



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Maestría y Doctorado en Ciencias Bioquímicas

ANÁLISIS DE UN mRNA BICISTRÓNICO EN *Drosophila melanogaster*

TESIS

QUE PARA OPTAR POR EL GRADO DE

Doctor en Ciencias

PRESENTA:

M. en C. ALYERI BUCIO MÉNDEZ

TUTOR PRINCIPAL

DR. MARIO ENRIQUE ZURITA ORTEGA

[Instituto de Biotecnología, UNAM](#)

COTUTORES

DRA. HILDA MARÍA LOMELÍ BUYOLI

[Instituto de Biotecnología, UNAM](#)

DR. RAMÓN GONZÁLEZ GARCÍA CONDE

[Centro de Investigación en Dinámica Celular, UAEM](#)

Cuernavaca, Morelos. Noviembre, 2018



Universidad Nacional
Autónoma de México



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.

"If people did not sometimes do silly things,
nothing intelligent would ever get done"

Ludwig Wittgenstein

Este trabajo se presenta como tesis doctoral dentro del programa de Maestría y Doctorado en Ciencias Bioquímicas de la Universidad Nacional Autónoma de México. Se llevó a cabo en el Instituto de Biotecnología de la UNAM, en el Departamento de Genética del Desarrollo y Fisiología Molecular bajo la dirección del Dr. Mario Enrique Zurita Ortega.

Agradecimientos

Gracias al laboratorio del Dr. Mario Zurita, incluyendo a todos los integrantes actuales y pasados. En general les debo mi formación profesional, aprendí tanto sobre el tipo de investigadora que quiero ser de todos ustedes. En particular a Zoraya, por avivar la llamita desde el principio. A Sara, y Ximena por ser cómplices en que es necesario relajarse de vez en cuando. A Maritere y Grisel por siempre apoyarme, enseñarme y tenerme paciencia. A Viviana y Martha por todas sus enseñanzas y sabiduría. A Mario por aguantarme tanto tiempo.

A la Dra. Tzvetanka Dimitrova por su ayuda en el proyecto.

Al Dr. Ramón González y la Dra. Hilda Lomelí por sus consejos a lo largo de todo el desarrollo del proyecto.

A los Dres. Enrique Merino, David Romero, Víctor Bustamente, Ignacio López, y la Dra. Denhi Schnabel por sus aportaciones para mejorar este escrito y consejos en mi presentación oral del trabajo.

A mi familia por estar ahí en todas las complicaciones que esta carrera implica. A Anyela por aguantar las explicaciones no solicitadas. A Rubén, Elena, y Nina por siempre estar al pendiente de mí en este largo proceso.

A Rodrigo, que aunque entra en la categoría de gente de conocí en el instituto, recibe mención honorífica por aceptarme tal cual soy y ayudarme a crecer en el ámbito personal y profesional.

A toda la gente que conocí en el instituto de biotecnología, puedo decir con certeza que muchos serán amigos de por vida. A Delia, Guille, Adriana, y demás personas que me enseñaron que hay vida más allá de *Drosophila*. Y a todos mis amigos fuera de la academia, que han sido igual de importantes en este camino.

A CONACyT, y a la UNAM por el apoyo y financiamiento para llevar a cabo este proyecto. Al Programa de Apoyo a los Estudio de Posgrado (PAEP) por permitirme asistir a un congreso internacional que en definitiva hizo crecer mi desarrollo profesional.

Contenido

| | |
|---|----|
| Resumen..... | 1 |
| Abstract..... | 2 |
| Índice de Figuras | 3 |
| Abreviaturas usadas..... | 4 |
| I. Introducción..... | 5 |
| 1.1 Transcripción en organismos eucariotes..... | 5 |
| 1.2 Factor de transcripción TFIIH | 7 |
| 1.3 Subunidad Dmp8..... | 8 |
| 1.4 Dmp18..... | 9 |
| 1.5 Traducción en eucariotes..... | 9 |
| 1.6 Policistrones en eucariotes..... | 12 |
| 1.7 Mecanismos de traducción no canónicos..... | 12 |
| II. Antecedentes | 15 |
| 2.1 <i>Drosophila melanogaster</i> como modelo de estudio | 15 |
| 2.2 Bicistrones en la mosca <i>Drosophila melanogaster</i> | 16 |
| 2.3 Bicistrón Dmp8-Dmp18 en la mosca <i>Drosophila melanogaster</i> | 16 |
| III. Justificación..... | 20 |
| IV. Objetivo..... | 20 |
| 3.1 Objetivo general | 20 |
| 3.2 Objetivos particulares | 20 |
| V. Materiales y Métodos..... | 21 |
| 5.1 Clonación del mensajero bicistrónico..... | 21 |
| 5.2 Inserción de poliA a la construcción | 22 |
| 5.3 Generación de construcciones a partir del mensajero bicistrónico..... | 22 |
| 5.4 Ensayos de Transcripción/Traducción <i>in vitro</i> | 23 |
| 5.5 Generación de vectores para transfectar células..... | 24 |

| | |
|---|----|
| 5.6 Cultivo celular..... | 26 |
| 5.7 Transfección para expresión transitoria de luciferasas en células S2R+ | 26 |
| 5.8 Medición intensidad de luciferasa en células transfectadas..... | 27 |
| 5.9 Detección de proteína en larvas de tercer instar | 27 |
| 5.10 Extracción de RNA de larvas para ensayo RT-PCR..... | 28 |
| 5.11 Extracción de proteína de células S2R+ | 28 |
| VI. Resultados | 30 |
| 6.1 Alineamiento de las regiones intercistrónicas de Artrópodos | 30 |
| 6.2 Análisis <i>in silico</i> de regiones promotoras del bicistrón Dmp8-Dmp18..... | 31 |
| 6.3 Ensayos de transcripción/traducción <i>in vitro</i> del mRNA de Dmp8-Dmp18..... | 31 |
| 6.4 Análisis de traducción <i>in vivo</i> del mRNA del bicistrón Dmp8-Dmp18..... | 36 |
| 6.5 Detección de Dmp8 en ensayos de transfección | 38 |
| 6.6 Detección de las proteínas Dmp8 y Dmp18 en larvas..... | 39 |
| 6.7 Análisis del transcrito en situaciones de estrés | 41 |
| 6.8 Análisis de traducción después de estrés celular | 42 |
| 6.9 Análisis de bicistrones en <i>Drosophila melanogaster</i> | 45 |
| 6.10 Búsqueda <i>in silico</i> de candidatos con una regulación similar..... | 46 |
| VII. Discusión..... | 48 |
| VIII. Conclusiones | 50 |
| VIX. Perspectivas..... | 51 |
| X. Bibliografía | 52 |
| Anexo 1. Tabla de policistrones reportados en <i>Drosophila melanogaster</i> | 56 |
| Anexo 2. Artículos publicados durante el doctorado..... | 64 |

Resumen

Es importante que las células puedan llevar a cabo una correcta expresión de sus genes en respuesta a los diferentes estímulos a los que está sometido un organismo. Los mecanismos mediante los cuales logra esto están finamente regulados para permitir la correcta expresión de genes y de proteínas de manera oportuna, especialmente cuando percibe estrés del ambiente.

Muchos mecanismos están a cargo de mantener la transcripción activa para que exista el templado y la maquinaria de traducción pueda llevar a cabo la producción de proteínas. Existen alternativas al mecanismo canónico de traducción dependiente de cap, que son usadas por virus y células, y que les permiten sintetizar proteínas en momentos de infecciones virales y estrés.

En *Drosophila* existen cientos de RNA mensajeros (mRNA) policistrónicos, y se conocen varios mecanismos que permiten la correcta traducción de ellos. Las proteínas que se encuentran en un mismo mRNA generalmente están involucradas en el mismo mecanismo celular. Pero este no es el caso de la proteína Dmp18, homóloga de Swc6/p18Hamlet (parte del complejo de remodelación de la cromatina SWR1/SRCAP) que está codificada en el mismo mRNA bicistrónico que la proteína Dmp8 (proteína estructural, parte del factor de transcripción TFIIH) dando lugar al bicistrón Dmp8-Dmp18.

El arreglo bicistrónico de estos dos genes se encuentra conservado en todos los artrópodos, y a pesar de que las regiones que codifican a ambas proteínas están muy bien conservadas, la región intercistrónica varía entre dos a noventa y dos bases, lo que sugiere un mecanismo inusual de traducción del segundo marco de lectura del mensajero.

En este trabajo encontramos que aunque la traducción de Dmp18 ocurre de manera independiente de aquella de Dmp8, es necesario que Dmp18 se encuentre en una conformación cercana a Dmp8, ya que los últimos veintisiete nucleótidos del marco de lectura de Dmp8 son necesarios para permitir la traducción de Dmp18 en respuesta a estrés celular ocasionado por choque térmico y radiación ultravioleta.

Abstract

It is important that cells are able to attain a correct gene expression in response to the different stimuli an organism is subjected. The mechanisms to achieve this are finely tuned in order to have a correct gene and protein expression in a timely manner, especially when subjected to environmental stress.

A lot of mechanisms are operating to maintain an active transcription state so that the translation machinery can produce proteins. Alternatives to the cap mechanism in translation are often used by viruses and cells to allow them to synthesize proteins in events of stress and viral infection.

In *Drosophila* there are hundreds of polycistronic messenger RNA (mRNA), and various mechanisms are known to achieve their translation. Proteins in a same mRNA often work in similar cellular mechanism, this is not the case for *Drosophila's* Swc6/p18Hamlet homolog Dmp18 (part of the SWR1/SRCAP chromatin remodeling complex) who is encoded in a bicistronic mRNA next to Dmp8 (a structural component of transcription factor TFIID) to form Dmp8-Dmp18 bicistronic mRNA.

The organization of these two genes as a bicistron is conserved in all arthropods, and even though the protein coding regions are fairly conserved, the length of the intercistronic sequence varies from two to more than ninety bases, suggesting an unusual translation mechanism for the second open reading frame.

In this work, we found that albeit translation of Dmp18 occurs independently from that of Dmp8, it is necessary for Dmp18 to be in a conformation near Dmp8, since the last 27 nucleotides form its open reading frame are the ones that allow its correct translation during cellular stress caused by damage via heat-shock and UV radiation.

Índice de Figuras

| | |
|---|-----------|
| Figura 1. El complejo de pre-iniciación de la transcripción..... | 6 |
| Figura 2. Esquema de las subunidades del factor TFIID..... | 8 |
| Figura 3. Inicio de traducción en eucariotes..... | 11 |
| Figura 4. Inicio de traducción mediado por Sitios Internos de Entrada del Ribosoma..... | 13 |
| Figura 5. Desarrollo de <i>Drosophila melanogaster</i> | 15 |
| Figura 6. Bicistrón Dmp8-Dmp18 en artrópodos..... | 17 |
| Figura 7. mRNA bicistrónico a lo largo del desarrollo..... | 18 |
| Figura 8. Presencia de las proteínas del bicistrón en todo el desarrollo de la mosca..... | 19 |
| Figura 9. Vector pBKSII..... | 22 |
| Figura 10. Oligos utilizados..... | 23 |
| Figura 11. Vector pAc5.1/V5-HisA..... | 24 |
| <i>Figura 12 Vector resultante para expresión de la Luciferasa de luciérnaga.....</i> | <i>25</i> |
| Figura 13. Vector resultante para expresión de la Luciferasa de Renilla..... | 26 |
| Figura 14. Conservación de regiones intercistrónicas en Artrópodos..... | 30 |
| Figura 15. Región regulatoria del bicistrón Dmp8-Dmp18..... | 31 |
| Figura 16. Ensayo de transcripción/traducción in vitro del mRNA Dmp8-Dmp18..... | 32 |
| Figura 17. Adición de poliA al mRNA de Dmp8-Dmp18..... | 33 |
| Figura 18. Cinética de aparición de ambas proteínas del bicistrón..... | 34 |
| Figura 19. Detección de las proteínas del bicistrón modificando al mensajero..... | 35 |
| Figura 20. Eficiencia de traducción de un reportero de luciferasa..... | 36 |
| Figura 21. Detección de Dmp8 en células transfectadas..... | 39 |
| Figura 22. Detección de las proteínas del bicistrón en situaciones de estrés..... | 41 |
| Figura 23. Análisis de transcripción del mensajero bicistrónico..... | 42 |
| Figura 24. Cinéticas de luminiscencia en células sometidas a estrés..... | 43 |
| Figura 25. Análisis de traducción en regiones acotadas río arriba del sitio de inicio de traducción de Dmp18..... | 44 |
| Figura 26. Cinética de respuesta a estrés por luz UV o choque térmico en regiones acotadas..... | 44 |
| Figura 27. Cinética de traducción con región intercistrónica de <i>P. humanus</i> | 45 |
| Figura 28. Posibles funciones de los segundos cistrones..... | 46 |
| Figura 29. Alineamiento de la región intercistrónica del bicistrón CG42374-CG9666..... | 47 |

Abreviaturas usadas

aa: aminoácido

cDNA: DNA complementario

DNA: Ácido Desoxirribonucleico, por sus siglas en inglés

dNTPs: deoxinucleótidos trifosfato

eIF: Factor de Iniciación eucariote, por sus siglas en inglés

FLuc: Luciferasa de Luciérnaga

IRES: Sitio Interno de Entrada de Ribosoma, por sus siglas en inglés

mRNA: RNA mensajero

nt: nucleótido

ORF: Marco de Lectura Abierto, por sus siglas en inglés

PIC: Complejo de Pre-Iniciación, por sus siglas en inglés

REN: Reparación por Escisión de Nucleótidos

RLuc: Luciferasa de Renilla

RNA: Ácido Ribonucleico, por sus siglas en inglés

RNAPII: RNA Polimerasa II

rRNA: RNA ribosomal

TBP: Proteína de Unión a caja TATA, por sus siglas en inglés

TF: Factor de Transcripción, por sus siglas en inglés

TTD: Tricotiodistrofia

UV: UltraVioleta

I. Introducción

Los mecanismos involucrados en que los organismos logren mantener una correcta homeostasis son importantes para permitir una adecuada transcripción de genes una vez que estos son activados, y son muchos. Esto quiere decir que es necesario que la célula sea capaz de percibir estímulos del ambiente y responder a estos de manera adecuada¹.

