas al uso de lipofectamina.

Las construcciones guía generadas por *PCR* son eficientes en la línea celular NB4-Cas9

Debido a la escasa eficiencia de transfección de las líneas celulares hematopoyéticas se decidió desarrollar líneas celulares con expresión constitutiva de Cas9. Mediante transducción por lentivirus se generaron tres líneas celulares con la inserción del gen codificante de la nucleasa. De esta forma, se empleó solamente un paso de transducción vírica, ya que los siguientes pasos de introducción de las construcciones guía se realizaron mediante nucleofección, evitando la inserción de más elementos de ADN. Estas líneas celulares nos servirán, además de para llevar a cabo una prueba de concepto en el presente trabajo, para poder abordar futuras investigaciones de manera más sencilla. En todas las líneas celulares transducidas se detectó la presencia de la nucleasa Cas9 mediante *PCR*, menos en la línea Katsumi-1, que se descartó. En el momento del desarrollo de los experimentos, la empresa suministradora de reactivos y protocolos de nucleofección no ofrecía un protocolo para la nucleofección de la línea MOLM-13. Por lo tanto, se continuó el trabajo con las líneas HL60-Cas9 y NB4-Cas9, cuyos protocolos estaban disponibles. Tras testar varias cantidades de construcción guía solamente se observó corte en el ensayo T7 El en la línea NB4-Cas9.

Para comprobar si la falta de edición se debía a un problema en la nucleofección, se evaluó la eficiencia de transfección mediante citometría de flujo. Para ello, las células HL60-Cas9 y NB4-Cas9 se nucleofectaron con la construcción producida por *PCR GFP-pU6-MYBL2*. Esta construcción codifica el guía para editar el gen *MYBL2* y la proteína fluorescente *GFP*. En estos experimentos se obtuvieron de media eficiencias de transfección mayores a las obtenidas en NB4-Cas9, llegando a la conclusión de que las construcciones eran capaces de introducirse en las células. A pesar de ser tipos celulares similares y presentar eficiencias de transfección semejantes, puede que la cantidad de construcción nucleofectada en las HL60-Cas9 no fuera suficiente para producir edición. Sería necesario realizar una optimización de la cantidad de construcción a transfectar, teniendo en cuenta que puede producir un aumento de la mortalidad celular por toxicidad.

Continuando con la línea NB4-Cas9 se probó que las construcciones guía producen mayor eficiencia de transfección que otros vectores ampliamente usados como el vector PX458 (de casi 10 kb) y el vector pMAXGFP (aproximadamente 3 kb). Los ratios de células GFP+ son similares a los obtenidos en el estudio de Wu *et al.* (2015), en el cual probaron que una construcción generada amplificando el gen codificante de *GFP* mediante *PCR* produce mayores ratios de transfección en células NB4 que el plásmido original pEGFP-N2 (K. Wu *et al.*, 2015).

Mediante las construcciones guía se obtuvieron eficiencias medias de 10,52 % de *NHEJ* y un 2,2 % de *HDR*. Estas eficiencias son cercanas a los rangos obtenidos en líneas celulares leucémicas de fácil manejo, como la línea K562. En el trabajo de Brabetz *et al.* (2017) los investigadores introdujeron la mutación R140 en el gen *IDH2*, obteniendo eficiencias de *NHEJ* en el rango de 24 al 46 % y de *HDR* del 3 al 10 %. Para ello, los autores emplearon los vectores PX458 y pBS-SK (+) (Brabetz *et al.*, 2017). Como se ha comentado anteriormente, el tipo celular y la secuencia diana son factores determinantes de la eficiencia de edición. A pesar de no ser el mismo tipo celular, al emplear una región cercana como diana es posible comparar las eficiencias obtenidas. Por otro lado, no sería posible emplear los mismos vectores en las células NB4 debido a su baja eficiencia de nucleofección, como se ha discutido en el párrafo anterior.

Por último, al realizar los ensayos de *HDR* con dos tipos de *ssODN*, observamos que el *ssODN* con sentido directo sí que produce edición, mientras el anti sentido no produce edición detectable en el ensayo de *RFLP*. Esto puede indicar diferencias en la accesibilidad a la región dependiendo del tipo de *ssODN* durante el proceso de reparación por *HDR*. En estudios previos, Okamoto y colaboradores describieron como los *ssODN* diseñados en la hebra complementaria al *sgRNA* son más eficientes (Okamoto *et al.*, 2019). En el caso presentado en esta Tesis, el *sgIDH2_8* está diseñado en la hebra codificante, y el *sgIDH2_3* en la complementaria, por lo que la regla de Okamoto y colaboradores se cumpliría para ambos *ssODN*. Puede que esta disparidad esté causada por el uso de dos complejos Cas9. Se ha descrito que la nucleasa permanece unida a la diana después de producirse el corte, pudiendo disminuir la accesibilidad del *ssODN* (Miyaoaka *et al.*, 2016; Sternberg *et al.*, 2014). Es posible que la cercanía de las dos secuencias diana y que se mantengan las Cas9 unidas tras el corte disminuya la accesibilidad del *ssODN* anti sentido.

La eficiencia de las construcciones CRISPR generadas por *PCR* es comparable a la de los complejos *RNPs*

El uso directo de los elementos CRISPR en formato de *RNPs* presenta numerosas ventajas, como su alta eficiencia de edición y pocos efectos *off-targets* comparados con plásmidos y mRNA codificantes de Cas9 (S. Zhang *et al.*, 2021). Esto es debido a que las *RNPs* no han de expresarse desde un vector de ADN o ARN mensajero, si no que actúan directamente una vez se han introducido en el núcleo celular. En consonancia con esto, en este trabajo se ha obtenido una mayor eficiencia de *NHEJ* empleando las *RNPs*. Otros estudios han investigado la diferencia de eficiencias empleando distintos formatos de los elementos CRISPR, sugiriendo que la introducción de proteínas exógenas puede ser menos tóxica para las células que la introducción de nucleótidos (Okamoto *et al.*, 2019). Por otro lado, a pesar de obtener un mayor porcentaje de corte con las *RNPs*, no observamos un aumento de la edición por *HDR*. Otros trabajos resaltaron observaciones similares (Miyaoaka *et al.*, 2016).

Los mecanismos principales de reparación de cortes en la doble cadena de ADN por *NHEJ* y *HDR* involucran vías moleculares distintas y la prevalencia de un mecanismo de reparación sobre otro puede determinar las eficiencias de edición. El *NHEJ* es

normalmente el mecanismo predominante de reparación ya que es un sistema robusto y rápido, a pesar de ser propenso a errores. En esta vía actúa un complejo de siete proteínas distintas, pudiendo alcanzar una eficiencia del 90% (M. Liu *et al.*, 2019). Al contrario, la vía de *HDR* es un mecanismo de reparación fiel y requiere de la intervención de más de 13 proteínas distintas (M. Liu *et al.*, 2019). Además, la región homóloga de la cromátida hermana o del molde usado para editar debe encontrarse próxima al sitio de rotura. Todo ello hace que la vía de *HDR* sea un método más complejo y menos eficiente. Asimismo, el empaquetamiento de la cromatina también influye en la accesibilidad de estos complejos de reparación (Chakrabarti *et al.*, 2019). Por todo ello, las características intrínsecas del mecanismo de *HDR* y la accesibilidad del gen *IDH2* en las células NB4 pueden producir el bloqueo de la eficiencia de edición mediante *HDR* a pesar de aumentar la eficiencia de *NHEJ* con el uso de *RNPs*.

Debido a la baja eficiencia de la reparación vía *HDR* que ocurre en prácticamente todos los sistemas estudiados, se han desarrollado distintas moléculas que bloquean el mecanismo de reparación por *NHEJ*. Ejemplos de estas moléculas son SCR7, E1B55K y E4orf6, que actúan inhibiendo la Ligasa IV (M. Liu *et al.*, 2019; Vicente *et al.*, 2021). También se han empleado pequeños ARNs interferencia para inhibir la proteína Ku, iniciadora del proceso de *NHEJ*. Otras moléculas inhiben el *NHEJ* afectando a vías moleculares generales, como Brefeldin A, que inhibe el transporte de proteínas del retículo endoplasmático al aparato de Golgi, y RS-1, aumentando el *HDR* dependiendo de la diana y la estabilización de Rad51 y el ADN (M. Liu *et al.*, 2019). Sin embargo, el uso de estas moléculas puede derivar en un aumento de la apoptosis celular, como han demostrado algunos autores (Acosta *et al.*, 2018; M. Liu *et al.*, 2019), por lo que no se consideró su uso en este trabajo. Otra estrategia que se ha implementado para aumentar la eficiencia del *HDR* es la introducción de modificaciones en las moléculas empleadas como donadores. Se ha comprobado que modificaciones químicas, como trietilenglicol, en el extremo 5' aumentan la eficiencia de edición cuando se introducen en donadores de gran longitud, mientras que en donadores cortos se consigue alcanzar el tope de eficiencia empleando concentraciones menores (Ghanta *et al.*, 2021). Al igual que otros aspectos de la edición mediante el sistema CRISPR/Cas9, estas modificaciones

producen distintos efectos dependiendo de la diana escogida y del modelo empleado (Ghanta *et al.*, 2021).

La secuenciación masiva confirma las eficiencias de corte y edición obtenidas.

El ensayo de detección de *mismatches* mediante T7 EI es un método sencillo y preciso para evaluar la eficiencia de la edición génica generada por la endonucleasa Cas9. Sin embargo, esta estrategia puede no ser fiable en determinadas circunstancias. La capacidad de evaluación de edición de esta enzima se basa en la formación de heterodúplex de ADN. Por ello, si el *sgRNA* empleado es muy eficiente, y se da de forma dominante un solo tipo de *indel* la formación de heterodúplex entre secuencias editadas y salvajes es menor, y la cuantificación por T7 EI no es fiable. Los resultados de los trabajos de Senmanat *et al.* sugieren que la enzima T7 EI presenta mayor sensibilidad en las muestras con alta complejidad de *indels* con rangos de distintos de tamaños (Sentmanat *et al.*, 2018). Por otro lado, mediante *NGS* es posible conseguir una estimación precisa del tipo, tamaño y frecuencia de los *indels* generados (Sentmanat *et al.*, 2018). Ello permite también estudiar los posibles cortes inespecíficos en las secuencias *off-targets* aunque se hayan producido con una frecuencia muy baja.

Para calcular la actividad de la nucleasa Cas9 a través de los datos de *NGS* se deben considerar los porcentajes obtenidos de los diferentes tipos de edición (*NHEJ*, *HDR*, *HDR* imperfecto y secuencias ambiguas), ya que el estudio se realiza sobre el *pool* de células editadas. De esta forma se obtuvieron porcentajes de corte muy similares a los obtenidos por T7 EI: 14,52 % por T7 y 13,82 % por *NGS* para NB4-Cas9 Edición y 29,8 % por T7 y 29,9 % por *NGS* para NB4 RNP. En cuanto a la eficiencia de edición, también se obtienen porcentajes similares mediante el ensayo de *RFLP* y la *NGS*. Por todo ello, en nuestro caso el ensayo de la endonucleasa T7 EI es un método muy fiable para determinar la eficiencia de *NHEJ*. El principal motivo de reparación tras los dos *DSB* producidos es la delección del fragmento de nucleótidos entre ambas secuencias guías, en el rango de 20 a 29 pb. Esta estrategia se ha empleado por otros autores para deleccionar fragmentos de ADN (Blayney *et al.*, 2020; Chen *et al.*, 2014; Duzgunes *et al.*, 2021).

Es interesante comentar la posible relación entre la conformación de las secuencias PAM con las deleciones producidas. En el trabajo de Shou y colaboradores tras dos eventos de *DSB* se detectaron inserciones no aleatorias en los puntos de unión de ADN tras producirse deleciones, inserciones y duplicaciones. Estudiando estos patrones han desarrollado un modelo de predicción del tipo de reparación y ligación del ADN teniendo en cuenta cuatro posibles configuraciones de las secuencias PAM. La configuración de las secuencias PAM produce que la nucleasa Cas9 medie la inserción de secuencias nucleotídicas cortas en extensión 5'-3', mediante la acción conjunta con ADN polimerasas y ADN ligasas. Estas secuencias condicionarán que ocurran procesos de inversión, deleción o duplicación de la región entre los puntos de *DSB*. La configuración de las secuencias PAM empleadas en este trabajo es 5' NGG y CCN 3'. En el modelo propuesto por los autores, esta configuración debería desencadenar la adición de los nucleótidos 5'-GC-3' en el punto de rotura mediado por la PAM NGG, y los nucleótidos 5'-CCC-3' en el punto de rotura mediado por la PAM CCN. Según los autores, esto debería producir una duplicación del fragmento de nucleótidos entre las secuencias PAM (Shou *et al.*, 2018). Esto contradice los resultados obtenidos en este trabajo, donde encontramos que la deleción es el principal motivo tras la reparación. Según Shou *et al.*, la deleción ocurriría con la conformación 5' CCN-NGG 3' de las secuencias PAM. Esto puede deberse a la poca distancia entre ambas secuencias PAM. En nuestro caso ambas se encuentran a ocho pb, mientras que en el trabajo comentado trabajan con guías que distan entre si entre 700 pb, 1 y 6 kb. También puede ser que la regla descrita por Shou y colaboradores sea específica de diana y no se pueda extrapolar a otros *loci*.

