



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

Edición de la región promotora del gen Ago1 mediante CRISPR-Cas

TESIS

QUE PARA OPTAR POR EL GRADO DE:

MAESTRO EN CIENCIAS

PRESENTA:

DIANA CHAVIRA DESALES

TUTOR PRINCIPAL

DR. CARLOS FABIÁN FLORES JASSO
INSTITUTO NACIONAL DE MEDICINA GENÓMICA

MIEMBROS DEL COMITÉ TUTOR

DR. ALFREDO TORRES LARIOS
INSTITUTO DE FISIOLÓGÍA CELULAR

DR. NOÉ VALENTÍN DURÁN FIGUEROA
UNIDAD PROFESIONAL INTERDISCIPLINARIA DE BIOTECNOLOGÍA

CIUDAD DE MÉXICO, OCTUBRE, 2019



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



UNAM – Dirección General de Bibliotecas
Tesis Digitales
Restricciones de uso

DERECHOS RESERVADOS ©
PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

PMDCB/1440/2019

Chavira Desales Diana
Estudiante de Maestría en Ciencias Bioquímicas
P r e s e n t e

Los miembros del Subcomité Académico en reunión ordinaria del día 10 de junio del presente año, conocieron su solicitud de asignación de **JURADO DE EXAMEN** para optar por el grado de **MAESTRA EN CIENCIAS**, con la réplica de la tesis "**Edición de la región promotora del gen Ago1 mediante CRISPR-Cas**", dirigida por el/la Dr(a). **Flores Jasso Carlos Fabián**.

De su análisis se acordó nombrar el siguiente jurado integrado por los doctores:

PRESIDENTE	Reyes Taboada José Luis
VOCAL	Becerra Bracho Arturo Carlos II
VOCAL	Valdés Rodríguez Víctor Julián
VOCAL	Licona Limón Paula
SECRETARIO	Navarro González Rosa Estela

Sin otro particular por el momento, aprovecho la ocasión para enviarle un cordial saludo.

Atentamente
"POR MI RAZA, HABLARÁ EL ESPÍRITU"
Cd. Universitaria, Cd. Mx., a 10 de junio de 2019.
COORDINADORA



Dra. ANA BRÍGIDA CLORINDA ARIAS ÁLVAREZ

*A la memoria de
José Desales*

Agradecimientos académicos

Al Dr. Fabián Flores Jasso por dejarme continuar en su grupo de investigación, por los consejos, la paciencia y la orientación durante mis estudios de maestría.

A la Dra. Eréndira Avendaño Vázquez por el apoyo durante la realización de este proyecto.

A mi comité tutorial Dr. Noé Valentín Durán Figueroa y Dr. Alfredo Torres Larios, por los comentarios y aportaciones al proyecto y por contribuir a mi formación académica.

Al Dr. José Luis Reyes Taboada, al Dr. Arturo Carlos Il Becerra Bracho, al Dr. Víctor Julián Valdés Rodríguez, a la Dra. Paula Licona Limón y a la Dra. Rosa Estela Navarro González por haber leído y corregido mi escrito.

Agradecimientos personales

A mis papás, por todo el amor y por siempre confiar en mí. A Daniela, por siempre apoyarme, consentirme y cuidarme. No importa dónde estemos o qué hagamos, si estamos juntos sé que todo estará bien.

A José Desales y a Julia García, porque me dieron el regalo más valioso que pude recibir. July, promete que serás eterna. Agradezco también al resto de los Desales, por ser una ohanita tan bonita. En especial a mis más favs, Dani, Alancin, Chucha, Peter y Aldin por hacer la vida mucho más divertida.

A mis amigos del laboratorio. A los Fabianators. Fili y Viri por los momentos que compartimos mientras intentábamos crecer juntos. A Omar por la sal en los bazos y por convertir cualquier situación en algo divertido. A la Dra. Mariana por tratar de apoyar siempre y por escuchar mis inquietudes. A Mario por despertarme todos los días y por tratar de ser fit juntos. A los Eréndiros. A Perli por las fotos, la música, el baile, la compañía, las porras y los nombres mal escritos. Al Dr. Bustos por escucharme y por los consejos. A Heber, por ser mi amigo, por los masajes y por ser tan divertido.

A las chicas que pasaron por el penthouse Viri, Jesse, Itza y Bere, y a Marian que siempre ha estado ahí. Porque de cada una de ellas pude aprender algo diferente. Especialmente a Itza por hacer nuestra vida más dinámica y tratar de alejarnos del consumismo. Y a Marian, por todos los "amiga date cuenta", aunque nunca me diera cuenta.

A las que siempre han estado ahí. Andrea, Cynthia y Fernanda. Por el apoyo que me han brindado, por escucharme, por aguantar el drama, por lo que hemos compartido y por permanecer.

Esta tesis de maestría se realizó bajo la dirección del Dr. Carlos Fabián Flores Jasso en el Consorcio de Metabolismo de RNA y Vesículas Extracelulares, del Instituto Nacional de Medicina Genómica, con el apoyo económico del proyecto INMEGEN #321, dictamen 05/2017/I; con título "Identificación sin falsos positivos de microRNAs de células en cultivo"; y del proyecto INMEGEN #407, dictamen D1-41-2019, con título "Implementación de un método de alta sensibilidad para la construcción de bibliotecas para RNAseq".

Los estudios de maestría se realizaron con una beca otorgada por CONACYT. Se agradece el apoyo recibido por parte del Programa de Apoyo a los Estudios del Posgrado (PAEP).

Índice

Resumen.....	8
Abstract	10
Abreviaturas.....	12
Introducción	13
RNAs pequeños	13
Biogénesis de los miRNAs	14
Las proteínas Argonata	17
Función de los miRNAs.....	18
Antecedentes	20
¿Qué determina la vida media de un miRNA?	20
Modificaciones postraduccionales en Argonata.....	20
<i>Drosophila melanogaster</i> como modelo de estudio de miRNAs	21
Planteamiento del problema	23
Hipótesis.....	26
Objetivos	26
Objetivo general.....	26
Objetivos específicos.....	26
Materiales y métodos	27
Cultivo de células S2.....	27
Extracción de DNA genómico y PCR de la región promotora	27
Alineado y rellenado de oligonucleótidos.....	28
Urea-PAGE.....	28
Transcripción <i>in vitro</i>	28
Construcción de plásmidos	28
Transformación bacteriana	29
Transfección y selección de células transfectadas.....	30
Secuenciación Sanger.....	30
Extracción y cuantificación de proteínas	30
SDS-PAGE y western blot	30
Resultados	34
Isoformas de Ago1 en <i>D. melanogaster</i>	34
Diseño de gRNAs	36
Análisis de los posibles sitios off-target de los gRNAs	39

Síntesis de sgRNAs mediante transcripción <i>in vitro</i>	41
Construcción de los plásmidos de expresión de Cas9 y sgRNAs	44
Comprobación de la edición a nivel genómico	46
Expresión de Ago1 en las células editadas	47
Proliferación de las células editadas	53
Identificación de los promotores de Ago1 activos en células S2 de <i>D. melanogaster</i>	54
Expresión de Ago1 durante el desarrollo de <i>D. melanogaster</i>	57
Discusión	59
En <i>D. melanogaster</i> se expresan cuatro isoformas de Ago1	59
Los sgRNAs diseñados son eficientes y específicos	59
La edición de los promotores se comprobó a nivel genómico en una población celular heterogénea	61
La edición de los posibles promotores del gen Ago1 produce el decaimiento de la proteína	63
El decaimiento de AGO1 conlleva a la muerte celular	66
Dos de los tres posibles promotores de Ago1 se encuentran activos en las células S2 de <i>D. melanogaster</i>	67
Las isoformas de Ago1 se expresan de manera diferencial durante el desarrollo de <i>D. melanogaster</i>	68
Conclusiones	71
Perspectivas	72
Referencias	73

Resumen

Los miRNAs son RNAs no codificantes de aproximadamente 22 nt de longitud. Participan en la regulación postranscripcional de los genes actuando en conjunto con proteínas Argonauta. El miRNA y Argonauta forman un complejo funcional llamado RISC. La función de los miRNAs es guiar a Argonauta hasta el mRNA sobre el que debe actuar. La manera en que se regula la expresión es mediante represión de la traducción y el decaimiento del mRNA. Hasta el momento se han descrito a detalle los mecanismos por los que los miRNAs se originan y por el cual actúan. Sin embargo, aún es poco lo que conocemos acerca del destino de estos y Argonauta después de que ha ocurrido el silenciamiento del mRNA blanco. Lo que sucede después de este punto es de suma importancia pues la intensidad y duración del efecto de silenciamiento de miRNAs sobre mRNAs depende completamente de la vida media activa del complejo RISC. La incógnita de cuáles son las señales que dirigen el decaimiento de estos RNAs pequeños permanece sin resolver. Es conocido que existen modificaciones postraduccionales en Argonauta que tienen influencia sobre el proceso de ensamblaje con el miRNA. Sin embargo, aún se desconoce si algunas de estas modificaciones podrían dirigir la vida media de RISC. Estas interrogantes permanecen sin respuesta debido principalmente a la dificultad actual del estudio bioquímico de los Argonautas *in vivo*, pues estas proteínas son esenciales para la sobrevivencia de las células y el desarrollo de organismos completos. Sería de suma importancia el poder monitorear los cambios ocurridos en Argonauta y el miRNA luego de su encuentro con el mRNA blanco. En este proyecto planteamos la edición del promotor del gen Ago1 en células S2 de *D. melanogaster*. En este organismo, Ago1 presenta cuatro isoformas cuya síntesis se encuentra regulada por tres posibles promotores. Mediante CRISPR-Cas9 deletamos el centro de cada uno de los promotores de Ago1. La delección del genoma de estos tres sitios propuestos como promotores resulta en el decaimiento de AGO1. A su vez, el decaimiento de la proteína provoca la muerte celular. Sorprendentemente, sólo dos de estos tres posibles promotores se encuentra activo en la línea celular utilizada en este proyecto. Es decir, la delección de los dos promotores activos provoca el decaimiento de la proteína, pero ocurre lo mismo al deletar el promotor inactivo. Además, de manera interesante, este promotor que en células S2 se encuentra inactivo, se encuentra transcripcionalmente activo en sólo dos

estadios de desarrollo de *D. melanogaster*: larva y pupa. Los resultados sugieren una regulación específica de la región promotora del gen, en la que posiblemente se necesite la presencia de los tres posibles promotores para lograr la transcripción del mRNA de Ago1. Además, las diferencias entre isoformas de Ago1 podrían conferirles características específicas en cuanto a especificidad, estabilidad y tiempo de vida media. Como perspectiva hemos planteado la sustitución de cada promotor endógeno de Ago1 por un promotor metabólicamente inducible, que permita controlar de manera temporal la síntesis de AGO1. El control de la expresión de AGO1 nos permitirá abordar las incógnitas antes mencionadas, en un sistema celular.

Abstract

miRNAs are ~22 nt non-coding RNAs that participate in the post-transcriptional regulation of genes acting in conjunction with Argonaute proteins. miRNA and Argonaute form a functional complex called RISC. The function of miRNAs is to guide Argonautes to the mRNA on which they must act. How expression is regulated is through repression of translation and decay of mRNA. So far, the mechanisms by which miRNAs are synthesized and by which they act have been described in detail. However, little is known about the fate of miRNAs and Argonautes after the silencing of the mRNA target. It is important to know what is happening after mRNA silencing, as the intensity and duration of this effect of miRNAs over mRNAs depend entirely on the active half-life of the RISC complex. The question of which are the signals that direct the decay of these small RNAs remains unsolved. It is known that there are post-translational modifications in Argonaute that influence the miRNA assembly process. However, it is still unknown if some of these modifications could lead to the half-life of RISC. These questions remain unanswered mainly due to the current difficulty of the biochemical study of Argonautes *in vivo*, since these proteins are essential for cell survival and for the development of whole organisms. It would be very useful to be able to monitor the changes that occurred in Argonaute and the miRNA after their encounter with the mRNA target. In this project we propose the edition of the Ago1 gene promoter in S2 cells of *D. melanogaster*. In this organism, Ago1 has four isoforms whose synthesis is regulated by three possible promoters. Using the CRISPR-Cas9 system, we delete the center of each of the Ago1 promoters. The genome deletion of these three sites proposed as promoters results in the decay of AGO1. Protein decay causes cell death. Surprisingly, only two of these three possible promoters are active in the cell line used in this project. That is, the deletion of the two active promoters causes the protein to decay, but the same occurs when we delete the inactive promoter. Besides, interestingly, this promoter, which is inactive in S2 cells, is transcriptionally active in only two stages of development of *D. melanogaster*: larva and pupa. The results suggest a specific regulation of the promoter region of the gene, in which the presence of the three possible promoters may be needed to achieve transcription of the Ago1 mRNA. Also, the differences between Ago1 isoforms could give them specific characteristics in terms of specificity, stability and interaction with other

proteins. As a perspective, we have proposed the replacement of each endogenous Ago1 promoter with metabolically inducible promoter, which allows the AGO1 synthesis to be temporarily controlled. The control of AGO1 expression will allow us to address the above-mentioned questions, in a cellular system.

Abreviaturas

aa	Aminoácidos
Ago	Gen Argonauta
AGO	Proteína Argonauta
Bp	Pares de bases
CAGE	Cap-analysis gene expression
Cas9	Proteína asociada a CRISPR 9
CFD	Cutting Frequency Determination
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeat
dsDNA	DNA de doble hebra
EPD	Base de datos de promotores eucariontes
gRNA	RNA guía
H3K4m3	Trimetilación de la lisina 4 de la histona 3
H3K27ac	Acetilación de la lisina 27 de la histona 3
IVT	Transcripción <i>in vitro</i>
miRNA	microRNA
mRNA	RNA mensajero
nt	Nucleótidos
PAM	Protospacer Adjacent Motif
PCR	Reacción en cadena de la polimerasa
pre-miRNA	microRNA precursor
Pre-RISC	Complejo RISC inmaduro
pri-miRNA	microRNA primario
RISC	RNA Induced Silencing Complex
RNAi	RNA de interferencia
sgRNA	Single guide RNA
siRNA	Small interfering RNA
TSS	Sitio de inicio de la transcripción
UTR	Región no traducible
WT	Wild type

Introducción

El DNA es la molécula que almacena la información necesaria para la síntesis de algunas de las macromoléculas que la célula necesita para sobrevivir, como las proteínas. Las proteínas en la célula pueden ser estructurales o funcionales. Se encargan tanto de crear estructuras altamente organizadas dentro de la célula como de participar en importantes reacciones bioquímicas¹.

La homeostasis de los organismos depende en gran medida del control espacial y temporal de la expresión de los genes¹. En estos procesos participan las diferentes clases de RNA presentes en las células eucariontes. El RNA mensajero (mRNA) es la molécula que dicta el orden en el que los aminoácidos deben ser agregados a una cadena peptídica, durante la síntesis de proteínas.

Existen otras clases de RNAs que participan en el proceso de traducción: el RNA ribosomal (rRNA) y el RNA de transferencia (tRNA). El rRNA forma parte de la estructura de los ribosomas, mientras que los tRNAs son moléculas transportadoras que interactúan con los aminoácidos para que puedan ser incorporados a la proteína que se sintetiza. Estas clases de RNA forman parte de los RNAs no codificantes (ncRNA)^{1,2}.

RNAs pequeños

Además de los tRNAs y rRNAs, existen otras clases de RNAs que forman parte del grupo de los RNAs no codificantes. Por ejemplo, los microRNAs (miRNAs), los RNAs asociados a Piwi (Piwi-RNAs o piRNAs), los RNAs largos no codificantes (lncRNAs), los RNAs pequeños nucleares (snRNAs) y los RNAs pequeños nucleolares (snoRNAs), entre otros. Cada una de estas clases de RNA cumple funciones importantes dentro de la célula. Por ejemplo, los miRNAs, piRNAs y lncRNAs participan en la regulación de la expresión génica, mientras que los snRNAs y snoRNAs cumplen su función durante la maduración del RNA².

Los RNAs no codificantes se pueden clasificar de acuerdo con su longitud. Los miRNAs, siRNAs y piRNAs se encuentran en la categoría de RNAs pequeños, por tener una longitud de entre 20 y 30 nucleótidos (nt). Estas tres clases de RNA comparten la

característica de participar en la regulación de la expresión génica, pero difieren entre sí en cuanto al mecanismo por el cual actúan, las moléculas precursoras y proteínas que participan en su biogénesis y las proteínas con las que interactúan para llevar a cabo su función. Los miRNAs, por ejemplo, son RNAs de aproximadamente 22 nt de longitud derivados de precursores de RNA largos sintetizados por la RNA pol II, actúan en conjunto con proteínas de la familia Argonauta y participan en la regulación postranscripcional de los genes³.

Biogénesis de los miRNAs

La biogénesis de los miRNAs inicia en el núcleo celular con la síntesis, por la RNA polimerasa II, de un precursor largo llamado miRNA primario (pri-miRNA). Este pri-miRNA es un RNA que cuenta con un cap y una cola de poly A, además toma una estructura de tallo-asa⁴. En el núcleo, esta estructura de tallo-asa es reconocida por el microprocesador, que es un complejo compuesto por la RNasa tipo III Drosha y la proteína de unión a RNA de doble hebra DGCR8 (humanos) o Pasha (*Drosophila*). La función de DGCR8 en el complejo microprocesador es incrementar la afinidad de Drosha por el sustrato sobre el que va a actuar. Drosha escinde la estructura de tallo-asa del pri-miRNA, generando un miRNA precursor (pre-miRNA). El pre-miRNA tiene una longitud de aproximadamente 60 nt y conserva la estructura de tallo-asa. Además, el corte de Drosha produce extremos asimétricos, de manera que, el extremo 3' del pre-miRNA cuenta con 2 nt no apareados⁵⁻⁸ (Figura 1). Un pri-miRNA puede ser monocistrónico si da origen a sólo un pre-miRNA o policistrónico si de él se originan más de un pre-miRNA. Los miRNAs que provienen del mismo pre-miRNA pueden o no actuar sobre blancos que participan en las mismas vías o cuya expresión se encuentra relacionada^{9,10}.

El pre-miRNA sintetizado en el núcleo es entonces exportado al citoplasma por acción del transportador nuclear exportina 5 (XPO5) que actúa en conjunto con la proteína Ran-GTP. El transporte del pre-miRNA del núcleo al citoplasma es un proceso dependiente de GTP^{11,12}. En el citoplasma, el asa del pre-miRNA es cortada y como resultado se obtiene un dúplex de RNA de aproximadamente 22 nt. El corte del pre-miRNA lo realiza

la RNasa tipo III Dicer que actúa en conjunto con la proteína de unión a RNA de doble hebra TRBP (humanos) o Loquacious (*Drosophila*). Al igual que el corte de Drosha, el corte de Dicer produce extremos asimétricos, por lo que el dúplex de RNA que se obtiene cuenta con 2 nt sobresalientes en los extremos 3' del RNA, además de grupos OH en los extremos 3' y fosfatos en los 5'¹³⁻¹⁵ (Figura 1).

El dúplex de RNA resultante del corte de Dicer es cargado en el complejo multiproteínico RISC (del inglés RNA Induced Silencing Complex) que está conformado mínimamente por una proteína Argonauta (AGO)¹⁶. El proceso de cargado del dúplex en AGO es un proceso dependiente de ATP en el que actúan las proteínas chaperonas Hsc70 y Hsp90, así como de sus co-chaperonas¹⁷. El complejo resultante de la interacción del dúplex de RNA y AGO (pre-RISC) es inactivo y requiere de un proceso de maduración para convertirse en su forma activa. La maduración implica la selección de una de las hebras del dúplex para permanecer cargada en AGO (hebra guía) y el desensamble de la otra hebra (hebra pasajera). La elección de cuál de las dos hebras será la hebra guía y cuál de ellas será la hebra pasajera se determina por la estabilidad de los extremos del dúplex, la hebra cuyo apareamiento en el extremo 5' sea el termodinámicamente menos estable será la hebra que permanezca cargada en AGO, es decir, la hebra guía¹⁸. En *Drosophila*, la asimetría en la estabilidad de los extremos 5' de las hebras es sensada por el complejo de cargado de RISC (RLC, RISC Loading Complex), que está compuesto por Dcr-2 y R2D2. R2D2 se une al extremo más estable del dúplex, mientras que Dcr-2 lo hace al menos estable¹⁹. Una vez que se ha determinado qué hebra será la guía, la hebra pasajera se desensambla del complejo pre-RISC, dando paso a la formación del complejo RISC maduro (RISC), este último complejo es funcional y puede entonces actuar sobre el mRNA blanco (Figura 1).