Cuando la célula percibe algún estímulo, se inicia un proceso en el cual se lee la información que se encuentra en el DNA y es capaz de producir a partir de este una molécula de RNA. Este mecanismo se denomina Transcripción, y es el primer paso en la formación de proteínas a partir del material genético. De esta manera se pueden formar proteínas con papeles estructurales y catalíticos que le permitirán a la célula desempeñar diversas funciones en las diferentes condiciones en las que se encuentra².

La transcripción es un proceso muy complejo. En organismos eucariotes involucra diferentes factores, así como el templado de cromatina (complejo conformado por proteínas, DNA y RNA). La regulación de la transcripción es finamente controlada por los factores de transcripción y otras proteínas accesorias dependiendo de la etapa de desarrollo, tejido, o condición de un organismo. En esta regulación es importante coordinar desde la estructura de la cromatina hasta los componentes de la maquinaria de transcripción, para que se puedan llevar a cabo o ejecutar de manera adecuada los mecanismos que darán lugar a la de síntesis de proteínas a partir del mensajero resultante³.

1.1 Transcripción en organismos eucariotes

Los primeros esfuerzos en entender el mecanismo de transcripción en organismos eucariotes se dieron gracias a la purificación e identificación de las RNA polimerasas. A partir de esto se logró identificar diversos factores esenciales o basales de transcripción *in vitro*. A pesar de que son necesarios más elementos para lograr una transcripción eficiente, estos factores de transcripción son lo único que se requiere para iniciarla³.

Una vez que se da el estímulo inicial que indica a las células que se debe iniciar la transcripción de ciertos genes, se reúnen un grupo de proteínas que se necesitan en el sitio de inicio de la transcripción, incluida la polimerasa II de RNA (RNAPII), diversos factores de transcripción, cofactores y agentes remodeladores de la cromatina. Para que sea posible una correcta transcripción es necesario que se forme un complejo de pre-iniciación (PIC por sus siglas en inglés) en la región promotora de los genes que se van a transcribir⁴.

Para la formación del PIC de transcripción, el primero en llegar al promotor, y que se encarga de reconocer los sitios de inicio de la transcripción es el factor de transcripción (TF por sus siglas en inglés) IID (TFIID), cuya proteína más importante es la proteína de unión a la caja TATA (TBP por sus siglas en inglés). Posteriormente se anexan secuencialmente los factores de transcripción TFIIA, TFIIIB que le dan estabilidad al complejo, y así permitir que se recluten RNAPII y TFIIIF. Es así como se inicia el reclutamiento de TFIIIE que es el encargado a su vez de reclutar al último complejo, TFIIH⁵ (Figura 1).

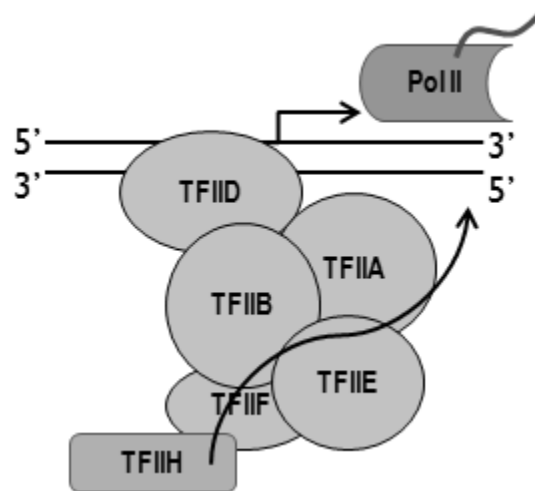


Figura 1. El complejo de pre-iniciación de la transcripción.

Se ensambla en la región promotora de genes que se van a transcribir, río arriba del sitio de inicio de transcripción. Se conforma por los factores TFII-A, B, D, E, F, H, y la RNA polimerasa II. Es necesario que exista un correcto ensamblaje de todos los factores para que la polimerasa inicie la transcripción.

Se sabe que el factor TFIIH es uno de los factores más importantes para comenzar la transcripción, ya que esta no inicia hasta que el factor TFIIH es reclutado a pesar de que los demás factores ya se encuentren posicionados en la región de inicio de transcripción².

1.2 Factor de transcripción TFIIH

TFIIH es un complejo multiprotéico conformado por 10 proteínas: XPD, XPB, p62, p52, p44, p34, Cdk7, la ciclina H, MAT-1, y p8 (Figura 2). Además de estar involucrado en la transcripción de genes, tiene un papel importante en la reparación del daño al DNA ocasionado por radiación UV (ultravioleta), mediante el mecanismo de Reparación por Escisión de Nucleótidos (REN). Esta reparación se lleva a cabo mediante la formación de un sub-complejo que abarca únicamente 7 de las subunidades de TFIIH. Se dice que es el factor limitante de la transcripción, ya que es el último complejo en unirse y es el factor que permite que la polimerasa ya posicionada lleve a cabo la elongación^{4,6}.

Una de las razones por las cuales TFIIH ayuda a iniciar la transcripción es gracias a que las proteínas XPB y XPD tienen una actividad de Helicasa-ATPasa. En el caso de XPB, esta actividad permite la translocación de una de las cadenas del DNA para permitir posicionar a la RNAPII en el sitio de inicio de la transcripción y así dar inicio a la transcripción⁷.

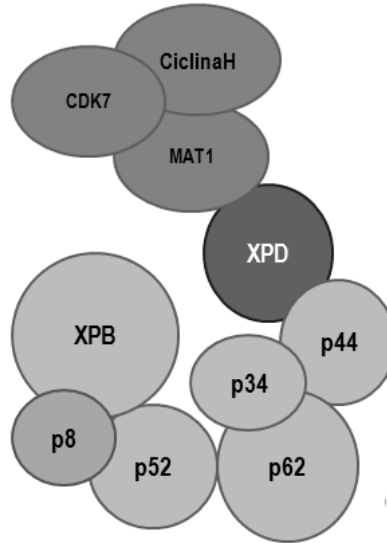


Figura 2. Esquema de las subunidades del factor TFIIH.

Las proteínas XPD y XPB tienen una actividad helicasa ATP-asa que se encarga de la apertura de las cadenas de ADN para la subsecuente entrada de la maquinaria de transcripción, y así permitir la transcripción de genes eucariotas. Cdk7 tiene actividad de cinasa, mientras que las demás subunidades sólo parecen tener una función estructural.

1.3 Subunidad Dmp8

Todas las proteínas del complejo TFIIH están conservadas en todos los organismos eucariotes, una subunidad del complejo en *Drosophila melanogaster* es Dmp8 (homóloga de p8), la cual no tiene actividad catalítica aparente, pero cuando se encuentra en heterodímero con Dmp52 se encarga de regular la actividad ATPasa de la helicasa 3'→5' XPB. Se identificó a la proteína en humano p8 como la responsable de una variante de la enfermedad Tricotodistrofia (TTD), nombrada TTD-A, un síndrome humano^{7,8}. La TTD es una enfermedad rara, y se ha visto que la versión de TTDA causada por mutaciones en p8 es una forma leve de la enfermedad, caracterizada por foto-sensibilidad, uñas y cabello quebradizo, y retraso de aprendizaje⁹.

Existen diversos reportes sobre los efectos de las mutaciones en p8/Dmp8, en algunos casos se ha demostrado que la falta de ella afecta más la reparación dependiente de NER que la transcripción⁸; o que sin esta proteína los defectos de transcripción, reparación, y viabilidad son tan fuertes que causan la letalidad del

embrión⁹; y finalmente otros reportes indican que no es esencial para que los niveles de transcripción se mantengan estables y por lo tanto los organismos llegan hasta adultos⁷.

Una particularidad del gen que codifica a Dmp8 en *D. melanogaster* es que se encuentra en un arreglo bicistrónico con la proteína Dmp18.

1.4 Dmp18

La proteína Swc6/p18^{Hamlet} en levadura y humanos es una subunidad del complejo de remodelación de la cromatina SWR1/SRCAP. Este complejo se encarga del intercambio de la histona canónica H2A por la variante H2AZ¹⁰. El intercambio de esta histona se ha visto cuando se requiere de una cromatina permisiva, la cual permite la entrada de la maquinaria de transcripción, así como en situaciones de estrés celular¹¹.

En estudios en humanos, se ha probado que hay acumulaciones de p18Hamlet en respuesta a daño causado por radiación UV, y esta acumulación puede llevar a la célula a apoptosis gracias a que recluta a la proteína p53 mediando su interacción con la proteína p38¹². Además, defectos en ella provocan que existan fallas en la regulación del ciclo celular y el desarrollo del organismo¹³. Sin embargo, la proteína Dmp18 no ha sido caracterizada aún en *Drosophila melanogaster*.

1.5 Traducción en eucariotes

Como se mencionó con anterioridad, en general, la transcripción en general culmina en la formación de un RNA mensajero (mRNA) formado por tripletes de nucleótidos (nt) denominados codones que van a codificar diferentes aminoácidos (aa). Estos aa son los bloques formadores de las proteínas².

El mRNA necesita estar maduro para que se lleve a cabo la traducción, esto se logra mediante la adición de un Cap en el extremo 5' del mensajero por parte del Complejo Enzimático de Capping, y mediante la interacción con la proteína de unión a poliA (PABP), que como lo indica su nombre, se ancla a la cola de poliA de los transcritos nacientes².

Los ribosomas son los organelos principales encargados de la traducción, y consisten en un complejo de proteínas, RNA especializado llamado RNA ribosomal (rRNA), y proteínas accesorias. Los ribosomas empiezan su ensamblaje en el nucléolo, resultando en un ensamblaje parcial que culmina en el citoplasma. En los organismos eucariotes existen una subunidad pequeña 40S, y una subunidad grande 60S².

En los primeros pasos de la traducción en organismos eucariotes es cuando se lleva a cabo la formación del PIC de traducción, denominado 43S el cual está formado por la subunidad 40S y los complejos de inicio de traducción (eIF por sus siglas en inglés) 1, 2, 3, y 5. Por otro lado, el mRNA maduro va a ser reconocido por los complejos eIF4 (conformado por eIF4E, eIF4G, y la helicasa eIF4A) y sus cinasas asociadas MNK. El anclaje de este grupo de proteínas al extremo 5' del mRNA permite desenrollar sus estructuras terciarias en él y de esta manera estabilizarlo^{2,14}.

Cuando esta subunidad 43S se encuentran bien posicionada en el extremo 5' del mRNA se inicia un escaneo de la región no codificante hasta que la maquinaria encuentre el primer sitio de síntesis de proteínas (generalmente dado por el codón AUG en conjunto con la secuencia Kozak¹⁵) lo que permite que la subunidad eIF2 hidrolice su GTP para promover su remoción parcial del complejo, formando así el complejo de iniciación^{2,14}.

Este complejo es ya capaz de reclutar a la unidad ribosomal 60S, al tRNA cargado de la metionina, y al factor eIF5B acoplado a GTP que termina de desplazar a eIF2, así como a eIF1, 3 y 5. Finalmente se hidroliza el GTP unido a eIF5B, lo que lleva a su liberación del complejo, así como de eIF1A. De esta manera se termina de armar al complejo de traducción, y permite que los ribosomas se encuentren ensamblados en el mensajero y la formación del complejo 80S del cual depende para continuar con la elongación de la traducción, es decir, iniciar la adición de aminoácidos^{2,9} (Figura 3).