Mediante secuenciación masiva no se detectan cambios significativos en los *off targets* analizados

Por último, en este trabajo se han estudiado los posibles cortes inespecíficos de la nucleasa Cas9. El sistema CRISPR/Cas9 admite variaciones entre 3 y 5 nucleótidos debido a que la secuencia *sgRNA* tolera esas diferencias en el ADN. A pesar de que esta unión es menos eficiente y con una eficiencia de corte de la nucleasa mucho más baja, es posible que se produzcan modificaciones indeseadas (Manghwar *et al.*, 2020). Por otra parte, existen otros factores que pueden producir que la Cas9 corte en sitios sin homología completa. El ADN o el *sgRNA* pueden formar pequeños bucles en los nucleótidos no homólogos entre ambos, corrigiendo el alineamiento y permitiendo el

corte de la nucleasa (Manghwar *et al.*, 2020). Dependiendo de la posición de estos bucles el sistema CRISPR/Cas9 puede tolerar bucles de hasta 4 pb (Lin *et al.*, 2014). Estas posibilidades se tuvieron en cuenta a la hora de escoger los *off-targets* a estudiar. Al realizar la predicción *in silico* con un bucle de ADN y/o ARN no se obtuvo ningún *off-target*. Sin embargo, al combinar la opción de presentar *mismatches* y bucles, los *off-targets* aumentaron en gran medida: la combinación de 2 *mismatches* y/o 1 bucle de ADN y/o ARN dio como resultado 100 posibles *off-targets* para el guía 3 y 81 para el guía 8; la combinación de 3 *mismatches* y 1 bucle de ADN y/o ARN mostró 563 para el guía 3 y 858 para el guía 8. No es posible abarcar un estudio de tantas regiones empleando la estrategia de secuenciación dirigida que se ha desarrollado. Para completar este estudio sería necesario complementar el análisis de NGS realizado con una secuenciación genómica más amplia con el fin de comprobar el posible efecto de los bucles de ADN y ARN en los experimentos realizados. Por ello, en este trabajo se escogió como criterio la selección de *off-targets* con máximo 3 nucleótidos distintos respecto a la secuencia de los guías.

Mediante el software CRISPResso2 se analizaron las lecturas obtenidas de los potenciales *off-targets* y con la variante CRISPResso2 Compare se compararon las lecturas obtenidas en las muestras NB4-Cas9 control y NB4-Cas9 editadas. No se obtuvieron diferencias entre ambas muestras. Por otro lado, no se obtuvieron lecturas suficientes de la muestra NB4 control para compararla con los resultados de la muestra NB4 editada. A pesar de ello, en el análisis simple de CRISPResso2 no se obtuvieron variaciones significativas en los 16 *off-targets*.

El sistema de edición génica CRISPR/Cas9 ha supuesto una gran revolución en los últimos años y ha sido adoptado rápidamente como herramienta de ingeniería genética por innumerables laboratorios de investigación. A pesar de ello, el conocimiento para predecir la precisión, eficacia y el resultado de la edición en distintas dianas del genoma es limitado. Es necesario profundizar en los mecanismos de interacción entre las nucleasas y los componentes celulares eucariotas, como la cromatina, regiones repetitivas complejas y las diferentes vías de reparación del ADN. Para ello es necesario caracterizar en profundidad las formas de edición en distintas dianas y tipos celulares y

compararlas entre sí, con el objetivo de detectar patrones generales o específicos de reparación. De esta forma será posible generar protocolos precisos para desarrollar diversos tipos de edición en distintas dianas de forma precisa y eficiente.

Detección del oncometabolito 2-HG en los modelos desarrollados en *Caenorhabditis elegans*

Las vías metabólicas de ácidos grasos y aminoácidos están muy estudiadas en *C. elegans*. Mediante el uso de moléculas de *RNAi* y la inactivación de genes se ha podido caracterizar los cambios metabólicos que sufre el gusano a lo largo de su ciclo vital y cómo responde ante condiciones adversas, como la restricción de nutrientes, variaciones de temperatura y anoxia. Al estudiar la tasa metabólica en los distintos estadios ocurre un aumento en estadio de larva L1, que continúa hasta los estadios de larva L2-L4. Este aumento tras la eclosión puede deberse a que los embriones obtienen energía principalmente del ciclo del glioxilato, empleando las reservas de lípidos y la gluconeogénesis. Tras la eclosión, la principal fuente de producción de energía pasa a ser el ciclo de los ácidos tricarbóxicos (CAT). Por otro lado, son muchos los parámetros metabólicos que disminuyen a lo largo de la vida adulta de los gusanos, como el consumo de oxígeno, generación de calor, mantenimiento de los niveles de ATP, ratio ATP: AMP, el potencial de producción de superóxido, la capacidad de bio-reducción y la producción de amonio (Braeckman et al., 2009).

Realizar estudios metabólicos en *C. elegans* presenta muchas ventajas pero también dificultades, como la obtención de una cantidad suficiente de individuos sincronizados y métodos eficientes de rotura de la cutícula (Braeckman et al., 2009). En este trabajo, hemos realizado una puesta a punto del número de gusanos necesario para hacer la cuantificación y del método de lisis para obtener una cantidad de muestra suficiente para su análisis. La disminución del metabolismo a mayor edad de los gusanos se ve reflejado en los resultados de la cuantificación de 2-HG. Al analizar los estadios L4 y adulto observamos que la concentración es menor en los adultos. Como se ha señalado, los estadios de larva presentan una mayor actividad del metabolismo, además de implicar principalmente al CAT para la obtención de energía (Edwards et al., 2013).

A lo largo del ciclo vital del gusano, la mitocondria es uno de los orgánulos donde más se acumulan moléculas reactivas de oxígeno (ROS) y, por lo tanto, donde más daño

oxidativo se produce. Con el avance de la edad, los mecanismos moleculares de defensa contra la acumulación de ROS disminuyen, causando poco a poco una alteración mitocondrial y de la homeostasis que afecta al CAT (Gao *et al.*, 2018). Esta alteración paulatina de la actividad mitocondrial produce la disminución de metabolitos, como el 2-HG. Alteraciones de los intermediarios del CAT altera el ciclo vital normal del gusano. Por ejemplo, su inhibición o bloqueo inhibe las primeras etapas de la embriogénesis. Por otro lado, la adición de malato, fumarato y oxalacetato produce un aumento de la esperanza de vida, del ratio NAD/NADH y del consumo de oxígeno (Edwards *et al.*, 2013).

Xudong Fu y colaboradores estudiaron el efecto de aumentar los niveles de los dos enantiómeros de 2-HG en el gusano. Para ello añadieron estos metabolitos en las placas de NGM. Observaron un aumento de la esperanza de vida del gusano similar al que previamente habían asociado al aumento de α -CG (Fu *et al.*, 2015). En estos experimentos los gusanos en estadio L4 presentaban una media de esperanza de vida adulta de 20 días con suplementación de R-2-HG, y de 20,7 días con L-2-HG, mayor que los controles, con una media de 14 días. Concluyeron que tanto el aumento del α -CG como el de los dos enantiómeros de 2-HG inhiben la enzima ATP sintasa, disminuyendo la respiración mitocondrial y la vía de mTOR (Fu *et al.*, 2015). Esto a su vez hace que se aumente la autofagia y otras vías de supervivencia celular (Konstantinidis *et al.*, 2021).

El CAT presenta una alta conservación evolutiva gracias a su papel esencial, existiendo una alta homología entre el gen *IDH2* y la proteína codificante con su ortólogo, el gen *idh-2*. Gracias a esta homología hemos encontrado conservados los residuos R140 en la posición R123, y el R172 en R155, permitiendo la introducción de las mutaciones homólogas descritas en LMA. Las mutaciones en *idh-2* se pueden mantener en homocigosis en el gusano, incluso la concurrencia de ambas mutaciones, de forma viable. A pesar de ello, las cepas desarrolladas no presentan el aumento esperado del oncometabolito ni la alteración de la esperanza de vida descrita por otros autores.

Por otro lado, existen otros modelos de mutaciones asociadas a desórdenes hematológicos en *C. elegans* que no desarrollan los fenotipos observados en pacientes. Serrat y colaboradores introdujeron mutaciones *missense* relacionadas con SMD en el gen *sftb-1*, ortólogo del factor de *splicing* *SF3B1*. Las mutaciones en *SF3B1* producen

alteraciones del transcriptoma, que conllevan a la alteración de múltiples funciones celulares que se asocian directamente con el desarrollo de la enfermedad (*Putowski et al.*, 2021). Estas alteraciones en el transcriptoma no se observaron en el modelo de *C. elegans* desarrollado. Por otro lado, fueron capaces de incluir mutaciones en otros genes que raramente se dan de forma conjunta con mutaciones en *SF3B1* (*U2AF1*^{S34P} y *SRSF2*^{P95H}) debido a su letalidad en otros organismos. De esta forma los autores pudieron estudiar nuevas vulnerabilidades e interacciones funcionales entre estos genes (*Serrat et al.*, 2019).

Muchos investigadores han observado que modelos desarrollados mediante edición génica con nucleasas no presentan los fenotipos esperados. En relación a esto, en la revisión de Salgana y en el trabajo de Serobyán, se describen numerosos casos de incoherencias entre modelos desarrollados con morfolinós y ARN anti sentido con modelos desarrollados mediante CRISPR en *C. elegans*, pez cebra y ratón (*Salanga et al.*, 2021; *Serobyán et al.*, 2020), donde los primeros sí presentan un fenotipo claro y los segundos no. En algunos casos se describe que el efecto de la mutación introducida con CRISPR se compensa con un aumento de expresión de otro gen con una función similar. En estos casos se habla de una adaptación transcripcional o compensación génica producida por la presencia del ARN mensajero mutante. La presencia de este ARN, junto a su degradación, están involucrados en el aumento de señales de expresión, como el aumento de la metilación H3K4me3. La no producción de estos ARNm mutantes en los modelos desarrollados con morfolinós y con ARN anti sentido hace que no se active esta vía, no se de la compensación y se produzca el fenotipo (*Salanga et al.*, 2021; *Serobyán et al.*, 2020).

Bajo este escenario, es posible que en los modelos desarrollados en este trabajo también se esté produciendo una compensación génica para evitar los efectos de la acumulación del 2-HG. Gracias a un mayor interés por el metabolito 2-HG recientemente se han descrito un mayor número de enzimas involucradas en su metabolismo. La enzima D-2-hidroxiglutarato deshidrogenasa (D-2-HGD) es una enzima mitocondrial dependiente de FAD encargada de catalizar la conversión del D-2-HG a α -CG (*Veiga-da-Cunha et al.*, 2020). Mutaciones *missense* o *nonsense* en esta enzima disminuyen o eliminan la capacidad de unión a FAD, disminuyendo la actividad de la proteína y

derivando en un aumento de los niveles de D-2-HG (Yang *et al.*, 2021). Esta acumulación desencadena un desorden neurológico raro llamado Aciduria D-2-hidrogutárica Tipo I (OMIM 600721). Esta alteración metabólica produce el aumento de los niveles de D-2-HG en orina, plasma y fluido cerebroespinal. Los fenotipos clínicos son muy diversos, pero se pueden agrupar en dos clases. Los fenotipos graves se caracterizan por encefalopatía epiléptica infantil de inicio temprano o neonatal con hipotonía marcada, fallo y retraso visual neurológica, y retraso serio del desarrollo. Por otro lado, estaría el fenotipo medio, que presenta mayor variabilidad y se caracteriza por hipotonía y retraso del crecimiento (Kranendijk *et al.*, 2010). Mutaciones en este gen también se asocian con linfoma difuso de células grandes B (Yang *et al.*, 2021). Se han encontrado ortólogos de esta enzima en otros organismos modelo, como *E. coli* (Zhao *et al.*, 1996), *Saccharomyces cerevisiae* (Struys, 2006), *Pseudomonas putida* (Kopchick *et al.*, 1979) y *Arabidopsis thaliana* (Araújo *et al.*, 2010). En *C. elegans* encontramos el ortólogo F54D5.12. Este gen presenta un porcentaje de identidad del 43,74 % respecto al gen humano (Ensembl). Hasta la fecha, no se han desarrollado estudios funcionales sobre este gen, ni cepas con modificaciones del mismo. Si la proteína codificada en F54D5.12 tiene la misma funcionalidad que D-2-HGD, sería la encargada de transformar el D-2-HG en α -CG en las mitocondrias del gusano y su acción podría regular el aumento de 2-HG producido por las mutaciones introducidas en *idh-2*.

El desarrollo de cepas con las mutaciones de *idh-2* y la inactivación de *F54D5.12* permitirían esclarecer la función del gen *F54D5.12*, su papel en la vía del 2-HG y su implicación en la compensación de las mutaciones de *idh-2*. Estos dobles mutantes servirían para producir el aumento del oncometabolito y estudiar las vías moleculares que se ven afectadas y que pueden ser comunes entre la LMA y la aciduria.