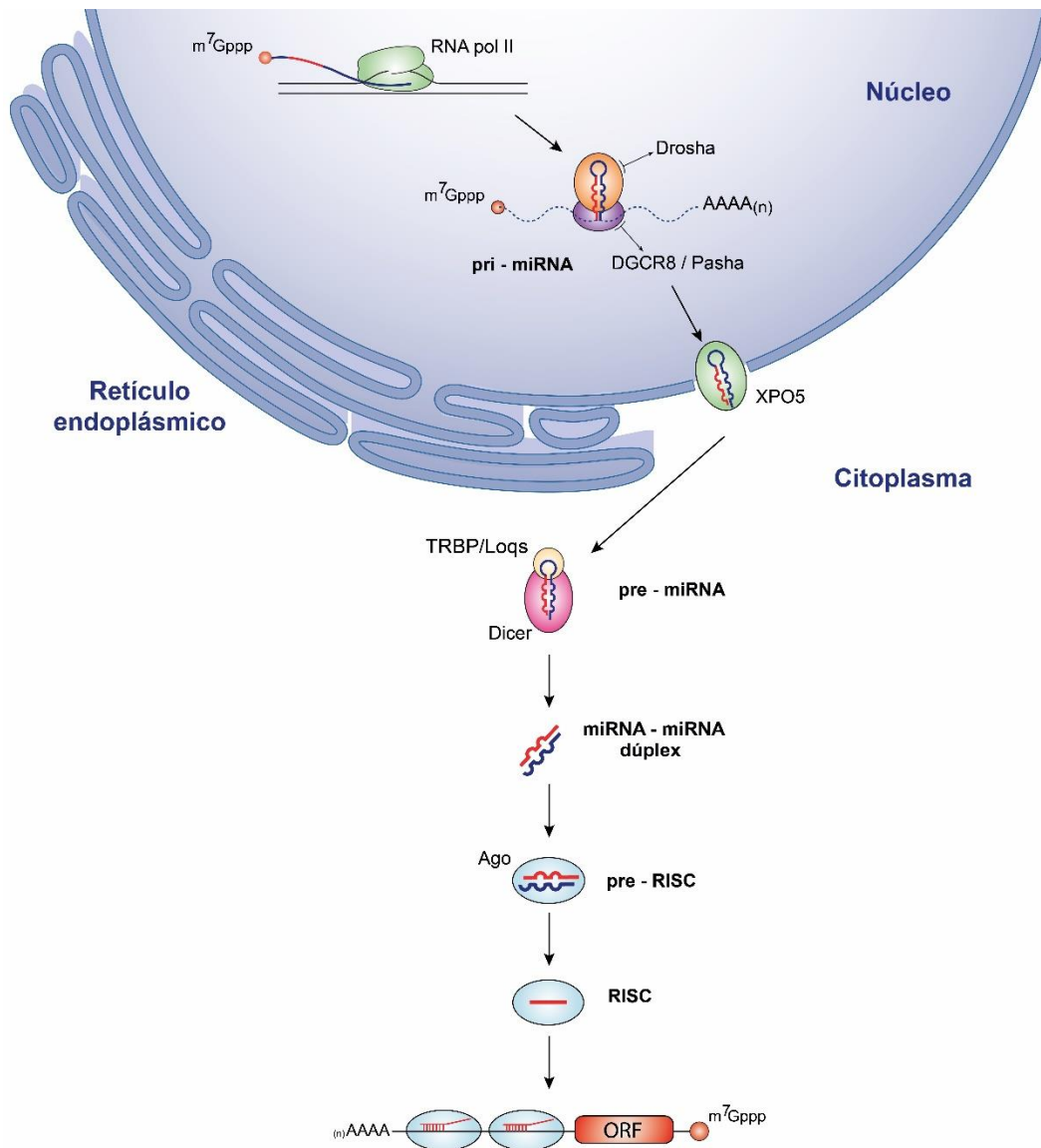


Figura 1. Biogénesis de los miRNAs. En el núcleo de la célula, la RNA polimerasa II (RNA pol II) transcribe un RNA largo denominado microRNA primario (pri-miRNA). Este RNA cuenta con un cap (m^7Gppp) y una cola de poly A. El pri-miRNA es procesado por Drosha y DGCR8/Pasha, generando un microRNA precursor (pre-miRNA). El pre-miRNA es exportado al citoplasma por acción del transportador nuclear Exportina 5 (XPO5). En el citoplasma, Dicer y TRBP/Loqs procesan el pre-miRNA, generando un dúplex de miRNA-miRNA. El dúplex es cargado en una proteína Argonauta (Ago) formando el complejo pre-RISC. Este complejo madura cuando una de las hebras del dúplex es elegida para permanecer cargada en Ago. RISC es un complejo activo y puede actuar sobre el mensajero blanco. Figura de elaboración propia.

Las proteínas Argonauta

Las proteínas Argonauta son una familia evolutivamente conservada cuyos miembros participan en la regulación postranscripcional mediante el silenciamiento de genes. Esta familia puede dividirse en dos subfamilias: proteínas AGO y proteínas PIWI. Los Argonautas realizan su función interactuando con miRNAs y siRNAs. Estas proteínas se expresan de manera ubicua en numerosos eucariontes, aunque el número de ellas puede variar de un organismo a otro. Por ejemplo, en mamíferos existen cuatro proteínas Argonauta: AGO1, AGO2, AGO3 y AGO4, mientras que en *Drosophila* existen sólo dos: AGO1 y AGO2^{20,21}.

Las proteínas AGO, están compuestas por una sola cadena peptídica y presentan cuatro dominios característicos: el dominio amino terminal (N-terminal), el dominio Piwi-Argonauta-Zwille (PAZ), el dominio medio (MID) y el dominio P-element induced wimpy testes (PIWI). Además, cuentan con dos linkers, el L1 conecta a los dominios N-terminal y PAZ y el L2 conecta a los dominios PAZ y MID^{22,23} (Figura 2).

El miRNA y AGO interactúan de una manera muy específica. El grupo fosfato del extremo 5' del miRNA se une a un "binding pocket" que estructuralmente se encuentra entre los dominios MID y PIWI. El dominio PAZ, por su parte, une al extremo 3' del miRNA. Como resultado de la interacción entre el miRNA y AGO, se crea una región dentro del miRNA llamada región semilla. Esta región está comprendida entre los nucleótidos 2 y 8 del miRNA, estos nucleótidos se pre-ordenan en una estructura de semi-hélice que ayuda a disminuir el gasto energético que conllevaría la unión del miRNA y el mRNA blanco. La región semilla es el mayor determinante para el reconocimiento del mRNA blanco por parte del miRNA en el contexto de RISC (Figura 2)²²⁻²⁴.

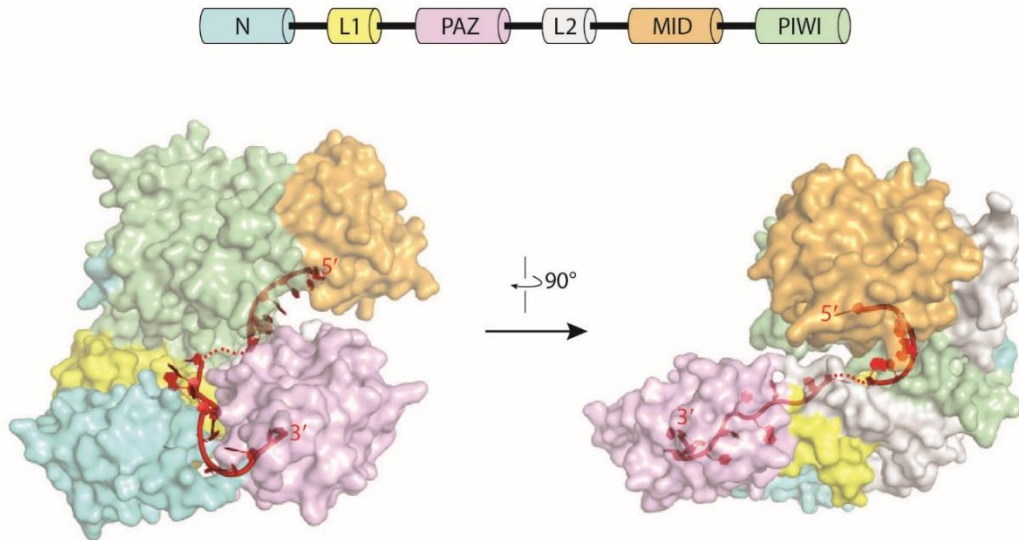


Figura 2. Dominios de Argonauta y su interacción con el miRNA. AGO cuenta con cuatro dominios funcionales: N, PAZ, MID y PIWI y con dos linkers: L1 y L2. Se muestra la estructura de AGO2 humano y la interacción con el miRNA (rojo), la parte punteada representa a los nucleótidos desordenados. (Imagen modificada de Nakanishi, K. 2016)²⁵.

Función de los miRNAs

La función de los miRNAs es guiar a Argonauta hasta el mRNA sobre el que debe actuar. El miRNA se une por complementariedad de bases a la región 3' UTR del mRNA blanco, provocando el silenciamiento de este mediante represión de la traducción y el decaimiento del mRNA^{16,26}. Algunos Argonautas, como AGO2 de mamífero y AGO2 de *D. melanogaster* son capaces de cortar mRNAs blancos altamente complementarios al RNA pequeño con el que están interactuando. Este corte es realizado por el dominio PIWI, que puede presentar actividad similar a la de una RNasa H. Sin embargo, la gran mayoría de los miRNAs hibridan sólo parcialmente con el mRNA blanco, lo que impide que este sea cortado, aun cuando RISC esté conformado por un Argonauta con actividad catalítica¹⁰.

La inhibición de la traducción del mRNA mediada por RISC ocurre durante el inicio de este proceso. RISC interfiere con la función de los factores de iniciación eIF4A-I y eIF4A-II, el mecanismo por el que esto ocurre podría implicar la disociación de estos factores,

impidiendo el ensamblaje del complejo de inicio de la traducción y el escaneo del mRNA por el ribosoma^{27,28}. En cuanto al decaimiento del mRNA, Argonauta recluta a algunas proteínas para llevarlo a cabo. Una de esas proteínas es TNRC6 (mamíferos) o GW182 (*Drosophila*)^{29,30}. A su vez, GW182 interactúa con PABPC para reclutar a los complejos PAN2-PAN3 y a CCR4 y promover la deadenilación del mensajero^{31,32}. La deadenilación del mRNA conlleva a que el complejo DCP1-DCP2 remueva el cap³³. Estos dos eventos hacen que el mRNA sea altamente susceptible a la acción de exoribonucleasas, como XRN1³⁴ (Figura 3).

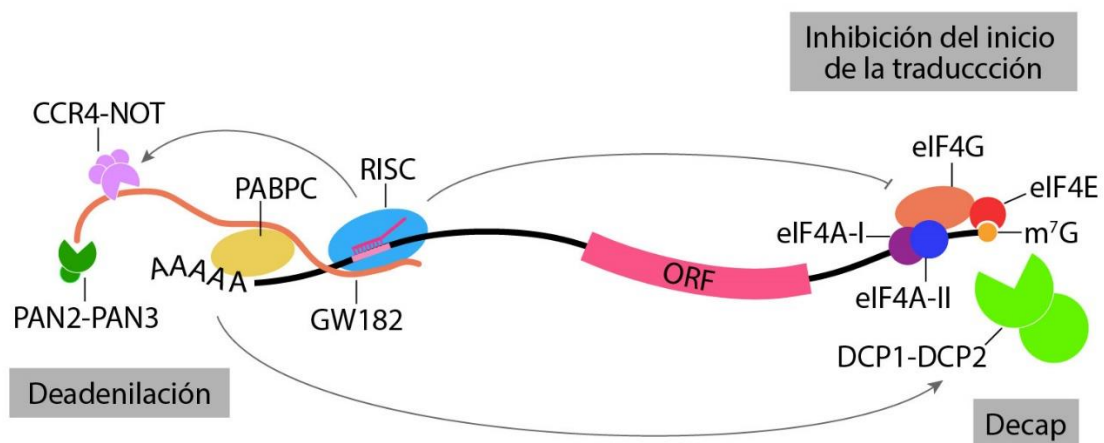


Figura 3. Los miRNAs y Argonauta regulan la expresión de los genes mediante la inhibición del inicio de la traducción y la desestabilización del mensajero. RISC reconoce una secuencia específica en la región 3' UTR del mRNA. La inhibición del inicio de la traducción ocurre mediante el desensamblaje de factores de inicio de la traducción (eIF4A-I, eIF4A-II, eIF4E y eIF4G). El mRNA es desestabilizado mediante el acortamiento de la cola de poly A, proceso en el que participan las proteínas GW182 y PABPC y los complejos CCR4-NOT y PAN2-PAN3. Para lograr la desestabilización del mensajero ocurre también la remoción del cap (m^7G). Figura de elaboración propia.

Antecedentes

¿Qué determina la vida media de un miRNA?

Los miRNAs presentan tiempos de vida media muy variados, que pueden ir desde algunos minutos hasta varios días³⁵. A pesar de que se desconoce aquello que dirige estos tiempos de vida media, conocemos algunos factores que podrían contribuir en su determinación, como la secuencia del miRNA³⁶, la etapa del ciclo celular³⁷, el estadio del desarrollo³⁸ o el tejido³⁹ en el que se expresa.

Puesto que el silenciamiento de un mRNA es dictado por su complementariedad con el miRNA en RISC, la vida media del miRNA puede también estar modulada por su interacción con el mRNA blanco. Se sabe que mRNAs blanco con un alto grado de complementariedad al miRNA promueven su degradación. El mecanismo por el que esto sucede se desconoce, pero se ha propuesto que pudiera suceder mediante la adición de nucleótidos al extremo 3' del miRNA, que al ser más extenso podría ser blanco de nucleasas que lo reconocerían y degradarían⁴⁰. Otra propuesta es que estos mRNAs blanco con alto grado de complementariedad con el miRNA podrían desestabilizar la interacción entre este y Argonauta, provocando el desensamble de RISC y la degradación del miRNA por nucleasas citoplásmicas⁴¹. De igual manera, se ha reportado que blancos que carecen de secuencia complementaria a la región semilla del miRNA, pero presentan secuencias complementarias a la región 3' pueden dirigir su degradación⁴².

Un hecho importante es que, por ser RNAs de hebra sencilla, los miRNAs son altamente vulnerables a la acción de RNAsas intracelulares. Una vez que ha madurado RISC, si el miRNA se disociara de la proteína Argonauta con la que interactúa, sería rápidamente degradado. Esto implica que la vida media de RISC – y su función – después del encuentro con un mRNA no puede ir más allá de la asociación física del miRNA con Argonauta.

Modificaciones postraduccionales en Argonauta

La fosforilación es la modificación postraduccional mejor caracterizada para Argonauta y ha sido descrita en tres principales regiones. La primera de ellas se encuentra en el

linker L2 y comprende los sitios Ser387 y Tyr393. Se ha visto que la fosforilación en la Ser387 promueve la interacción de AGO2 humano con las proteínas TNRC6A y la proteína LIM domain-containing protein 1, ambas implicadas en la actividad de RISC sobre el mRNA⁴³⁻⁴⁵. Por otra parte, la fosforilación de la Tyr393 impide la interacción de AGO2 con Dicer y, por lo tanto, el cargado de miRNAs^{45,46}. Se ha propuesto que el receptor del factor de crecimiento epidérmico (EGFR) es el encargado de la fosforilación de AGO2 en la Tyr393⁴⁶; mientras que la fosforilación en Ser387 se llevaría a cabo por la cinasa Akt³⁴⁴.

La segunda región de fosforilación descrita para Argonauta se encuentra en el dominio MID y corresponde a la Tyr529⁴⁵. Este residuo participa en la unión del primer nucleótido y del grupo fosfato en el 5' del miRNA a Argonauta²². Se ha propuesto que la fosforilación en este sitio bloquea el cargado del miRNA en Argonauta⁴⁵.

La tercera región de fosforilación de Argonauta se encuentra en el dominio PIWI, en un bucle expuesto al exterior de la proteína. Este bucle se conoce como el "cluster S824-S834" y en él se encuentran cinco residuos que pueden ser fosforilados, Ser824, Ser828, Thr830, Ser831 y Ser834. La fosforilación de estos residuos interviene en la interacción de AGO2 y el mRNA blanco. Luego de que RISC reconoce al mRNA, los residuos S824-S834 se fosforilan, reduciendo la afinidad de AGO2 por el mRNA y promoviendo la liberación de este. Posteriormente, los residuos son defosforilados, lo que permite que AGO2 interactúe con un nuevo mRNA blanco. CSNK1A1 es la cinasa identificada como la encargada de fosforilar a AGO2 en esta región⁴⁷.

***Drosophila melanogaster* como modelo de estudio de miRNAs**

A partir del descubrimiento de los miRNAs y los siRNAs, la bioquímica y el mecanismo de acción de estos RNAs pequeños han sido estudiados en diversos organismos. Uno de los modelos animales más empleados es *D. melanogaster*. En este organismo, la biogénesis de miRNAs y siRNAs se encuentra separada y biológicamente definida en dos grandes vías moleculares: los precursores de miRNAs son procesados en el citoplasma por la RNasa tipo III Dicer-1 y su compañera Loqs1; el producto es cargado en AGO1.

Por otra parte, los siRNAs son procesados en el citoplasma por Dicer-2 y R2D2 y son cargados en AGO2⁴⁸⁻⁵⁰.

En cuanto al mecanismo de acción, los miRNAs regulan la expresión de los genes mediante el arresto traduccional y el decaimiento del mRNA, mientras que los siRNAs actúan escindiendo al mRNA y es parte del mecanismo de defensa contra virus.

Otra de las razones por las que *D. melanogaster* es frecuentemente elegida como modelo de estudio de miRNAs, es que las proteínas efectoras y RNasas que participan en la biogénesis de estos RNAs pequeños cuentan con sus contrapartes homólogas en mamíferos, mostrando un alto grado de identidad.

Planteamiento del problema

A pesar de los grandes avances que se tienen hasta el momento en cuanto al entendimiento de la biogénesis y el mecanismo de acción de los miRNAs, es poco lo que se conoce acerca del destino de estos y de sus proteínas efectoras después de que ha ocurrido el silenciamiento de los mRNAs blanco. El conocimiento actual sobre la función de los miRNAs abarca sólo hasta el encuentro del complejo RISC con su primer mRNA blanco^{51,52}. Lo que sucede después de este punto es de gran importancia ya que la intensidad y duración del efecto de silenciamiento de miRNAs sobre mRNAs depende de la vida media activa del complejo RISC.

Puesto que la actividad es también dependiente de la vida media de los miRNAs, un punto importante es el desconocimiento de aquello que determina el tiempo en el que cada miRNA se encuentra disponible en la célula. Lo único que es claro hasta el momento es que este tiempo de vida media es altamente variable entre miRNAs, lo que obliga a preguntarse cuáles son las señales que dirigen el decaimiento de estos RNAs pequeños.

Dado que el decaimiento del miRNA ocurre de manera inevitable, es también obligatorio preguntarse qué ocurre con la proteína Argonauta a la cual se encontraba asociado. Hasta el momento, se han encontrado diferentes sitios de fosforilación en Argonauta 2, pero sólo se ha descrito la función de algunos de ellos, principalmente durante el ensamblaje de RISC^{47,53,54}. Aún se desconoce si estas fosforilaciones tienen alguna influencia sobre los pasos posteriores al ensamblaje y si estas u otras modificaciones postraduccionales podrían dirigir la vida media de RISC y por lo tanto de los miRNAs después de su encuentro con el mRNA blanco. De ser así, es posible que existan códigos de modificación post-traduccionales que dirijan y posiblemente limiten los ciclos funcionales de la proteína Argonauta, su asociación con miRNAs y la formación de múltiples complejos RISC. De esta manera, la asociación consecutiva de Argonauta con múltiples miRNAs establecería un recambio de estos.

De acuerdo con el conocimiento actual, podríamos plantear diferentes escenarios biológicos sobre lo que podría ocurrir con Argonauta y el miRNA que conforman a RISC después de que este ha actuado sobre un mRNA. Es posible que RISC permanezca ensamblado y sea capaz de encontrar un nuevo mRNA blanco del miRNA con el que se

encuentra asociado. De esta manera, un RISC ensamblado podría realizar múltiples ciclos de silenciamiento sobre más de un mRNA, siempre y cuando estos contengan la secuencia de reconocimiento de este miRNA en su 3' UTR. Por otra parte, es posible que el miRNA sea desensamblado de Argonauta luego de actuar sobre un mRNA. En este caso, el miRNA se degradaría por acción de RNasas citoplasmáticas. Argonauta podría entonces interactuar con un nuevo miRNA y formar un nuevo RISC activo. Sin embargo, al desensamblarse RISC, tanto el miRNA como el Argonauta que lo conformaban podrían ser degradados.

Estas incógnitas permanecen sin resolver debido, en gran medida, a que las proteínas Argonauta son esenciales para la sobrevivencia de las células y el desarrollo de organismos completos, lo que ha dificultado su estudio *in vivo*. Por ejemplo, en modelos de ratón el knock-out de Argonauta 2 es letal y el desarrollo embrionario se detiene⁵⁵. Intentar disminuir la expresión de alguna de las proteínas Argonauta utilizando del mismo RNAi como herramienta silenciadora, propicia que los Argonautas no silenciados se sobre expresen, tal vez para suplir la función perdida por aquel que ha sido silenciado.

Sería de gran importancia el poder estudiar lo que sucede con los miRNAs y Argonauta post silenciamiento, por ejemplo, mediante experimentos de "pulso y caza", donde después de un estímulo fisiológico, se puedan monitorear con distintas técnicas moleculares los cambios ocurridos tanto en las proteínas Argonauta como en los miRNAs. De tal manera que al añadir cierto metabolito al medio de cultivo sea posible controlar la expresión de Ago1 bajo condiciones experimentales controladas y por tiempos específicos. Para realizar dichos estudios de pulso y caza controlados se pueden utilizar células S2 de *D. melanogaster* que cuentan con sólo dos proteínas Argonauta y en quienes las vías de silenciamiento de miRNAs y siRNAs se encuentran biológicamente separadas, teniendo a AGO1 confinado a la función exclusiva de los miRNAs.

Para resolver las incógnitas antes propuestas, proponemos la construcción de un sistema que nos permita controlar la expresión de Ago1 en células en cultivo. Esto podría lograrse al editar el genoma de células S2, sustituyendo el promotor endógeno del gen por un promotor metabólicamente inducible. De esta manera, la expresión de Ago1 en las células podría controlarse mediante estímulos en el medio de cultivo.