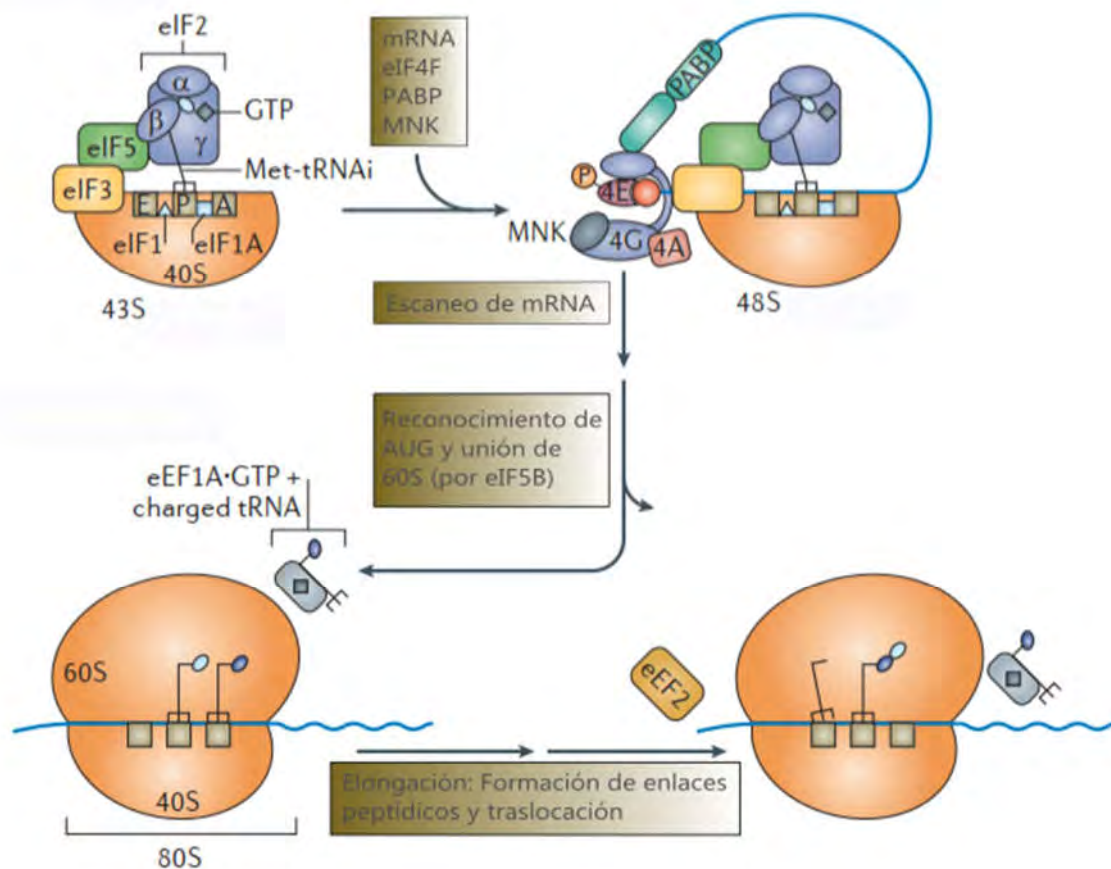


Figura 3. Inicio de traducción en eucariotes.

La traducción en eucariotes inicia con la formación del complejo ribosomal 43S, conformado por la subunidad 40S del ribosoma y los complejos eIF1, 2, 3 y 5. Esto permite que se una al mensajero previamente cargado y estabilizado por la acción del complejo eIF4 y PABP, dando lugar a la subunidad 48S. Gracias a su unión en el extremo 5' UTR se inicia el escaneo hasta encontrar el sitio de inicio de traducción, AUG. El reconocimiento permite la unión del tRNA de metionina cargado, y una parcial remoción del complejo eIF2 para así permitir la unión de la subunidad 60S ribosomal, culminando con la formación del complejo completo 80S que seguirá con la incorporación de tRNAs mediante formación de enlaces peptídicos, lo que terminará de remover a eIF2. Modificado de Walsh *et. al.* 2011¹⁶.

En la mayoría de los casos, esta adición de aa continúa hasta que la maquinaria se encuentra con un codón de paro, que provoca la liberación de la proteína completa y desensambla por completo al complejo de síntesis de proteínas². Sin embargo, existen mecanismos que permiten que en algunos casos

durante el inicio de la traducción, algunos AUGs se pueden saltar durante el escaneo por parte de los ribosomas¹⁷, o que los ribosomas no se desensamblen al finalizar la traducción de la primera proteína¹⁸. Esto ocurre en particular en los casos de traducción de policistrones.

1.6 Policistrones en eucariotes

Los policistrones son mRNAs con la información para codificar a más de una proteína en la misma unidad. Se creía que los mRNA policistrónicos eran comunes únicamente en bacterias y virus, que deben ser capaces de codificar una gran diversidad de proteínas a partir de genomas más pequeños¹⁹.

Sin embargo ahora se sabe que son comunes en eucariotes e incluso se conocen algunos mecanismos mediante los cuales es posible la traducción de proteínas que son parte de mensajeros policistrónicos.

1.7 Mecanismos de traducción no canónicos

Existen diversas maneras en las que las células llevan a cabo la traducción de las proteínas miembros de los mensajeros policistrónicos, quizá la más estudiada es la presencia de estructuras que se forman en ciertas secuencias de RNA a lo largo del mensajero, denominadas Sitios de Entrada Interna del Ribosoma (IRES por siglas en inglés de Internal Ribosome Entry Site)²⁰ (Figura 4). Los IRES generalmente están conformados por varios cientos de pares de nt y su estructura está especializada para reclutar a alguna de las proteínas involucradas en la formación del PIC de traducción²¹. Existen diversos reportes de diferentes proteínas del PIC de traducción reclutadas sin necesidad de tener un mensajero con cap, lo cual permite que exista traducción en ciertos mensajeros sin necesidad de tener un mRNA maduro. Este mecanismo le brinda a la célula la capacidad de traducir proteínas cuando no existen las condiciones basales de traducción²²⁻²⁴. Este mecanismo es el más estudiado, y es el más empleado por los virus^{25,26}.

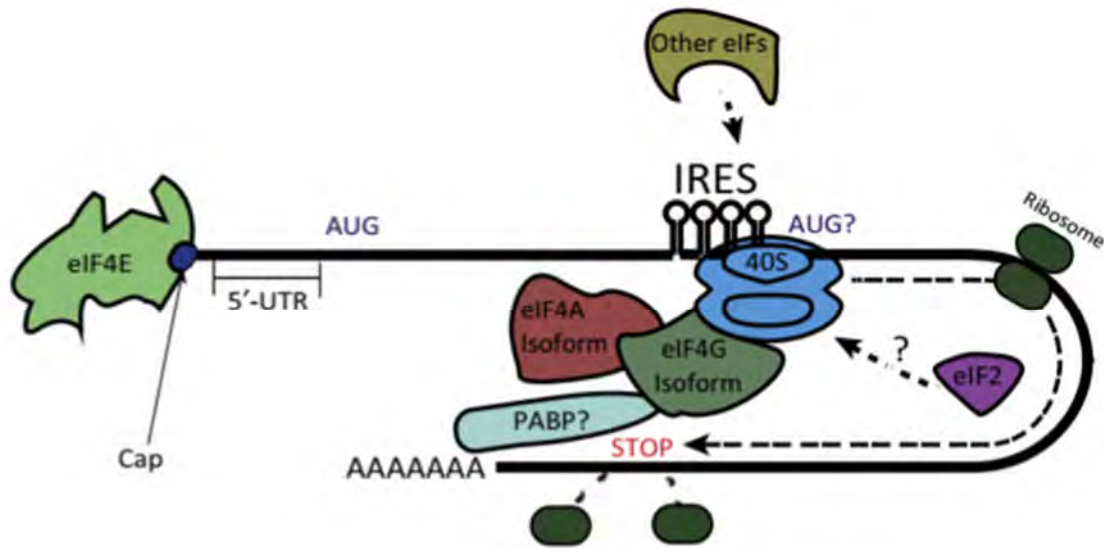


Figura 4. Inicio de traducción mediado por Sitios Internos de Entrada del Ribosoma (IRES, por sus siglas en inglés). Los mensajeros pueden contar con secuencias específicas río arriba o río abajo del sitio de inicio de traducción, que pueden o no formar estructuras secundarias específicas. Estas secuencias se denominan Sitios Internos de Entrada del Ribosoma porque son capaces de reclutar a diferentes proteínas de los complejos de inicio de traducción sin que necesariamente esté formando el complejo completo (isoformas), o incluso reclutar a la subunidad pequeña del ribosoma. De esta manera se puede iniciar la traducción sin el previo reconocimiento del cap en el mensajero. Modificado de Karginov *et. al.* 2017¹⁹.

También en algunos casos durante el inicio de la traducción, los ribosomas se pueden saltar algunos AUGs durante el escaneo y empezar su traducción en otro sitio de inicio. Este mecanismo es denominado escaneo con fuga o *leaky scanning*. En el caso de los mensajeros policistrónicos, se ha reportado este mecanismo cuando las regiones intercistrónicas son cortas (alrededor de 10 nt), y permiten que el ribosoma encuentre el siguiente sitio de inicio de traducción antes de poder desensamblarse y permitir la traducción de la siguiente proteína del mRNA. De esta manera se garantiza que las proteínas de un policistrón se sintetice en una relación 1:1, ya que es una traducción secuencial¹⁸.

En particular, en *Drosophila* se han clasificado diversos mensajeros que dependen de un mecanismo conocido como sobrepaso del codón de paro o *stop*

codon readthrough, en el cual el ribosoma reconoce secuencias específicas posteriores al sitio de término de la primera proteína, que evitan que se desensamble y siga leyendo en la región intercistronica. El ribosoma entonces comienza la adición de aminoácidos que lee de esta región, y continúa hasta terminar la segunda proteína del mensajero. Esto resulta en una proteína con aa extra que requieren de una proteólisis para quedarse con la proteína completa, ya sea que la primera proteína tenga estos aa extra en su extremo C-terminal, o que la segunda proteína los tenga en su extremo N-terminal²⁷.

II. Antecedentes

2.1 *Drosophila melanogaster* como modelo de estudio

A lo largo de los años se ha trabajado con la mosca de la fruta, aunque en realidad la traducción literal sería mosca del vinagre, *Drosophila melanogaster*, ya que presenta muchas ventajas para trabajar en el laboratorio. Además de que es sencillo y económico mantenerlas en el laboratorio, su genoma es pequeño y está completamente secuenciado. Su desarrollo está muy bien caracterizado desde embrión a organismos adultos (Figura 5) y por ello existen muchas técnicas para trabajar con ella. Una característica destacada es que existe una conservación en los procesos biológicos entre este organismo y los mamíferos. Aproximadamente 75% de los genes relacionados a enfermedades humanas tienen un gen equivalente en *D. melanogaster*²⁸.

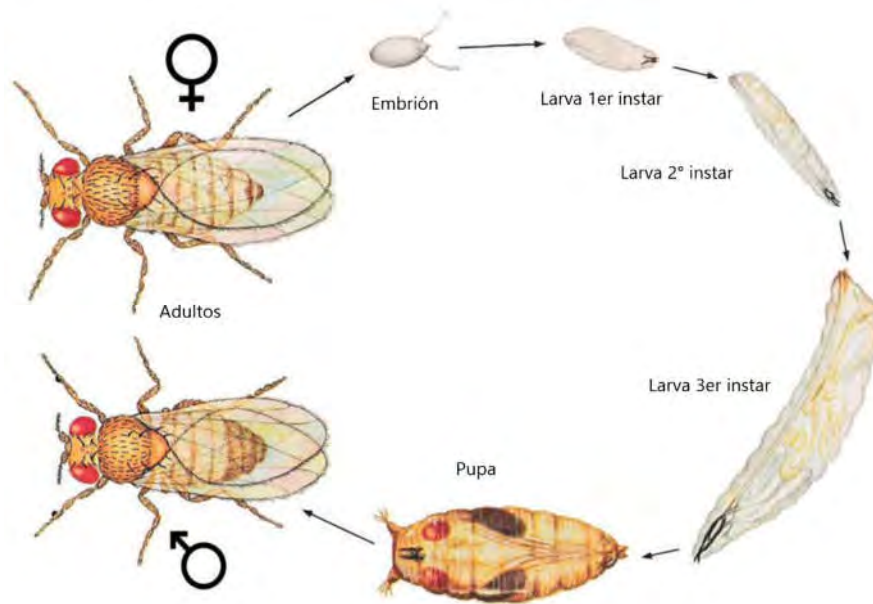


Figura 5. Desarrollo de *Drosophila melanogaster*.

Después de la cópula entre hembras y machos, la hembra deposita embriones fértiles que van a dar lugar a la larva. Existen 3 estadios larvarios (instars), en los que ocurren diferentes mecanismos de maduración del individuo. La larva de 3er instar es la encargada de dar lugar a un capullo que le permite llegar a ser pupa, en el cual se lleva a cabo la metamorfosis a organismos adultos. Modificado de Flagg R.O. 1979²⁹.