CONCLUSIONES

CONCLUSIONES

El trabajo desarrollado en esta Tesis nos lleva a las siguientes conclusiones:

1. Mediante el sistema de *PCRs* desarrollado sobre el vector pEGR1 como molde es posible producir construcciones guía de forma rápida, sencilla y eficiente.
2. Las construcciones guía generadas por *PCR* producen una eficiencia de edición mayor que otros vectores CRISPR ampliamente utilizados en las células HEK293.
3. Las construcciones guía generadas por *PCR* presentan eficiencias de transfección mayores en las líneas celulares leucémicas NB4 y HL60 que otros vectores CRISPR.
4. El uso de las construcciones guía desarrolladas en células NB4 con expresión constitutiva de Cas9 producen eficiencias de corte menores que los complejos de ribonucleoproteínas, pero eficiencias de edición de la mutación *IDH2*^{R172} similares.
5. El empleo de dos secuencias guía adyacentes produce que los principales cambios introducidos en la edición sean deleciones de los nucleótidos entre ambas posiciones.
6. La inserción del gen codificante de la nucleasa Cas9 en el genoma de las células NB4 no produce cambios significativos en los potenciales *off targets* de los guías *IDH2_3* e *IDH2_8*.
7. El gen *idh-2* de *Caenorhabditis elegans* presenta una gran homología con el gen *IDH2* humano, y los dominios del producto que codifica, afectados por las mutaciones en pacientes, están conservados al 100 %. Gracias a esto, se han generado las cepas RVM(479), RVM(480) y RVM(481) con las mutaciones ortólogas a R140, R172 y ambas, respectivamente.
8. La introducción de las mutaciones ortólogas a R140 y R172 en *idh-2* por si solas no reproducen el aumento del oncometabolito R-2-hidroxiglutarato observado en pacientes. Es necesario realizar una caracterización funcional más exhaustiva para estudiar los cambios que estas mutaciones podrían producir en *C. elegans* y estudiar otras vías que puedan estar involucradas en la acumulación del oncometabolito.

ANEXOS

ANEXO I

Tabla 2.1: Secuencias guía empleadas en los ensayos *in vitro*.

Secuencia guía	Secuencia (5'– 3')
MYBL2	TTGGCAGCTGGCTCCCTC
IDH2_1	GCGTGCCTGCCAATGGTGAT
IDH2_2	TCCAGCCAGGGACTAGGCGT
IDH2_3	CTGGCCTACCTGGTCGCCAT
IDH2_4	GCTTGGTCCAGCCAGGGACT
IDH2_5	GTGATGGGCTTGGTCCAGCC
IDH2_6	GCGTGCCTGCCAATGGTGA
IDH2_7	GAGAGGGGATCCACTGACCT
IDH2_8	TGGACCAAGCCCATCACCAT

Tabla 2.2. Secuencia del ssODN empleado para introducir la mutación R172 en células.

ssODN	Secuencia (5'-3')
IDH2_R172K	tgcaaaaacatcccacgcctagtcctggctggactaa accgataacgatcgga agcacgcgcatggcgac caggtaggccagggtggagaggggat
IDH2_R172K anti sentido	atcccctctccaccctggcctacctggtcgccatg cgcgctcttccgatcggttatcggttagtccagccagggga ctaggcgtgggatgttttgca

Los nucleótidos resaltados en negrita y en color naranja indican los cambios producen la mutación R172K en células. En rojo se marcan los cambios introducidos para evitar el corte por la nucleasa Cas9.

Tabla 2.3: Potenciales *off targets* de los guías *IDH2_3* e *IDH2_8*.

	Posición cromosómica	Gen	Secuencia	Región
sgIDH2_3				
3_off1	chr5: 173428070	<i>CTB-32H22</i>	CTGaCCTgCCTGGTCcCCATTGG	Intrón
3_off2	chr1: 205830643 (-)	<i>PM20D1</i>	tTGGCCTcCCTGGTCGtCATGGG	Intrón
3_off3	chr2: 241745044	<i>D2HGD</i>	CTGGCCTtCCTGGTgGtCATGGG	Intrón
3_off4	chr17: 2879971(-)	<i>RAP1GAP2</i>	CaGGCCTACCTGGTCCcCATTGG	Intrón
sgIDH2_8				
8_off1	chr3: 76082901	<i>ROBO2</i>	cGGACCAAGgCgATCACCATGGG	Intrón
8_off2	chr4: 130711158 (-)	-	TGtACCAAGctCATCAaCATTGG	Intergénico
8_off3	chr1: 2405874	<i>PEX10</i>	TGGgCCatGCCCATcCCATCGG	Intrón
8_off4	chr1: 54465152	-	TGGACCAAGCCCcTCACctTGGG	Intergénico
8_off5	chr15: 70078666	<i>TLE3</i>	TGGcCCAAGCCctTCACCAaCGG	Intrón
8_off6	chr17: 83115842 (-)	-	TGGACCAAGtCCAcTtCCATGGG	Intergénico
8_off7	chr10: 127909863 (-)	<i>PTPRE</i>	TGGACCATGCCCATCcaCATCGG	Intrón
8_off8	chr6: 31629230	<i>PRRC2A</i>	gGGACCAAtCCCATCACCCtTTGG	Exón
8_off9 (*)	chr14: 99534805	<i>CCNK/CCDC85</i> C	TGGcCaAAGCCctTCACCATAGG	Exón/Intrón
8_off10	chr9: 38523669	<i>RP11-103F21.4</i>	gGGACCAgGCCctTCACCATTGG	Pseudogen
8_off11	chr9: 97112588	-	gGGACCAgGCCctTCACCATTGG	Intergénico
8_off12	chr18: 58014733 (-)	-	TGaACCAAGCCCATaACCcTTGG	Intergénico

(*) La posición del *off-target* 8.9 coincide con el final de la secuencia del exón 11 del gen *CCNK* y con el primer intrón del gen *CCDC85C*.

Los nucleótidos en rojo y en minúscula indican las disparidades entre los *off targets* y las secuencias guía.

Tabla 2.4. Cepas de *C. elegans* empleadas

Nombre	Genotipo	Procedencia
N2	Cepa salvaje	CGC
RMV479	<i>idh-2</i> (vlt24)	Este trabajo
RMV480	<i>idh-2</i> (vlt25)	Este trabajo
RMV481	<i>idh-2</i> (vlt26)	Este trabajo

Tabla 2.5: Secuencias guía empleadas en *C. elegans*.

Secuencia guía	Secuencia (5'– 3')
<i>idh-2</i> _R123Q	AAGATGTGGCTCTCACCAA
<i>idh-2</i> _R155K	ACTGATCTCCAAAAGCGTGG
<i>dpy-64</i>	GCTACCATAGGCACCACGAG

Tabla 2.6. Secuencias de los ssODNs empleados en *C. elegans*.

<i>ssODN</i>	Secuencia (5'– 3')
<i>idh-2</i> _R123Q	cggacgaggctcgcatcaaagagtttaatcttaagaaatgtggctt cttccgaatggaactattc aaaatattc taggaggaactgtatttcgtgagccaat
<i>idh-2</i> _R155K	cacgactagtgcctggatggactcagccgatcacaattgg aaaacatgc atttgagatcagtataaatgcac agatttagttattccaa
<i>dpy-10</i> (cn64)	cacttgaactcaatacggcaagatgagaatgactggaaaccgtaccgc atg cggtgcctatggtagcggag cttcacatggcttcagaccaacagcctat

Los nucleótidos resaltados en negrita y en color naranja indican los cambios que introducen las mutaciones R123Q, R155K y *dpy* (cn64). En rojo se marcan los cambios introducidos para evitar el corte por la nucleasa Cas9. Se tuvo en cuenta el patrón de uso de codones característico en *C. elegans* para introducir todos los cambios.

Tabla 2.7. Potenciales *off targets* de los guías *idh-2*_R123Q e *idh-2*_R155K

	Posición cromosómica	Gen	Secuencia	Región
<i>idh2</i> _R123Q				
R123Q_off1	chrIII:13302221-13302243 (+)	<i>Y43F4B.5</i>	A A t A a G T G G a T C T g ACCAAA	Exón
<i>idh2</i> _R155K				
R155K_off1	chrV:13359735-13359757 (+)	<i>F32H5.3</i>	A T t G A T g c C C g AAAGCGTGG	Exón
R155K_off2	chrV: 4862871-4862893	<i>C37H5.13</i>	A g a G A T C T C t AAAA a CGTGG	Exón
R155K_off3	chrII: 7574514-7574536	<i>F35D2.2</i>	A t T G A T t a C AAAA t CGTGG	Exón

Los nucleótidos en rojo y en minúscula indican las diferencias entre los *off targets* y las secuencias guía.

Tabla 2.9. *Primers* empleados para la amplificación de la secuencia U6-sgRNA en el vector PX458.

	Secuencia (5'-3')
pU6_sgRNA_F_AfIII	AGCTGC C T T A A GGGTTCTGCTTTT G CTGG
pU6_sgRNA_R_AfIII	AGCTGC C T T A A GGGTACCTCTAGAGCC A TTTG

En cursiva y negrita se muestra el sitio de reconocimiento de la enzima AfIII.

Tabla 2.10. Primers diseñados para generar las construcciones sgRNA mediante PCRs.

	Secuencia (5'-3')
FPCR_1F	GTTTTAGAGCTAGAAATAGCAAG
FPCR_1R	CCGGCAGACGACGATTTATGACGGAAAAAAGCACCGACTCGGTGCCAC
FPCR_2F	AAAAAACCTCCCACACCTCC
FPCR_2R	GGTGTTCGTCCTTTCCACAA
FPCR_3F	GTTGTGGTTTGTCCAAACTC
FPCR_3R	CCGGCAGACGACGATTTATGACGG
FPCR_4F	GGGTCATTAGTTCATAGCCC
FPCR_4R	GGTGTTCGTCCTTTCCACAA
FPCR_5F	AGCCCATATATGGAGTTCCG

La secuencia guía en forma *directo* debe añadirse al *primer* FPCR_1F y en forma reversa al *primer* FPCR_2R, en 5'

Tabla 2.11. Primers empleados para amplificar el casete de expresión de Cas9 del vector hCas9

Nombre	Secuencia (5'-3')
Cas9_F	TCTGCTCTGATGCCGCATAG
Cas9_R	CTGCGCAGATCTGCTATGGC

Tabla 2.12. Primers empleados para amplificar la región editada de los genes MYBL2 e IDH2 en las líneas celulares

Diana	Directo (5'-3')	Reverso (5'-3')
MYBL2	CTTCCCACCCTGTAGCAGG	CCACATGCTGCAAATGTCTCAG
IDH2	ACTCCAGAGCCCACACATTT	ATGCCGGGATTACAAGTGTC

Tabla 2.13. *Primers* usados en la preparación de la librería (PCR 1).

<i>Off target</i>	Directo (5'-3')	Reverso (5'-3')
<i>sgRNA3_1</i>	TGTGTGTGGTTGGAGGATG	AGCCATTCTTCTGTACAGC
<i>sgRNA3_2</i>	CTGAAAGCGTATAGCTATCC	GGCCTTGGGTATGAAATGAAG
<i>sgRNA3_3</i>	CCACTGTTGGTGTGAGGAAG	GCATGTCCCTCAACCTTACTC
<i>sgRNA3_4</i>	AGCCCTACTGTAATGCTTGG	ATGTCATCCATCAGTCCCCG
<i>sgRNA8_1</i>	GATGATTGCCATAGCCTCAG	ATCCTATCTCCCACACTGTG
<i>sgRNA8_2</i>	GTACTCTAGCCTGGTGACAG	AGAATGGGAGGGGGAATAAG
<i>sgRNA8_3</i>	AGAGGTCATCTGTGTCCAGG	CTGGGACTCGGCCTTCTGAA
<i>sgRNA8_4</i>	TGAAAGGAGGGTGAGCCG	GAAGGTTCCCCCTCATTG
<i>sgRNA8_5</i>	AGTATCACTGGCATCTCTGC	CTGAGGAGGCCAGAATGACT
<i>sgRNA8_6</i>	CCATATGGGAGGGCGGTAGC	CAGAACAGATCCCCCACAG
<i>sgRNA8_7</i>	CCTACCAGTGGAGCCTTGAG	GCTTCGATCCTGTCCCTCC
<i>sgRNA8_8</i>	TGTTGGACTAGATCACTCTG	GCCCTTAGAGGGTAAACAAC
<i>sgRNA8_9</i>	CGCCAGGTCTGCAATGGAAG	GCCTCTTGTAACCTCTCCTC
<i>sgRNA8_10</i>	GCCCTGTCAATGAAGTGAAG	CCACAAATGGCTGCAGAAAC
<i>sgRNA8_11</i>	AAGGCCCTGTCAATGAACTG	TTCACTGCTAGGTTTGCCAC
<i>sgRNA8_12</i>	GTTGGGATTACAGGTATGAG	TAGTCAAGGTCTTATTGCAC
<i>On target</i>	TTATCTCTGTCCTCACAGAG	CAAAGTCTGTGGCCTTGTAC
Secuencia del adaptador de primer directo	TACTCTTTCCCTACACGACGCTCTTCCGATCT	
Secuencia del adaptador del primer reverso	GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT	

Tabla 2.14. *Primers* usados en la preparación de la librería (PCR 2).

	Secuencia <i>primers</i> PCR 2 (5'-3')
TruSeq Universal Adapter_Directo	AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT
TruSeq_NB4Cas9-Control	CAAGCAGAAGACGGCATAACGAGATACATCGGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TruSeq_NB4Cas9-Edición	CAAGCAGAAGACGGCATAACGAGATTGGTCA GTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TruSeq_NB4-Control	CAAGCAGAAGACGGCATAACGAGATCTGATCGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TruSeq_NB4_RNP	CAAGCAGAAGACGGCATAACGAGATGATCTGGTACTGGAGTTCAGACGTGTGCTCTTCCGATCT
TruSeq_NB4PCR	CAAGCAGAAGACGGCATAACGAGATCAAGTGTGACTGGAGTTCAGACGTGTGCTCTTCCGATCT

Los nucleótidos verdes representan la zona de homología con la secuencia de los *primers* usados en la PCR 1. En rojo se muestran las etiquetas escogidas para diferenciar las muestras.