Como primer paso para lograr el control de la expresión de Ago1, en este proyecto hemos planteado la edición del promotor de este gen. Se editará el genoma de células S2 de *D. melanogaster*, específicamente la región promotora del gen Ago1 por medio del sistema CRISPR-Cas9. La delección del promotor del gen Ago1 podría tener repercusión sobre la expresión de AGO1 en las células en cultivo.

El sistema CRISPR-Cas9 (del inglés Clustered Regularly Interspaced Short Palindromic Repeat) está compuesto por una endonucleasa que puede ser programada por un RNA guía para generar cortes en secuencias específicas de DNA. Este sistema funciona con una nucleasa, Cas9, que puede realizar cortes de doble hebra en el DNA blanco gracias a los dos diferentes dominios de nucleasa que posee (HNH y RuvC)⁵⁶.

De manera endógena, CRISPR requiere de dos RNAs para actuar: el CRISPR RNA (crRNA) y el crRNA transactivador (tracrRNA). El crRNA contiene la secuencia específica (secuencia spacer), de 20 nt de longitud, complementaria al sitio del genoma sobre el que la nucleasa actuará y algunos nucleótidos complementarios al tracrRNA. Sin embargo, el sistema también puede ser programado utilizando un solo RNA, resultado de la fusión del crRNA y el tracrRNA, llamado sgRNA (del inglés single guide RNA)^{57,58}.

Para que la nucleasa Cas9 corte al DNA, adyacente a la secuencia blanco debe haber un trinucleótido NGG, a esta secuencia se le conoce como PAM (del inglés Protospacer Adjacent Motif)⁵⁶.

Los cortes de doble hebra que Cas9 realiza en el genoma eucarionte pueden ser reparados siguiendo mecanismos diferentes. Si ocurre mediante reparación de extremos no homólogos (NHEJ, Non-homologous end joining) se podrían generar inserciones aleatorias o delecciones. Si el corte se repara mediante recombinación homóloga (HDR, homology directed repair) es posible insertar secuencias específicas utilizando un DNA donador con secuencias homólogas a las adyacentes al sitio de corte⁵⁹.

Hipótesis

La edición del promotor del gen Ago1 en células S2 de *D. melanogaster* tendrá repercusión sobre la expresión de la proteína.

Objetivos

Objetivo general

Editar mediante CRISPR la región promotora del gen Ago1 en células S2 de *D. melanogaster*.

Objetivos específicos

1. Construir las herramientas moleculares necesarias para llevar a cabo la edición genómica con CRISPR-Cas9.
2. Transfectar y seleccionar las células editadas por medio de antibióticos.
3. Verificar la edición a nivel genómico por medio de PCR.
4. Monitorear la expresión de AGO1 luego de la edición de su región promotora.

Materiales y métodos

Los plásmidos utilizados en este proyecto se obtuvieron de Addgene, para la expresión de Cas9 y los sgRNAs se utilizó el plásmido pAc-sgRNA-Cas9⁶⁰, AGO1 se expresó con el plásmido pAFW-Ago1⁶¹ y AGO2 mediante el plásmido pAFW-Ago2⁶¹. Los primers se obtuvieron de IDT y se enlistan en la Tabla 1 al final de esta sección. Para las inmunodetecciones se utilizaron los anticuerpos Anti-Ago1 ab5070 (Abcam) y Anti-beta Actin ab8224 (Abcam).

Cultivo de células S2

Las células S2 fueron cultivadas en suspensión a temperatura ambiente en Schneider's Drosophila Media (Biowest) suplementado al 10% con suero fetal bovino (SFB, Biowest) inactivado a 56 °C durante 30 minutos, penicilina 50 U/mL y estreptomycin 50 ug/mL. Durante el cultivo celular, las células se mantuvieron a una densidad aproximada de entre 1×10^6 y 1×10^7 células/mL. El conteo de las células para el análisis de proliferación celular se realizó utilizando una cámara de Neubauer, siguiendo las especificaciones del proveedor.

Extracción de DNA genómico y PCR de la región promotora

El DNA genómico se obtuvo mediante el método de fenol-cloroformo. Las PCRs de la región promotora se realizaron con la enzima Taq DNA polymerase (NEB) con buffer Thermopol (NEB) siguiendo las especificaciones del proveedor, se utilizaron los primers listados en el Tabla 1, Región promotora. En cuanto a las condiciones de PCR, se realizó una desnaturalización inicial a 95 °C durante 1 minuto; posteriormente se realizaron 30 ciclos de desnaturalización a 95 °C por 30 segundos, alineamiento de oligonucleótidos a 66 °C durante 45 segundos y extensión a 68 °C por 1 minuto; la extensión final se realizó a 68 °C por durante 5 minutos.

Alineado y relleno de oligonucleótidos

El relleno de los oligonucleótidos se realizó utilizando el fragmento Klenow de la DNA polimerasa I (Thermo). La reacción de alineado se preparó utilizando cantidades equimolares de los oligonucleótidos a alinear, 2 μ L de buffer de reacción 10X, 0.5 μ L de dNTPs 10 mM y 15 μ L de H₂O grado biología molecular. La reacción se calentó a 94 °C y la temperatura se disminuyó lentamente hasta llegar a los 37 °C. Se añadieron 0.5 μ L de Klenow a la reacción y se incubó a 37 °C durante 1 hora. La síntesis de los dsDNAs molde para la transcripción *in vitro* de los gRNAs se realizó con los oligonucleótidos listados en la Tabla 1, Síntesis del dsDNA molde para IVT.

Urea-PAGE

La electroforesis desnaturalizante de DNA se realizó utilizando geles de poliacrilamida 8% - urea 8 M – TBE 1X. El buffer de carga se preparó utilizando urea 8M y azul de bromofenol. Las muestras se mezclaron 1:1 en buffer de carga, se incubaron por 5 minutos en un baño de agua hirviendo y se cargaron en el gel. La electroforesis se realizó a 90 V y los geles se revelaron con bromuro de etidio.

Transcripción *in vitro*

La reacción de transcripción *in vitro* de los gRNAs se realizó utilizando 50 U de la enzima T7 RNA pol (NEB), 2 μ L de T7 transcription buffer 10X, concentraciones finales de 5 mM de ATP, CTP y UTP y de 8 mM de GTP, 0.5 μ L de DTT 1M y la reacción se llevó a un volumen final de 20 μ L con agua grado biología molecular. La reacción se incubó a 37 °C durante 2 horas, posteriormente se centrifugó a 13,000 $\times g$ durante 1 minuto, se recuperó el sobrenadante y a este último se le añadió MgCl₂ hasta una concentración final de 6 mM. El proceso de incubación y centrifugación se repitió dos veces más.

Construcción de plásmidos

Los plásmidos de expresión de cada una de las sgRNAs y Cas9 se obtuvieron a partir de la modificación mediante PCR del plásmido pAc-sgRNA-Cas9. Para cada sgRNA se utilizó un par de primers diferente (Tabla 1, Modificar secuencia spacer en pAc-sgRNA-Cas9).

La reacción se llevó a cabo en un volumen total de 50 μ L, utilizando 50 ng del plásmido pAc-sgRNA-Cas9, 1 μ L de dNTPs 10 mM, 1 μ L de cada primer 10 μ M, 1 U de Phusion Hot Start II DNA Polymerase (Thermo) y 10 μ L de 5x Phusion HF Buffer (Thermo), el volumen final se ajustó con agua grado biología molecular. Se tomaron 8 μ L de la reacción de edición y se digirieron con la enzima DpnI (NEB) utilizando 10 U de enzima y 1 μ L de buffer CutSmart (NEB). La reacción se incubó a 37 °C durante una hora y se inactivó incubando a 80 °C por 20 minutos. Posteriormente, se transformaron bacterias DH5 α competentes con 2 μ L de la reacción de digestión. La construcción de los plásmidos se comprobó mediante PCR, utilizando primers específicos para cada plásmido construido (Tabla 1, Comprobar edición de pAc-sgRNA-Cas9) la enzima Taq DNA Polymerase (NEB) y el buffer ThermoPol (NEB) siguiendo las instrucciones del proveedor.

Transformación bacteriana

Las bacterias DH5 α competentes utilizadas se prepararon con el Mix & Go E. coli Buffer Set (Zymo Research) siguiendo las instrucciones del proveedor y se almacenaron a -80 °C hasta su uso. La transformación de los plásmidos se realizó mediante choque térmico utilizando 15 ng de plásmido purificado o 2 μ L de la reacción de edición del plásmido y 50 μ L de bacterias competentes a una densidad óptica a 600 nm de entre 0.4 y 0.6. Se añadió el plásmido a la suspensión bacteriana y se incubaron en hielo durante 30 minutos. Posteriormente se mantuvieron en baño de agua a 42 °C por 45 segundos y se colocaron nuevamente en hielo durante 2 minutos. Se añadieron 500 μ L de medio Luria-Bertani (LB, Mo Bio Laboratories) sin antibiótico y se incubaron a 37 °C durante 90 minutos en agitación constante. Al término de la incubación se centrifugaron a 3,000 xg por 30 segundos, se retiraron 450 μ L de sobrenadante y el pellet se resuspendió en los 100 μ L de medio restantes. Las bacterias se sembraron en una placa de LB agar con ampicilina (100 μ g/mL) y se incubaron a 37 °C durante 12 horas. Las colonias obtenidas se resembraron en una placa de LB agar nueva, con la misma concentración de antibiótico y se realizó PCR para comprobar si contenían el plásmido esperado. Los plásmidos se purificaron en el kit QIAprep Spin Miniprep (Qiagen) siguiendo las recomendaciones del proveedor.

Transfección y selección de células transfectadas

Los plásmidos se transfectaron en las células S2 en cultivo mediante el método de coprecipitación con fosfato de calcio. Las células se sembraron a una densidad de aproximadamente 1×10^6 células/mL y se incubaron a temperatura ambiente, 24 horas después se realizó un cambio de medio de cultivo, por medio fresco sin antibiótico y luego de 2 horas se realizó la transfección. La reacción de transfección se preparó utilizando 10 μ L de CaCl_2 2.5 M, 3 μ g de cada uno de los plásmidos a transfectar y se llevó a un volumen total de 100 μ L con agua grado biología molecular. A la reacción se le añadieron 100 μ L de HBS 2X, en agitación constante y se incubó a temperatura ambiente por 30 minutos. A las células en cultivo se les añadieron 200 μ L de la reacción de transfección y se incubaron durante 24 horas a temperatura ambiente. Posteriormente se cambió el medio de cultivo por medio fresco. La selección de las células transfectadas se inició 48 horas después con dihidrocloruro de Puomicina (Sigma) a una concentración final de 2 μ g/mL.

Secuenciación Sanger

Los plásmidos construidos se secuenciaron en la Unidad de Secuenciación del Instituto Nacional de Medicina Genómica (INMEGEN).

Extracción y cuantificación de proteínas

Las células se lisaron en buffer Tritón 100 (150 mM NaCl, 1% Tritón X-100, 50 mM Tris-HCl pH 8.0) suplementado con inhibidor de proteasas cOmplete EDTA-free (Sigma), incubando a 4 °C durante 30 minutos en agitación constante. Las células lisadas se centrifugaron a 12,000 rpm por 20 minutos a 4 °C, se recolectó el sobrenadante y se almacenó a -80 °C. La cuantificación de proteínas se realizó mediante un ensayo colorimétrico de Bradford.

SDS-PAGE y western blot

Para el SDS-PAGE se prepararon geles de acrilamida al 7.5% con SDS al 10%. A las muestras de proteína se les añadió buffer de carga 1X (1:1, Tris-HCl 0.5 M pH 6.8, SDS

10%, 2-mercaptoetanol, glicerol y azul de bromofenol), se incubaron durante 5 minutos en un baño de agua hirviendo y se cargaron en el gel. La electroforesis se realizó utilizando buffer de corrida 1X (Tris-base 25 mM, glicina 192 mM, SDS 0.5%) y el marcador de peso molecular Precision Plus Protein Unstained Standards (Bio-Rad). La corrida electroforética se realizó a 80 V durante 120 minutos. Luego del SDS-PAGE, las proteínas se transfirieron a la membrana de PVDF (Bio-Rad) a 300 mA por 2 horas y se tiñeron con rojo de Ponceau para comprobar la transferencia. Las membranas se bloquearon en una solución de PBS 1X, leche 5% en agitación constante durante 2 horas a temperatura ambiente. La inmunodetección se realizó incubando la membrana con la dilución correspondiente de anticuerpo primario en PBS 1X, Leche 1%, Tween 20 0.1%. Se colocaron en agitación a 4 °C durante toda la noche. Posteriormente, la membrana se lavó tres veces con PBS 1X, Tween 20 al 0.3% durante 10 minutos en agitación. Se realizó la incubación con el anticuerpo secundario en PBS 1X, Leche 5 %, Tween 20 0.1% a 4 °C, en agitación durante una hora. La membrana se lavó dos veces con una solución de PBS 1X, Tween 20 0.3 %. El tercer lavado se realizó con PBS 1X durante 10 minutos en agitación. Por último, la quimioluminiscencia se detectó utilizando Luminata Forte Western HRP substrate (Merck Millipore). Para realizar la inmunodetección de una segunda proteína en la misma membrana se realizaron tres lavados con PBS 1X, Tween 20 0.3 % en agitación durante 10 minutos y se prosiguió con la inmunodetección. Para la cuantificación, las imágenes capturadas se analizaron utilizando el software ImageJ (NIH).

Región promotora	
Ago1_A_FW	GCAATTGTATGCGTGTCTGTCGC
Ago1_A_RV	TTCACGCTCTTCAGCTGCTTGC
Ago1_B_FW	GCGCGAGGAGATCACCAACAAG
Ago1_B_RV	GCATGGCCCATTTGGTCGACTT
Ago1_CD_FW	TGTCTCACTGCCCTCTCTCTTCG
Ago1_CD_RV	CCCCCTAACAAAAGCTCGCTGC
Síntesis del dsDNA molde para IVT	
Scaffold_FW	GTTTTAGAGCTAGAAATAGCAAGTTAAAATAAGGCTAGTCCGTTA TCAAC

Scaffold_RV	GCACCGACTCGGTGCCACTTTTTCAAGTTGATAACGGACTAGCCT TAT
Spacer_A5'	TAATACGACTCACTATAGGGCAGCGCTCGATTGATTGTTTTAGA GCTAGAAATAGCAA
Spacer_A3'	TAATACGACTCACTATAGGGCTTGATGGTGATGAACCGTTTTAGA GCTAGAAATAGCAA
Spacer_B5'	TAATACGACTCACTATAGCTTTGCCCTCATTTAAGTGGTTTTAGAG CTAGAAATAGCAA
Spacer_B3'	TAATACGACTCACTATAGGTTTGTGCATAAGAGCGGGGTTTTAGA GCTAGAAATAGCAA
Spacer_CD5'	TAATACGACTCACTATAGCTCTCTCATCGCAACACACGTTTTAGA GCTAGAAATAGCAA
Spacer_CD3'	TAATACGACTCACTATAGAGAAGGTGTGTGTGTCTAGGTTTTAGA GCTAGAAATAGCAA
gRNA_RV	GCACCGACTCGGTGC

Modificar secuencia spacer en pAc-sgRNA-Cas9

Ago1A5'_gRNA_FW	AATCGAATCGAGCGCTGCCCGAAGTATTGAGGAAAACATACCT
Ago1A5'_gRNA_RV	GGGCAGCGCTCGATTGATTGTTTTAGAGCTAGAAATAGCAAG
Ago1A3'_gRNA_FW	GGTTCATCACCATCAAGCCCGAAGTATTGAGGAAAACATACCT
Ago1A3'_gRNA_RV	GGGCTTGATGGTGATGAACCGTTTTAGAGCTAGAAATAGCAAG
Ago1B5'_gRNA_FW	CACTTAAATGAGGGCAAAGCGAAGTATTGAGGAAAACATACCT
Ago1B5'_gRNA_RV	GCTTTGCCCTCATTTAAGTGGTTTTAGAGCTAGAAATAGCAAG
Ago1B3'_gRNA_FW	CCCGCTTATGCACAAACCGAAGTATTGAGGAAAACATACCT
Ago1B3'_gRNA_RV	GGTTTGTGCATAAGAGCGGGGTTTTAGAGCTAGAAATAGCAAG
Ago1CD5'_gRNA_FW	GTGTGTTGCGATGAGAGAGCGAAGTATTGAGGAAAACATACCT
Ago1CD5'_gRNA_RV	GCTCTCTCATCGCAACACACGTTTTAGAGCTAGAAATAGCAAG
Ago1CD3'_gRNA_FW	CTAGACACACACCTTCTCGAAGTATTGAGGAAAACATACCT
Ago1CD3'_gRNA_RV	GAGAAGGTGTGTGTGTCTAGGTTTTAGAGCTAGAAATAGCAAG

Comprobar edición de pAc-sgRNA-Cas9

Amp-R_FW	ATAATACCGCGCCACATAGC
dU6-2_sgRNA_RV	TCCTCAATACTTCGGAAGAGCG
dU6-2_A5'_RV	CTCAATACTTCGGGCAGC
dU6-2_A3'_RV	CCTCAATACTTCGGGCTTGA

dU6-2_B5'_RV	TTCCTCAATACTTCGCTTTGC
dU6-2_B3'_RV	CCTCAATACTTCGGTTTGTGC
dU6-2_CD5'_RV	TCCTCAATACTTCGCTCTCTCA
dU6-2_CD3'_RV	TCCTCAATACTTCGAGAAGGTG

Tabla 1. Listado de oligonucleótidos utilizados. La secuencia de cada oligonucleótido se muestra en sentido 5' – 3'.

Resultados

Puesto que el objetivo del presente proyecto es lograr la edición de la región promotora de Ago1 en células S2 de *D. melanogaster*, lo primero que hicimos fue la búsqueda de datos que nos ayudaran a determinar el locus del promotor de Ago1 en este organismo. Luego, se diseñaron y sintetizaron los componentes del sistema CRISPR que nos permitieron la edición de esta región del genoma. Posteriormente se determinó el efecto de la edición genómica sobre la expresión de la proteína. Por último, se realizó un análisis para determinar la actividad de cada posible promotor en células S2 y se analizó la expresión de cada isoforma de Ago1 durante el desarrollo de *D. melanogaster*.

Isoformas de Ago1 en *D. melanogaster*

En *D. melanogaster* el gen Ago1 se localiza en la hebra antisentido del cromosoma 2R. En el mismo locus, pero en la hebra sentido, se localiza el gen mRpL53, que codifica para una proteína ribosomal de mitocondria.

De acuerdo con información recopilada por algunas bases de datos genómicos, este gen cuenta con cuatro isoformas a nivel de transcritos: Ago1_A, Ago1_B, Ago1_C y Ago1_D. De acuerdo con la información obtenida de bases de datos genómicos, las isoformas de Ago1 podrían transcribirse a partir de sitios de inicio de la transcripción (TSS) diferentes.

Las isoformas difieren entre sí en la secuencia y longitud de los extremos 5' y 3' UTR. Además, las isoformas A, C y D cuentan con un exón traducible extra comparado con la isoforma B. Esto hace que se originen dos proteínas diferentes. Mientras que Ago1_A, Ago1_C y Ago1_D resultan en el mismo producto de 984 aminoácidos (aa) de longitud, Ago1_B produce una proteína más corta, de 950 aa. Los 34 aa que las diferencian se localizan en el extremo N-terminal de la proteína. El resto de los aa son idénticos para ambas proteínas. Los dominios implicados en la función enzimática de Argonauta (N, MID, PIWI y PAZ) están presentes en las cuatro isoformas. Por lo tanto, es de esperarse que cualquiera de ellas pueda ser funcional; incluyendo a Ago1_B, que a pesar de ser la proteína de menor tamaño comparada con el resto de las isoformas ha sido utilizada en estudios sobre la bioquímica de RISC⁶².