2.2 Bicistrones en la mosca *Drosophila melanogaster*

En un escaneo de los posibles transcritos bicistrónicos en *Drosophila* se encontró que existen varios *loci* capaces de dar lugar a dos o más proteínas a partir de un solo mRNA. De los mecanismos conocidos capaces de permitir esta traducción previamente descritos, todos se han visto en la mosca, permitiendo que cuando termine la traducción del primer marco de lectura abierto (ORF por sus siglas en inglés) se pueda reiniciar en el segundo inicio de traducción inmediatamente después en el segundo sitio de inicio de traducción^{27,30}.

Se han descrito secuencias similares a las IRES, como es el caso de la proteína *dMyc* que se puede traducir gracias a la acción de la unión de CNBP, regulando así el desarrollo del ala en *Drosophila melanogaster*³¹; hasta un salto en el escaneo ribosomal debido a la presencia de otra secuencia AUG muy cerca del codón de paro del primer ORF, lo que no permite que se desensamble la maquinaria de traducción, como es el caso del gen *stoned*, involucrado en el funcionamiento neuronal de la mosca^{32,33}.

En años recientes se describió un mecanismo utilizado por todos los metazoarios para permitir la traducción de un mensajero después del codón de paro, al que se le denomina *Stop Codon Readthrough* que fue descrito con anterioridad. En *Drosophila* se identificaron alrededor de 283 candidatos a llevar a cabo este proceso, lo que además parece indicar que es mucho más común de lo que se esperaba en organismos eucariotes²⁷.

También se ha visto que en algunas proteínas, cuando la célula es sometida a algún tipo de estrés, es posible que la célula lleve a cabo la traducción de un mensajero sin necesidad de la presencia de un Cap en su extremo 5'. Este mecanismo puede resultar ventajoso porque en situaciones de estrés, los niveles de eIF4F disminuyen considerablemente³⁴.

2.3 Bicistrón Dmp8-Dmp18 en la mosca *Drosophila melanogaster*

Se desconoce cómo se puedan estar llevando a cabo la traducción de estas proteínas, su región intercistrónica abarca 54 nt, lo que impediría que el ribosoma continúe escaneando el mensajero al terminar la traducción de la primera proteína.

Además de que se ha visto que la traducción de Dmp18 no depende de que se produzca Dmp8 de manera correcta, esto se vio en un trabajo previo de nuestro laboratorio gracias a que una mosca mutante que tenía un transposón insertado en la región codificante de Dmp8, y que resultaba en fallas para traducir a Dmp8, no afectaba la traducción de Dmp18⁷.

Aunque ambas proteínas se encuentran presentes en todos los organismos, el arreglo bicistrónico se mantiene únicamente en organismos del filo de los artrópodos⁷ (Figura 6).

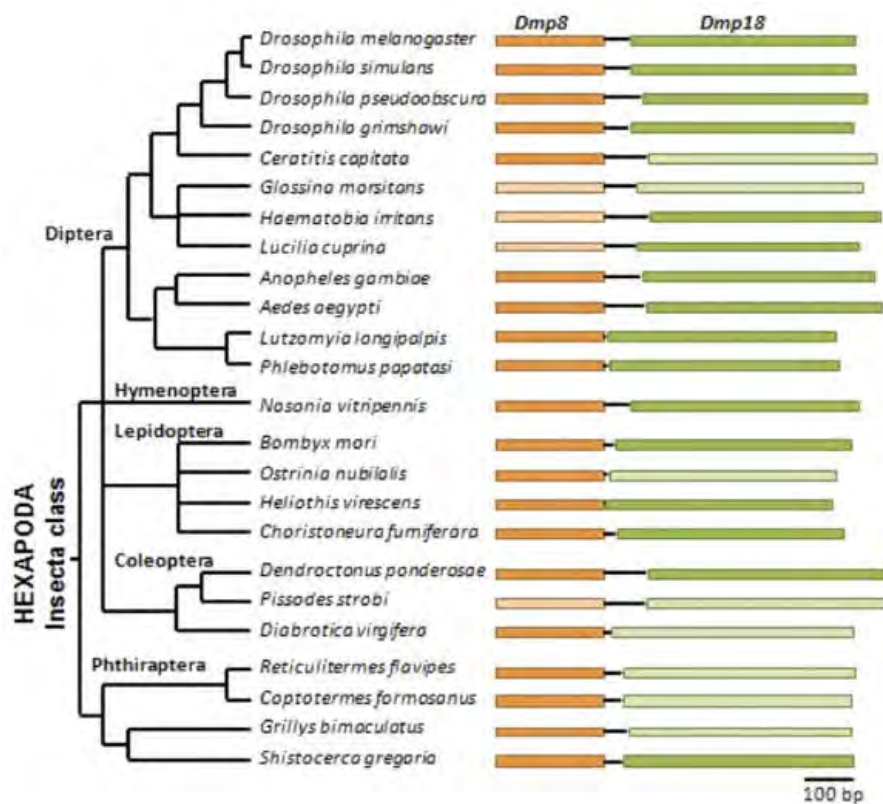


Figura 6. Bicistrón Dmp8-Dmp18 en artrópodos.

Histograma que representa la configuración del bicistrón Dmp8-Dmp18 conservada en varios insectos. Se puede observar que ambas proteínas están muy bien conservadas (p8 en cajas naranjas y p18 en cajas verdes, en color claro se resaltan las proteínas que no han sido anotadas), pero no así la región intercistrónica (representada por una línea negra continua), que varía considerablemente en tamaño. Tomado de Herrera-Cruz M., *et al.* 2012⁷.

Lo que llama la atención sobre las proteínas de mensajero bicistrónico Dmp8-Dmp18 es que, a diferencia de la mayoría de las proteínas policistrónicas, no están llevando a cabo su función en la misma vía de señalización. Aunque ambas están involucradas en mecanismos de reparación, y se ha demostrado una interacción genética entre ambas. Existe sólo un mRNA a lo largo de todo el desarrollo de *D. melanogaster* que permite la traducción de ambas proteínas (Figura 7)⁷.

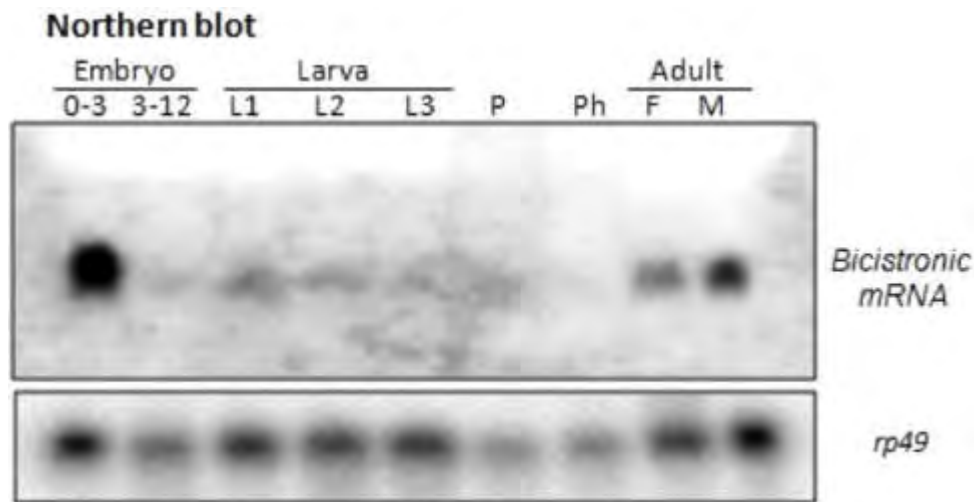


Figura 7. mRNA bicistrónico a lo largo del desarrollo.

Ensayo tipo Northern Blot que muestra la presencia del mRNA bicistrónico Dmp8-Dmp18 en todos los estadios del desarrollo de la mosca (2 momentos del desarrollo embrionario, 3 estadios larvarios, pupa, farato, y hembras y machos adultos). Se puede apreciar la presencia del transcrito en todas las etapas, así como que sólo existe un transcrito. Como control se utiliza rp49, mRNA de una subunidad ribosomal. Tomado de Herrera-Cruz et. al. 2012⁷.

Además se vio que ambas proteínas se expresan en todo el desarrollo de la mosca, lo cual no es de sorprender porque ambas están involucradas en procesos importantes para el mantenimiento de la homeostasis celular (Figura 8)³⁵.

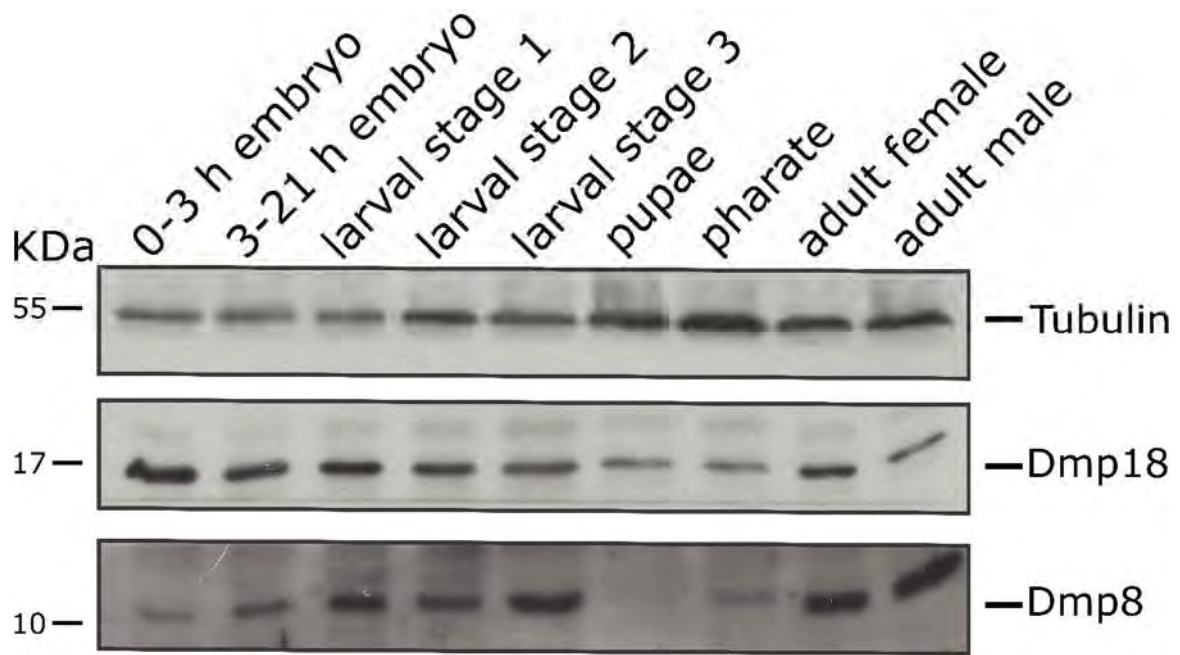


Figura 8. Presencia de las proteínas del bicistrón en todo el desarrollo de la mosca.

Ensayo tipo Western Blot para detectar la presencia de las proteínas del bicistrón Dmp8 y Dmp18 que muestra que ambas proteínas se expresan durante todo el desarrollo de la mosca. Se utiliza Tubulina como control de carga. Bucio-Mendez A et. al. 2018.

III. Justificación

Dado que Dmp8 y Dmp18 se encuentran en el mismo mensajero, y existen evidencias de que la traducción de Dmp18 es independiente de aquella de Dmp8, se podría inferir que la regulación de estos sea independiente una de la otra. Por lo tanto, su estudio podría dar evidencia sobre la forma en que la traducción de estas proteínas se está regulando.

IV. Objetivo

3.1 Objetivo general

Comprobar si existe una regulación diferencial en la traducción de las proteínas Dmp8 y Dmp18.

3.2 Objetivos particulares

- i.** Determinar si es necesario que ambas proteínas se encuentren en el mismo elemento genético para permitir la traducción de ambas
- ii.** Determinar si la región río arriba del sitio de inicio de traducción de Dmp18 es importante para su traducción.
- iii.** Establecer si existe una ventaja biológica para alguna de las proteínas del bicistrón Dmp8-Dmp18 el estar en esta organización.

V. Materiales y Métodos

5.1 Clonación del mensajero bicistrónico

Se extrajo RNA total a partir de 500 mg de moscas adultas de tipo silvestre, utilizando TRIZOL (Invitrogen) según se indica en el manual de uso. Posteriormente se obtuvo cDNA mediante una reacción de retro-transcripción, a partir de 50 ng del RNA total. En un volumen final de 20 uL se prepara el RNA con 1mM de dNTPs, 10 pM de oligo dT, inhibidores de RNAsa, 10 mM DTT y enzima M-MuLV (Invitrogen). Se incubó durante 1 h a 37° C. Posteriormente se desactiva la reacción durante 15 minutos a 65° C.

Para obtener el mensajero a partir del extremo 5' UTR, se utilizaron los siguientes oligos con sitios de clonación NotI-SalI: 5'-GATATAGCGGCCGCATTCAGCTCTGAGAATTGGT-3' (Forward) y 5'-GATCCGGTCGACCAGTTATGCCTAATGTTTG-3' (Reverse). Se purificó el fragmento en un gel de agarosa al 1% con el Kit QIAEX II de Qiagen según se indica en el manual.