Tabla 2.15: Primers utilizados para realizar el screening de las mutaciones de *idh-2* en *C. elegans*

	Directo (5'-3')	Reverso (5'-3')
<i>idh-2</i> _R123 alelo salvaje	GAATAGGAATTATCCTCCAG	TGATTGTTCCGTTTGGTGAG
<i>idh-2</i> _R123 alelo mutante	GAATAGGAATTATCCTCCAG	GAATAGTTCCATTCGGAGAA
<i>idh-2</i> _R155 alelo salvaje	GTGTTGCCACTACCAGATGC	GATCTCCAAAAGCGTGCGGG
<i>idh-2</i> _R155 alelo mutante	GTGTTGCCACTACCAGATGC	CTGATCTCCAAATGCATGTTTT

Tabla 2.16: Primers empleados para secuenciación Sanger en *C. elegans*.

	Directo (5'-3')	Reverso (5'-3')
<i>idh-2</i>	TTTGCAGGTCACCATTGACG	CACATCTCCGTCGTAGTTCT
<i>Y43F4B.5</i>	ATGGATGGGAAATCGAGCCG	CAGTCCAAATTGCTGTGGTC
<i>F32H5.3</i>	AGCAGCTCCAATTGTGACGG	CCCGTGTTTTATCCTGTGTT
<i>C37H5.13</i>	GAAAGTGACGTTAATTCTG	CTGACAACACTGTTTAGCCT
<i>F35D2.2</i>	TAGGAGTTTATGTGGGCCAG	CTCCTCCAATATGACGAGAA

Tabla 2.17. Cepas bacterianas empleadas en este trabajo

Nombre cepa	Descripción	Fuente
TOP10	Cepa <i>E. coli</i> electrocompetentes empleadas para transformar y amplificar plásmidos	Addgene
OP50-1	Cepa de <i>E. coli</i> empleada como fuente de alimento para <i>C. elegans</i>	Addgene

Tabla 2.18. Vectores empleados en este trabajo

Nombre	Fuente
pX458	Addgene
hCas9	Addgene
pEGFP-N1	Addgene
pEGR1	Este trabajo
lentiCRISPR V.2	Addgene
psPAX2	Addgene
pMD2.G	Addgene
PMAXGFP	Lonza

ANEXO II

CRISPResso Compare NB4-Cas9 Control VS NB4-Cas9 Edición

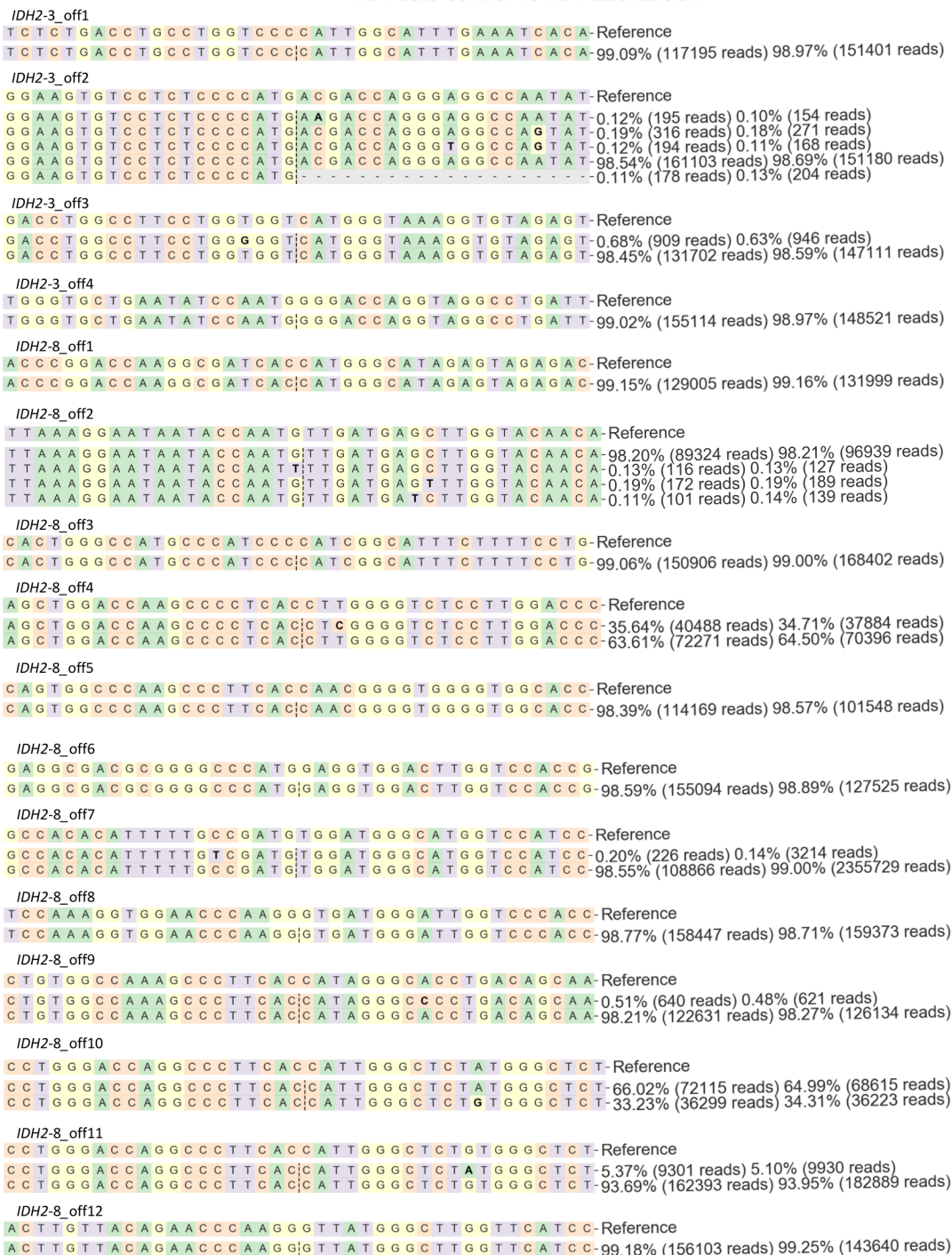


Figura 3.15: Análisis de los off targets mediante CRISPResso2 Compare. A la derecha de cada secuencia se especifica el porcentaje y número de lecturas detectadas en NB4-Cas9 control y NB4-Cas9 Edición, respectivamente. Los porcentajes de lecturas con estas secuencias son similares en ambas muestras, lo que indica que pueden tratarse de variantes presentes en la línea celular. Con una línea de puntos se indica el punto de corte esperado de la nucleasa Cas9.

REFERENCIAS

REFERENCIAS

- Acosta, S., Fiore, L., Carota, I. A., & Oliver, G. (2018). Use of two gRNAs for CRISPR/Cas9 improves bi-allelic homologous recombination efficiency in mouse embryonic stem cells. *Genesis*, *56*(5), e23212. doi: 10.1002/dvg.23212
- Ahmadmehrabi, K., Haque, A. R., Aleem, A., Griffiths, E. A., & Roloff, G. W. (2021). Targeted therapies for the evolving molecular landscape of acute myeloid leukemia. In *Cancers* (Vol. 13, Issue 18, p. 4646). doi: 10.3390/cancers13184646
- Apfeld, J., & Alper, S. (2018). What Can We Learn About Human Disease from the Nematode *C. elegans*? In *Methods in Molecular Biology* (Vol. 1706, pp. 53–75). doi: 10.1007/978-1-4939-7471-9_4
- Aquino-Jarquín, G. (2019). CRISPR-Cas14 is now part of the artillery for gene editing and molecular diagnostic. *Nanomedicine: Nanotechnology, Biology and Medicine*, *18*, 428–431. doi: 10.1016/j.nano.2019.03.006
- Aquino-Jarquín, G. (2021). Current advances in overcoming obstacles of CRISPR/Cas9 off-target genome editing. *Molecular Genetics and Metabolism*, *134*(1–2), 77–86. doi: 10.1016/j.ymgme.2021.08.002
- Araújo, W. L., Ishizaki, K., Nunes-Nesi, A., Larson, T. R., Tohge, T., Krahnert, I., Witt, S., Obata, T., Schauer, N., Graham, I. A., Leaver, C. J., & Fernie, A. R. (2010). Identification of the 2-Hydroxyglutarate and Isovaleryl-CoA Dehydrogenases as Alternative Electron Donors Linking Lysine Catabolism to the Electron Transport Chain of Arabidopsis Mitochondria. *The Plant Cell*, *22*(5), 1549–1563. doi: 10.1105/tpc.110.075630
- Arber, D. A., Orazi, A., Hasserjian, R., Thiele, J., Borowitz, M. J., Le Beau, M. M., Bloomfield, C. D., Cazzola, M., & Vardiman, J. W. (2016). The 2016 revision to the World Health Organization classification of myeloid neoplasms and acute leukemia. *Blood*, *127*(20), 2391–2405. doi: 10.1182/blood-2016-03-643544
- Arribere, J. A., Bell, R. T., Fu, B. X. H., Artiles, K. L., Hartman, P. S., & Fire, A. Z. (2014). Efficient Marker-Free Recovery of Custom Genetic Modifications with CRISPR/Cas9 in *Caenorhabditis elegans*. *Genetics*, *198*(3), 837–846. doi: 10.1534/genetics.114.169730
- Asou, H., Gombart, A. F., Takeuchi, S., Tanaka, H., Tanioka, M., Matsui, H., Kimura, A., Inaba, T., & Koeffler, H. P. (2003). Establishment of the acute myeloid leukemia cell line Kasumi-6 from a patient with a dominant-negative mutation in the DNA-binding region of the *C/EBPα* gene. *Genes Chromosomes and Cancer*, *36*(2), 167–174. doi: 10.1002/gcc.10161
- Bae, S., Park, J., & Kim, J. S. (2014). Cas-OFFinder: A fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics*, *30*(10), 1473–1475. doi: 10.1093/bioinformatics/btu048
- Becker, S., & Boch, J. (2021). TALE and TALEN genome editing technologies. *Gene and Genome Editing*, *2*, 100007. doi: 10.1016/j.ggedit.2021.100007
- Bhardwaj, A., & Nain, V. (2021). TALENs—an indispensable tool in the era of CRISPR: a mini review. *Journal of Genetic Engineering and Biotechnology*, *19*(1), 125. doi: 10.1186/s43141-021-00225-z
- Bibikova, M., Golic, M., Golic, K. G., & Carroll, D. (2002). Targeted Chromosomal Cleavage and Mutagenesis in *Drosophila* Using Zinc-Finger Nucleases. *Genetics*, *161*(3), 1169–1175. doi: 10.1093/genetics/161.3.1169
- Blayney, J., Foster, E., Jagielowicz, M., Kreuzer, M., Morotti, M., Reglinski, K., Xiao, J., & Hublitz, P. (2020). Unexpectedly High Levels of Inverted Re-Insertions Using Paired sgRNAs for Genomic Deletions. *Methods and Protocols*, *3*(3), 53. doi: 10.3390/mps3030053
- Boch, J., Scholze, H., Schornack, S., Landgraf, A., Hahn, S., Kay, S., Lahaye, T., Nickstadt, A., & Bonas, U. (2009). Breaking the Code of DNA Binding Specificity of TAL-Type III Effectors. *Science*, *326*(5959), 1509–1512. doi: 10.1126/science.1178811
- Brabetz, O., Alla, V., Angenendt, L., Schliemann, C., Berdel, W. E., Arteaga, M. F., & Mikesch, J. H. (2017). RNA-Guided CRISPR-Cas9 System-Mediated Engineering of Acute Myeloid Leukemia Mutations. *Molecular Therapy - Nucleic Acids*, *6*(March), 243–248. doi: 10.1016/j.omtn.2016.12.012
- Braeckman, B. P., Houthoofd, K., & Vanfleteren, J. R. (2009). Intermediary metabolism. *WormBook : The Online Review of C. Elegans Biology*, 1–24. doi: 10.1895/wormbook.1.146.1
- Brenner, S. (1974). The Genetics of *Caenorhabditis elegans*. *Genetics*, *77*(1), 71–94. Retrieved from <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1213120&tool=pmcentrez&rendertype=abstract>