Existen algunos repositorios y algoritmos diseñados para la predicción e identificación de posibles promotores en el genoma de organismos eucariontes. Uno de ellos es la base de datos de promotores eucariontes (EPD – Eukaryotic Promoter Database)⁶³. De acuerdo con la información recopilada en esta base de datos, la expresión las isoformas de Ago1 se encontraría regulada por tres promotores. Ago1_A y Ago1_B se expresarían cada una bajo la regulación de un promotor diferente, mientras que la expresión de Ago1_C y Ago1_D dependería de un mismo promotor (Figura 4)

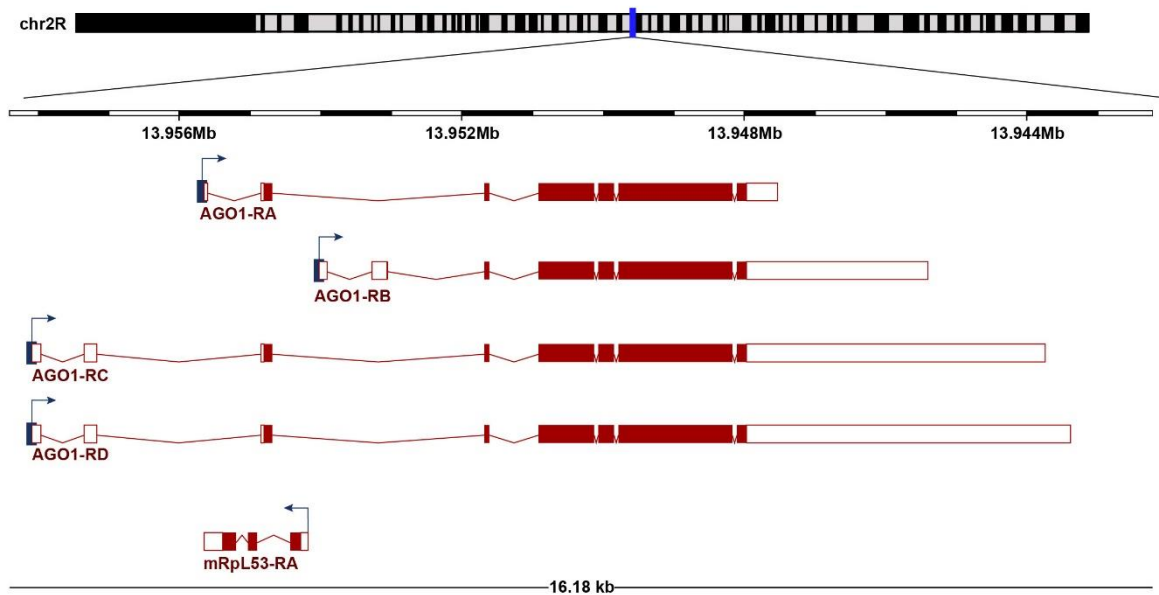


Figura 4. Locus de Ago1 en *D. melanogaster*. Ago1 se encuentra en el cromosoma 2R, en la hebra antisentido. Este gen cuenta con cuatro isoformas, ubicadas en el mismo locus. En la parte superior se muestra una representación del cromosoma 2R y se indica el sitio en el que se encuentra el gen de Ago1. Se muestra la estructura y localización de cada una de las isoformas. Los exones que conforman a las regiones 5' y 3' UTR se representan con rectángulos blancos y los rectángulos rojos corresponden a los exones que se traducen. Las líneas indican intrones. Los rectángulos azules representan el centro de cada posible promotor y el sitio de inicio de la transcripción se indica con una flecha. Figura generada en el visualizador Ensembl (www.ensembl.org).

Al no tener evidencia sobre cuál de los promotores se encuentra activo en las células S2, decidimos editar cada uno de ellos por separado. A pesar de que es posible que sólo uno, dos o los tres promotores se encuentren activos, no descartamos la posibilidad de que al deletar cualquiera de estos, aquel o aquellos que no se encuentran activos puedan activarse para suplir la función del que ha sido inactivado. Esta estrategia podría ayudarnos, además, a identificar cuál de estos se encuentra activo en la línea celular que utilizamos.

La edición de la región promotora se diseñó de manera que se lograra la delección del centro de cada posible promotor. Durante el diseño consideramos al centro de un promotor como la región de aproximadamente 50 nt río arriba y río abajo del sitio de inicio de la transcripción (TSS) de cada mRNA⁶⁴. De esta manera, se diseñaron un par de gRNAs para cada posible promotor, cada uno flanqueando al centro de cada posible promotor.

Diseño de gRNAs

El éxito en la edición del genoma utilizando el sistema CRISPR depende principalmente de la correcta elección de la secuencia spacer del gRNA. El objetivo es encontrar un gRNA que sea lo más eficiente y específico posible. La eficiencia se refiere a la probabilidad de que el gRNA reconozca la secuencia de DNA blanco y se realice un corte en ella (on-target), mientras que la especificidad es la probabilidad de que el gRNA reconozca sitios del genoma que no son su blanco y se realicen cortes en estas secuencias (off-target).

Existen numerosas herramientas diseñadas para encontrar las secuencias spacer más eficientes y específicas. Para el diseño de los gRNAs a utilizar en este proyecto nos apoyamos de la herramienta CRISPOR (www.crispor.tefor.net)⁶⁵. CRISPOR permite encontrar sitios dentro de un genoma que puedan ser editados mediante CRISPR. Esta herramienta muestra posibles gRNAs y los clasifica de acuerdo a una predicción sobre su eficiencia y especificidad. CRISPOR utiliza dos algoritmos para predecir la eficiencia de un gRNA, clasificando estos últimos según el score de Doench o el Score de Moreno-Mateos. El score de Doench es utilizado para medir eficiencia de gRNAs expresados en

células bajo la regulación de algún promotor de la RNA polimerasa III. Este algoritmo considera la eficiencia predicha para todos los gRNAs propuestos para cierto organismo y los compara entre sí. Los gRNAs que elegimos tienen una eficiencia según Doench mayor o igual a la predicha para el 63% del total de posibles gRNAs en *D. melanogaster*.

El MIT Guide Specificity Score es el parámetro que CRISPOR utiliza para predecir la especificidad de los gRNAs. A cada uno de ellos se le asigna un puntaje dentro de un intervalo de 0 a 100 puntos. A un menor número de potenciales sitios off-target corresponde una mayor puntuación dentro del MIT Guide Specificity Score. Para que un gRNA sea considerado como específico debe tener un score de por lo menos 50. Las secuencias spacer que se eligieron para este proyecto tienen un score ≥ 96 . Las secuencias spacer de los gRNAs a utilizar, las secuencias PAM que reconocerá la nucleasa, así como el valor predicho de eficiencia y especificidad para cada uno de ellos se muestran en la Tabla 2.

Isoforma	Spacer	PAM	Eficiencia (%)	Especificidad	Sitio de corte
Ago1_A	5' GGCAGCGCTCGATTCGATT	CGG	63	98	13,955,725
	3' GGGCTTGATGGTGATGAACC	GGG	72	99	13,955,570
Ago1_B	5' GCTTTGCCCTCATTTAAGTG	TGG	78	98	13,954,055
	3' GGTTTGTGCATAAGAGCGGG	CGG	97	100	13,953,860
Ago1_CD	5' GCTCTCTCATCGCAACACAC	TGG	89	98	13,958,210
	3' GAGAAGGTGTGTGTGTCTAG	TGG	94	96	13,957,811

Tabla 2. Secuencias spacer diseñadas para la edición de los promotores de Ago1. Se muestran las secuencias spacer, las secuencias PAM y los valores asignados de eficiencia y especificidad para cada gRNA. Para cada spacer se indica el sitio de corte esperado en el cromosoma 2R de *D. melanogaster*, de acuerdo con la versión del genoma dm6.

Para asegurarnos de que los gRNAs que diseñamos eran los adecuados para editar el genoma de las células con las que contamos en el laboratorio, amplificamos por PCR y secuenciamos la región del genoma que contiene a cada promotor. De esta manera,

comprobamos que las secuencias son las mismas que las reportadas que utilizamos para el diseño de los gRNAs (Figura 5).

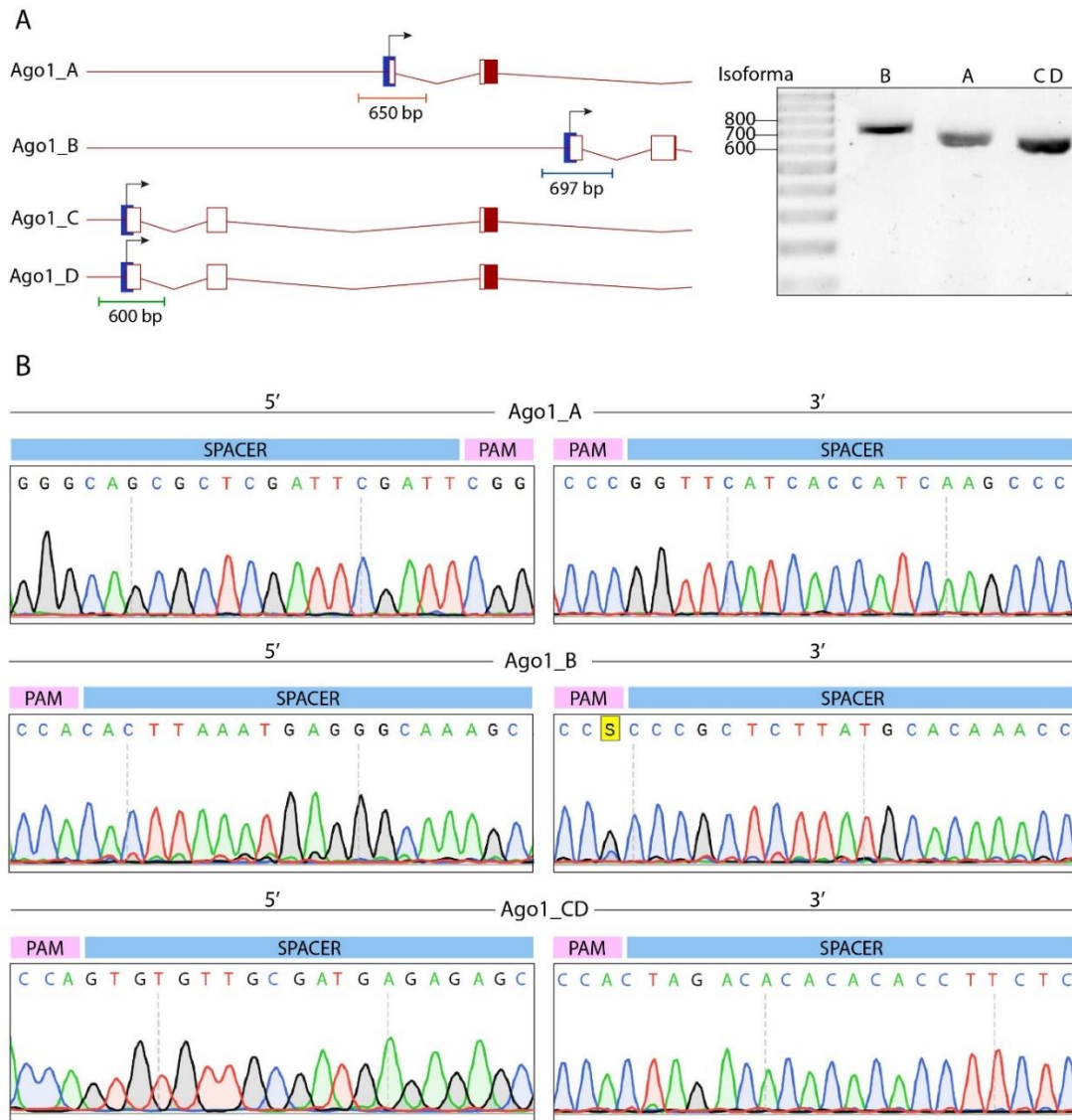


Figura 5. Secuenciación de los posibles promotores de Ago1. **A.** Las regiones adyacentes a los centros de los promotores propuestos se amplificaron por PCR. Se muestra la región amplificada y la electroforesis de los productos obtenidos. **B.** Secuenciación de los productos de PCR. Se secuenció la hebra sentido de cada región amplificada. Dependiendo de si el gRNA hibrida con la hebra sentido o la antisentido, la imagen muestra la secuencia de DNA complementaria al gRNA o la secuencia de DNA correspondiente al gRNA, respectivamente. De igual manera, se indica la secuencia PAM o su complementaria, en el extremo 3' de la secuencia spacer.

Una vez que comprobamos que los sgRNAs diseñados podían dirigir a Cas9 hasta el sitio previsto en el genoma de *D. melanogaster* realizamos un análisis para conocer si estos gRNAs pudieran también dirigir a la nucleasa hasta otros sitios en el genoma.

Análisis de los posibles sitios off-target de los gRNAs

El corte de Cas9 en sitios off-target depende de algunas variables en el reconocimiento de la secuencia spacer del gRNA y la secuencia de DNA. Entre mayor sea el número de mismatches (bases no complementarias) entre el sitio off-target y el gRNA, la probabilidad de que el DNA sea cortado en ese sitio es menor. Además del número, la posición de los mismatches también influye en la probabilidad de cortes off-target, mismatches cercanos a la secuencia PAM disminuyen la probabilidad del corte inespecífico.

Para la determinación de los posibles sitios off-target de los gRNAs diseñados, se analizó todo el genoma de *D. melanogaster*, incluyendo regiones repetidas, y se recopiló aquellos sitios que presentaban hasta cuatro mismatches con la secuencia spacer (datos experimentales muestran que el 90% de los sitios off-target que ocurren durante la edición del genoma mediante CRISPR tienen menos de cuatro mismatches). Además, en el análisis se incluyeron aquellos posibles sitios off-target con secuencias PAM menos comunes pero funcionales, como NGA y NAG.

A cada uno de los sitios off-target se le asignó un valor, que puede ir de 0 a 1.0, dentro del score Cutting Frequency Determination (CFD)⁶⁶. Este score determina la probabilidad de que ocurra el corte en cierta secuencia off-target con base en el número de mismatches, la identidad de los nucleótidos que participan en estos y la posición de cada uno. A un menor valor asignado dentro del score CFD corresponde una menor probabilidad de que ocurra un corte inespecífico. Sitios off-target con un score menor a 0.023 son, muy probablemente, falsos positivos dentro de la identificación de secuencias off-target⁶⁵.

En la Tabla 3 se muestran los posibles sitios off-target para cada uno de los gRNAs diseñados. Para cada secuencia de DNA se indican los mismatches que tendría al hibridar

al gRNA, la secuencia PAM que podría ser reconocida por la nucleasa, el valor de CFD que se le asignó y el locus en el que se encuentra. Cuando el sitio off-target se encuentra en una región intergénica se menciona el gen más cercano. A pesar de que todos los gRNAs tienen posibles sitios off-target con menos de cuatro mismatches, sólo tres de estos sitios tienen un valor dentro del CFD score mayor a 0.023.

gRNA	Secuencia off-target (mismatches)	CFD Score	Locus
Ago1_A 5'	•C•A•TG•••••••• TGG	0.014	Región intergénica CR45377-pyr
	•••••GTG••C••••••• GGG	0.009	Exón Dyrk2
	CT•••T••••••••••A• GGG	0.008	Región intergénica CR43314/bft- CR43314
	•A•••AA••••••••G•••• GGG	0.008	Intrón pb
	•C••G•••G•••C••••••• TGG	0.006	Exón son
	••CG•A••••••••C••••• GGG	0.005	Exón CG4679
	••CAC••••••••C•••••• TGG	0.004	Región intergénica Syt1-daw
Ago1_A 3'	AT••C••••••••••A• CGG	0.024	Región intergénica CR44894- CR44894/RhoU
	•••••C•A••••GA••••• TGG	0.017	Exón Pabp2/Obp44a
	•••••AC••C••••••T•• TGG	0.009	Exón CG4830
	•••••A••T•••T•••C•• CGG	0.008	Intrón RhoL
	•••A•G•••••C•••••C•• CGG	0.003	Exón Plod
	C•••G••••••••••C•A• TGG	0.003	Exón brk
Ago1_B 5'	•••••T•••••A••••• AGA	0.003	Región intergénica CR43624- CG11588
	•••••A•••A•••••G•C• TGG	0.002	Región intergénica CG32982- CG13108
Ago1_B 3'	C•••••G••••TG••••• GGG	0.007	Intrón Vsx1
	•T•••••G••C••••T••• TGG	0.004	Intrón CG32521
Ago1_CD 5'	••••TC•••A••C••••• TGG	0.016	Exón Mhcl
	••T••G•T•••T•••••• TGG	0.015	Exón CR42862
	C•G••••••••A••A•••• CGG	0.013	Exón CG14551

Ago1_CD 3'	A•T••••••••C••••A AGG	0.009	Región Intergénica CR44526-SrpRbeta
	••••••••G••G••A•T•• TGG	0.004	Región intergénica otk-otk2
	••••GTA••••••••••T GGG	0.033	Región intergénica Octbeta2R-CR46017
	•T•TG•••••••••••••• TGG	0.033	Exón CG13594
	AC••C••••••••••••A•• TGG	0.023	Intrón CNMaR
	••T••A••••C•••••C•• GGG	0.020	Intrón CG42666
	A••••T•••C••••••••G•• TGG	0.010	Región intergénica NijA-CG43245
	••••••••CA••••G•••C AGG	0.007	Región intergénica CR43640-wb
	••••CT••••••••••••T•G• TGG	0.005	Intrón Gfrl
•C••••••C•••••••T•C• GGG	0.004	Intrón sNPF	

Tabla 3. Potenciales sitios off-target para cada gRNA diseñado. Se muestran los posibles sitios off-target para cada secuencia spacer, los mismatches que ocurrirían en la interacción DNA-RNA, el valor de CFD asignado a cada sitio y el locus en el que se encuentra. Los sitios en los que sí existe probabilidad de corte se muestran en rojo.

Decidimos editar las células utilizando estos gRNAs pues, a pesar de que dos de los seis gRNAs diseñados tienen sitios off-target en los que sí existe la probabilidad de corte, esta probabilidad es muy baja y podrían tratarse de eventos que experimentalmente no ocurrirían. El siguiente paso fue la síntesis de cada uno de los componentes del sistema CRISPR.

Síntesis de sgRNAs mediante transcripción *in vitro*

El gRNA y la nucleasa que componen al sistema CRISPR-Cas9 pueden obtenerse de diferentes maneras, como es la síntesis *in vitro* del gRNA y el mRNA de Cas9, el uso de complejos ribonucleoproteicos o el uso de plásmidos de expresión. La primera estrategia que seguimos para la síntesis de los sgRNAs diseñados fue la síntesis por transcripción *in vitro*. Para lograrlo, construimos un DNA de doble hebra (dsDNA) molde para cada sgRNA a partir de tres oligonucleótidos de DNA.

El primer paso para la síntesis del molde a transcribir fue el ensamblado y rellenado de un par de oligonucleótidos correspondientes a la secuencia del scaffold del sgRNA. Puesto que la secuencia scaffold es la misma para todos los sgRNAs, este dsDNA se utilizó para la síntesis de los seis sgRNAs diseñados. Posteriormente, se realizaron reacciones de extensión entre el dsDNA scaffold y seis oligonucleótidos con la secuencia del promotor de la T7 RNA pol y las seis diferentes secuencias spacer diseñadas. De estas últimas reacciones se obtuvieron seis dsDNAs correspondientes a los moldes para la transcripción de cada uno de los sgRNAs (Figura 6).

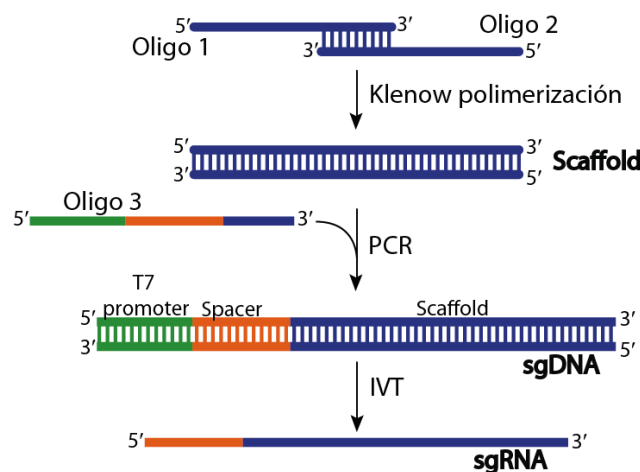


Figura 6. Síntesis de sgRNAs mediante transcripción *in vitro*. Para la obtención de los dsDNA molde, el primer paso fue el ensamblado de dos oligonucleótidos (oligo 1 y oligo 2) que conforman la secuencia scaffold. Posteriormente, se ensambló un tercer oligonucleótido (oligo 3) que contiene la secuencia spacer y el promotor de T7. El dsDNA obtenido puede ser utilizado como molde para la transcripción del sgRNA.

Al realizar la reacción de hibridación y polimerización de los oligonucleótidos 1 y 2 se obtuvo un producto de 76 nt, correspondiente a la secuencia de la región scaffold del sgRNA. Posteriormente, al realizar la reacción de extensión entre el dsDNA scaffold y uno de los oligos spacer, se obtuvo un producto de 113 nt (Figura 7, A). Este dsDNA contiene la secuencia del promotor de la T7 RNA pol, la secuencia spacer y el scaffold del sgRNA. Para obtener suficiente DNA molde para la reacción de transcripción *in vitro*, se hizo una PCR, con la que se logró incrementar la cantidad de DNA molde, aunque

también se obtuvieron productos inespecíficos (Figura 7, B). Con el propósito de disminuir la cantidad de productos inespecíficos en la reacción, purificamos la banda correspondiente a la longitud del producto esperado (113 nt) y repetimos la reacción de amplificación, sin embargo, los productos inespecíficos continuaron apareciendo (Figura 7, C). A pesar de tener una cantidad considerable de productos inespecíficos, la reacción nos permitió obtener una buena cantidad del producto de la longitud esperada, tal vez suficiente para realizar la síntesis del sgRNA mediante transcripción *in vitro*. De esta manera, decidimos sintetizar el resto de los moldes de DNA siguiendo la misma estrategia, obteniendo resultados similares (Figura 7, D).