Posteriormente se clonó en el plásmido pBluescript II KS entre los sitios NotI y SalI (Figura 9), para posteriormente hacer la transformación en células electrocompetentes y conseguir clones con el inserto deseado.

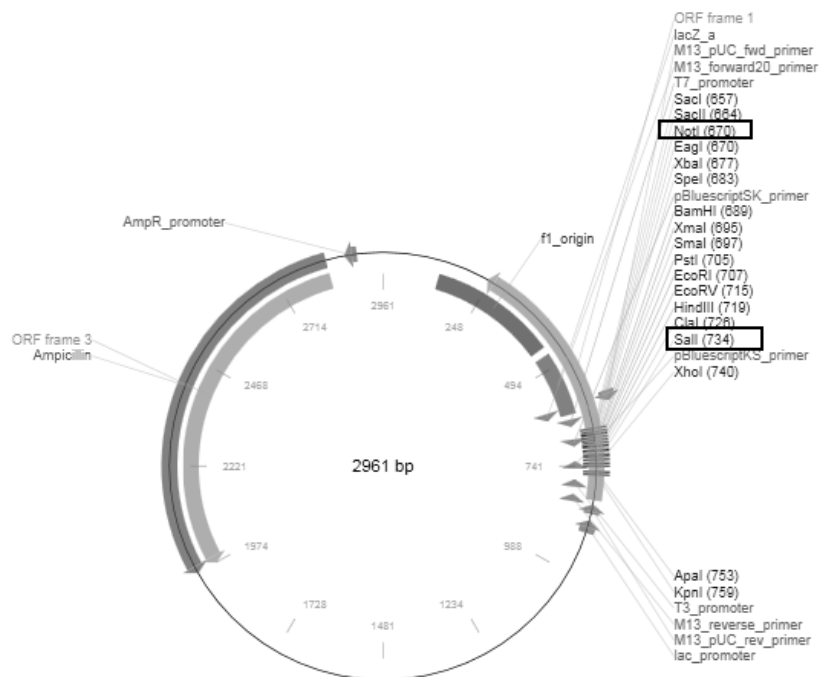


Figura 9. Vector pBKSII.

Representación del vector en el cual se clonó la región mensajera del bicistrón desde el inicio del extremo 5' UTR hasta el codón de paro de Dmp18. Remarcados se encuentran los sitios donde se clonó el mensajero del bicistrón.

La colonia seleccionada fue comprobada por secuenciación del inserto, y fue la utilizada para los ensayos *in vitro*.

5.2 Inserción de poliA a la construcción

Para insertar poliA a la construcción del bicistrón, justo en el final del mensajero, se sintetizaron oligos con poliA flanqueado por sitios de corte de la enzima SalI: 5'-CGCGCCGTCAAAAAAAAAAAAAAAAAAAGTCGACGCGGCG-3' y 5'-CGCCGCGTCGACTTTTTTTTTTTTTTTTTTTTGGACGCGCG-3'.

Se hizo una mezcla de ambos oligos a una concentración de 10 µM y se dejó incubar durante 1 hora a 74°C. Posteriormente se cortó con la enzima SalI, tanto a los oligos alineados, como al vector con la construcción del bi-cistrón. El vector se trató con Fosfatasa Alcalina de Camarón (SAP), y el fragmento de oligos con T4 Polinucleótido Cinasa, esto para evitar que el vector se ligara consigo mismo otra vez. Posterior al tratamiento se hizo la ligación para hacer la transformación de los vectores resultantes.

Las colonias obtenidas se secuenciaron para confirmar la inserción de la secuencia de poliA en el vector. Esta nueva construcción es la que se utilizó para los experimentos *in vitro*.

5.3 Generación de construcciones a partir del mensajero bicistrónico

Para analizar qué regiones eran importantes para la traducción de ambas proteínas, se generaron construcciones a partir del mensajero bicistrónico con los oligos: **A** 5'-GATGCGGCCGCGAGAACCATTTGTTTA-3', **B** 5'-CTAAGCGTCCTTGTCGGAATTCGAGACA-3', **C** 5'-GCTACAGAATTCATGACGGGTCGCGAAT-3', y **D** 5'-GTTGCCAGCATTGTTGAATTCGGCCCA-3'. La posición de los oligos se muestra en la Figura 10.

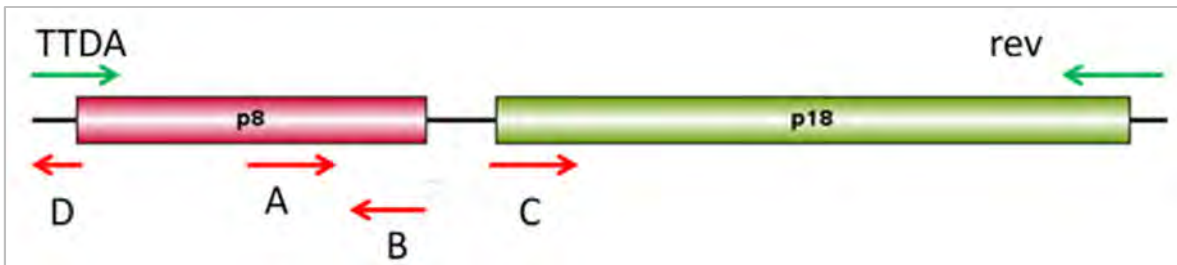


Figura 10. Oligos utilizados.

Representación de la localización de los oligos para generar diferentes versiones del bicistrón.

Con los oligos A y rev se generó la construcción que contenía sólo la última parte de Dmp8 hasta el final del bicistrón. Otras construcciones se generaron a partir de clonar a Dmp18 con los oligos C y rev en el vector pBKSII, para posteriormente a la construcción añadir el fragmento que sólo incluye Dmp8 que se clonó con los oligos TTDA y B; por otra parte, a una última construcción se le unió el 5' UTR que se clonó con los oligos TTDA y D.

5.4 Ensayos de Transcripción/Traducción *in vitro*

Una vez que se contaba con las construcciones en el vector pBKSII, se llevaron a cabo ensayos de transcripción acoplada a traducción *in vitro* con el kit TNT Coupled Reticulocyte Lysate System, de Promega. Siguiendo las instrucciones del manual, e incorporando EasyTag™ L-[35S]-Methionine de Perkin-Elmer.

Posteriormente, se tomaron 5 µl de muestra y se hirvieron con Buffer Laemmli durante 10 minutos. Las muestras fueron corridas en un gel SDS-PAGE al 16% para permitir la correcta separación de las fracciones de 8 y 18 KDa. El gel se secó en una cámara Slab Gel Dryer durante 4 horas a 65°C. Los geles secos se sometieron a autoradiografía durante 1 semana a -70°C para revelarlos posteriormente.

Para la determinación de la cinética de tiempo, se colocaron reacciones por duplicado y se analizaron de acuerdo con el manual de uso del producto. 5 µl de la muestra fueron sacados a diferentes tiempos (20, 40, 60, 90, y 120 minutos). Estas muestras se hirvieron con Buffer Laemmli durante 10 minutos para proseguir con la separación de proteínas en el gel SDS-PAGE como se menciona en el párrafo anterior.

5.5 Generación de vectores para transfectar células

Las construcciones se hicieron en el vector pAc5.1 /V5-His A. Para facilitar la clonación de los derivados bicistrónicos, fue necesario modificar el polylinker del vector. Con este objetivo insertamos el sitio de corte de la enzima SacII. Se decidió eliminar uno de los sitios BstXI, ya que el vector contaba con otro de estos sitios dentro de su secuencia, y el objetivo era dejar un sitio único para esta enzima de restricción. En la figura 11 se muestran las características del vector elegido, y se señala el cambio en la región polylinker.

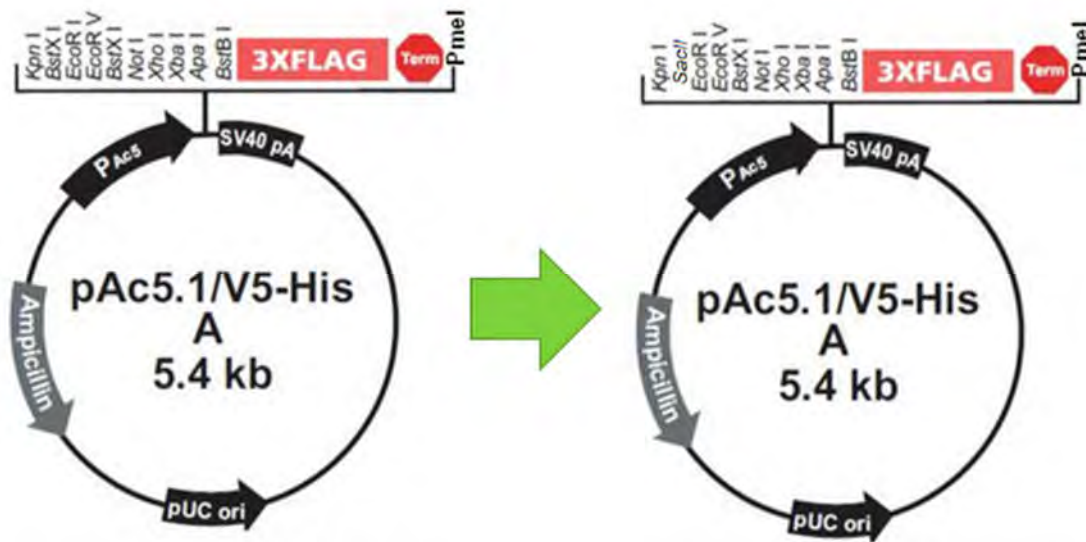


Figura 11. Vector pAc5.1/V5-HisA

Esquema de las características del vector pAc5.1/V5-His A. Se muestra el cambio en la región polylinker del vector para poder clonar las construcciones del vector pBKSII

La luciferasa se clonó a partir del vector pGL4.13 (Figura 12). Se quitó el sitio de inicio de traducción de la luciferasa, con la intención de asegurar que la traducción fuera inducida por la secuencia del bicistrón. El bicistrón se clonó entre los sitios EcoRI y EcoRV del vector.

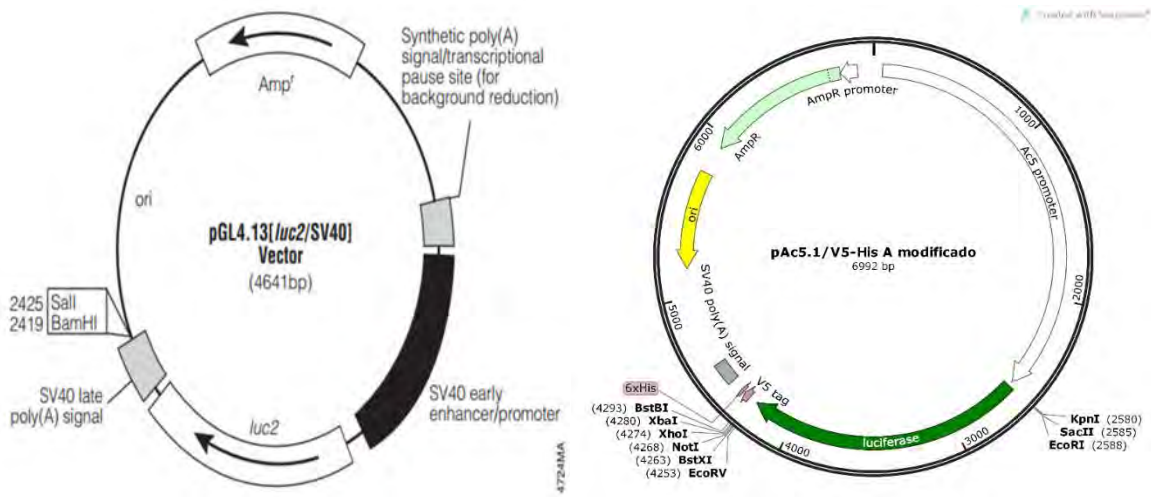


Figura 12 Vector resultante para expresión de la Luciferasa de luciérnaga
 En la izquierda se muestra el esquema del vector a partir del cual se clonó a la Luciferasa de luciérnaga en el vector para la transfección. A la derecha se muestra el vector obtenido con el reportero de la luciferasa de luciérnaga para la transfección.

Para poder analizar la eficiencia en la transfección, es necesario contar con un control de la misma. Es por ello que se decidió usar la Luciferasa de Renilla (RLuc) como control. Esta también se clonó entre los sitios EcoRI y EcoRV. En la figura 13 se muestran al vector a partir del cual clonamos a la RLuc y cómo quedó el vector de transfección con la RLuc clonada.