- Brezgin, S., Kostyusheva, A., Kostyushev, D., & Chulanov, V. (2019). Dead Cas Systems: Types, Principles, and Applications. *International Journal of Molecular Sciences*, 20(23), 6041. doi: 10.3390/ijms20236041
- Brown, G., & Marcinkowska, E. (2017). Acute myeloid leukaemia: New targets and therapies. *International Journal of Molecular Sciences*, 18(12), 3–6. doi: 10.3390/ijms18122577
- Capecchi, M. R. (2005). Gene targeting in mice: functional analysis of the mammalian genome for the twenty-first century. *Nature Reviews Genetics*, 6(6), 507–512. doi: 10.1038/nrg1619
- Castañó, J., Herrero, A. B., Bursen, A., González, F., Marschalek, R., Gutiérrez, N. C., & Menendez, P. (2016). Expression of MLL-AF4 or AF4-MLL fusions does not impact the efficiency of DNA damage repair. *Oncotarget*, 7(21), 30440–30452. doi: 10.18632/oncotarget.8938
- Castanotto, D., & Rossi, J. J. (2004). Construction and transfection of PCR products expressing siRNAs or shRNAs in mammalian cells. *Methods in Molecular Biology (Clifton, N.J.)*, 252(8), 509–514. doi: 10.1385/1-59259-746-7:509
- Cathelin, S., Sharon, D., Subedi, A., Cojocari, D., Phillips, D. C., Levenson, J. D., MacBeth, K., Nicolay, B., Narayanaswamy, R., Ronseaux, S., Liu, G., & Chan, S. M. (2018). Combination of Enasidenib and Venetoclax Shows Superior Anti-Leukemic Activity Against IDH2 Mutated AML in Patient-Derived Xenograft Models. *Blood*, 132(Supplement 1), 562–562. doi: 10.1182/blood-2018-99-119688
- Cerchione, C., Romano, A., Daver, N., DiNardo, C., Jabbour, E. J., Konopleva, M., Ravandi-Kashani, F., Kadia, T., Martelli, M. P., Isidori, A., Martinelli, G., & Kantarjian, H. (2021). IDH1/IDH2 Inhibition in Acute Myeloid Leukemia. *Frontiers in Oncology*, 11(March), 1–9. doi: 10.3389/fonc.2021.639387
- Chakrabarti, A. M., Henser-Brownhill, T., Monserrat, J., Poetsch, A. R., Luscombe, N. M., & Scaffidi, P. (2019). Target-Specific Precision of CRISPR-Mediated Genome Editing. *Molecular Cell*, 73(4), 699–713.e6. doi: 10.1016/j.molcel.2018.11.031
- Charlesworth, C. T., Deshpande, P. S., Dever, D. P., Camarena, J., Lemgart, V. T., Cromer, M. K., Vakulskas, C. A., Collingwood, M. A., Zhang, L., Bode, N. M., Behlke, M. A., Dejene, B., Cieniewicz, B., Romano, R., Lesch, B. J., Gomez-Ospina, N., Mantri, S., Pavel-Dinu, M., Weinberg, K. I., & Porteus, M. H. (2019). Identification of preexisting adaptive immunity to Cas9 proteins in humans. *Nature Medicine*, 25(2), 249–254. doi: 10.1038/s41591-018-0326-x
- Chen, X., Xu, F., Zhu, C., Ji, J., Zhou, X., Feng, X., & Guang, S. (2014). Dual sgRNA-directed gene knockout using CRISPR/Cas9 technology in *Caenorhabditis elegans*. *Scientific Reports*, 4(1), 7581. doi: 10.1038/srep07581
- Clark, O., Yen, K., & Mellinghoff, I. K. (2016). Molecular Pathways: Isocitrate Dehydrogenase Mutations in Cancer. *Clinical Cancer Research*, 22(8), 1837–1842. doi: 10.1158/1078-0432.CCR-13-1333
- Clement, K., Rees, H., Canver, M. C., Gehrke, J. M., Farouni, R., Hsu, J. Y., Cole, M. A., Liu, D. R., Joung, J. K., Bauer, D. E., & Pinello, L. (2019). CRISPResso2 provides accurate and rapid genome editing sequence analysis. *Nature Biotechnology*, 37(3), 224–226. doi: 10.1038/s41587-019-0032-3
- Coelho, M. A., De Braekeleer, E., Firth, M., Bista, M., Lukasiak, S., Cuomo, M. E., & Taylor, B. J. M. (2020). CRISPR GUARD protects off-target sites from Cas9 nuclease activity using short guide RNAs. *Nature Communications*, 11(1), 4132. doi: 10.1038/s41467-020-17952-5
- Culetto, E. (2000). A role for *Caenorhabditis elegans* in understanding the function and interactions of human disease genes. *Human Molecular Genetics*, 9(6), 869–877. doi: 10.1093/hmg/9.6.869
- Cyranoski, D. (2019). The CRISPR-baby scandal: what's next for human gene-editing. *Nature*, 566(7745), 440–442. doi: 10.1038/d41586-019-00673-1
- Dang, L., & Su, S.-S. M. (2017). Isocitrate Dehydrogenase Mutation and (R)-2-Hydroxyglutarate: From Basic Discovery to Therapeutics Development. *Annual Review of Biochemistry*, 86(1), 305–331. doi: 10.1146/annurev-biochem-061516-044732
- Di Stazio, M., Foschi, N., Athanasakis, E., Gasparini, P., & D'Adamo, A. P. (2021). Systematic analysis of factors that improve homologous direct repair (HDR) efficiency in CRISPR/Cas9 technique. *PLOS ONE*, 16(3), e0247603. doi: 10.1371/journal.pone.0247603
- Doench, J. G., Hartenian, E., Graham, D. B., Tothova, Z., Hegde, M., Smith, I., Sullender, M., Ebert, B. L., Xavier, R. J., & Root, D. E. (2014). Rational design of highly active sgRNAs for CRISPR-Cas9-mediated gene inactivation. *Nature Biotechnology*, 32(12), 1262–1267. doi: 10.1038/nbt.3026
- Döhner, H., Estey, E., Grimwade, D., Amadori, S., Appelbaum, F. R., Büchner, T., Dombret, H., Ebert, B. L., Fenaux, P., Larson, R. A., Levine, R. L., Lo-Coco, F., Naoe, T., Niederwieser, D., Ossenkoppele, G. J., Sanz, M., Sierra, J., Tallman, M. S., Tien, H.-F., ... Bloomfield, C. D. (2017). Diagnosis and management of AML in adults: 2017 ELN recommendations from an international expert panel. *Blood*, 129(4), 424–447. doi: 10.1182/blood-2016-08-733196

- Döhner, H., Weisdorf, D. J., & Bloomfield, C. D. (2015). Acute Myeloid Leukemia. *New England Journal of Medicine*, 373(12), 1136–1152. doi: 10.1056/NEJMra1406184
- Doyon, Y., McCammon, J. M., Miller, J. C., Faraji, F., Ngo, C., Katibah, G. E., Amora, R., Hocking, T. D., Zhang, L., Rebar, E. J., Gregory, P. D., Urnov, F. D., & Amacher, S. L. (2008). Heritable targeted gene disruption in zebrafish using designed zinc-finger nucleases. *Nature Biotechnology*, 26(6), 702–708. doi: 10.1038/nbt1409
- Durand, S., & Cimarelli, A. (2011). The Inside out of Lentiviral Vectors. *Viruses*, 3(2), 132–159. doi: 10.3390/v3020132
- Duzgunes, N., Singh, V., Wu, N., Xu, Y., Peng, X., Zheng, Y., Jin, C., Lu, X., Han, D., Fu, H., & Chen, C. (2021). Inactivation of Latent HIV-1 Proviral DNA Using Clustered Regularly Interspaced Short Palindromic Repeats/Cas9 Treatment and the Assessment of Off-Target Effects. *Frontiers in Microbiology | Www.Frontiersin.Org*, 1, 629153. doi: 10.3389/fmicb.2021.629153
- Dyson, M. R. (2016). Fundamentals of Expression in Mammalian Cells. In *Advances in Experimental Medicine and Biology* (Vol. 896, pp. 217–224). doi: 10.1007/978-3-319-27216-0_14
- Edwards, C. B., Copes, N., Brito, A. G., Canfield, J., & Bradshaw, P. C. (2013). Malate and Fumarate Extend Lifespan in *Caenorhabditis elegans*. *PLoS ONE*, 8(3), e58345. doi: 10.1371/journal.pone.0058345
- Farboud, B., & Meyer, B. J. (2015). Dramatic Enhancement of Genome Editing by CRISPR/Cas9 Through Improved Guide RNA Design. *Genetics*, 199(4), 959–971. doi: 10.1534/genetics.115.175166
- Fathi, A. T., DiNardo, C. D., Kline, I., Kenvin, L., Gupta, I., Attar, E. C., Stein, E. M., & De Botton, S. (2018). Differentiation syndrome associated with enasidenib, a selective inhibitor of mutant isocitrate dehydrogenase 2 analysis of a phase 1/2 study. *JAMA Oncology*, 4(8), 1106–1110. doi: 10.1001/jamaoncol.2017.4695
- Fu, X., Chin, R. M., Vergnes, L., Hwang, H., Deng, G., Xing, Y., Pai, M. Y., Li, S., Ta, L., Fazlollahi, F., Chen, C., Prins, R. M., Teitell, M. A., Nathanson, D. A., Lai, A., Faull, K. F., Jiang, M., Clarke, S. G., Cloughesy, T. F., ... Huang, J. (2015). 2-Hydroxyglutarate Inhibits ATP Synthase and mTOR Signaling. *Cell Metabolism*, 22(3), 508–515. doi: 10.1016/j.cmet.2015.06.009
- Fuster, Ó., Llop, M., Dolz, S., García, P., Such, E., Ibáñez, M., Luna, I., Gómez, I., López, M., Cervera, J., Montesinos, P., Moscardó, F., Cordón, L., Solves, P., de Juan, I., Palanca, S., Bolufer, P., Sanz, M. Á., & Barragán, E. (2013). Adverse prognostic value of MYBL2 overexpression and association with microRNA-30 family in acute myeloid leukemia patients. *Leukemia Research*, 37(12), 1690–1696. doi: 10.1016/j.leukres.2013.09.015
- Gao, A. W., uit de Bos, J., Sterken, M. G., Kammenga, J. E., Smith, R. L., & Houtkooper, R. H. (2018). Forward and reverse genetics approaches to uncover metabolic aging pathways in *Caenorhabditis elegans*. *Biochimica et Biophysica Acta (BBA) - Molecular Basis of Disease*, 1864(9), 2697–2706. doi: 10.1016/j.bbadis.2017.09.006
- Ghanta, K. S., Chen, Z., Mir, A., Dokshin, G. A., Krishnamurthy, P. M., Yoon, Y., Gallant, J., Xu, P., Zhang, X.-O., Ozturk, A. R., Shin, M., Idrizi, F., Liu, P., Gneid, H., Edraki, A., Lawson, N. D., Rivera-Pérez, J. A., Sontheimer, E. J., Watts, J. K., & Mello, C. C. (2021). 5'-Modifications improve potency and efficacy of DNA donors for precision genome editing. *ELife*, 10, 1–20. doi: 10.7554/elife.72216
- González-Romero, E., Martínez-Valiente, C., García-Ruiz, C., Vázquez-Manrique, R. P., Cervera, J., & Sanjuan-Pla, A. (2019). CRISPR to fix bad blood: a new tool in basic and clinical hematology. *Haematologica*, 104(5), 881–893. doi: 10.3324/haematol.2018.211359
- Gootenberg, J. S., Abudayyeh, O. O., Lee, J. W., Dy, A. J., Joung, J., Verdine, V., Daringer, N. M., Freije, C. A., Myhrvold, C., Roby, P., Livny, J., Regev, A., Koonin, E. V., & Hung, D. T. (2017). *Nucleic acid detection with CRISPR-Cas13a / C2c2*. 356(6336), 438–442. doi: 10.1126/science.aam9321.Nucleic
- Grad, Y., Aach, J., Hayes, G. D., Reinhart, B. J., Church, G. M., Ruvkun, G., & Kim, J. (2003). Computational and Experimental Identification of *C. elegans* microRNAs. *Molecular Cell*, 11(5), 1253–1263. doi: 10.1016/S1097-2765(03)00153-9
- Graham, F. L., Smiley, J., Russell, W. C., & Nairn, R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *Journal of General Virology*, 36(1), 59–72. doi: 10.1099/0022-1317-36-1-59
- Granados-Riveron, J. T., & Aquino-Jarquín, G. (2018). CRISPR–Cas13 Precision Transcriptome Engineering in Cancer. *Cancer Research*, 1–8. doi: 10.1158/0008-5472.CAN-18-0785
- Guo, Q., Mintier, G., Storton, D., Wang, X., Xiao, X., & Kienzle, B. (2018). ‘Cold shock’ increases the frequency of homology directed repair gene editing in induced pluripotent stem cells. *Scientific Reports, February 2017*, 1–11. doi: 10.1038/s41598-018-20358-5