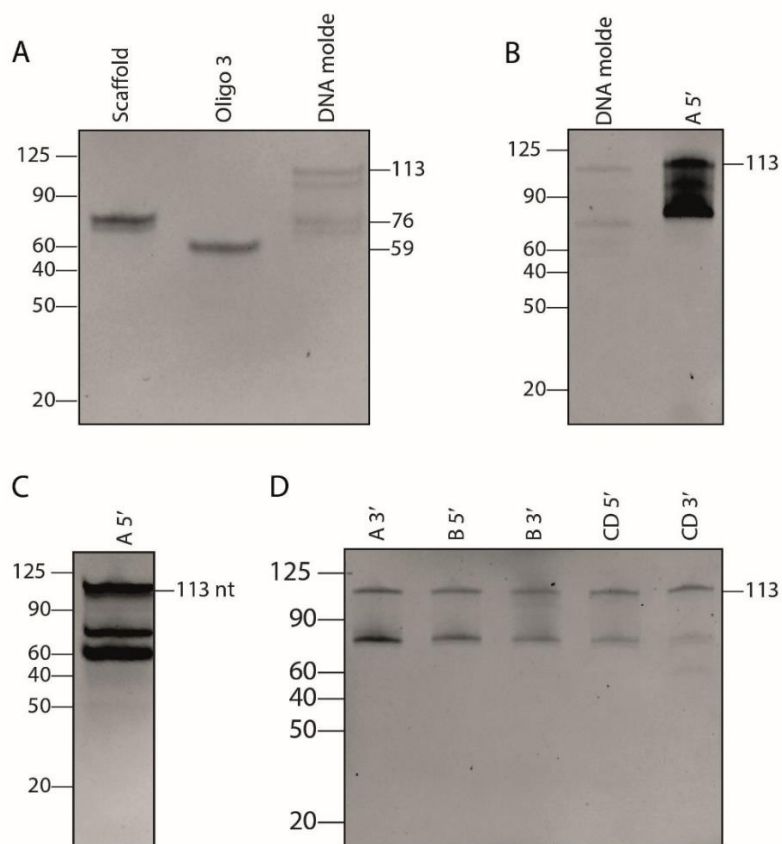


Figura 7. Reacciones de síntesis del DNA molde. **A.** Producto del alineado y rellenado del oligo 1 y oligo 2 (Scaffold, 76 pb) y reacción de ensamblaje y extensión del scaffold y el oligo 3 (DNA molde, 113 pb). **B.** Producto de la reacción de amplificación del DNA molde del sgRNA A 5' (A 5', 113 pb). **C, D.** El producto de la longitud esperada en la reacción de amplificación del DNA molde para cada sgRNA se purificó y reamplificó (113 pb). Los productos se corrieron en geles de urea 8M – acrilamida 8%.

A partir de los DNA molde obtenidos se realizaron las reacciones de transcripción *in vitro* de los sgRNAs, sin embargo, no logramos obtener una cantidad de producto suficiente para ser visualizado en un gel de acrilamida. Decidimos cambiar la estrategia a seguir para la expresión de los sgRNAs y Cas9 en las células S2 en cultivo.

Construcción de los plásmidos de expresión de Cas9 y sgRNAs

Cas9 y cada uno de los sgRNAs se expresaron en las células S2R+ a partir de vectores de expresión que construimos modificando el plásmido pAc-sgRNA-Cas9 (Addgene #49330). Este plásmido expresa al gRNA bajo la regulación del promotor dU6-2 y expresa a Cas9 y al gen de resistencia a puomicina bajo la regulación del promotor de Actina 5C. Las células transfectadas con este plásmido pueden entonces ser seleccionadas por resistencia a puomicina. Para la obtención de los seis diferentes plásmidos de expresión de los gRNAs y Cas9, mediante PCR insertamos cada una de las secuencias spacer que diseñamos, downstream al promotor dU6-2 y en marco con la secuencia scaffold del gRNA (Figura 8).

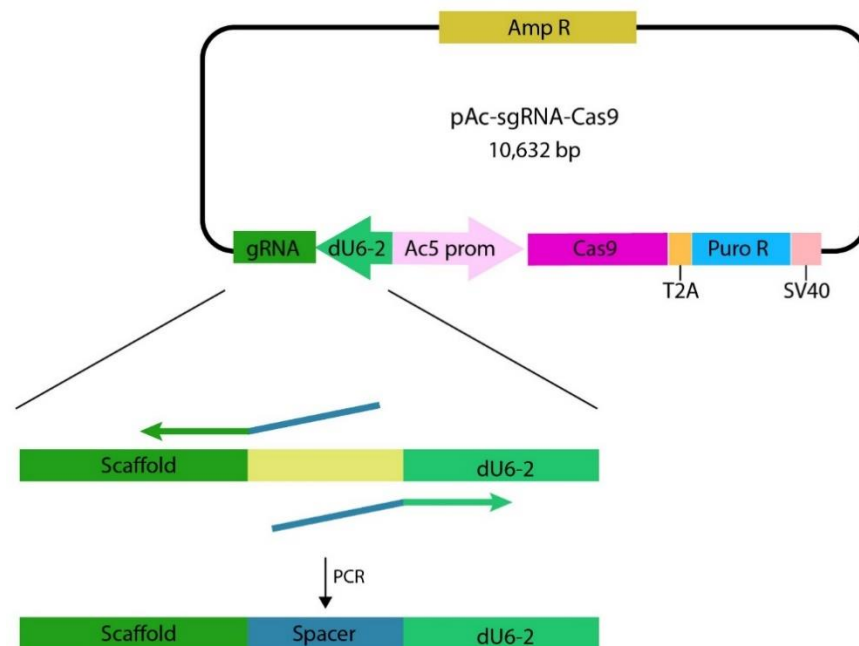


Figura 8. Modificación del plásmido pAc-sgRNA-Cas9. Cada una de las secuencias spacer diseñadas se insertó en el plásmido pAc-sgRNA-Cas9 mediante PCR. Se muestra el diseño de los primers para la inserción de la secuencia spacer.

De esta manera, obtuvimos seis plásmidos, cada uno expresando un gRNA diferente bajo la regulación del promotor dU6-2 y a Cas9 bajo la regulación del promotor de β -actina. Las secuencias spacer se comprobaron mediante secuenciación Sanger (Figura 9).

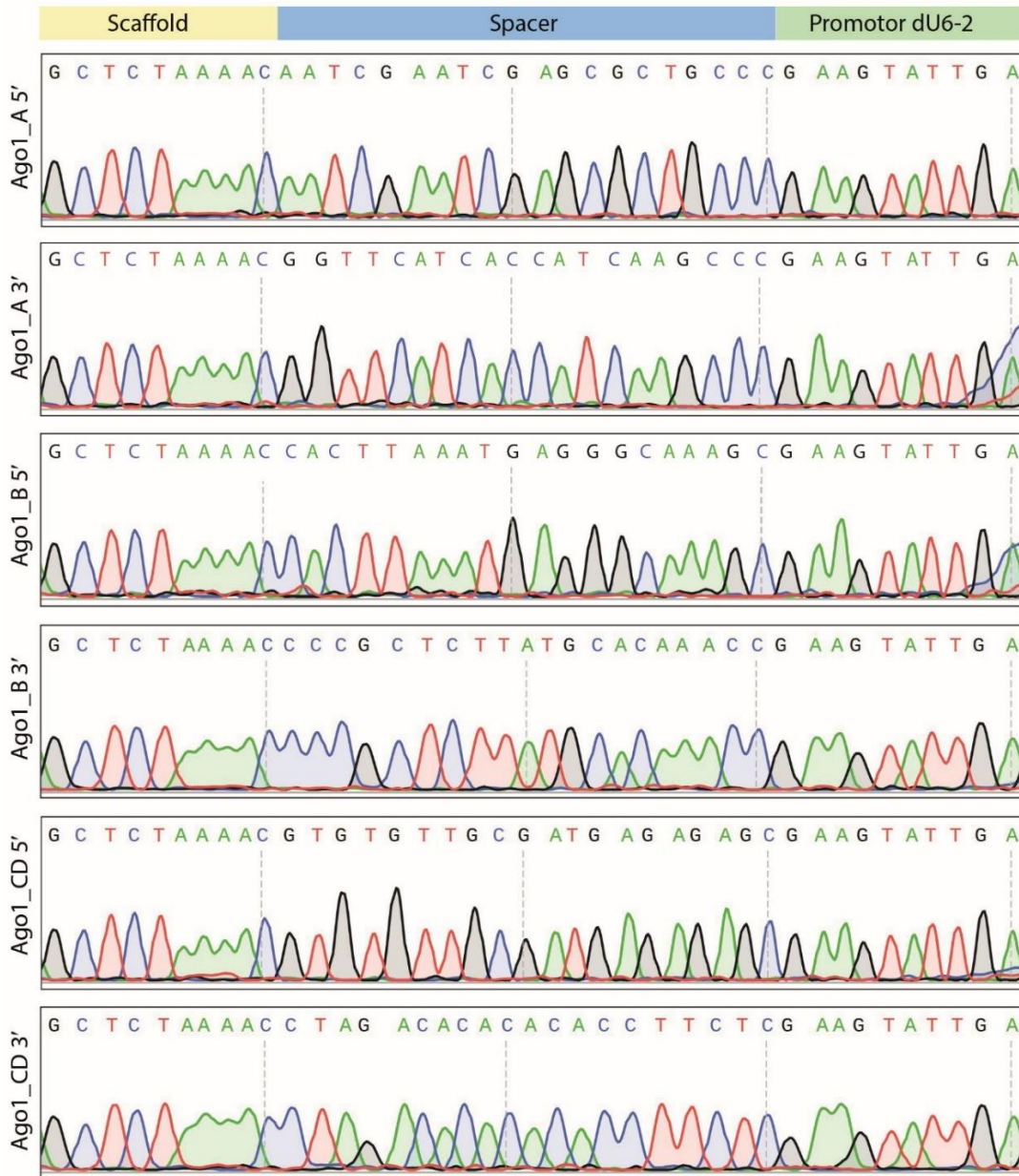


Figura 9. Secuencias spacer insertadas en el plásmido pAc-sgRNA-Cas9. La modificación del plásmido de expresión de los sgRNAs y Cas9 se comprobó mediante secuenciación de Sanger. En los plásmidos que construimos, el promotor y la secuencia del gRNA se encuentran en la hebra antisentido. La figura muestra la secuencia sentido de la región correspondiente a la secuencia spacer de cada uno de los plásmidos construidos.

Luego de comprobar la inserción de cada secuencia spacer en los plásmidos de expresión de Cas9, continuamos con la edición de la región promotora de Ago1 en la línea celular S2.

Comprobación de la edición a nivel genómico

Con los plásmidos construidos se editó cada uno de los promotores de Ago1 en las células S2. Las células se transfectaron con los plásmidos de expresión del par de gRNAs específico para cada promotor y Cas9. Luego de la transfección, las células se seleccionaron por resistencia a puromicina y se amplificaron por PCR los sitios editados. La disminución de la longitud del amplicón obtenido es indicativo de que la edición se llevó a cabo. La longitud del fragmento deletado al editar cada promotor se calculó considerando que Cas9 realiza el corte de DNA de doble hebra a 3 bp de distancia de la secuencia PAM^{57,67}. De esta manera, en las células en las que se editó el promotor de la isoforma A se esperaba un amplicón de 495 bp para las células editadas y de 650 bp para las células WT; para la isoforma B el amplicón esperado para las células editadas es de 502 bp y el de las células WT es de 697 bp; por último, el amplicón de las isoformas CD con el promotor deletado se esperaba de 201 bp y el de las células WT de 600 bp. La edición de los promotores se pudo comprobar para las variantes A y CD, aunque no para la variante B (Figura 10).

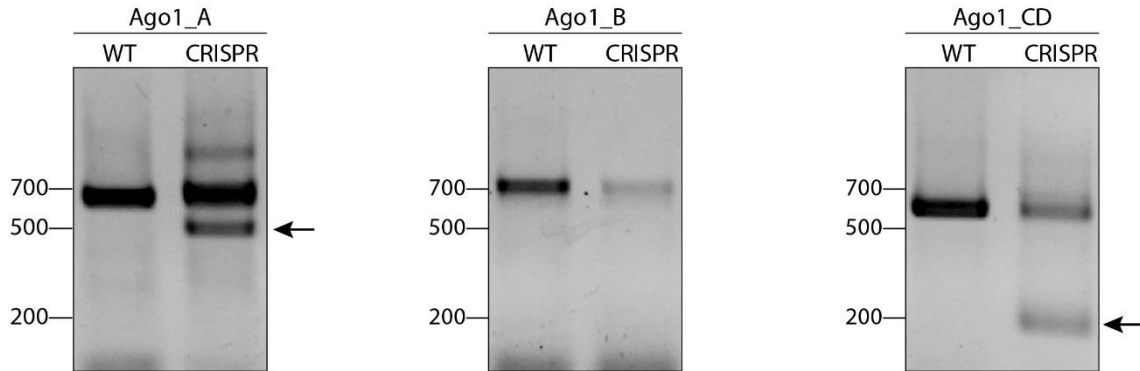


Figura 10. Comprobación de la edición de los promotores de Ago1 en células S2. La delección de los promotores de cada isoforma de Ago1 se comprobó por PCR, a los siete días de haber realizado la transfección de los plásmidos de expresión de Cas9 y los sgRNAs. Para cada isoforma se compara el producto de PCR de las células sin editar (WT) y el de las células editadas (CRISPR). Para los controles de las células sin editar se utilizó DNA de células transfectadas con el plásmido de expresión de Cas9 y el sgRNA con una secuencia spacer que no tiene un sitio blanco en el genoma de *D. melanogaster*. Las flechas indican el producto esperado para los promotores editados.

Luego de realizar la delección de los promotores de Ago1, monitoreamos las células editadas en cultivo y observamos una respuesta similar al editar cualquiera de los tres promotores. Lo que observamos fue que, sin importar qué promotor se hubiera editado, había un periodo en el que un porcentaje importante de las células en cultivo moría. La viabilidad de las células en cultivo se determinó mediante un ensayo con azul de tripán. Al observar la misma respuesta celular al editar cualquiera de los tres promotores decidimos continuar con el análisis a pesar de no haber comprobado aún la edición del promotor de la isoforma B.

Expresión de Ago1 en las células editadas

Lo siguiente fue conocer el efecto de la edición de cada promotor sobre la expresión de Ago1. Para esto, lo primero que hicimos fue comprobar que podíamos detectar cambios en la cantidad de AGO1 en una muestra de proteínas, de manera que construimos una

curva de concentración para la inmunodetección de AGO1 en proteína total obtenida de células S2 (Figura 11).

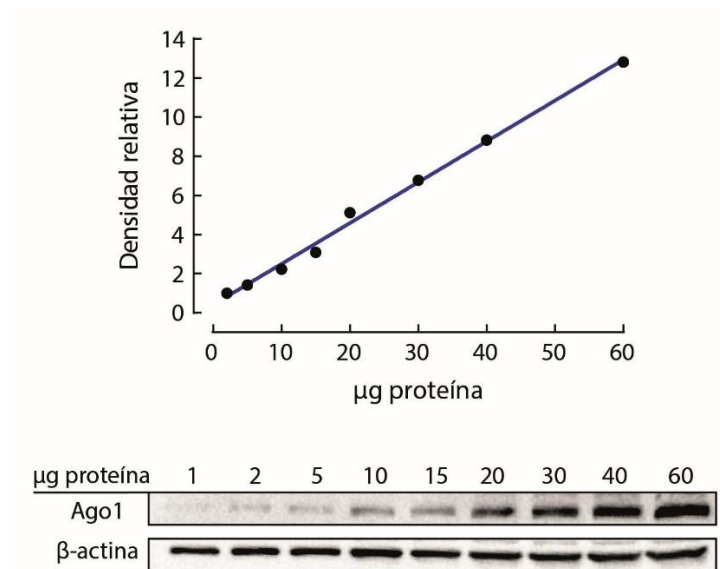


Figura 11. Inmunodetección de AGO1 en células S2. AGO1 y β -actina se detectaron en diferentes cantidades de proteína total. Se grafica la densitometría de las bandas obtenidas para AGO1 normalizada contra lo detectado en 1 μ g de proteína total.

De igual manera, realizamos la inmunodetección de AGO1 en proteína de células S2 transfectadas con el plásmido pAFW-Ago1, para comprobar si podía expresarse en nuestras células y ser utilizado en experimentos de recuperación de AGO1 luego de realizar la edición de los promotores (Figura 12).

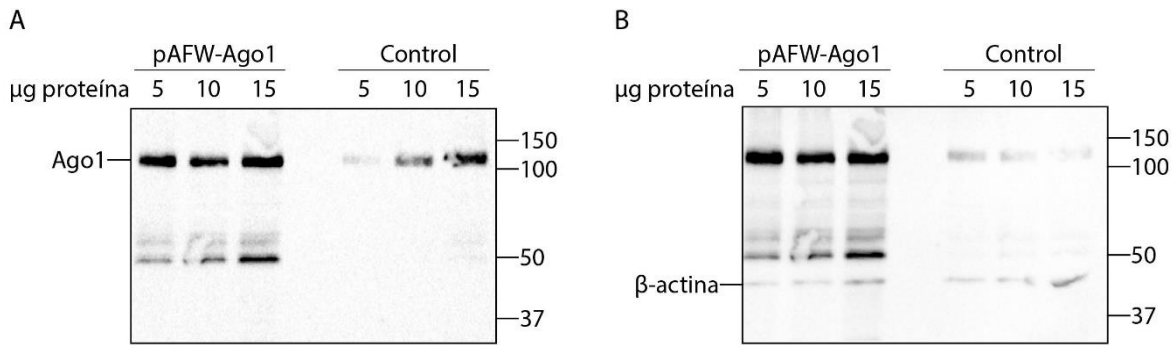


Figura 12. Sobreexpresión de AGO1 en células S2 con el plásmido pAFW-Ago1. Las células se transfectaron con el plásmido de expresión de AGO1 y 48 horas después se realizó la obtención de proteína total. La Inmunodetección se realizó en tres diferentes cantidades de proteína total: 5, 10 y 15 µg. Se compara la Inmunodetección en células transfectadas (pAFW-Ago1) y la de las células no transfectadas (Control) **A.** Inmunodetección de AGO1. **B.** Inmunodetección de β-actina. La Inmunodetección de ambas proteínas se realizó en la misma membrana.

Para comprobar si la edición de los promotores de las isoformas de Ago1 provoca el decaimiento de la proteína, construimos un curso temporal en el que medimos la presencia de AGO1 en las células editadas a los ocho y doce días luego de la transfección. Durante el experimento, además de realizar la edición por CRISPR, en un grupo de células se rescató la presencia de la proteína cotransfectando el plásmido de expresión de Ago1 (pAFW-Ago1) y los plásmidos de CRISPR, el otro grupo de células fue transfectado con los plásmidos de CRISPR y el plásmido de expresión de Ago2 (pAFW-Ago2) como control de expresión del plásmido pAFW-Ago1.

En los resultados de la inmunodetección podemos observar una disminución en la cantidad de AGO1 ocho días después de la edición con CRISPR; a los doce días, la cantidad de AGO1 parece aumentar. Las células en las que la presencia de AGO1 fue rescatada con la cotransfección del plásmido pAFW-Ago1 muestran, tal como se esperaba, un aumento en la expresión de esta proteína (Figura 13).

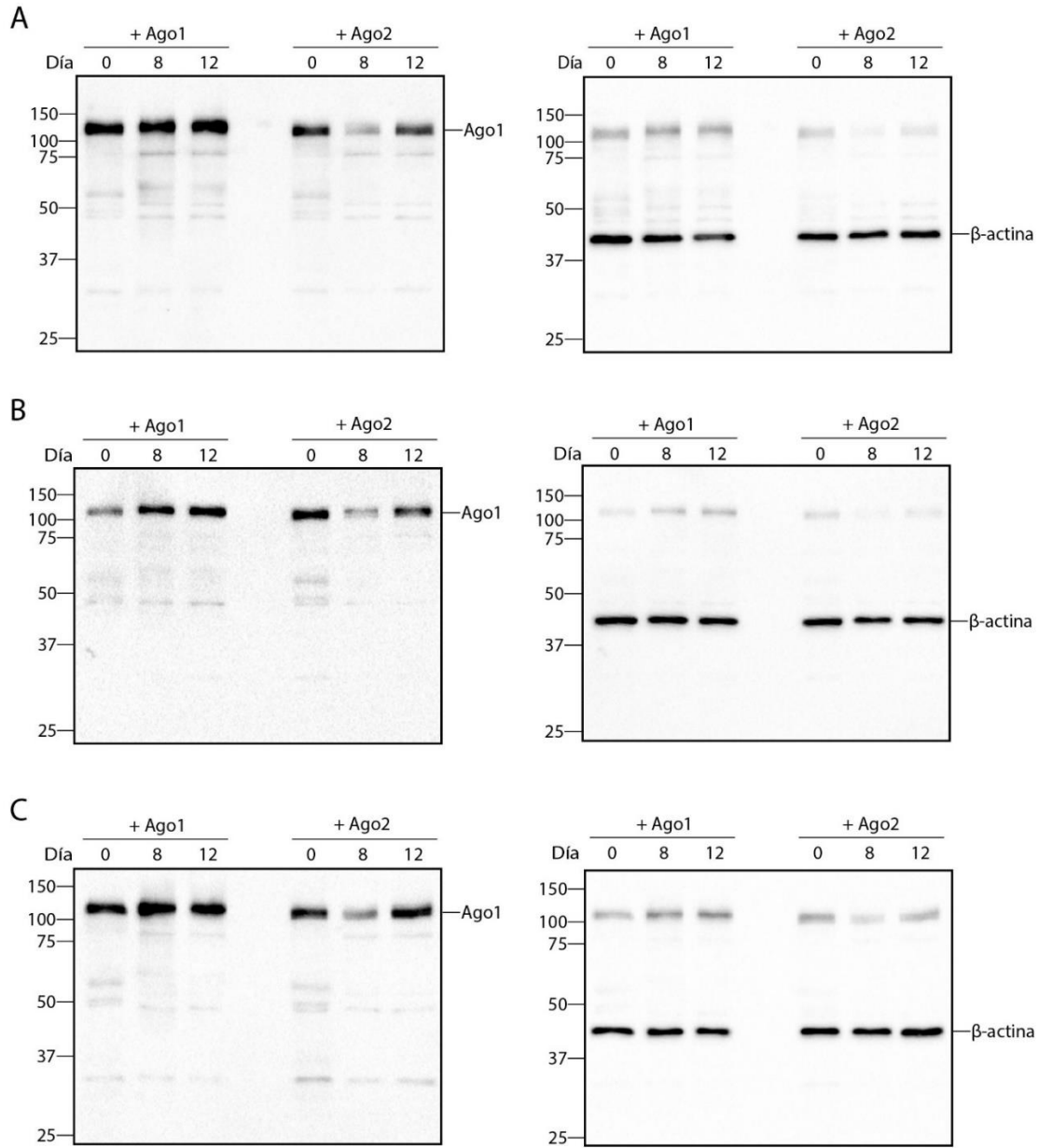


Figura 13. Inmunodetección de AGO1 en células S2 editadas. Se muestra la inmunodetección de AGO1 (columna izquierda) y de β -actina (columna derecha) en las células en las que se editaron los promotores de Ago1. Las detecciones de ambas proteínas se realizaron en la misma membrana. En cada caso se muestra la detección de las proteínas en células editadas y transfectadas con el plásmido de expresión de Ago1 (+Ago1) y la detección de las proteínas en células editadas y transfectadas con el plásmido de expresión de Ago2 (+Ago2). **A.** Edición del promotor de la isoforma A. **B.** Edición del promotor de la isoforma B. **C.** Edición del promotor de las isoformas C y D.