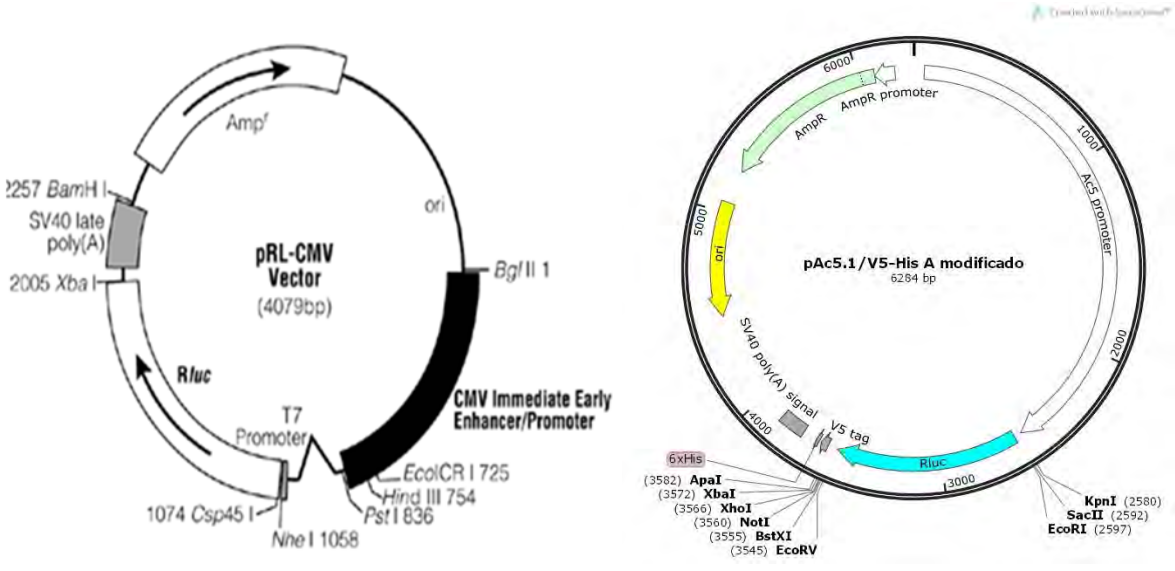


Figura 13. Vector resultante para expresión de la Luciferasa de Renilla

En la izquierda se muestra el esquema del vector a partir del cual se clonó a la Luciferasa de luciérnaga en el vector para la transfección. A la derecha se muestra el vector obtenido con el reportero de la luciferasa de luciérnaga para la transfección

5.6 Cultivo celular

Se utilizaron células de origen embrionario de *Drosophila melanogaster* S2R+³⁶. Se descongelaron los viales que contenían 1×10^7 células y se atemperaron a 25° C. Las células se lavaron 2 veces con 5 ml de medio Schneider complementado con suero fetal bovino al 20%, seguido de una centrifugación a 3000 rpm durante 5 minutos.

Finalmente, las células se incubaron en 10 ml de medio Schneider complementado con suero fetal bovino al 20%, y se incubaron a 26° C para su cultivo. Cuando el cultivo alcanzaba el 60% de confluencia, se realizaba nuevos cultivos celulares (pases).

5.7 Transfección para expresión transitoria de luciferasas en células S2R+

Para la expresión transitoria de la Luciferasa de Renilla transfectamos cultivos celulares que presentaban una confluencia del 70%, y contenían un total de 4×10^6 células por cajas de cultivo, en medio Schneider complementado con 10% suero fetal bovino.

La solución de transfección se preparó añadiendo la solución A (2M CaCl₂, 1 µg de vector seleccionado, 1 µg del vector para expresar Renilla Luciferasa, en agua inyectable estéril) lentamente por goteo y con agitación constante a la solución B (50 mM HEPES, 1.5 mM Na₂HPO₄, 280 mM NaCl, pH 7.1) para formar un precipitado salino en el tubo. La mezcla de ambas soluciones se dejó reposar 30 minutos para que se terminara de formar el precipitado.

La mezcla de la solución de transfección se agregó lentamente a la caja de cultivo, por goteo. Después de cada gota se agitaba suavemente la caja de cultivo. Las células se dejaron recuperar a 26° C durante 48 horas, antes de proceder al tratamiento experimental que se iba a evaluar.

5.8 Medición intensidad de luciferasa en células transfectadas

Se colectaron las células transfectadas de las cajas en un tubo de 15 ml, y se centrifugaron a 3000 rpm durante 5 minutos. Se removió todo el medio con una pipeta y se resuspendió el botón de células en 30 μ L de buffer PBS.

Para medir la intensidad de luminiscencia de ambas luciferasas, se utilizó el kit Dual-GLO de Promega según las instrucciones del fabricante, en el Luminómetro Monolight 3010.

Se consideró la intensidad de la RLuc como control de transfección, y en base a ella se midió el porcentaje de luminiscencia de la Luciferasa de luciérnaga, tomando aquella completa, la que tenía su sitio de inicio de traducción como el valor de 100%.

5.9 Detección de proteína en larvas de tercer instar

Se hizo una extracción de proteínas con TRIZOL, siguiendo las instrucciones del manual para extracción de proteína. El pellet resultante se hirvió en Buffer de laemmli 2X (4% SDS, 20% glicerol, Tris 0.125 M pH6.8, 0.004% azul de bromofenol, 1% de β - mercaptoetanol) para cargar en gel de acrilamida al 16%.

Se puso el gel con una membrana de nitrocelulosa en una cámara de transferencia inmerso en buffer de transferencia a pH8.3 (Tris-HCl 25 mM, Glicina 192mM, Metanol 20%) y se corrió a 80mA durante toda la noche a 4° C. Antes de sacarlo se subió el amperaje a 200mA durante 2 horas. Se trató la membrana con Ponceau (0.2% Ponceau, 3% ácido acético) y se enjuagó con agua destilada (Maniatis *et al.*, 1987).

Se bloquearon las membranas en leche diluida en PBST (PBS1X con Tween al 0.1%), al 10% para Dmp18, Tubulina y Actina, y al 20% para la proteína Dmp8. Se dejó en agitación durante toda la noche, y posteriormente 1 hora a temperatura ambiente. Se lavó la membrana con PBST. Se puso a incubar el anticuerpo que reconoce a Dmp8 y Dmp18 (NEB) a una concentración de 1:1,000 en leche al 5% en PBST durante 2 horas a temperatura ambiente. Se utilizaron los anticuerpos que reconocen a β -tubulina y Actina (E7 y JLA-20 respectivamente, del Banco del hibridoma) a una concentración 1:3,000 en leche al 5% en PBST durante 2 h a

temperatura ambiente. Se hicieron 5 lavados en PBST durante 10 minutos cada uno en agitación a temperatura ambiente antes de poner el anticuerpo que conjugado a peroxidasa de rábano HRP correspondiente, a una concentración 1:3,000 en leche al 5% en PBST, durante 1 hora en agitación leve a temperatura ambiente. Se lavaron las membranas 5 veces con PBST durante 10 minutos cada lavado en agitación leve a temperatura ambiente (Maniatis *et al.*, 1987).

Para revelar las membranas se siguió el protocolo de SuperSignal® West Pico Chemiluminescent Substrate de ThermoScientific.

5.10 Extracción de RNA de larvas para ensayo RT-PCR

De las larvas sometidas a los diferentes estímulos que fueron tratadas con TRIZOL, se extrajo también el RNA como indica el manual del reactivo. Se tomaron 100 ng de este RNA para preparar cDNA como se explica en el punto 5.1, y se usaron los oligos que amplifican a todo el mensajero bicistrónico en un PCR semicuantitativo.

5.11 Extracción de proteína de células S2R+

De las cajas con células transfectadas, después de 48 horas de la transfección, se juntaron en un tubo de 15 ml y se centrifugan a 2000 rpm durante 5 minutos. Se lavaron las células 2 veces con buffer PBS. Se resuspendió el botón celular en 200 µL de Buffer A (HEPES 10mM, KCl 10 mM, EDTA 0.1 mM, EGTA 0.1 mM, Espermidina 0.75 mM, Espermina 0.15 mM, inhibidores de proteasas) y se dejó incubar a 4°C durante 15 minutos. Se agregó Igepal para que quede en una concentración final de 1.25% y se mezcló con la pipeta.

Las células se centrifugaron a 1,200 rpm durante 30 segundos y se apartó el sobrenadante que incluye la fracción citoplásmica. Al botón celular se le agregaron 100 µL de Buffer B (HEPES 20 mM, NaCl 0.4 mM, EDTA 1 mM, EGTA 1 mM, inhibidores de proteasas) y se incubó en agitación fuerte durante 15 minutos a 4° C.

Se centrifugó a 12,000 rpm durante 5 minutos, y se recuperó el sobrenadante que tiene la fracción nuclear en el mismo tubo en el que se encontraba la fracción

citoplásmica. Se les agregó Buffer Laemmli para que quedara a una concentración final de 1:1 y se hirvieron durante minutos para poder cargar en un el SDS-PAGE.

Se procedió a analizar la proteína como se explica en el apartado 5.9.

VI. Resultados

6.1 Alineamiento de las regiones intercistrónicas de Artrópodos

Para intentar tener una aproximación teórica al posible mecanismo de traducción de Dmp18, se realizó un análisis *in silico*. En *D. melanogaster* la región intercistrónica es de 54 pb, lo cual nos hizo pensar que el mecanismo de *Leaky scanning* era el menos probable. Como se mencionó anteriormente, el mecanismo de *stop codon readthrough* requiere de secuencias conservadas, y en *Drosophila* la manera en la que fueron identificadas fue gracias a que en un alineamiento de todas las secuencias de estos genes policistrónicos se ve claramente esta conservación²⁷. Por ello, y porque ya se sabía que la secuencia de aminoácidos de ambas proteínas está altamente conservada⁷, realizamos un alineamiento de las regiones intercistrónicas en varios representantes del filo de los artrópodos utilizando MUSCLE (MULTiple Sequence Comparison by Log-Expectation)³⁷ (Figura 14). Estas secuencias varían en longitud desde 4 pares de bases, hasta 92.

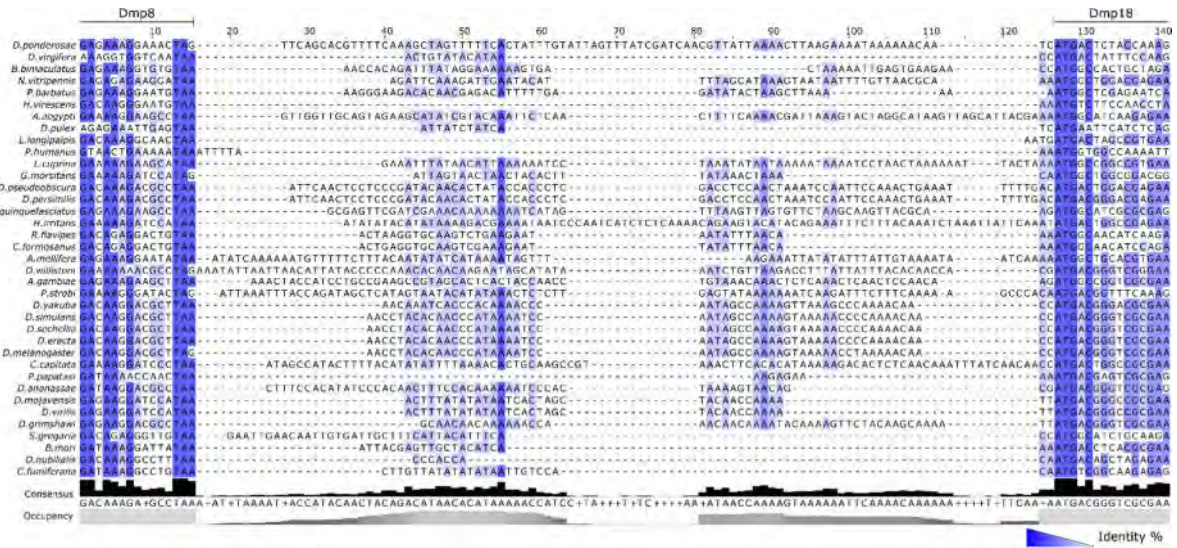


Figura 14. Conservación de regiones intercistrónicas en Artrópodos

Alineamiento de las regiones intercistrónicas del bicistrón Dmp8-Dmp18 en artrópodos. Se dejaron bases de las regiones codificantes de ambas proteínas del bicistrón como referencia. En escala de azules se puede ver el nivel de conservación de las secuencias. Se puede ver que no existe una conservación de esta región en Artrópodos. El alineamiento se hizo usando MUSCLE (Multiple Sequence Comparison by Log-Expectation).

Pudimos observar que no existen regiones altamente conservadas entre ellas, y que además ninguna coincidía con lo previamente reportado de secuencias que permiten un mecanismo de *stop codon readthrough*²⁷. Además gracias a este alineamiento podemos asumir que la composición de la región intercistónica parece ser irrelevante para la traducción de la segunda proteína del bicistrón, Dmp18.

6.2 Análisis *in silico* de regiones promotoras del bicistrón Dmp8-Dmp18

A continuación llevamos a cabo un análisis *in silico* en la base de datos de modENCODE, esto con el fin de averiguar si existe más de una región promotora, lo que podría explicar cómo se lleva a cabo la traducción de ambas proteínas, y que su regulación sea diferente.