- Harrington, L. B., Burstein, D., Chen, J. S., Paez-Espino, D., Ma, E., Witte, I. P., Cofsky, J. C., Kyrpides, N. C., Banfield, J. F., & Doudna, J. A. (2018). Programmed DNA destruction by miniature CRISPR-Cas14 enzymes. *Science*, *362*(6416), 839–842. doi: 10.1126/science.aav4294
- Heigwer, F., Kerr, G., & Boutros, M. (2014). E-CRISP: fast CRISPR target site identification. *Nature Methods*, *11*(2), 122–123. doi: 10.1038/nmeth.2812
- Hirakawa, M. P., Krishnakumar, R., Timlin, J. A., Carney, J. P., & Butler, K. S. (2020). Gene editing and CRISPR in the clinic: current and future perspectives. *Bioscience Reports*, *40*(4). doi: 10.1042/BSR20200127
- Hirko, A., Tang, F., & Hughes, J. (2003). Cationic Lipid Vectors for Plasmid DNA Delivery. *Current Medicinal Chemistry*, *10*(14), 1185–1193. doi: 10.2174/0929867033457412
- Hitoshi, N., Ken-ichi, Y., & Jun-ichi, M. (1991). Efficient selection for high-expression transfectants with a novel eukaryotic vector. *Gene*, *108*(2), 193–199. doi: 10.1016/0378-1119(91)90434-D
- Horiuchi, K., Perez-Cerezales, S., Papasaikas, P., Ramos-Ibeas, P., López-Cardona, A. P., Laguna-Barraza, R., Fonseca Balvís, N., Pericuesta, E., Fernández-González, R., Planells, B., Viera, A., Suja, J. A., Ross, P. J., Alén, F., Orío, L., Rodríguez de Fonseca, F., Pintado, B., Valcárcel, J., & Gutiérrez-Adán, A. (2018). Impaired Spermatogenesis, Muscle, and Erythrocyte Function in U12 Intron Splicing-Defective Zrsr1 Mutant Mice. *Cell Reports*, *23*(1), 143–155. doi: 10.1016/j.celrep.2018.03.028
- Horlbeck, M. A., Witkowsky, L. B., Guglielmi, B., Replogle, J. M., Gilbert, L. A., Villalta, J. E., Torigoe, S. E., Tjian, R., & Weissman, J. S. (2016). Nucleosomes impede Cas9 access to DNA in vivo and in vitro. *ELife*, *5*(MARCH2016), 1–21. doi: 10.7554/eLife.12677
- Howe, K. L., Achuthan, P., Allen, J., Allen, J., Alvarez-Jarreta, J., Amode, M. R., Armean, I. M., Azov, A. G., Bennett, R., Bhai, J., Billis, K., Boddu, S., Charkhchi, M., Cummins, C., Da Rin Fioretto, L., Davidson, C., Dodiya, K., El Houdaigui, B., Fatima, R., ... Flicek, P. (2021). Ensembl 2021. *Nucleic Acids Research*, *49*(D1), D884–D891. doi: 10.1093/nar/gkaa942
- Hu, Y., Flockhart, I., Vinayagam, A., Bergwitz, C., Berger, B., Perrimon, N., & Mohr, S. E. (2011). An integrative approach to ortholog prediction for disease-focused and other functional studies. *BMC Bioinformatics*, *12*(1), 357. doi: 10.1186/1471-2105-12-357
- Intlekofer, A. M., Shih, A. H., Wang, B., Nazir, A., Rustenburg, A. S., Albanese, S. K., Patel, M., Famulare, C., Correa, F. M., Takemoto, N., Durani, V., Liu, H., Taylor, J., Farnoud, N., Papaemmanuil, E., Cross, J. R., Tallman, M. S., Arcila, M. E., Roshal, M., ... Stein, E. M. (2018). Acquired resistance to IDH inhibition through trans or cis dimer-interface mutations. *Nature*, *559*(7712), 125–129. doi: 10.1038/s41586-018-0251-7
- Isaac, R. S., Jiang, F., Doudna, J. A., Lim, W. A., Narlikar, G. J., & Almeida, R. (2016). Nucleosome breathing and remodeling constrain CRISPR-Cas9 function. *ELife*, *5*(APRIL2016). doi: 10.7554/eLife.13450
- Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., & Nakata, A. (1987). Nucleotide sequence of the iap gene, responsible for alkaline phosphatase isozyme conversion in *Escherichia coli*, and identification of the gene product. *Journal of Bacteriology*, *169*(12), 5429–5433. doi: 10.1128/jb.169.12.5429-5433.1987
- Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., & Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science*, *337*(6096), 816–821. doi: 10.1126/science.1225829
- Kaeuferle, T., Deisenberger, L., Jablonowski, L., Stief, T. A., Blaeschke, F., Willier, S., & Feuchtinger, T. (2020). CRISPR-Cas9-Mediated Glucocorticoid Resistance in Virus-Specific T Cells for Adoptive T Cell Therapy Posttransplantation. *Molecular Therapy*, *28*(9), 1965–1973. doi: 10.1016/j.ymthe.2020.06.002
- Kernytsky, A., Wang, F., Hansen, E., Schalm, S., Straley, K., Gliser, C., Yang, H., Travins, J., Murray, S., Dorsch, M., Agresta, S., Schenkein, D. P., Biller, S. A., Su, S. M., Liu, W., & Yen, K. E. (2015). IDH2 mutation-induced histone and DNA hypermethylation is progressively reversed by small-molecule inhibition. *Blood*, *125*(2), 296–303. doi: 10.1182/blood-2013-10-533604
- Kim, H., Ishidate, T., Ghanta, K. S., Seth, M., Conte, D., Shirayama, M., & Mello, C. C. (2014). A Co-CRISPR Strategy for Efficient Genome Editing in *Caenorhabditis elegans*. *Genetics*, *197*(4), 1069–1080. doi: 10.1534/genetics.114.166389
- Kim, H. J., Lee, H. J., Kim, H., Cho, S. W., & Kim, J.-S. (2009). Targeted genome editing in human cells with zinc finger nucleases constructed via modular assembly. *Genome Research*, *19*(7), 1279–1288. doi: 10.1101/gr.089417.108
- Kim, J.-S. (2016). Genome editing comes of age. *Nature Protocols*, *11*(9), 1573–1578. doi:

- 10.1038/nprot.2016.104
- Kim, S., Ji, S., & Koh, H. R. (2021). CRISPR as a Diagnostic Tool. *Biomolecules*, *11*(8), 1162. doi: 10.3390/biom11081162
- Kim, S., Kim, D., Cho, S. W., Kim, J., & Kim, J.-S. (2014). Highly efficient RNA-guided genome editing in human cells via delivery of purified Cas9 ribonucleoproteins. *Genome Research*, *24*(6), 1012–1019. doi: 10.1101/gr.171322.113
- Kim, Y. G., Cha, J., & Chandrasegaran, S. (1996). Hybrid restriction enzymes: zinc finger fusions to Fok I cleavage domain. *Proceedings of the National Academy of Sciences*, *93*(3), 1156–1160. doi: 10.1073/pnas.93.3.1156
- Kleinstiver, B. P., Pattanayak, V., Prew, M. S., Tsai, S. Q., Nguyen, N. T., Zheng, Z., & Joung, J. K. (2016). High-fidelity CRISPR-Cas9 nucleases with no detectable genome-wide off-target effects. *Nature*, *529*(7587), 490–495. doi: 10.1038/nature16526
- Konstantinidis, G., & Tavernarakis, N. (2021). Molecular Basis of Neuronal Autophagy in Ageing: Insights from *Caenorhabditis elegans*. *Cells*, *10*(3), 694. doi: 10.3390/cells10030694
- Kopchick, J. J., & Hartline, R. A. (1979). alpha-Hydroxyglutarate as an intermediate in the catabolism of alpha-amino adipate by *Pseudomonas putida*. *Journal of Biological Chemistry*, *254*(9), 3259–3263. doi: 10.1016/S0021-9258(18)50753-5
- Kranendijk, M., Struys, E. A., Schaftingen, E. Van, Gibson, K. M., Kanhai, W. A., Knaap, M. S. Van Der, Amiel, J., Buist, N. R., Das, A. M., Klerk, J. B. De, Feigenbaum, A. S., Grange, D. K., Hofstede, F. C., Holme, E., Kirk, E. P., Korman, S. H., Morava, E., Morris, A., Smeitink, J., ... Salomons, G. S. (2010). DH2 mutations in patients with D-2-hydroxyglutaric aciduria. *Science*, *330*(October), 7.
- Kwon, Y.-W., Ahn, H.-S., Lee, J.-W., Yang, H.-M., Cho, H.-J., Kim, S. J., Lee, S.-H., Yang, H.-M., Jang, H.-D., Kim, S. J., & Kim, H.-S. (2021). HLA DR Genome Editing with TALENs in Human iPSCs Produced Immune-Tolerant Dendritic Cells. *Stem Cells International*, *2021*, 1–14. doi: 10.1155/2021/8873383
- Labun, K., Montague, T. G., Krause, M., Torres Cleuren, Y. N., Tjeldnes, H., & Valen, E. (2019). CHOPCHOP v3: expanding the CRISPR web toolbox beyond genome editing. *Nucleic Acids Research*, *47*(W1), W171–W174. doi: 10.1093/nar/gkz365
- Lander, E. S., Baylis, F., Zhang, F., Charpentier, E., Berg, P., Bourgain, C., Friedrich, B., Joung, J. K., Li, J., Liu, D., Naldini, L., Nie, J.-B., Qiu, R., Schoene-Seifert, B., Shao, F., Terry, S., Wei, W., & Winnacker, E.-L. (2019). Adopt a moratorium on heritable genome editing. *Nature*, *567*(7747), 165–168. doi: 10.1038/d41586-019-00726-5
- Lanotte, M., Martin-Thouvenin, V., Najman, S., Balerini, P., Valensi, F., & Berger, R. (1991). NB4, a maturation inducible cell line with t(15;17) marker isolated from a human acute promyelocytic leukemia (M3). *Blood*, *77*(5), 1080–1086. doi: 10.1182/blood.v77.5.1080.bloodjournal7751080
- Lee, R. C., Feinbaum, R. L., & Ambros, V. (1993). The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell*, *75*(5), 843–854. doi: 10.1016/0092-8674(93)90529-Y
- Lehner, R., Wang, X., & Hunziker, P. (2013). Plasmid linearization changes shape and efficiency of transfection complexes. *European Journal of Nanomedicine*, *5*(4), 205–212. doi: 10.1515/ejnm-2013-0028
- Li, Q., & Engebrecht, J. A. (2021). BRCA1 and BRCA2 Tumor Suppressor Function in Meiosis. *Frontiers in Cell and Developmental Biology*, *9*(April), 1–14. doi: 10.3389/fcell.2021.668309
- Liang, X., Potter, J., Kumar, S., Zou, Y., Quintanilla, R., Sridharan, M., Carte, J., Chen, W., Roark, N., Ranganathan, S., Ravinder, N., & Chesnut, J. D. (2015). Rapid and highly efficient mammalian cell engineering via Cas9 protein transfection. *Journal of Biotechnology*, *208*, 44–53. doi: 10.1016/j.jbiotec.2015.04.024
- Lim, L. P., Lau, C. N., Weintein, E., Abdelhakim, A., Yekta, S., Rhoades, M. W., Burge, C. B., & Bartel, D. (2003). The microRNAs of *Caenorhabditis elegans*. *Genes & Development*, *17*(8), 991–1008. doi: 10.1101/gad.1074403
- Lin, Y., Cradick, T. J., Brown, M. T., Deshmukh, H., Ranjan, P., Sarode, N., Wile, B. M., Vertino, P. M., Stewart, F. J., & Bao, G. (2014). CRISPR/Cas9 systems have off-target activity with insertions or deletions between target DNA and guide RNA sequences. *Nucleic Acids Research*, *42*(11), 7473–7485. doi: 10.1093/nar/gku402
- Liu, M., Rehman, S., Tang, X., Gu, K., Fan, Q., Chen, D., & Ma, W. (2019). Methodologies for improving HDR efficiency. *Frontiers in Genetics*, *10*(JAN), 1–9. doi: 10.3389/fgene.2018.00691
- Liu, W., Li, L., Jiang, J., Wu, M., & Lin, P. (2021). Applications and challenges of CRISPR-Cas gene-editing

- to disease treatment in clinics. *Precision Clinical Medicine*, 4(3), 179–191. doi: 10.1093/pcmedi/pbab014
- Lloyd, A., Plaisier, C. L., Carroll, D., & Drews, G. N. (2005). Targeted mutagenesis using zinc-finger nucleases in Arabidopsis. *Proceedings of the National Academy of Sciences of the United States of America*, 102(6), 2232–2237. doi: 10.1073/pnas.0409339102
- Ma, Y., Shen, N., Wicha, M. S., & Luo, M. (2021). The Roles of the Let-7 Family of MicroRNAs in the Regulation of Cancer Stemness. *Cells*, 10(9), 2415. doi: 10.3390/cells10092415
- Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., Norville, J. E., & Church, G. M. (2013). RNA-guided human genome engineering via Cas9. *Science*, 339(6121), 823–826. doi: 10.1126/science.1232033
- Manghwar, H., Li, B., Ding, X., Hussain, A., Lindsey, K., Zhang, X., & Jin, S. (2020). CRISPR/Cas Systems in Genome Editing: Methodologies and Tools for sgRNA Design, Off-Target Evaluation, and Strategies to Mitigate Off-Target Effects. In *Advanced Science* (Vol. 7, Issue 6). John Wiley and Sons Inc. doi: 10.1002/advs.201902312
- Markaki, M., & Tavernarakis, N. (2010). Modeling human diseases in *Caenorhabditis elegans*. *Biotechnology Journal*, 5(12), 1261–1276. doi: 10.1002/biot.201000183
- Maruyama, J. (2021). Genome Editing Technology and Its Application Potentials in the Industrial Filamentous Fungus *Aspergillus oryzae*. *Journal of Fungi*, 7(8), 638. doi: 10.3390/jof7080638
- Matres, J. M., Hilscher, J., Datta, A., Armario-Nájera, V., Baysal, C., He, W., Huang, X., Zhu, C., Valizadeh-Kamran, R., Trijatmiko, K. R., Capell, T., Christou, P., Stoger, E., & Slamet-Loedin, I. H. (2021). Genome editing in cereal crops: an overview. *Transgenic Research*, 30(4), 461–498. doi: 10.1007/s11248-021-00259-6
- Matsuo, Y., MacLeod, R. A. F., Uphoff, C. C., Drexler, H. G., Nishizaki, C., Katayama, Y., Kimura, G., Fujii, N., Omoto, E., Harada, M., & Orita, K. (1997). Two acute monocytic leukemia (AML-M5a) cell lines (MOLM-13 and MOLM-14) with interclonal phenotypic heterogeneity showing MLL-AF9 fusion resulting from an occult chromosome insertion, *ins(11;9)(q23;p22p23)*. *Leukemia*, 11(9), 1469–1477. doi: 10.1038/sj.leu.2400768
- Medeiros, B. C., Fathi, A. T., DiNardo, C. D., Pollyea, D. A., Chan, S. M., & Swords, R. (2017). Isocitrate dehydrogenase mutations in myeloid malignancies. *Leukemia*, 31(2), 272–281. doi: 10.1038/leu.2016.275
- Miller, J. C., Tan, S., Qiao, G., Barlow, K. A., Wang, J., Xia, D. F., Meng, X., Paschon, D. E., Leung, E., Hinkley, S. J., Dulay, G. P., Hua, K. L., Ankoudinova, I., Cost, G. J., Urnov, F. D., Zhang, H. S., Holmes, M. C., Zhang, L., Gregory, P. D., & Rebar, E. J. (2011). A TALE nuclease architecture for efficient genome editing. *Nature Biotechnology*, 29(2), 143–150. doi: 10.1038/nbt.1755
- Mitreva, M., Wendl, M. C., Martin, J., Wylie, T., Yin, Y., Larson, A., Parkinson, J., Waterston, R. H., & McCarter, J. P. (2006). Codon usage patterns in Nematoda: Analysis based on over 25 million codons in thirty-two species. *Genome Biology*, 7(8), R75. doi: 10.1186/gb-2006-7-8-r75
- Miyaoka, Y., Berman, J. R., Cooper, S. B., Mayerl, S. J., Chan, A. H., Zhang, B., Karlin-Neumann, G. A., & Conklin, B. R. (2016). Systematic quantification of HDR and NHEJ reveals effects of locus, nuclease, and cell type on genome-editing. *Scientific Reports*, 6(1), 23549. doi: 10.1038/srep23549
- Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Almendros, C. (2009). Short motif sequences determine the targets of the prokaryotic CRISPR defence system. *Microbiology*, 155(3), 733–740. doi: 10.1099/mic.0.023960-0
- Mojica, F. J. M., Díez-Villaseñor, C., García-Martínez, J., & Soria, E. (2005). Intervening Sequences of Regularly Spaced Prokaryotic Repeats Derive from Foreign Genetic Elements. *Journal of Molecular Evolution*, 60(2), 174–182. doi: 10.1007/s00239-004-0046-3
- Mojica, F. J. M., Ferrer, C., Juez, G., & Rodríguez-Valera, F. (1995). Long stretches of short tandem repeats are present in the largest replicons of the Archaea *Haloferax mediterranei* and *Haloferax volcanii* and could be involved in replicon partitioning. *Molecular Microbiology*, 17(1), 85–93. doi: 10.1111/j.1365-2958.1995.mmi_17010085.x
- Mojica, F. J. M., & Montoliu, L. (2016). On the Origin of CRISPR-Cas Technology: From Prokaryotes to Mammals. *Trends in Microbiology*, 24(10), 811–820. doi: 10.1016/j.tim.2016.06.005
- Morton, J., Davis, M. W., Jorgensen, E. M., & Carroll, D. (2006). Induction and repair of zinc-finger nuclease-targeted double-strand breaks in *Caenorhabditis elegans* somatic cells. *Proceedings of the National Academy of Sciences*, 103(44), 16370–16375. doi: 10.1073/pnas.0605633103
- Musa, J., Aynaud, M.-M., Mirabeau, O., Delattre, O., & Grünewald, T. G. (2017). MYBL2 (B-Myb): a central regulator of cell proliferation, cell survival and differentiation involved in tumorigenesis.