Para observar de manera cuantitativa la disminución de AGO1 en las células editadas, construimos gráficas a partir de determinaciones densitométricas de los resultados de la inmunodetección mostrados en la Figura 13. Al construir las gráficas, cada una de las detecciones se normalizó contra su control de β -actina y la detección de AGO1 al día 0 del experimento se igualó a 1. Los resultados nos indican que la mayor disminución en la cantidad de AGO1 detectado se observó al editar el promotor de la isoforma B, pues ocho días después de la edición la cantidad de esta proteína en las células disminuyó cerca de siete veces. Cuando se editó el promotor de la isoforma A, la disminución observada fue casi cuatro veces respecto a las células sin editar. Por último, el cambio menos drástico se observó en las células en las que se editó el promotor de las isoformas C y D, en tal caso se detectó casi tres veces menos AGO1 que en las células no editadas. El análisis densitométrico de las inmunodetecciones de AGO1 en células editadas también nos mostró que, a los doce días de la edición, detectamos una cantidad de AGO1 similar a la observada antes de realizar la edición de los promotores (Figura 14).

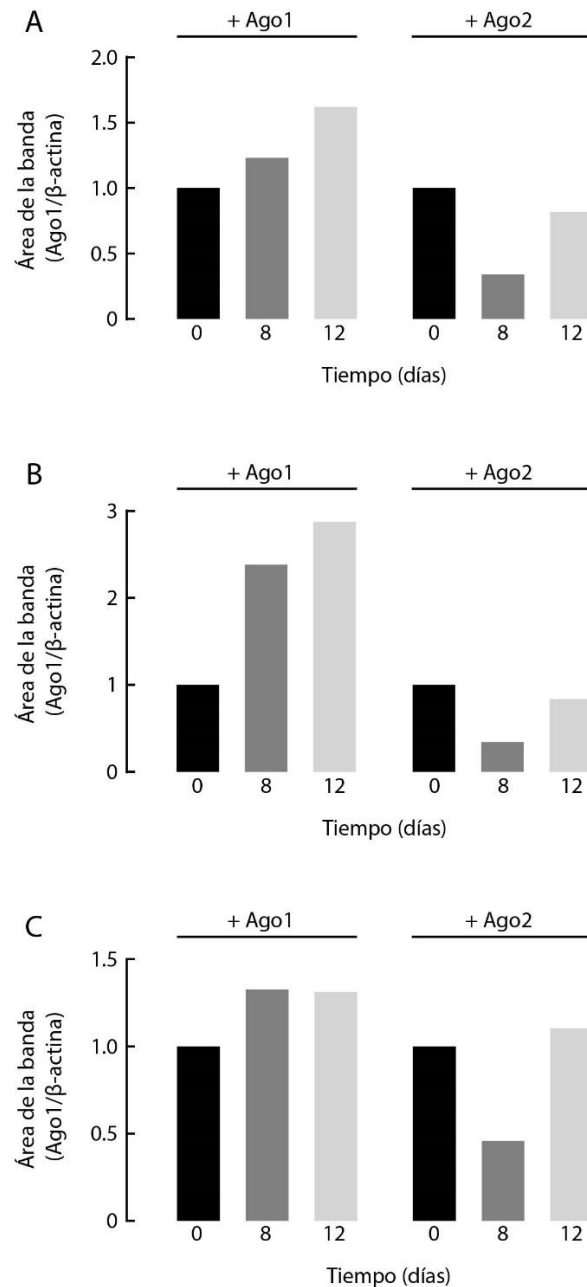


Figura 14. Análisis densitométrico de la inmunodetección de AGO1 en células S2 editadas. Se muestra el análisis por densitometría de la inmunodetección de AGO1 luego de la edición de cada uno de los promotores, mostrada en la Figura 14. Cada una de las detecciones se normalizó contra su control de β -actina y la detección de AGO1 al día 0 del experimento se igualó a 1. **A.** Edición del promotor de la isoforma A. **B.** Edición del promotor de la isoforma B. **C.** edición del promotor de las isoformas C y D. Experimento representativo.

El análisis de la presencia de AGO1 en las células editadas nos mostró que la delección de cualquiera de los tres promotores de Ago1 afecta la expresión de la proteína. Lo siguiente que analizamos fue si este decaimiento en la cantidad de AGO1 en las células afectaba la proliferación.

Proliferación de las células editadas

Para determinar la relación entre el decaimiento de AGO1 y la proliferación de las células editadas, cuantificamos el número de células que se encontraban viables luego de 12 días de haber editado los promotores. Realizamos una comparación de estas cuantificaciones contra células transfectadas con los plásmidos de CRISPR, pero con una secuencia spacer sin blanco en el genoma de *D. melanogaster* (Figura 15, Cas9) y contra células que sí habían sido editadas pero cuya expresión de Ago1 fue rescatada al cotransfectar el plásmido pAFW-Ago1 y los plásmidos de CRISPR (Figura 15, +Ago1). A pesar de que en las células transfectadas con el plásmido de expresión de Ago1 la proliferación de las células editadas no es la misma que la de las células no editadas (Cas9), encontramos una menor cantidad de células en los casos en los que se editaron los promotores y no se rescató la presencia de AGO1 en las células.

Los resultados de las figuras 13, 14 y 15 nos muestran que al deletar los promotores de Ago1 la cantidad de esta proteína en las células decrece al tiempo que la proliferación celular disminuye. Cuando rescatamos la presencia de AGO1 transfectando el plásmido de expresión para esta proteína al tiempo que realizamos CRISPR, observamos que la proliferación de las células es mayor que la de las células editadas en las que no se reestableció la presencia de AGO1, aunque es menor que la de las células transfectadas con el plásmido de expresión de Cas9 y un gRNA sin blanco en el genoma de *D. melanogaster*. Estos resultados indican que la delección de cualquiera de los tres promotores de las isoformas de Ago1 conlleva al decaimiento de esta proteína y afecta la proliferación celular.

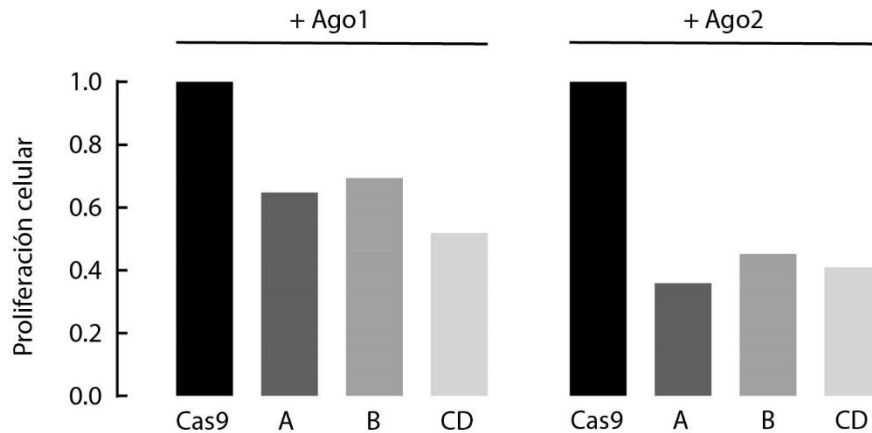


Figura 15. Proliferación de células editadas. La gráfica muestra la cuantificación de células luego de 12 días de haber realizado la edición de los promotores de Ago1. **Izquierda.** Células editadas y transfectadas con el plásmido de expresión de Ago1. **Derecha.** Células editadas y transfectadas con el plásmido de expresión de Ago2. En ambos casos Cas9 representa células transfectadas con los plásmidos de expresión de Cas9 y los gRNAs, pero con una secuencia spacer sin blanco en el genoma de *D. melanogaster*. Experimento representativo.

Al obtener el mismo resultado al deletar los tres promotores de Ago1 decidimos analizar la región promotora del gen, para saber si cada una de las diferentes isoformas de Ago1 podrían estar transcribiéndose a partir de un promotor activo diferente y si todas ellas se encontraban presentes en las células S2.

Identificación de los promotores de Ago1 activos en células S2 de *D. melanogaster*

Para comprobar si los tres promotores propuestos se encuentran activos en las células S2 de *D. melanogaster*, analizamos aquellas características propias de una región promotora activa. Por ejemplo, el estado de metilación y acetilación de la cromatina tanto en la región río arriba como en la río abajo con respecto al sitio de inicio de la transcripción, la probabilidad de unión de la RNA pol II a esta región y la presencia de transcritos cuyo sitio de inicio de la transcripción se encuentre en el centro de esta región⁶⁸.

Analizamos datos publicados⁶⁹ de experimentos de ChIP-seq que identifican sitios de unión de la RNA pol II y el estado de metilación y acetilación de histonas⁷⁰, específicamente muestra la trimetilación de la lisina 4 de la histona 3 (H3K4m3) y la acetilación de la lisina 27 de la histona 3 (H3K27ac). De igual manera, analizamos datos de secuenciación de transcritos con cap⁷¹ (Cap-analysis gene expression, CAGE). Puesto que en esta técnica de secuenciación se analizan sólo los primeros 50 nt del mRNA⁷², el análisis nos permitió diferenciar la expresión de las isoformas A, B y CD pero no fue posible diferenciar entre las isoformas C y D, pues estas últimas sólo difieren en la longitud del 3' UTR. Sin embargo, el análisis nos aproximó a conocer cuál de los tres promotores de Ago1 se encuentra activo en estas células. Además, esta técnica de secuenciación permite identificar con resolución a nivel de nucleótido el sitio de inicio de la transcripción de genes específicos⁷².

Al analizar los resultados, encontramos que los promotores de las isoformas B, C y D, pero no el promotor de la isoforma A presentan las marcas H3K4m3 y H3K27ac río abajo al TSS, ambas marcas son características de regiones promotoras activas⁷³. De igual manera, observamos que la RNA polimerasa II se posiciona en la región inmediata río abajo al TSS propuesto para los promotores de las isoformas B, C y D, siendo la región río abajo al TSS de la isoforma B el sitio al que la polimerasa se une con mayor frecuencia. Los resultados de CAGE, además de mostrarnos el TSS más frecuente, nos indican la abundancia de cada isoforma. De acuerdo con los resultados, en las células S2 la isoforma más abundante sería Ago1_B, seguida de la suma de las isoformas C y D, mientras que Ago1_A no se expresaría en estas células (Figura 16).

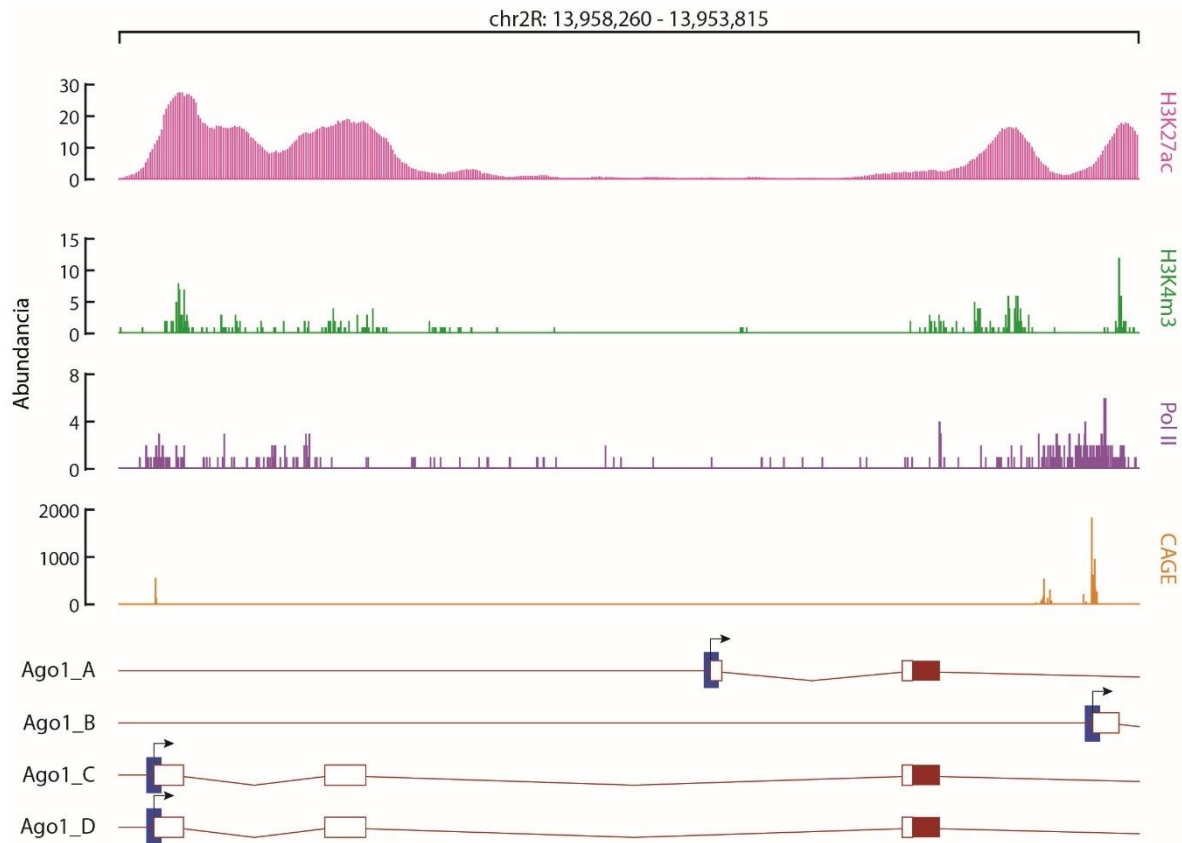


Figura 16. Región promotora de Ago1 en células S2. Se muestra el análisis de las principales características de un promotor activo. La acetilación de la lisina 27 de la histona 3 (H3K27ac, rosa), la trimetilación de la lisina 4 en la histona 3 (H3K4m3, verde), la probabilidad de unión de la RNA polimerasa II (pol II, morado) y la secuenciación de transcritos con cap (CAGE, naranja). En la parte inferior se representa la región promotora del gen, el sitio de inicio de la transcripción se indica con una flecha y la región propuesta como centro del promotor se muestra en azul. Los rectángulos blancos representan exones que formarán parte de la región 5' UTR de cada mRNA, los rectángulos guinda representan exones traducibles y las líneas indican intrones y regiones intergénicas. Los datos para el análisis de H3K4m3 y sitios de unión de la RNA pol II se tomaron de Gan, Q. 2010⁷⁴, el análisis de H3K27ac se realizó a partir de datos recuperados de Mueller, B. 2017⁷⁰ y los datos analizados de CAGE se tomaron de Rennie, S. 2018⁷¹.

Estos resultados indican que, de manera inesperada, a pesar de que la isoforma A no se transcribe en las células S2, deletar el centro de su promotor afecta la expresión de AGO1 de manera similar que al editar cualquiera de los promotores activos. Puesto que el promotor de la isoforma A no presenta características propias de un promotor activo y el mRNA de esta isoforma no se encuentra en células S2, nos preguntamos si este promotor podría estar transcribiendo el mRNA de la isoforma A en algún tejido o estadio del desarrollo de *D. melanogaster*.

Expresión de Ago1 durante el desarrollo de *D. melanogaster*

Para comprobar si el promotor de la isoforma A podría encontrarse activo en algún momento durante el desarrollo de *D. melanogaster*, analizamos resultados de CAGE⁷⁵ de muestras correspondientes a cada estadio del desarrollo de este organismo.

Al analizar los datos encontramos que tanto el promotor de la isoforma B como el de las isoformas C y D se encuentra activo en todos los estadios del desarrollo de *D. melanogaster*. Además, la isoforma B es la más abundante en todas las muestras. De manera interesante, los datos indican que el promotor de la isoforma A se encuentra activo sólo en los estadios de larva y pupa, generando una menor cantidad de transcrito que el resto de los promotores (Figura 17).

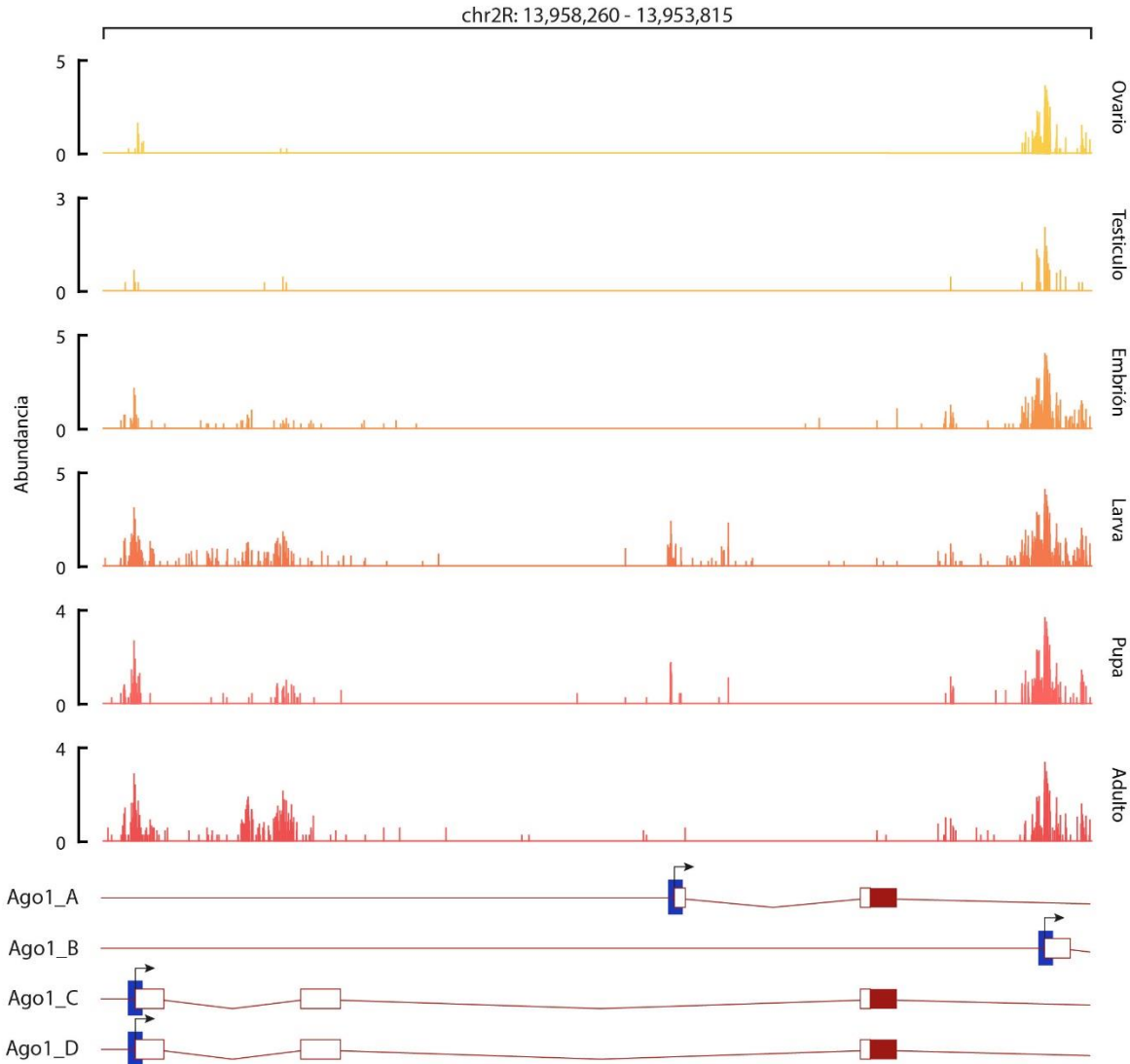


Figura 17. Región promotora de Ago1 en diferentes estadios del desarrollo de *D. melanogaster*. Se analizaron datos de secuenciación de transcritos con cap (CAGE). La gráfica muestra el sitio de inicio de la transcripción de las isoformas de Ago1, así como la abundancia de transcrito generado a partir de cada uno de los promotores. El estadio del desarrollo "adulto" corresponde a cabeza de mosca adulta. En la parte inferior se representa la región promotora del gen, el sitio de inicio de la transcripción se indica con una flecha y la región propuesta como centro del promotor se muestra en azul. Los rectángulos blancos representan exones que formarán parte de la región 5' UTR de cada mRNA, los rectángulos guinda representan exones traducibles y las líneas indican intrones y regiones intergénicas. Los datos utilizados para realizar este análisis de recuperaron de Chen, ZX. 2014⁷⁵.