Se encontró que sólo hay una región predicha para el promotor, y se localizaba 100 pb río arriba del inicio de transcripción de la proteína Dmp8 (Figura 15). Así mismo el reporte más exacto que existe sobre el mapeo de promotores en *Drosophila* sólo logra detectar un promotor para este mRNA bicistrónico³⁸

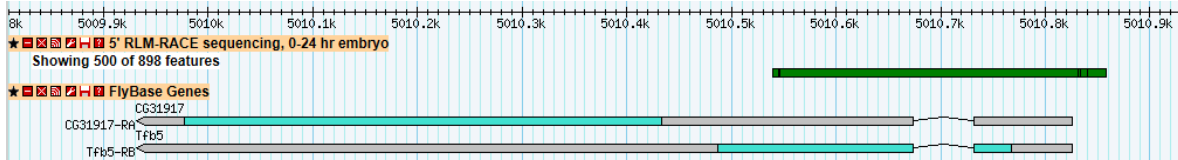


Figura 15. Región regulatoria del bicistrón Dmp8-Dmp18.

Esquema de la presencia de sólo una región regulatoria para el transcrito bicistrónico en *D. melanogaster*. En verde se muestra un solo resultado para la identificación de sitios regulatorios en la proximidad del transcrito bicistrónico. En azul se pueden observar ambas proteínas del bicistrón reportadas en la base de datos. Tomado de la base de datos de modENCODE de datos depositados por S. Celniker³⁸.

Esto quiere decir que una sola región se encarga de la transcripción del mRNA bicistrónico y por lo tanto de la traducción de ambas proteínas.

6.3 Ensayos de transcripción/traducción *in vitro* del mRNA de Dmp8-Dmp18

Como ya se había reportado la presencia de un solo transcrito a lo largo de todo el desarrollo⁷ y vimos que solo existía una región promotora para el mensajero,

procedimos a clonar estos genes desde su 5' UTR hasta el final de la proteína Dmp18 en el vector pBKSII para llevar a cabo ensayos de transcripción-traducción *in vitro*, empleando el sistema heterólogo de reticulocitos de conejo, con metionina marcada radioactivamente. Como control utilizamos a ambas proteínas marcadas con una etiqueta previamente utilizada para ensayos de doble híbrido. En la Figura 16 podemos observar que el transcrito bicistrónico por sí mismo es capaz de llevar a cabo la traducción de ambas proteínas (carril 4).

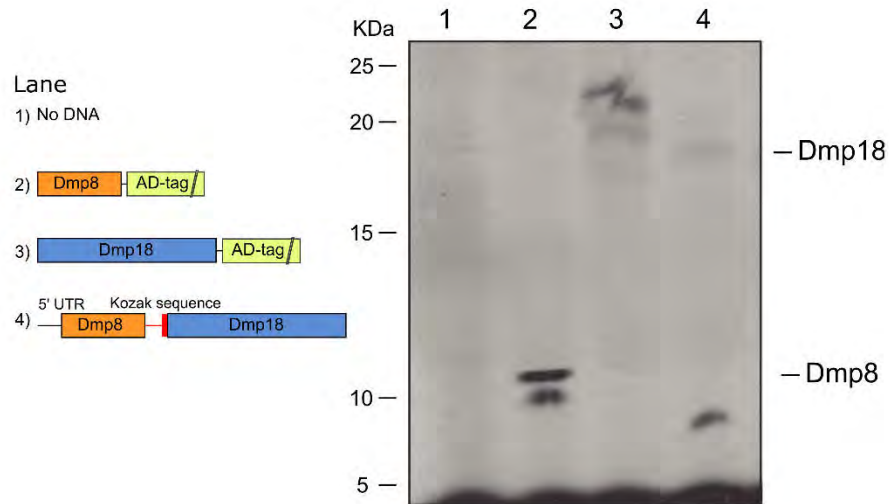


Figura 16. Ensayo de transcripción/traducción *in vitro* del mRNA Dmp8-Dmp18.

Dmp8-Dmp18. En el primer carril está el control negativo sin vector, en el que no se espera ningún producto. En los carriles 2 y 3 se encuentran ambas proteínas con la etiqueta AD, como control individual. En el carril 4 podemos apreciar la presencia de ambas proteínas a partir del mensajero completo

Cabe mencionar que el vector en el que están clonadas las proteínas por separado son vectores con aa extra porque se utilizan para realizar ensayos de doble híbrido.

Para tratar de hacer más eficiente la transcripción/traducción *in vitro*, se insertó en la clonación del mensajero una secuencia poliA. Se hizo el ensayo con esta nueva construcción para comparar si mejoraba el rendimiento (Figura 17).

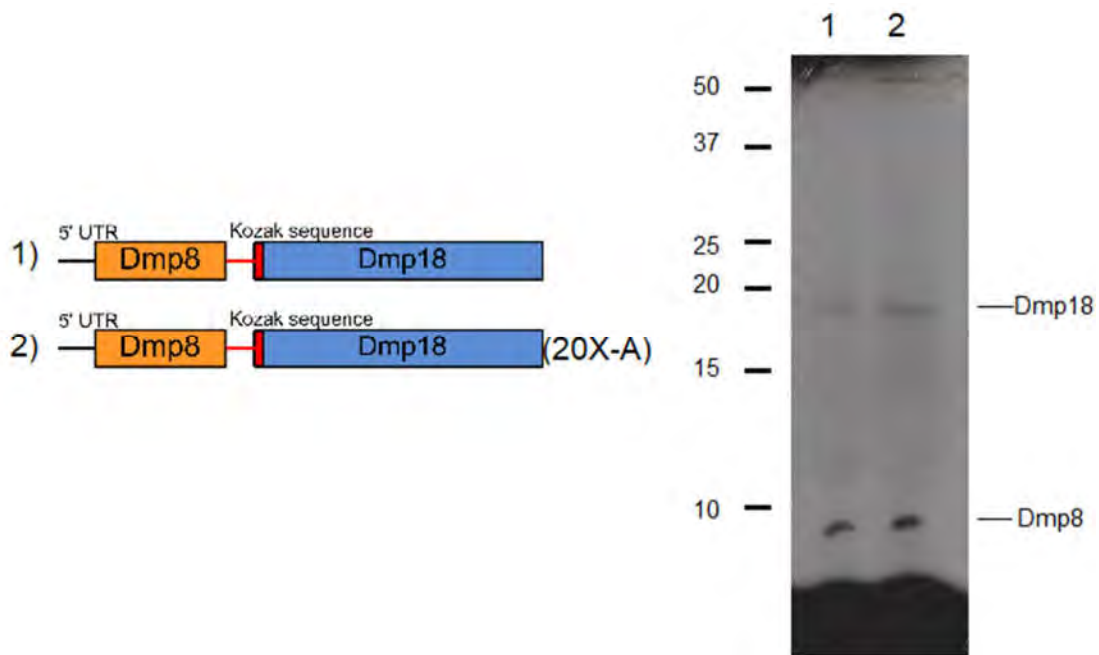


Figura 17. Adición de poliA al mRNA de *Dmp8-Dmp18*.

Ensayo de transcripción/traducción *in vitro* de construcciones con el mensajero bicistrónico (carril 1) comparado con la traducción cuando se le adiciona una cola poliA al final, para darle estabilidad al mensajero (carril 2). No se apreciaron diferencias en la traducción de las proteínas del bicistrón.

El agregar la cola poliA al mensajero mejoró la traducción de ambas proteínas del bi-cistrón. De aquí en adelante se decidió utilizar esa construcción con el poliA incorporado, esto ocurre porque le da más estabilidad al mensajero.

Una pregunta que surgió con esto, era ver si la traducción de ambas proteínas ocurría de manera paralela. Es por eso que se decidió hacer una cinética de tiempo de transcripción-traducción de la cual se tomaba muestra a los 20, 40, 60, 90, y 120 minutos después de iniciar la reacción (Figura 18).

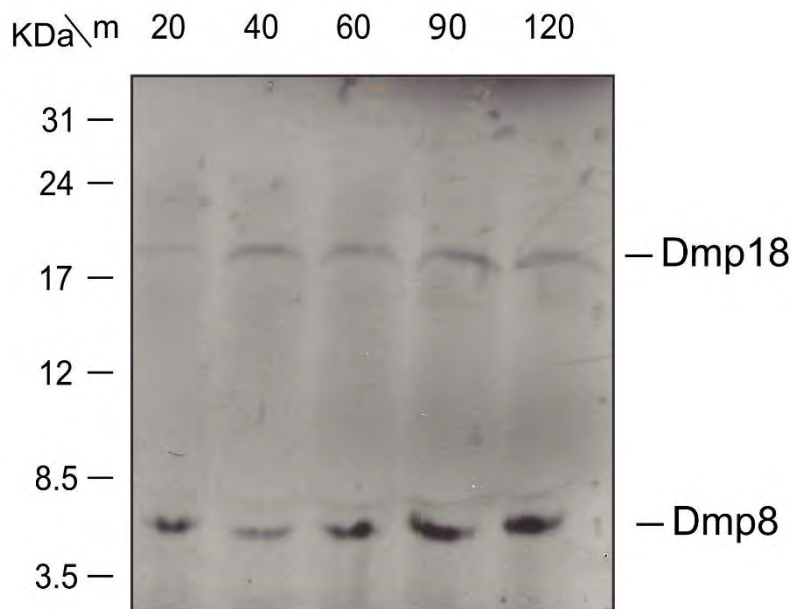


Figura 18. Cinética de aparición de ambas proteínas del bicistrón.

Dinámica de aparición de las proteínas del bi-cistrón. Se muestra la presencia de las proteínas a 20, 40, 60, 90, y 120 minutos después de iniciar la reacción. Se puede apreciar que ambas proteínas aparecen en los primeros 20 minutos sin una aparente traducción retrasada de una con respecto a la otra.

En la cinética se pudo apreciar que ambas proteínas aparecen a partir de los 20 minutos, sin aparentes diferencias a lo largo del tiempo. Esto nos puede sugerir que la traducción de ambas proteínas no es secuencial. Sin embargo, el experimento tiene sus limitaciones para decir esto porque lo que se observa es la incorporación de Metionina, y Dmp8 tiene 4 metioninas, mientras Dmp18 tiene tan solo 2.

En un trabajo previo se había demostrado que una mosca que tenía un transposón insertado en la secuencia de *Dmp8*, no traducía eficientemente a la proteína Dmp8, pero no parecía afectar la traducción de Dmp18⁷. Por ello, se hizo una construcción que dejara a Dmp18 con los 54 nt de Dmp8 que se conserva de un posible transcrito resultante de esta mosca transgénica. En la Figura 19 podemos ver que esa construcción (carril 4) efectivamente permite la correcta traducción de Dmp18. Considerando el resultado del análisis de la secuencia intercistrónica, en el que vimos que no había ninguna región conservada, decidimos generar una construcción en la que abolimos a la región intercistrónica, bajo la presunción de

que esta región es irrelevante para la traducción de ambas proteínas. Para nuestra sorpresa pudimos observar que Dmp8 se traduce de manera normal (carril 5), pero no así Dmp18.

Como al quitar la región intercistrónica completa también eliminamos la secuencia Kozak de Dmp18, evaluamos su importancia analizando la traducción de Dmp18 solo y con su secuencia Kozak. Pudimos observar que si bien Dmp18 se puede traducir cuando se encuentra desde su AUG (carril 6), la adición de su secuencia Kozak (carril 7) mejora visiblemente su traducción. Por ello, generamos otra construcción en la que la secuencia Kozak fue reconstituida en el espacio intercistrónico sin contar con elementos de esta secuencia (carril 8). Para nuestra sorpresa, observamos que a pesar de contar con su secuencia Kozak, no hay una correcta traducción de Dmp18. Finalmente generamos una construcción en la que se le añadía una secuencia irrelevante antes de la secuencia Kozak de Dmp18 para medir su traducción (carril 9), y vimos que no hay cambios notables.

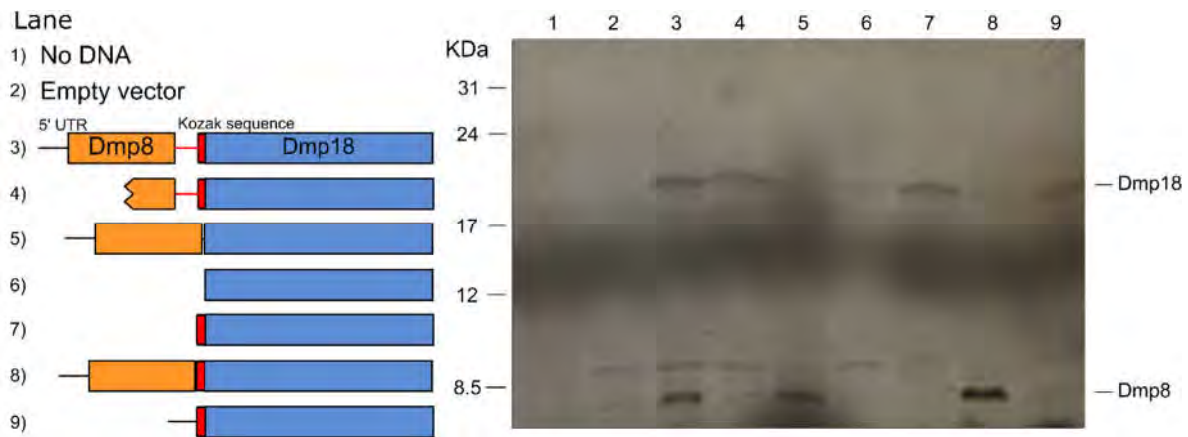


Figura 19. Detección de las proteínas del bicistrón modificando al mensajero.