- Cell Death & Disease*, 8(6), e2895–e2895. doi: 10.1038/cddis.2017.244
- Okamoto, S., Amaishi, Y., Maki, I., Enoki, T., & Mineno, J. (2019). Highly efficient genome editing for single-base substitutions using optimized ssODNs with Cas9-RNPs. *Scientific Reports*, 9(1), 4811. doi: 10.1038/s41598-019-41121-4
- Orioli, A., Pascali, C., Pagano, A., Teichmann, M., & Dieci, G. (2012). RNA polymerase III transcription control elements: Themes and variations. *Gene*, 493(2), 185–194. doi: 10.1016/j.gene.2011.06.015
- Özcan, A., Krajcski, R., Ioannidi, E., Lee, B., Gardner, A., Makarova, K. S., Koonin, E. V., Abudayyeh, O. O., & Gootenberg, J. S. (2021). Programmable RNA targeting with the single-protein CRISPR effector Cas7-11. *Nature*, 597(7878), 720–725. doi: 10.1038/s41586-021-03886-5
- Paix, A., Folkmann, A., & Seydoux, G. (2017). Precision genome editing using CRISPR-Cas9 and linear repair templates in *C. elegans*. *Methods*, 121–122, 86–93. doi: 10.1016/j.ymeth.2017.03.023
- Papaemmanuil, E., Gerstung, M., Bullinger, L., Gaidzik, V. I., Paschka, P., Roberts, N. D., Potter, N. E., Heuser, M., Thol, F., Bolli, N., Gundem, G., Van Loo, P., Martincorena, I., Ganly, P., Mudie, L., McLaren, S., O’Meara, S., Raine, K., Jones, D. R., ... Campbell, P. J. (2016). Genomic Classification and Prognosis in Acute Myeloid Leukemia. *New England Journal of Medicine*, 374(23), 2209–2221. doi: 10.1056/NEJMoa1516192
- Peixoto, C. A., de Melo, J. V., Kramer, J. M., & de Souza, W. (1998). Ultrastructural analyses of the *Caenorhabditis elegans* rol-6 (su1006) mutant, which produces abnormal cuticle collagen. *The Journal of Parasitology*, 84(1), 45–49. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9488336>
- Peng, B., Theng, P. Y., & Le, M. T. N. (2021). Essential functions of miR-125b in cancer. *Cell Proliferation*, 54(2). doi: 10.1111/cpr.12913
- Peterson, L., Bloomfield, C., Döhner, H., Niemeyer, C., Godley, & L. (2016). *Myeloid neoplasms with germline predisposition*. In: Swerdlow S, Campo E, HarrisNL, et al, eds. *World Health Organization Classification of Tumours of Haematopoietic and Lymphoid Tissues*. (Update to). Lyon, France: World Health Organization.
- Pinello, L., Canver, M. C., Hoban, M. D., Orkin, S. H., Kohn, D. B., Bauer, D. E., & Yuan, G. C. (2016). Analyzing CRISPR genome-editing experiments with CRISPResso. *Nature Biotechnology*, 34(7), 695–697. doi: 10.1038/nbt.3583
- Pirozzi, C. J., & Yan, H. (2021). The implications of IDH mutations for cancer development and therapy. *Nature Reviews Clinical Oncology*, 18(10), 645–661. doi: 10.1038/s41571-021-00521-0
- Pittermann, E., Lachmann, N., Maclean, G., Emmrich, S., Ackermann, M., Ohring, G., Schlegelberger, B., Welte, K., Schambach, A., Heckl, D., Orkin, S. H., Cantz, T., & Klusmann, J.-H. (2017). Gene correction of HAX1 reversed Kostmann disease phenotype in patient-specific induced pluripotent stem cells. *Blood Advances*, 1(14), 903–914. doi: 10.1182/bloodadvances.2016003798
- Porteus, M. H. (2003). Chimeric Nucleases Stimulate Gene Targeting in Human Cells. *Science*, 300(5620), 763–763. doi: 10.1126/science.1078395
- Putowski, M., & Giannopoulos, K. (2021). Perspectives on Precision Medicine in Chronic Lymphocytic Leukemia: Targeting Recurrent Mutations—NOTCH1, SF3B1, MYD88, BIRC3. *Journal of Clinical Medicine*, 10(16), 3735. doi: 10.3390/jcm10163735
- Qu, H., Shan, K., Tang, C., Cui, G., Fu, G., Qi, Y., Cui, J., Li, J., Wang, R., Feng, N., & Chen, Y. Q. (2021). A novel small-molecule fatty acid synthase inhibitor with antitumor activity by cell cycle arrest and cell division inhibition. *European Journal of Medicinal Chemistry*, 219(2), 113407. doi: 10.1016/j.ejmech.2021.113407
- Ran, F. A., Hsu, P. D., Wright, J., Agarwala, V., Scott, D. A., & Zhang, F. (2013). Genome engineering using the CRISPR-Cas9 system. *Nature Protocols*, 8(11), 2281–2308. doi: 10.1038/nprot.2013.143
- Rathe, S. K., Moriarity, B. S., Stoltenberg, C. B., Kurata, M., Aumann, N. K., Rahrman, E. P., Bailey, N. J., Melrose, E. G., Beckmann, D. A., Liska, C. R., & Largaespada, D. A. (2015). Using RNA-seq and targeted nucleases to identify mechanisms of drug resistance in acute myeloid leukemia. *Scientific Reports*, 4(1), 6048. doi: 10.1038/srep06048
- Reinhart, B. J., Slack, F. J., Basson, M., Pasquinelli, A. E., Bettinger, J. C., Rougvie, A. E., Horvitz, H. R., & Ruvkun, G. (2000). The 21-nucleotide let-7 RNA regulates developmental timing in *Caenorhabditis elegans*. *Nature*, 403(6772), 901–906. doi: 10.1038/35002607
- Reyon, D., Tsai, S. Q., Khgayer, C., Foden, J. A., Sander, J. D., & Joung, J. K. (2012). FLASH assembly of TALENs for high-throughput genome editing. *Nature Biotechnology*, 30(5), 460–465. doi: 10.1038/nbt.2170
- Roboz, G. J., Rosenblatt, T., Arellano, M., Gobbi, M., Altman, J. K., Montesinos, P., O’Connell, C., Solomon,

- S. R., Pigneux, A., Vey, N., Hills, R., Jacobsen, T. F., Gianella-Borradori, A., Foss, Ø., Vetrhusand, S., & Giles, F. J. (2014). International randomized phase III study of elacytarabine versus investigator choice in patients with relapsed/refractory acute myeloid leukemia. *Journal of Clinical Oncology*, *32*(18), 1919–1926. doi: 10.1200/JCO.2013.52.8562
- Rouet, P., Smih, F., & Jasin, M. (1994). Introduction of double-strand breaks into the genome of mouse cells by expression of a rare-cutting endonuclease. *Molecular and Cellular Biology*, *14*(12), 8096–8106. doi: 10.1128/mcb.14.12.8096
- Rudin, N., & Haber, J. E. (1988). Efficient repair of HO-induced chromosomal breaks in *Saccharomyces cerevisiae* by recombination between flanking homologous sequences. *Molecular and Cellular Biology*, *8*(9), 3918–3928. doi: 10.1128/mcb.8.9.3918-3928.1988
- Salanga, C. M., & Salanga, M. C. (2021). Genotype to phenotype: Crispr gene editing reveals genetic compensation as a mechanism for phenotypic disjunction of morphants and mutants. In *International Journal of Molecular Sciences* (Vol. 22, Issue 7, p. 3472). doi: 10.3390/ijms22073472
- Sanjana, N. E., Shalem, O., & Zhang, F. (2014). Improved vectors and genome-wide libraries for CRISPR screening. *Nature Medicine*, *11*(8), 783–784. doi: 10.1038/nmeth.3047.Improved
- Schwartzman, J. M., Reuter, V. P., Koche, R. P., & Thompson, C. B. (2019). 2-hydroxyglutarate inhibits MyoD-mediated differentiation by preventing H3K9 demethylation. *Proceedings of the National Academy of Sciences of the United States of America*, *116*(26), 12851–12856. doi: 10.1073/pnas.1817662116
- Sentmanat, M. F., Peters, S. T., Florian, C. P., Connelly, J. P., & Pruett-Miller, S. M. (2018). A Survey of Validation Strategies for CRISPR-Cas9 Editing. *Scientific Reports*, *8*(1), 888. doi: 10.1038/s41598-018-19441-8
- Seroby, V., Kontarakis, Z., El-Brolosy, M. A., Welker, J. M., Tolstenkov, O., Saadeldein, A. M., Retzer, N., Gottschalk, A., Wehman, A. M., & Stainier, D. Y. R. (2020). Transcriptional adaptation in *Caenorhabditis elegans*. *eLife*, *9*, 1–20. doi: 10.7554/eLife.50014
- Serrat, X., Kukhtar, D., Cornes, E., Esteve-Codina, A., Benlloch, H., Cecere, G., & Cerón, J. (2019). CRISPR editing of *sftb-1/SF3B1* in *Caenorhabditis elegans* allows the identification of synthetic interactions with cancer-related mutations and the chemical inhibition of splicing. *PLoS Genetics*, *15*(10), e1008464. doi: 10.1371/journal.pgen.1008464
- Shahzad, U., Taccone, M. S., Kumar, S. A., Okura, H., Krumholtz, S., Ishida, J., Mine, C., Gouveia, K., Edgar, J., Smith, C., Hayes, M., Huang, X., Derry, W. B., Taylor, M. D., & Rutka, J. T. (2021). Modeling human brain tumors in flies, worms, and zebrafish: From proof of principle to novel therapeutic targets. *Neuro-Oncology*, *23*(5), 718–731. doi: 10.1093/neuonc/noaa306
- Shi, X., Kitano, A., Jiang, Y., Luu, V., Hoegenauer, K. A., & Nakada, D. (2018). Clonal expansion and myeloid leukemia progression modeled by multiplex gene editing of murine hematopoietic progenitor cells. *Experimental Hematology*, *64*, 33–44.e5. doi: 10.1016/j.exphem.2018.04.009
- Shou, J., Li, J., Liu, Y., & Wu, Q. (2018). Precise and Predictable CRISPR Chromosomal Rearrangements Reveal Principles of Cas9-Mediated Nucleotide Insertion. *Molecular Cell*, *71*(4), 498–509.e4. doi: 10.1016/j.molcel.2018.06.021
- Silverman, G. A., Luke, C. J., Bhatia, S. R., Long, O. S., Vetica, A. C., Perlmutter, D. H., & Pak, S. C. (2009). Modeling Molecular and Cellular Aspects of Human Disease Using the Nematode *Caenorhabditis elegans*. *Pediatric Research*, *65*(1), 10–18. doi: 10.1203/PDR.0b013e31819009b0
- Stein, E. M., DiNardo, C. D., Fathi, A. T., Pollyea, D. A., Stone, R. M., Altman, J. K., Roboz, G. J., Patel, M. R., Collins, R., Flinn, I. W., Sekeres, M. A., Stein, A. S., Kantarjian, H. M., Levine, R. L., Vyas, P., MacBeth, K. J., Tosolini, A., VanOostendorp, J., Xu, Q., ... de Botton, S. (2019). Molecular remission and response patterns in patients with mutant-IDH2 acute myeloid leukemia treated with enasidenib. *Blood*, *133*(7), 676–687. doi: 10.1182/blood-2018-08-869008
- Stemmer, M., Thumberger, T., del Sol Keyer, M., Wittbrodt, J., & Mateo, J. L. (2015). CCTop: An Intuitive, Flexible and Reliable CRISPR/Cas9 Target Prediction Tool. *PLOS ONE*, *10*(4), e0124633. doi: 10.1371/journal.pone.0124633
- Stepanenko, A. A., & Dmitrenko, V. V. (2015). HEK293 in cell biology and cancer research: Phenotype, karyotype, tumorigenicity, and stress-induced genome-phenotype evolution. *Gene*, *569*(2), 182–190. doi: 10.1016/j.gene.2015.05.065
- Sternberg, S. H., Redding, S., Jinek, M., Greene, E. C., Jennifer, A., Biophysics, M., Biology, C., Division, B., & Berkeley, L. (2014). DNA interrogation by the CRISPR RNA-guided endonuclease Cas9. *Nature*, *507*(7490), 62–67. doi: 10.1038/nature13011.DNA
- Struys, E. A. (2006). D-2-Hydroxyglutaric aciduria: unravelling the biochemical pathway and the genetic