Discusión

El objetivo del presente proyecto es editar la región promotora del gen Ago1 en células S2 de *D. melanogaster*. El gen Ago1 en *D. melanogaster* presenta cuatro isoformas cuya síntesis se encuentra regulada por tres posibles promotores. Utilizando CRISPR-Cas9 deletamos el centro de cada uno de estos posibles promotores. El resultado de cada delección fue el decaimiento de AGO1 y la disminución de la proliferación celular.

En *D. melanogaster* se expresan cuatro isoformas de Ago1

De acuerdo con la información obtenida de bases de datos genómicos, Ago1 cuenta con cuatro isoformas a nivel de transcritos. Estos se diferencian unos de otros en la secuencia y longitud de las regiones 5' y 3' UTR (Figura 4). A pesar de que no contamos con información detallada sobre cada una de las isoformas, pues aún no han sido estudiadas, las diferencias que existen entre ellas podrían conferirles ciertas características. Por ejemplo, cada uno de los mRNAs de Ago1 podrían expresarse en diferentes tejidos, etapas del desarrollo o fase del ciclo celular. Además, las diferencias entre ellos podrían determinar el tiempo de vida media que presentan en las células. Las dos proteínas que se originan a partir de los cuatro mRNAs podrían también contar con características específicas que determinen su función, especificidad o estabilidad en la célula.

Puesto que aún no se ha reportado cuál o cuáles de las isoformas de Ago1 se expresan en las células S2, la estrategia que seguimos fue la de editar cada uno de los tres posibles promotores, por separado. La respuesta celular a cada edición nos podría indicar cuál de estos promotores se encuentra activo y, por lo tanto, cuál de los mRNAs se expresa en nuestra línea celular. La delección de los centros de los promotores se realizó utilizando el sistema CRISPR-Cas9.

Los sgRNAs diseñados son eficientes y específicos

Al diseñar los RNAs guías elegimos utilizar aquellos que presentaran un mayor puntaje en los scores que determinan la eficiencia y especificidad de las secuencias spacer. Para la determinación de la eficiencia se utilizó el score de Doench y para la especificidad el

MIT Guide Specificity score. El resultado al utilizar ambos algoritmos fue que los gRNAs diseñados son lo suficientemente eficientes para dirigir a Cas9 hacia las regiones del genoma previstas y tan específicos que es poco probable que ocurran cortes en sitios off-target durante la edición del genoma (Tabla 2).

Al realizar el análisis de los posibles sitios off-target para cada una de las guías encontramos que sólo los gRNAs diseñados para el reconocimiento de las secuencias río abajo al promotor de las isoformas A, C y D podrían dirigir cortes en sitios no específicos (Tabla 3). El corte off-target que podría dirigir el sgRNA de la isoforma A y uno de los que podrían ser dirigidos por el gRNA de las isoformas C y D ocurrirían en regiones intergénicas. Al no ocurrir dentro de regiones codificantes, existe una menor probabilidad de que estos cortes produzcan un fenotipo que pueda interferir en nuestro estudio. Sin embargo, no se descarta la posibilidad de que se traten de regiones reguladoras o que al reparar el corte al DNA se produzcan reordenamientos que afecten a los genes que se encuentran cercanos, lo que tal vez sí podría ocasionar algún cambio en el fenotipo celular. El segundo sitio off-target predicho para el gRNA de las isoformas C y D ocurriría en un exón del gen CG13594. Este gen codifica para una proteína cuya función aún no ha sido anotada, por lo que no podemos saber si el corte en este sitio podría o no afectar algún proceso celular.

A pesar de que el análisis mediante diferentes algoritmos señala que los sgRNAs que diseñamos y utilizamos durante este proyecto pueden guiar a Cas9 hasta los sitios previstos para el corte del DNA con una probabilidad muy baja de generar cortes en sitios no específicos del genoma, es importante mencionar que los algoritmos que se encuentran disponibles para este análisis han sido desarrollados a partir de predicciones y datos experimentales recabados de genomas de mamíferos. De manera que no hay garantía de que los gRNAs se comporten o no de la misma manera en células de insecto. Sin embargo, la elección de los algoritmos utilizados para medir eficiencia y especificidad de los gRNAs se realizó de manera minuciosa y se eligieron aquellos que podrían realizar la mejor aportación al proyecto.

El diseño de los sgRNAs se realizó también de manera tal que dirigieran el corte de Cas9 a la posición más cercana posible al centro del promotor para que el fragmento de DNA

que se deletara del genoma de *D. melanogaster* contuviera sólo al centro del promotor, o la menor cantidad de región intergénica o 5' UTR posible.

El gen Ago1 se encuentra ubicado en el cromosoma 2R, en la hebra antisentido. En este locus, pero en la hebra sentido, se encuentra co-localizado el gen mRpL53. Este gen se localiza entre el segundo y tercer intrón de las isoformas C y D, justo antes del promotor de la isoforma A (Figura 4). La edición del promotor de las isoformas C y D no afectan la expresión de este gen, sin embargo, la edición del promotor de la isoforma A implica el corte de aproximadamente 140 nt de la región 3' UTR del gen mRpL53. Este gen codifica para la proteína L53 que forma parte de los ribosomas en la mitocondria. El knock-out de algunas proteínas ribosomales da como resultado el síndrome "Minute", caracterizado por una serie de defectos celulares y de desarrollo de los organismos que lo presentan. En *D. melanogaster* este síndrome incluye un desarrollo tardío y una baja fertilidad y viabilidad. El análisis de las proteínas ribosomales, tanto de citoplasma como de mitocondria indica que dicho síndrome es causado por una síntesis deficiente de proteínas ribosomales de citoplasma, que implica una baja disponibilidad de citoribosomas; sin embargo, este síndrome no está asociado a una deficiencia de proteínas ribosomales en mitocondria⁷⁶. Por lo tanto, si al editar el promotor de la isoforma A se afectara la expresión del gen mRpL53, esto no tendría repercusión sobre la viabilidad celular o algún proceso metabólico que representara un sesgo para nuestro análisis. Es por esto que decidimos continuar con la edición de los promotores utilizando las guías ya diseñadas.

La edición de los promotores se comprobó a nivel genómico en una población celular heterogénea

La edición de los promotores fue comprobada a nivel genómico mediante PCR a partir de DNA obtenido después de siete días de haber realizado la transfección de los plásmidos de CRISPR, justo al terminar la selección de las células transfectadas mediante resistencia a puromicina (Figura 10). La edición de los promotores de las isoformas A y CD se comprobó mediante PCR. En ambos casos se obtuvo el producto del tamaño esperado luego de la deleción, sin embargo, la edición no pudo ser comprobada para

el promotor de la isoforma B. Al ser la isoforma B la más abundante en las células S2 y si la edición del promotor de esta isoforma conlleva al decaimiento de la proteína y por lo tanto afecta la viabilidad celular, es posible que, comparado con las células en las que se editaron los otros dos promotores, las células en las que se editó el promotor de la isoforma B tengan una menor viabilidad. Por lo tanto, de este grupo de células se pudo haber obtenido un DNA de menor calidad, lo que explicaría la baja eficiencia de la reacción de PCR y el no haber podido comprobar la edición de este promotor.

Las células en las que se editó el promotor de la isoforma A muestra un producto de PCR no esperado de aproximadamente 900 bp. Al reparar los cortes de doble hebra realizados durante la edición por CRISPR pueden ocurrir rearrreglos que podrían incluir sólo regiones cercanas al sitio de corte o regiones distales, produciendo cambios importantes en el cromosoma en cuestión⁷⁷. Esto pudo ocurrir durante la edición de la isoforma A, aunque también podría tratarse de la inserción de otro fragmento de DNA en el sitio de corte. Cualquiera de estas dos alternativas podría ser detectada mediante la PCR que realizamos.

Por otra parte, tanto las células correspondientes a la condición editada como las de la condición no editada contaban con DNA sin editar. Esto se puede explicar analizando las diferentes condiciones celulares que tenemos al momento de realizar la extracción de DNA. Para lograr la edición de cada promotor se necesitan dos gRNAs, cada uno de ellos expresado en un plásmido diferente (Figura 8). Al tener cada uno de los plásmidos el gen de resistencia a puromicina, si uno de ellos se transfectó en la célula, ésta es ahora resistente al antibiótico de selección, pero se necesita la co-transfección de ambos plásmidos para que la edición del promotor pueda ocurrir de la manera esperada. Es por esto que, luego de la transfección tenemos una población heterogénea de tres tipos de células diferentes: aquellas que se transfectaron con plásmidos de expresión para sólo uno de los dos gRNAs necesarios para la edición, aquellas en las que se logró la co-transfección de los plásmidos y por lo tanto puede ocurrir la edición y aquellas células que no fueron transfectadas. Después de la transfección se realizó la selección con puromicina, al seleccionar las células con este antibiótico logramos deshacernos de una de las poblaciones celulares, aquellas células que no fueron transfectadas y por lo tanto

no presentan resistencia al antibiótico de selección. Sin embargo, la población de células resistentes es aún heterogénea, pues el hecho de que una célula sea resistente al antibiótico de selección no implica que haya sido editada. El producto de PCR que corresponde al DNA no editado pudo provenir de estas células que son resistentes a puromicina, pero no fueron editadas. Por otra parte, las células S2 de *D. melanogaster* son diploides⁷⁸, por lo que, durante la edición del promotor de Ago1, es posible que se editara sólo uno de los alelos, generando heterocigosis para la edición. Estas células heterocigotas podrían haber contribuido a la detección tanto del DNA editado como del no editado.

La edición de los posibles promotores del gen Ago1 produce el decaimiento de la proteína

En primera instancia, puesto que Ago1 es un gen necesario para la supervivencia celular y debido a que se tiene más de una región promotora propuesta, es de esperarse que al deletar sólo uno de los tres posibles promotores, la expresión de la proteína no se vea afectada si los dos promotores que permanecen intactos se activan de manera tal que suplan la función de aquel que fue retirado del genoma. Esto podría suceder al deletar cualquiera de los tres promotores, si es que estos se activan y actúan de manera independiente y si existe la posibilidad de que se transcriba el mRNA de Ago1 a partir de cualquiera de ellos en las células en las que realizamos esta investigación. Sin embargo, la deleción de cualquiera de los tres posibles promotores produce el decaimiento de la proteína (Figuras 13 y 14).

En el día número ocho del experimento, luego de siete días de haber transfectado los plásmidos de CRISPR se realizó la extracción de proteínas para la inmunodetección de AGO1. El resultado que obtuvimos fue el mismo al deletar cualquiera de los tres posibles promotores, una clara disminución de la cantidad de AGO1 presente en las células (Figuras 13 14). Al igual que al comprobar la edición a nivel genómico, tenemos una población heterogénea de células y podemos detectar la proteína tanto de las que fueron editadas como de las que no. En aquellas células en las que se logró la deleción del promotor de Ago1 la proteína debió detener su síntesis y pudimos observar el

decaimiento de aquella que ya se encontraba presente en las células antes de la edición. Sin embargo, al mismo tiempo podemos detectar la proteína AGO1 de las células que no fueron editadas y que continúan con el metabolismo de manera normal.

La inmunodetección de AGO1 también se realizó a los doce días de haber comenzado el experimento, es decir, once días después de la transfección de los plásmidos de CRISPR. Para cada uno de los promotores editados, la cantidad detectada de AGO1 se reestablece y es cercana a la detectada antes de realizar la edición (Figuras 13 y 14). Esto podría deberse a diferentes motivos. Por ejemplo, es posible que mientras que las células editadas en las que la síntesis de AGO1 se ha detenido mueran, aquellas que no están editadas continúen proliferando a una velocidad normal. Si esto sucede, a los doce días, el AGO1 que detectamos podría provenir de estas células, que no fueron editadas y que en ese momento ya superan en número a las células que sí se editaron. Sin embargo, otra posibilidad es que la falta de uno de los posibles promotores provoque que alguno de los otros dos promotores se active en días posteriores a la edición. Esto podríamos saberlo si al día doce del experimento, o en días posteriores, se repitiera la PCR con la que detectamos a las células editadas. Si podemos detectar células editadas cuando la cantidad de AGO1 en las células se ha reestablecido entonces podría ser que el mRNA del gen se esté sintetizando a partir de otro promotor y a partir de este transcrito se genere la proteína. Otra alternativa para saber si alguno de los promotores no editados se activa sería realizar un análisis cuantitativo de cada isoforma de Ago1 a nivel de transcrito. De esta manera podríamos monitorear cambios en la expresión de cada isoforma y detectar si alguna de ellas se sobre expresa luego de la edición de los posibles promotores.

El análisis densitométrico de las bandas obtenidas en la inmunodetección de AGO1 mostró que la disminución de la proteína es diferente al deletar cada uno de los posibles promotores (Figura 14). De acuerdo con el análisis, la disminución es mayor al editar el promotor de la isoforma B, se observa una disminución media si se edita el promotor de la isoforma A y la menor disminución es observada al deletar el promotor de las isoformas C y D. Puesto que el promotor de la isoforma A se encuentra en el segundo intrón de las isoformas C y D, es posible que la edición de esta región afectara de alguna

manera la síntesis de estas isoformas, tal vez afectando el splicing de esta región. De la misma manera, la edición del promotor de la isoforma B pudo afectar la expresión de las isoformas A, C y D pues este se encuentra en regiones intrónicas de dichas isoformas. Puesto que el promotor de las isoformas C y D se ubica upstream al de las demás isoformas, su edición no afectaría la expresión de estas. Esto podría explicar que al editar los promotores se obtengan diferentes niveles de disminución de la proteína. Sin embargo, las diferencias en el decremento de la proteína entre las diferentes deleciones de los promotores podrían deberse a factores experimentales, como la eficiencia de los sgRNAs o el número de células que logran ser co-transfectadas y la proporción de estas en las que se logra la edición de uno o los dos alelos de Ago1. Es posible también, que los mRNAs de las isoformas tengan tiempos de vida media diferentes. Esto podría hacer que a partir de ellos se generen diferentes cantidades de proteína y esta podría ser la razón de que se obtengan diferentes niveles de disminución de AGO1 luego de la edición de cada posible promotor.

Que AGO1 decayera al editar cualquiera de los posibles promotores fue un resultado completamente inesperado. Durante el diseño experimental del proyecto se había considerado que al editar cada uno de los promotores se pudiera identificar aquel que se encuentra activo en las células S2 si la edición de sólo uno o dos de ellos producía el decaimiento de la proteína. Otra alternativa que consideramos fue que al editar cualquiera de los posibles promotores, la presencia de AGO1 continuara de manera normal. Esto último si Ago1 fuera un gen que presenta robustez, entendiendo este concepto como la ausencia de variaciones fenotípicas ante la presencia de variación en el genotipo⁷⁹. Es decir, que al deletar el promotor que se encuentra activo en las células, alguno de los dos que se encontraban inactivos pudiera activarse para suplir la función del que ha sido inactivado; o que a partir de algún otro mecanismo se asegurara la síntesis de la proteína a pesar de haber ocurrido un cambio en el genotipo celular.

Desde el punto de vista evolutivo, Ago1 en *D. melanogaster* pudo haber desarrollado tres diferentes promotores y sitios de inicio de la transcripción para asegurar la expresión del gen frente a posibles mutaciones que pudieran afectarlo. Esto por ser un gen necesario para el desarrollo de organismos completos. Sin embargo, Ago1 no sólo no

muestra robustez, sino que, además los resultados parecen indicar que los promotores propuestos no actúan de manera independiente, sino que se necesita la presencia de las tres secuencias para lograr la transcripción del gen. Esto último operaría en contra de asegurar la expresión del gen, pues ahora la probabilidad de que una mutación afecte la expresión de Ago1 se triplica. Sin embargo, hace falta realizar un análisis más extenso de la región promotora de Ago1 para poder obtener conclusiones acerca de la manera en la que se regula.

El decaimiento de AGO1 conlleva a la muerte celular

Para obtener un resultado cuantitativo de las observaciones realizadas sobre la proliferación durante el cultivo y puesto que AGO1 es una proteína necesaria para la supervivencia celular, cuantificamos las células de cada condición al último día del experimento. Lo que encontramos fue que al editar los promotores de Ago1 y al tiempo que ocurre el decaimiento de la proteína, la proliferación celular disminuye. Además, observamos que la proliferación celular mejora si en las células editadas se rescata la presencia de AGO1 mediante un plásmido de expresión de esta proteína. La disminución de la proliferación es similar al editar cualquiera de los tres promotores (Figura 15). Aunque, en realidad, no es posible realizar una comparación entre las células editadas debido a los factores experimentales que influyen sobre la variabilidad del éxito de la reacción de edición. Sin embargo, el resultado obtenido de este último experimento indica que la edición de cualquiera de los tres posibles promotores propicia el decaimiento de la proteína y la disminución de la proliferación celular.

Para resolver algunas de las preguntas que continúan abiertas deberíamos obtener monoclonas de las células editadas. A partir de estas podríamos comprobar la edición de los promotores en ambos alelos del gen Ago1, asegurando tener una población de células homocigotas para la edición. En estas células se realizaría también el análisis de posibles cortes en sitios off-target para asegurarnos de haber editado sólo la región esperada. La obtención de estas monoclonas nos permitiría hacer comparaciones entre lo observado al deletar cada uno de los promotores, pues eliminaríamos el sesgo introducido por factores experimentales. Luego de comprobar la edición on-target y

descartar cortes off-target se podría realizar una cuantificación de la expresión de cada isoforma de Ago1 para conocer el origen de las diferencias entre las detecciones de la proteína posterior a la edición de los posibles promotores. Esta cuantificación podría realizarse a nivel de transcrito mediante PCR de tiempo real, aprovechando la diferencia en la secuencia 5' UTR y el primer exón de los mRNAs de las isoformas de Ago1. Sin embargo, en este punto de la investigación no es posible la obtención de estas monoclonas debido a que, al editar los posibles promotores, la síntesis de AGO1 se detiene y las células no pueden sobrevivir. Para resolver esta dificultad se plantea la implementación de un sistema inducible que permita controlar temporalmente la síntesis de AGO1. De esta manera podríamos sustituir cada uno de los posibles promotores endógenos por un promotor metabólicamente inducible, comprobar las ediciones y continuar con el estudio de Argonauta, miRNAs y RISC.

Dos de los tres posibles promotores de Ago1 se encuentran activos en las células S2 de *D. melanogaster*.

Los resultados obtenidos de la edición de cada promotor parecen indicar que los tres promotores se encuentran activos en las células S2. Para comprobar esto, analizamos la región promotora del gen, buscando las características propias de una región promotora activa.

Para la identificación del o los promotores de genes activos se deben tomar en cuenta factores como el estado de condensación de la cromatina y marcas en las proteínas histonas que ayudan a potenciar o inhibir la transcripción de los genes. A pesar de que existen secuencias consenso que permiten la unión de los factores de la transcripción a un promotor, generalmente la unión de estos factores al promotor genera una actividad basal muy baja que puede ser suprimida o potenciada por modificaciones en la cromatina o por elementos reguladores distales^{80,81}. Además, la regulación de las regiones promotoras eucariontes, así como las marcas epigenéticas en las proteínas histonas, son procesos dinámicos que pueden ser influenciados por factores endógenos o estímulos externos. De esta manera, una región que en cierto tejido se encuentra epigenética y transcripcionalmente activa puede, en un tejido diferente, presentar marcas

epigenéticas distintas y ser transcripcionalmente inactiva⁶⁸. Es por esto que, para la identificación de qué promotor de Ago1 se encuentra activo en las células S2 analizamos datos de H3K27ac, H3K4m3, unión de la RNA pol II y CAGE generados específicamente a partir de esta línea celular (Figura 16).

De acuerdo con el análisis realizado, en las células S2 de *D. melanogaster* se encuentran activos sólo dos de los tres posibles promotores de Ago1. Los promotores de las isoformas B, C y D se encuentran transcripcionalmente activos. Pero lo más interesante es que, a pesar de que el promotor de la isoforma A, por las evidencias de H3K4m3, unión de la RNA pol II y presencia de transcritos cuyo sitio de inicio corresponda a esta isoforma, parece no estar activo en las células S2, la delección de esa región del genoma propicia el decaimiento de la proteína y la disminución de la proliferación celular de manera similar a la delección de los otros dos posibles promotores.

Los resultados obtenidos nos llevan a preguntarnos si las tres regiones propuestas como promotores para Ago1 son realmente centros de promotores y si cada uno de ellos representa un sitio de inicio de la transcripción diferente, así como si la actividad de estos es independiente o se trata de una región promotora en la que las tres secuencias de los promotores propuestos son necesarias para lograr la expresión del mensajero. En *D. melanogaster* existen ejemplos de genes que contienen más de un promotor basal, en los que los motivos presentes en ambos promotores son necesarios para lograr la expresión del gen. Además, la expresión de esos genes podría estar mediada también por elementos reguladores distales, como los enhancers⁸². Los posibles promotores de Ago1 podrían contener motivos específicos necesarios para la expresión del gen, a través de la interacción con elementos reguladores distales o mediante cambios en la estructura de la cromatina.

Las isoformas de Ago1 se expresan de manera diferencial durante el desarrollo de *D. melanogaster*

Puesto que los resultados anteriores sugieren que el mRNA de la isoforma A no se encuentra presente en las células S2, que pertenecen al estadio de desarrollo embrionario de *D. melanogaster*, nos preguntamos si esta isoforma estaría presente en

alguna otra etapa del desarrollo de este organismo. Para responder esto, analizamos datos de CAGE obtenidos de muestras de las diferentes etapas del desarrollo de *D. melanogaster*.