- defect. *Journal of Inherited Metabolic Disease*, 29(1), 21–29. doi: 10.1007/s10545-006-0317-9
- Teah, Y. F., Abduraman, M. A., Amanah, A., Adenan, M. I., Fariza Sulaiman, S., & Tan, M. L. (2017). Data on the construction of a recombinant HEK293 cell line overexpressing hERG potassium channel and examining the presence of hERG mRNA and protein expression. *Data in Brief*, 14, 584–591. doi: 10.1016/j.dib.2017.08.008
- Tommasini-Ghelfi, S., Murnan, K., Kouri, F. M., Mahajan, A. S., May, J. L., & Stegh, A. H. (2019a). Cancer-associated mutation and beyond: The emerging biology of isocitrate dehydrogenases in human disease. *Science Advances*, 5(5), 1–17. doi: 10.1126/sciadv.aaw4543
- Tommasini-Ghelfi, S., Murnan, K., Kouri, F. M., Mahajan, A. S., May, J. L., & Stegh, A. H. (2019b). Cancer-associated mutation and beyond: The emerging biology of isocitrate dehydrogenases in human disease. *Science Advances*, 5(5), eaaw4543. doi: 10.1126/sciadv.aaw4543
- Torres, R., Martin, M. C., Garcia, A., Cigudosa, J. C., Ramirez, J. C., & Rodriguez-Perales, S. (2014). Engineering human tumour-associated chromosomal translocations with the RNA-guided CRISPR–Cas9 system. *Nature Communications*, 5(1), 3964. doi: 10.1038/ncomms4964
- Townsend, J. A., Wright, D. A., Winfrey, R. J., Fu, F., Maeder, M. L., Joung, J. K., & Voytas, D. F. (2009). High-frequency modification of plant genes using engineered zinc-finger nucleases. *Nature*, 459(7245), 442–445. doi: 10.1038/nature07845
- Trevino, A. E., & Zhang, F. (2014). Genome Editing Using Cas9 Nickases. In *Methods in Enzymology* (1st ed., Vol. 546, Issue C, pp. 161–174). Elsevier Inc. doi: 10.1016/B978-0-12-801185-0.00008-8
- Vakulskas, C. A., Dever, D. P., Rettig, G. R., Turk, R., Jacobi, A. M., Collingwood, M. A., Bode, N. M., McNeill, M. S., Yan, S., Camarena, J., Lee, C. M., Park, S. H., Wiebking, V., Bak, R. O., Gomez-Ospina, N., Pavel-Dinu, M., Sun, W., Bao, G., Porteus, M. H., & Behlke, M. A. (2018). A high-fidelity Cas9 mutant delivered as a ribonucleoprotein complex enables efficient gene editing in human hematopoietic stem and progenitor cells. *Nature Medicine*, 24(8), 1216–1224. doi: 10.1038/s41591-018-0137-0
- Valletta, S., Dolatshad, H., Bartenstein, M., Yip, B. H., Bello, E., Gordon, S., Yu, Y., Shaw, J., Roy, S., Scifo, L., Schuh, A., Pellagatti, A., Fulga, T. A., Verma, A., & Boultonwood, J. (2015). ASXL1 mutation correction by CRISPR/Cas9 restores gene function in leukemia cells and increases survival in mouse xenografts Simona. *Oncotarget*, 6(42), 44061–44071. doi: 10.18632/oncotarget.6392
- Vats, P., Kaushik, R., Rawat, N., Sharma, A., Sharma, T., Dua, D., Singh, M. K., Palta, P., Singla, S. K., Manik, R. S., & Chauhan, M. S. (2021). Production of Transgenic Handmade Cloned Goat (*Capra hircus*) Embryos by Targeted Integration into Rosa 26 Locus Using Transcription Activator-like Effector Nucleases. *Cellular Reprogramming*, 23(4), 250–262. doi: 10.1089/cell.2021.0011
- Veiga-da-Cunha, M., Van Schaftingen, E., & Bommer, G. T. (2020). Inborn errors of metabolite repair. *Journal of Inherited Metabolic Disease*, 43(1), 14–24. doi: 10.1002/jimd.12187
- Vicente, M. M., Chaves-Ferreira, M., Jorge, J. M. P., Proença, J. T., & Barreto, V. M. (2021). The Off-Targets of Clustered Regularly Interspaced Short Palindromic Repeats Gene Editing. *Frontiers in Cell and Developmental Biology*, 9(September), 1–16. doi: 10.3389/fcell.2021.718466
- vonGroll, A., Levin, Y., Barbosa, M. C., & Ravazzolo, A. P. (2006). Linear DNA Low Efficiency Transfection by Liposome Can Be Improved by the Use of Cationic Lipid as Charge Neutralizer. *Biotechnology Progress*, 22(4), 1220–1224. doi: 10.1021/bp060029s
- Wang, F., Morita, K., DiNardo, C. D., Furudate, K., Tanaka, T., Yan, Y., Patel, K. P., MacBeth, K. J., Wu, B., Liu, G., Frattini, M., Matthews, J. A., Little, L. D., Gumbs, C., Song, X., Zhang, J., Thompson, E. J., Kadia, T. M., Garcia-Manero, G., ... Takahashi, K. (2021). Leukemia stemness and co-occurring mutations drive resistance to IDH inhibitors in acute myeloid leukemia. *Nature Communications*, 12(1). doi: 10.1038/s41467-021-22874-x
- Ward, P. S., Lu, C., Cross, J. R., Abdel-Wahab, O., Levine, R. L., Schwartz, G. K., & Thompson, C. B. (2013). The Potential for Isocitrate Dehydrogenase Mutations to Produce 2-Hydroxyglutarate Depends on Allele Specificity and Subcellular Compartmentalization. *Journal of Biological Chemistry*, 288(6), 3804–3815. doi: 10.1074/jbc.M112.435495
- Wu, K., Zhao, X. J., Wong, K. W., & Fan, X. Y. (2015). Comparison of plasmid DNA versus PCR amplified gene of insert DNA for nucleofection in Kasumi-1 cells. *Cytotechnology*, 67(2), 275–283. doi: 10.1007/s10616-013-9683-y
- Wu, Y., Vazquez-Prada, K. X., Liu, Y., Whittaker, A. K., Zhang, R., & Ta, H. T. (2021). Recent Advances in the Development of Theranostic Nanoparticles for Cardiovascular Diseases. *Nanotheranostics*, 5(4), 499–514. doi: 10.7150/ntno.62730
- Wu, Z. J., Zhao, X., Banaszak, L. G., Gutierrez-Rodrigues, F., Keyvanfar, K., Gao, S. G., Raffo, D. Q.,

- Kajigaya, S., & Young, N. S. (2018). CRISPR/Cas9-mediated ASXL1 mutations in U937 cells disrupt myeloid differentiation. *International Journal of Oncology*, *52*(4), 1209–1223. doi: 10.3892/ijo.2018.4290
- Xia, A. L., He, Q. F., Wang, J. C., Zhu, J., Sha, Y. Q., Sun, B., & Lu, X. J. (2019). Applications and advances of CRISPR-Cas9 in cancer immunotherapy. *Journal of Medical Genetics*, *56*(1), 4–9. doi: 10.1136/jmedgenet-2018-105422
- Xu, Q., Li, Y., Lv, N., Jing, Y., Xu, Y., Li, Y., Li, W., Yao, Z., Chen, X., Huang, S., Wang, L., Li, Y., & Yu, L. (2017). Correlation between isocitrate dehydrogenase gene aberrations and prognosis of patients with acute myeloid leukemia: A systematic review and meta-analysis. *Clinical Cancer Research*, *23*(15), 4511–4522. doi: 10.1158/1078-0432.CCR-16-2628
- Xu, Z. L., Mizuguchi, H., Ishii-Watabe, A., Uchida, E., Mayumi, T., & Hayakawa, T. (2001). Optimization of transcriptional regulatory elements for constructing plasmid vectors. *Gene*, *272*(1–2), 149–156. doi: 10.1016/S0378-1119(01)00550-9
- Yang, J., Zhu, H., Zhang, T., & Ding, J. (2021). Structure, substrate specificity, and catalytic mechanism of human D-2-HGDH and insights into pathogenicity of disease-associated mutations. *Cell Discovery*, *7*(1), 3. doi: 10.1038/s41421-020-00227-0
- Yarrington, R. M., Verma, S., Schwartz, S., Trautman, J. K., & Carroll, D. (2018). Nucleosomes inhibit target cleavage by CRISPR-Cas9 in vivo. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(38), 9351–9358. doi: 10.1073/pnas.1810062115
- Yen, C.-T., Fan, M.-N., Yang, Y.-L., Chou, S.-C., Yu, I.-S., & Lin, S.-W. (2016). Current animal models of hemophilia: the state of the art. *Thrombosis Journal*, *14*(S1), 22. doi: 10.1186/s12959-016-0106-0
- Yen, K., Travins, J., Wang, F., David, M. D., Artin, E., Straley, K., Padyana, A., Gross, S., DeLaBarre, B., Tobin, E., Chen, Y., Nagaraja, R., Choe, S., Jin, L., Konteatis, Z., Cianchetta, G., Saunders, J. O., Salituro, F. G., Quivoron, C., ... Su, S.-S. M. (2017). AG-221, a First-in-Class Therapy Targeting Acute Myeloid Leukemia Harboring Oncogenic IDH2 Mutations. *Cancer Discovery*, *7*(5), 478–493. doi: 10.1158/2159-8290.CD-16-1034
- Zetsche, B., Gootenberg, J. S., Abudayyeh, O. O., Slaymaker, I. M., Makarova, K. S., Essletzbichler, P., Volz, S. E., Joung, J., Van Der Oost, J., Regev, A., Koonin, E. V., & Zhang, F. (2015). Cpf1 Is a Single RNA-Guided Endonuclease of a Class 2 CRISPR-Cas System. *Cell*, *163*(3), 759–771. doi: 10.1016/j.cell.2015.09.038
- Zhang, J., Lieu, Y. K., Ali, A. M., Penson, A., Reggio, K. S., Rabadan, R., Raza, A., Mukherjee, S., & Manley, J. L. (2015). Disease-associated mutation in *SRSF2* misregulates splicing by altering RNA-binding affinities. *Proceedings of the National Academy of Sciences*, *112*(34), E4726–E4734. doi: 10.1073/pnas.1514105112
- Zhang, S., Shen, J., Li, D., & Cheng, Y. (2021). Strategies in the delivery of Cas9 ribonucleoprotein for CRISPR/Cas9 genome editing. *Theranostics*, *11*(2), 614–648. doi: 10.7150/thno.47007
- Zhao, G., & Winkler, M. E. (1996). A novel α -ketoglutarate reductase activity of the sera-encoded 3-phosphoglycerate dehydrogenase of *Escherichia coli* K-12 and its possible implications for human 2-hydroxyglutaric aciduria. *Journal of Bacteriology*, *178*(1), 232–239. doi: 10.1128/jb.178.1.232-239.1996
- Zheng, F., Kelly, M. R., Ramms, D. J., Heintschel, M. L., Tao, K., Tutuncuoglu, B., Lee, J. J., Ono, K., Foussard, H., Chen, M., Herrington, K. A., Silva, E., Liu, S. N., Chen, J., Churas, C., Wilson, N., Kratz, A., Pillich, R. T., Patel, D. N., ... Ideker, T. (2021). Interpretation of cancer mutations using a multiscale map of protein systems. *Science*, *374*(6563). doi: 10.1126/science.abf3067
- Zhou, Z. L., Sun, X. X., Ma, J., Man, C. H., Wong, A. S. T., Leung, A. Y., & Ngan, A. H. W. (2016). Mechanical oscillations enhance gene delivery into suspended cells. *Scientific Reports*, *6*(1), 22824. doi: 10.1038/srep22824