El análisis de datos de CAGE indica que el promotor de la isoforma A sí se encuentra activo en ciertas etapas. De manera interesante, encontramos que esta isoforma se expresa sólo en los estadios de larva y pupa y no así en la línea germinal, en el estadio embrionario o en tejido adulto. Los promotores de las isoformas B, C y D se encuentran activos durante todo el desarrollo de *D. melanogaster*, y las isoformas presentan diferente abundancia, siendo la isoforma B la más representada en todas las muestras analizadas (Figura 17).

En cada tejido o en cada estadio del desarrollo de *D. melanogaster* podría haber diferentes factores que activen los promotores de las isoformas de manera diferencial, a pesar de que de ellos se obtienen sólo dos proteínas diferentes. Existen algunos factores endógenos que podrían repercutir en la activación de los promotores, como son, la actividad de factores de transcripción, el nivel de compactación de la cromatina y la estructura tridimensional del genoma, entre otros⁶⁸.

A pesar de que el producto de la traducción de los mRNAs de las isoformas A, C y D es el mismo, estos transcritos difieren entre sí en las secuencias 5' y 3' UTR, así como en el sitio de poliadenilación. Estas diferencias podrían afectar la estabilidad, tiempo de vida media y la abundancia de la proteína que se sintetiza a partir de ellos. Además, que la isoforma A se exprese sólo en etapas específicas del desarrollo de *D. melanogaster* podría indicar una especialización en la función de esta proteína.

A partir de los cuatro mRNAs diferentes que se sintetizan para Ago1 se traducen dos proteínas, que difieren entre sí sólo en 34 aminoácidos situados en el extremo N-terminal. Al realizar búsquedas en bases de datos encontramos que esta región no es similar a algún otro dominio específico presente en otras proteínas de *D. melanogaster*. Mas bien, podría ser clasificada como una región de baja complejidad. Estas regiones pueden tener diferentes implicaciones dentro de la función de una proteína dependiendo del sitio en el que se encuentren. Cuando se encuentran en los extremos de las proteínas, se ha propuesto que funcionen como sitios de interacción con nuevas proteínas⁸³. Por

lo tanto, la proteína que se sintetiza a partir de las isoformas A, C y D podría tener interacciones con proteínas específicas que no podrían interactuar con el AGO1 más corto, sintetizado a partir de la isoforma B. Además, esta región podría darle cierta especificidad a la proteína sobre algunos miRNAs, quizá específicos de los estadios del desarrollo en los que se expresa. Es posible también que en esta región se encuentren sitios de reconocimiento de enzimas que realicen cambios postraduccionales sobre la proteína larga de AGO1, afectando su estabilidad, especificidad o función.

El estudio de las proteínas larga y corta de AGO1 podría realizarse luego de la obtención de monoclonas de células en las que se haya sustituido alguno de los promotores del gen por un promotor inducible. Estas células permitirían detener la síntesis de los transcritos que originan a la proteína corta o a la proteína larga, provocando que en las células se exprese sólo una de ellas. De esta manera se podrían realizar ensayos con genes reporteros y con diferentes miRNAs sintéticos para conocer si los aminoácidos que las diferencian impactan en la actividad o especificidad de las proteínas. De igual manera, mediante inmunoprecipitación y espectrometría de masas, se podrían estudiar las posibles interacciones con otras proteínas durante procesos como el ensamblado de RISC o la actividad sobre mRNAs blanco.

Conclusiones

En *D. melanogaster* el gen Ago1 transcribe cuatro isoformas a partir de tres posibles promotores. En la línea celular S2 de este organismo se encuentran activos sólo dos de estos promotores.

En las células S2 de *D. melanogaster*, la delección de cualquiera de los tres posibles promotores de Ago1 provoca el decaimiento de la proteína. La disminución en la cantidad de esta proteína se observa también al deletar el promotor que no muestra actividad transcripcional en estas células. El decaimiento de AGO1 conlleva a la muerte celular.

Los posibles promotores de Ago1 presentan actividad específica en ciertas etapas del desarrollo de *D. melanogaster*. Los promotores de las isoformas B, C y D se encuentran activos durante todo el desarrollo de este organismo. Sin embargo, el promotor propuesto para la isoforma A se encuentra activo sólo en los estadios del desarrollo de larva y pupa.

Los resultados obtenidos durante esta investigación sugieren una regulación específica en la región promotora del gen Ago1, en la que posiblemente la presencia de los tres promotores identificados hasta el momento sea necesaria para lograr la transcripción del mRNA y con ello la síntesis de la proteína y la supervivencia de las células. Sin embargo, es necesario un análisis cuantitativo del transcrito de cada isoforma luego de la edición de cada uno de los promotores para tener certeza de la manera en la que cada uno de los promotores contribuye a la expresión de las isoformas.

Perspectivas

Para poder llegar a una conclusión sobre la regulación del inicio de la transcripción y los promotores de Ago1 en *D. melanogaster*, es necesaria la obtención de monoclonas de células en las que se haya realizado la edición de cada uno de los promotores de las isoformas de este gen. Puesto que AGO1 es una proteína necesaria para la supervivencia celular, proponemos el uso de un sistema inducible para lograr el control metabólico de la expresión de Ago1 y de esta manera obtener líneas celulares cuyos promotores hayan sido editados.

La obtención de monoclonas editadas permitiría también el análisis cuantitativo del mRNA de cada isoforma luego de la edición de los promotores. Además, la obtención de un anticuerpo específico para la proteína más larga permitiría estudiar si esta proteína tiene especificidad por ciertos miRNAs o si presenta interacciones con nuevas proteínas, quizá asociadas al desarrollo durante el estadio de larva o pupa. La implementación del sistema inducible nos permitiría además estudiar si la isoforma larga de AGO1 o la isoforma corta son suficientes para mantener la viabilidad celular.

Para la continuación del presente proyecto se podría sustituir, en un grupo de células, el promotor de la isoforma B por el promotor del sistema inducible y deletar el promotor de las isoformas C y D, y en otro grupo sustituir el promotor de las isoformas C y D y deletar el promotor de la isoforma B. Luego de la edición, obtener monoclonas de las células editadas y comprobar cuantitativamente la síntesis de cada isoforma, además de realizar el correspondiente análisis de cortes en sitios off-target. Posteriormente se continuaría con el estudio del complejo RISC luego de su interacción con el mRNA blanco.

Referencias

1. Watson, J. *et al.* *Molecular Biology of the Gene*. (2013).
2. Cech, T. R. & Steitz, J. A. The Noncoding RNA Revolution—Trashing Old Rules to Forge New Ones. *Cell* **157**, 77–94 (2014).
3. Ghildiyal, M. & Zamore, P. D. Small silencing RNAs: an expanding universe. *Nat. Rev. Genet.* **10**, 94–108 (2009).
4. Lee, Y. *et al.* MicroRNA genes are transcribed by RNA polymerase II. *EMBO J.* **23**, 4051–4060 (2004).
5. Lee, Y. *et al.* The nuclear RNase III Drosha initiates microRNA processing. *Nature* **425**, 415–419 (2003).
6. Denli, A. M., Tops, B. B. J., Plasterk, R. H. A., Ketting, R. F. & Hannon, G. J. Processing of primary microRNAs by the Microprocessor complex. *Nature* **432**, 231–235 (2004).
7. Landthaler, M., Yalcin, A. & Tuschl, T. The Human DiGeorge Syndrome Critical Region Gene 8 and Its D. melanogaster Homolog Are Required for miRNA Biogenesis. *Curr. Biol.* **14**, 2162–2167 (2004).
8. Han, J. *et al.* The Drosha-DGCR8 complex in primary microRNA processing. *Genes Dev.* **18**, 3016–3027 (2004).
9. Lee, Y., Jeon, K., Lee, J.-T., Kim, S. & Kim, V. N. MicroRNA maturation: stepwise processing and subcellular localization. *EMBO J.* **21**, 4663–70 (2002).
10. Ameres, S. L. & Zamore, P. D. Diversifying microRNA sequence and function. *Nat. Rev. Mol. Cell Biol.* **14**, 475–488 (2013).
11. Bohnsack, M., Czaplinski, K. & Görlich, D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. *RNA* **10**, 185–191 (2004).
12. Lund, E., Güttinger, S., Calado, A., Dahlberg, J. E. & Kutay, U. Nuclear Export of MicroRNA Precursors. *Science (80-.).* **303**, 95–98 (2004).
13. Bernstein, E., Caudy, A. A., Hammond, S. M. & Hannon, G. J. Role for a bidentate ribonuclease in the initiation step of RNA interference. *Nature* **409**, 363–366 (2001).
14. Chendrimada, T. P. *et al.* TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. *Nature* **436**, 740–744 (2005).
15. Förstemann, K. *et al.* Normal microRNA Maturation and Germ-Line Stem Cell Maintenance

- Requires Loquacious, a Double-Stranded RNA-Binding Domain Protein. *PLoS Biol.* **3**, e236 (2005).
16. Jonas, S. & Izaurralde, E. Towards a molecular understanding of microRNA-mediated gene silencing. *Nat. Rev. Genet.* **16**, 421–433 (2015).
 17. Iwasaki, S. *et al.* Hsc70/Hsp90 chaperone machinery mediates ATP-dependent RISC loading of small RNA duplexes. *Mol. Cell* **39**, 292–9 (2010).
 18. Matranga, C., Tomari, Y., Shin, C., Bartel, D. P. & Zamore, P. D. Passenger-Strand Cleavage Facilitates Assembly of siRNA into Ago2-Containing RNAi Enzyme Complexes. *Cell* **123**, 607–620 (2005).
 19. Liu, Q. *et al.* R2D2, a Bridge Between the Initiation and Effector Steps of the Drosophila RNAi Pathway. *Science (80-.).* **301**, 1921–1925 (2003).
 20. Hutvagner, G. & Simard, M. J. Argonaute proteins: key players in RNA silencing. *Nat. Rev. Mol. Cell Biol.* **9**, 22–32 (2008).
 21. Höck, J. & Meister, G. The Argonaute protein family. *Genome Biol.* **9**, 210 (2008).
 22. Schirle, N. T. & MacRae, I. J. The crystal structure of human Argonaute2. *Science* **336**, 1037–40 (2012).
 23. Wang, Y., Sheng, G., Juranek, S., Tuschl, T. & Patel, D. J. Structure of the guide-strand-containing argonaute silencing complex. *Nature* **456**, 209–213 (2008).
 24. Bartel, D. P. MicroRNAs: Target Recognition and Regulatory Functions. *Cell* **136**, 215–233 (2009).
 25. Nakanishi, K. Anatomy of RISC: how do small RNAs and chaperones activate Argonaute proteins? *Wiley Interdiscip. Rev. RNA* **7**, 637–660 (2016).
 26. Guo, H., Ingolia, N. T., Weissman, J. S. & Bartel, D. P. Mammalian microRNAs predominantly act to decrease target mRNA levels. *Nature* **466**, 835–840 (2010).
 27. Fukaya, T., Iwakawa, H. & Tomari, Y. MicroRNAs Block Assembly of eIF4F Translation Initiation Complex in Drosophila. *Mol. Cell* **56**, 67–78 (2014).
 28. Meijer, H. A. *et al.* Translational Repression and eIF4A2 Activity Are Critical for MicroRNA-Mediated Gene Regulation. *Science (80-.).* **340**, 82–85 (2013).
 29. Liu, J. *et al.* A role for the P-body component GW182 in microRNA function. *Nat. Cell Biol.* **7**, 1261–1266 (2005).

30. Meister, G. *et al.* Identification of Novel Argonaute-Associated Proteins. *Curr. Biol.* **15**, 2149–2155 (2005).
31. Behm-Ansmant, I. *et al.* mRNA degradation by miRNAs and GW182 requires both CCR4:NOT deadenylase and DCP1:DCP2 decapping complexes. *Genes Dev.* **20**, 1885–98 (2006).
32. Braun, J. E., Huntzinger, E., Fauser, M. & Izaurralde, E. GW182 Proteins Directly Recruit Cytoplasmic Deadenylase Complexes to miRNA Targets. *Mol. Cell* **44**, 120–133 (2011).
33. Chen, C.-Y. A., Zheng, D., Xia, Z. & Shyu, A.-B. Ago–TNRC6 triggers microRNA-mediated decay by promoting two deadenylation steps. *Nat. Struct. Mol. Biol.* **16**, 1160–1166 (2009).
34. Braun, J. E. *et al.* A direct interaction between DCP1 and XRN1 couples mRNA decapping to 5' exonucleolytic degradation. *Nat. Struct. Mol. Biol.* **19**, 1324–1331 (2012).
35. Rügger, S. & Großhans, H. MicroRNA turnover: when, how, and why. *Trends Biochem. Sci.* **37**, 436–446 (2012).
36. Guo, Y. *et al.* Characterization of the mammalian miRNA turnover landscape. *Nucleic Acids Res.* **43**, 2326–2341 (2015).
37. Rissland, O. S., Hong, S.-J. & Bartel, D. P. MicroRNA Destabilization Enables Dynamic Regulation of the miR-16 Family in Response to Cell-Cycle Changes. *Mol. Cell* **43**, 993–1004 (2011).
38. Monticelli, S. *et al.* MicroRNA profiling of the murine hematopoietic system. *Genome Biol.* **6**, R71 (2005).
39. Krol, J. *et al.* Characterizing Light-Regulated Retinal MicroRNAs Reveals Rapid Turnover as a Common Property of Neuronal MicroRNAs. *Cell* **141**, 618–631 (2010).
40. Ameres, S. L. *et al.* Target RNA-Directed Trimming and Tailing of Small Silencing RNAs. *Science (80-.).* **328**, 1534–1539 (2010).
41. De, N. *et al.* Highly complementary target RNAs promote release of guide RNAs from human Argonaute2. *Mol. Cell* **50**, 344–55 (2013).
42. Park, J. H., Shin, S.-Y. & Shin, C. Non-canonical targets destabilize microRNAs in human Argonautes. *Nucleic Acids Res.* **45**, gkx029 (2017).
43. Horman, S. R. *et al.* Akt-mediated phosphorylation of argonaute 2 downregulates cleavage and upregulates translational repression of MicroRNA targets. *Mol. Cell* **50**, 356–67 (2013).
44. Bridge, K. S. *et al.* Argonaute Utilization for miRNA Silencing Is Determined by

- Phosphorylation-Dependent Recruitment of LIM-Domain-Containing Proteins. *Cell Rep.* **20**, 173–187 (2017).
45. Rü Del, S. *et al.* Phosphorylation of human Argonaute proteins affects small RNA binding. doi:10.1093/nar/gkq1032
 46. Shen, J. *et al.* EGFR modulates microRNA maturation in response to hypoxia through phosphorylation of AGO2. *Nature* **497**, 383–387 (2013).
 47. Golden, R. J. *et al.* An Argonaute phosphorylation cycle promotes microRNA-mediated silencing. *Nature* **542**, 197–202 (2017).
 48. Tomari, Y., Du, T. & Zamore, P. D. Sorting of *Drosophila* Small Silencing RNAs. *Cell* **130**, 299–308 (2007).
 49. Lee, Y. S. *et al.* Distinct roles for *Drosophila* Dicer-1 and Dicer-2 in the siRNA/miRNA silencing pathways. *Cell* **117**, 69–81 (2004).
 50. Förstemann, K., Horwich, M. D., Wee, L., Tomari, Y. & Zamore, P. D. *Drosophila* microRNAs Are Sorted into Functionally Distinct Argonaute Complexes after Production by Dicer-1. *Cell* **130**, 287–297 (2007).
 51. Ha, M. & Narry Kim, V. Regulation of microRNA biogenesis. *Nat. Rev. Mol. Cell Biol.* 509–524 (2014). doi:10.1038/nrm3838
 52. Jonas, S. & Izaurralde, E. Towards a molecular understanding of microRNA-mediated gene silencing. *Nat. Rev. Genet.* **16**, 421–433 (2015).
 53. Lopez-Orozco, J. *et al.* Functional analyses of phosphorylation events in human Argonaute 2. *RNA* **21**, 2030–2038 (2015).
 54. Quévillon Huberdeau, M. *et al.* Phosphorylation of Argonaute proteins affects mRNA binding and is essential for microRNA-guided gene silencing *in vivo*. *EMBO J.* **36**, 2088–2106 (2017).
 55. Liu, J. *et al.* Argonaute2 is the catalytic engine of mammalian RNAi. *Science* **305**, 1437–41 (2004).
 56. Sapranaukas, R. *et al.* The *Streptococcus thermophilus* CRISPR/Cas system provides immunity in *Escherichia coli*. *Nucleic Acids Res.* **39**, 9275–82 (2011).
 57. Jinek, M. *et al.* A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science (80-.).* **337**, 816–821 (2012).
 58. Deltcheva, E. *et al.* CRISPR RNA maturation by trans-encoded small RNA and host factor

- RNase III. *Nature* **471**, 602–607 (2011).
59. Jiang, F. & Doudna, J. A. CRISPR-Cas9 Structures and Mechanisms. (2017). doi:10.1146/annurev-biophys
 60. Lee, R. C., Feinbaum, R. L. & Ambros, V. The *C. elegans* heterochronic gene *lin-4* encodes small RNAs with antisense complementarity to *lin-14*. *Cell* **75**, 843–54 (1993).
 61. Kawamata, T., Seitz, H. & Tomari, Y. Structural determinants of miRNAs for RISC loading and slicer-independent unwinding. *Nat. Struct. Mol. Biol.* **16**, 953–960 (2009).
 62. Kobayashi, H., Shoji, K., Kiyokawa, K., Negishi, L. & Tomari, Y. Iruka Eliminates Dysfunctional Argonaute by Selective Ubiquitination of Its Empty State. *Mol. Cell* **73**, 119-129.e5 (2019).
 63. Dreos, R., Ambrosini, G., Cavin Périer, R. & Bucher, P. EPD and EPDnew, high-quality promoter resources in the next-generation sequencing era. *Nucleic Acids Res.* **41**, D157–D164 (2013).
 64. Ohler, U. Identification of core promoter modules in *Drosophila* and their application in accurate transcription start site prediction. doi:10.1093/nar/gkl608
 65. Haeussler, M. *et al.* Evaluation of off-target and on-target scoring algorithms and integration into the guide RNA selection tool CRISPOR. *Genome Biol.* **17**, 148 (2016).
 66. Doench, J. G. *et al.* Optimized sgRNA design to maximize activity and minimize off-target effects of CRISPR-Cas9. *Nat. Biotechnol.* **34**, 184 (2016).
 67. Gasiunas, G., Barrangou, R., Horvath, P. & Siksnys, V. Cas9-crRNA ribonucleoprotein complex mediates specific DNA cleavage for adaptive immunity in bacteria. *Proc. Natl. Acad. Sci. U. S. A.* **109**, E2579-86 (2012).
 68. Lenhard, B., Sandelin, A. & Carninci, P. Transcription start sites Metazoan promoters: emerging characteristics and insights into transcriptional regulation. *Nat. Rev. / Genet.* **13**, 233 (2012).
 69. Gan, Q. *et al.* Monovalent and unpoised status of most genes in undifferentiated cell-enriched *Drosophila* testis. *Genome Biol.* **11**, R42 (2010).
 70. Mueller, B. *et al.* Widespread changes in nucleosome accessibility without changes in nucleosome occupancy during a rapid transcriptional induction. *Genes Dev.* **31**, 451–462 (2017).
 71. Rennie, S. *et al.* Transcription start site analysis reveals widespread divergent transcription in *D. melanogaster* and core promoter-encoded enhancer activities. *Nucleic Acids Res.* **46**,

- 5455–5469 (2018).
72. Shiraki, T. *et al.* Cap analysis gene expression for high-throughput analysis of transcriptional starting point and identification of promoter usage. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 15776–81 (2003).
 73. Schuettengruber, B. *et al.* Functional Anatomy of Polycomb and Trithorax Chromatin Landscapes in *Drosophila* Embryos. *PLoS Biol.* **7**, e1000013 (2009).
 74. Gan, Q. *et al.* Monovalent and unpoised status of most genes in undifferentiated cell-enriched *Drosophila* testis. *Genome Biol.* **11**, R42 (2010).
 75. Chen, Z.-X. *et al.* Comparative validation of the *D. melanogaster* modENCODE transcriptome annotation. *Genome Res.* **24**, 1209–23 (2014).
 76. Marygold, S. J. *et al.* The ribosomal protein genes and Minute loci of *Drosophila melanogaster*. *Genome Biol.* **8**, R216 (2007).
 77. Kosicki, M., Tomberg, K. & Bradley, A. Repair of double-strand breaks induced by CRISPR–Cas9 leads to large deletions and complex rearrangements. *Nat. Biotechnol.* **36**, 765–771 (2018).
 78. Lee, H. *et al.* DNA copy number evolution in *Drosophila* cell lines. *Genome Biol.* **15**, R70 (2014).
 79. Payne, J. L. & Wagner, A. The causes of evolvability and their evolution. doi:10.1038/s41576-018-0069-z
 80. Kadonaga, J. T. Perspectives on the RNA polymerase II core promoter. *Wiley Interdiscip. Rev. Dev. Biol.* **1**, 40–51 (2012).
 81. Zabidi, M. A. & Stark, A. Regulatory Enhancer–Core–Promoter Communication via Transcription Factors and Cofactors. *Trends Genet.* **32**, 801–814 (2016).
 82. Ling, J., Umezawa, K. Y., Scott, T. & Small, S. Bicoid-Dependent Activation of the Target Gene hunchback Requires a Two-Motif Sequence Code in a Specific Basal Promoter. *Mol. Cell* (2019). doi:10.1016/j.molcel.2019.06.038
 83. Coletta, A. *et al.* Low-complexity regions within protein sequences have position-dependent roles. *BMC Syst. Biol.* **4**, 43 (2010).