Ensayos de transcripción-traducción in vitro modificando lo que se encuentra río arriba del AUG de Dmp18. Los primeros 2 carriles corresponden a controles de reacción sin vector (1) o del vector de clonación vacío (2). (3) El bicistrón completo es capaz de traducir a ambas proteínas. (4) Una construcción que abarca desde las últimas 54 pares de bases de Dmp8 permite la correcta traducción de Dmp18. (5) Cuando la región intercistrónica es removida, no se afecta la traducción de Dmp8, pero sí de Dmp18. (6) Cuando se trata de traducir a Dmp18 desde su AUG hay traducción, que es significativamente mejorada cuando se le agrega su secuencia Kozak (7). (8) Sin embargo, la adición de esta secuencia Kozak a la construcción sin

región intercistrónica no recupera la traducción de Dm18. (9) La adición de la región 5' del mensajero a Dmp18 no cambia significativamente su traducción.

6.4 Análisis de traducción *in vivo* del mRNA del bicistrón Dmp8-Dmp18

Como todos los experimentos previos se llevaron a cabo en un sistema heterólogo *in vitro*, el siguiente paso era extrapolar los resultados a una situación *in vivo*. Hicimos construcciones en un vector para transfectar células de mosca de la línea S2R+, acoplando las construcciones a un reportero de luciferasa de luciérnaga mediante la creación de una proteína fusión que tuviera los primeros 4 amino ácidos de Dmp18, y de esta manera hacer también algo que no se podía en los experimentos previos, cuantificar la eficiencia de traducción de cada construcción. Lo primero que hicimos fue emular las construcciones usadas en el sistema de transcripción-traducción *in vitro*. En la figura 20 se muestra la intensidad de luminiscencia como una medida directa de la traducción de Dmp18.

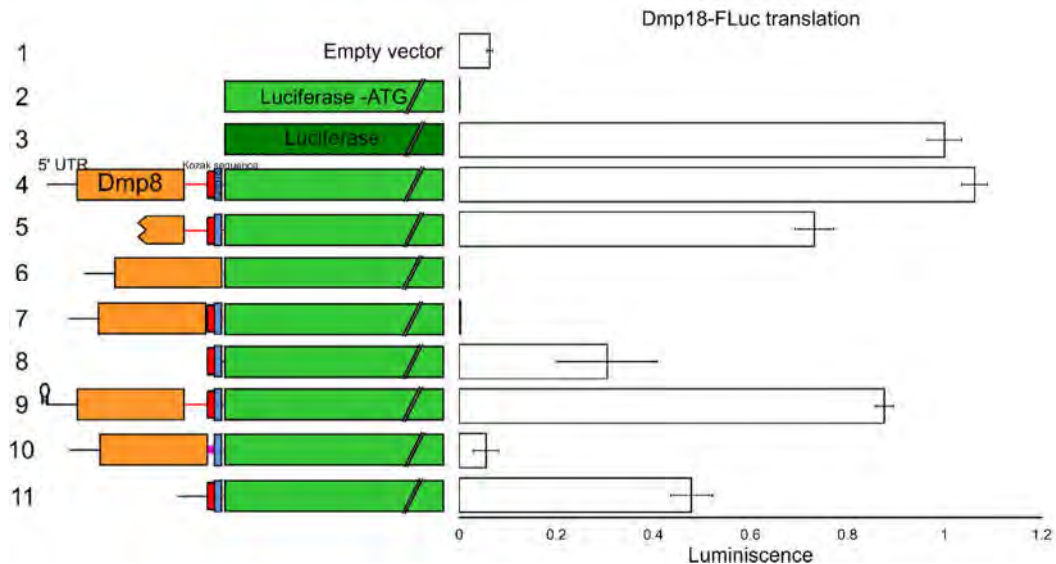


Figura 20. Eficiencia de traducción de un reportero de luciferasa.

Ensayos de transfección en células de *Drosophila* midiendo la intensidad de la luminiscencia de fusiones de Dmp18 fusionada a luciferasa de luciérnaga para medir la eficiencia de traducción al modificar la secuencia río arriba de Dmp18. En el carril 2 se muestra que al quitar el sitio de inicio de traducción de la Luciferasa de luciérnaga no se detecta su actividad en el luminómetro. El carril 3 representa a la Luciferasa de luciérnaga, considerando su intensidad como el 100% de actividad detectable. En el carril 4 se ve la representación del mensajero bicistrónico desde el

extremo 5' UTR y se puede apreciar que los niveles alcanzados son comparables a los de la luciferasa completa. En el carril 5, una construcción con tan solo 54 pares de bases de la región codificante de Dmp8 es suficiente para permitir la traducción. Carril 6: Dmp8 fusionada a la luciferasa no resulta en traducción de esta y eso no se ve mejorado cuando se le adiciona la secuencia Kozak y los 4 amino aa de Dmp18 (carril 7). Carril 8: Dmp18 puede traducirse sin problema cuando se le deja con únicamente su secuencia Kozak, aunque la eficiencia de traducción parece disminuir. Carril 9: la presencia de una secuencia que resulta en una estructura tipo tallo-asa que inhibe la traducción 5' dependiente no afecta la traducción de Dmp18. Carril 10: Insertar la región intercistrónica de *Pediculus humanus* no resulta en una traducción eficiente. Carril 11: La presencia del extremo 5' UTR antes de la secuencia Kozak de Dmp18 no afecta su traducción.

En el primer carril se encuentra el control con células transfectadas con el vector vacío. El segundo carril muestra el resultado de la luminiscencia de la luciferasa a la que se le removió el ATG, para de esta manera asegurarnos de que la luminiscencia se diera por traducción de los primeros 4 aa de Dmp18 y no porque estuviera presente el sitio de inicio de la luciferasa. En el carril 3 se aprecia la intensidad de luminiscencia de la luciferasa completa. La primera construcción consistió en verificar que desde el 5' UTR del mensajero bicistrónico existiera una eficiente traducción de Dmp18. A continuación (carril 5), podemos apreciar que al dejar los últimos 54 pares de bases de Dmp8, Dmp18 es capaz de traducirse correctamente aunque con una eficiencia menor. En el carril 6 vemos que al quitar la región intercistrónica y dejando el ORF de Dmp8, se pudo corroborar que cuando la región intercistrónica es casi inexistente sin importar la presencia de la secuencia Kozak de Dmp18 (carril 7) no es posible su traducción. También se confirmó que Dmp18 sólo necesita de su secuencia Kozak para traducirse, aunque esta construcción muestra una baja en la eficiencia de traducción. La traducción de Dmp8 de los carriles 4, 6, y 7 fue medida con ensayos tipo western Blot de las células transfectadas (Figura 21, carriles 1, 2, y 3).

Como parecía que Dmp18 sólo requiere de la presencia de la secuencia de Dmp8, pero no depende de que ocurra traducción de ella, se le añadió al extremo 5' una secuencia que genera un tallo-asa que inhibe la traducción dependiente de cap en los mensajeros³⁹. Se reportó que esta estructura inhibe la traducción dependiente de cap en el extremo 5' UTR, por lo que esperábamos que la traducción de Dmp18

no se viera afectada en esta construcción. Como se puede ver el carril 9 de la figura 17, la traducción de Dmp18 no se ve impedida por la falla de traducción de Dmp8.

Finalmente, consideramos que si bien la región intercistrónica no parecía tener relevancia en la traducción de Dmp18, claramente era importante su longitud. Eso nos llevó a fijarnos en la región intercistrónica de otros artrópodos, ya que como se mencionó anteriormente, existen organismos con tan sólo 4 pares de bases de longitud en esta región. Se decidió insertar la región intercistrónica de *Pediculus humanus*, con una longitud de 8 pares de bases (Figura 20, carril 10). Sorprendentemente, esta construcción no es capaz de traducir a Dmp18, pero no tuvo problemas en permitir la traducción de Dmp8 (Figura 21, carril 4). Finalmente introducimos el extremo 5' UTR del bicistrón inmediatamente antes de la secuencia Kozak de Dmp18, y no observamos una diferencia significativa de traducción con respecto a la construcción que sólo contaba con la secuencia Kozak.

El hecho de que al poner sólo 8 pares de bases en la región intercistrónica y que esto resultara en una falla en la traducción de Dmp18 nos ayuda a descartar el mecanismo de *leaky scanning* en el caso de este bicistrón. Eso nos deja con que pueda existir una situación de IRES o IRES-*like* en el mecanismo de traducción de este bicistrón. Las evidencias también apuntan a que la traducción de Dmp18 ocurre independientemente al cap del mensajero.

6.5 Detección de Dmp8 en ensayos de transfección

Como el reportero de luciferasa sólo nos permitía ver lo que estaba pasando con la traducción de Dmp18, decidimos hacer un análisis tipo western Blot de células transfectadas con diferentes construcciones que tenían a Dmp8 (Figura 20).

Se pudo observar que las construcciones que debían expresar a Dmp8 de manera adecuada lo hicieron (Figura 21, construcciones 1, 2, 3, y 4), mientras que la construcción con la secuencia que resultará en un tallo-asa en el extremo 5' UTR y que por lo tanto impide la traducción de Dmp8 (construcción 5) se puede ver una disminución en los niveles de Dmp8. No fue posible ver a la proteína Dmp8 en cantidades basales (Construcción 6), sólo en las sobreexpresiones.

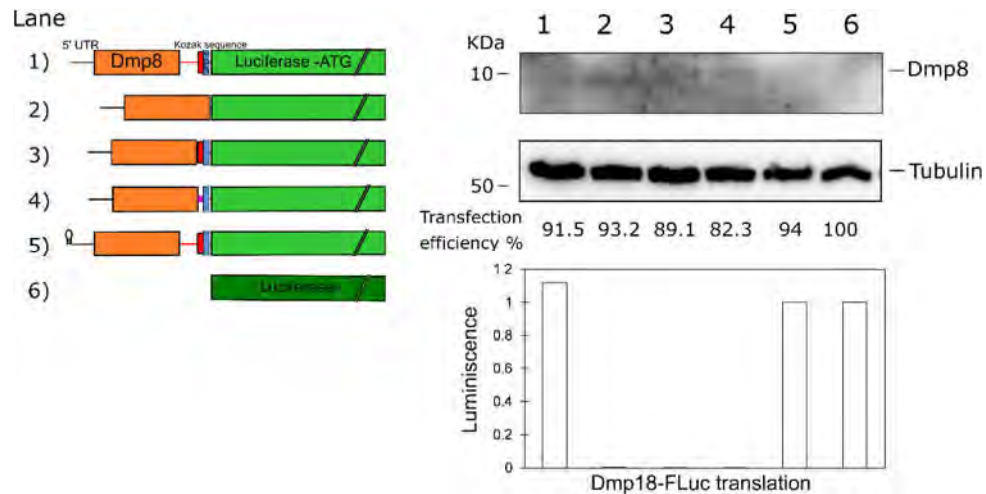


Figura 21. Detección de Dmp8 en células transfectadas.

Análisis de la expresión de Dmp8 en células transfectadas con las construcciones del reportero de luminiscencia. Ensayos tipo western Blot de extractos celulares posteriores a la transfección demuestran la presencia de Dmp8 cuando se transfectan las construcciones que emulan al bicistrón completo (carril 1), la proteína pegada a la luciferasa sin ATG (carril 2), la construcción que solamente tiene la secuencia Kozak separando a ambos cistrones (construcción 3) y la que tiene la región intercistronica de *P. humanus* (carril 4); pero no así aquella con la secuencia que genera un talloasa en el extremo 5' UTR inhibiendo la traducción dependiente de cap. En el carril 6 vemos células transfectadas con la luciferasa completa, mostrando que no fue posible detectar la proteína endógena. En la gráfica de abajo se muestra la luminiscencia de estas mismas muestras únicamente en el carril 1 y 5 cómo se había mencionado anteriormente.

6.6 Detección de las proteínas Dmp8 y Dmp18 en larvas

Considerando que Dmp18 podía prescindir de la presencia de un cap para lograr su traducción eficiente, nos preguntamos si era esta la razón por la cual resultaba ventajoso para su traducción encontrarse junto a Dmp8.

Se sabe que en situaciones de estrés celular la proteína encargada de reconocer al cap de los mensajeros es secuestrada⁴⁰. De esta manera se inhibe la traducción dependiente de cap. En esas condiciones extremas a las que fueron sometidas las larvas, las proteínas que son necesarias para contender con el daño recurren a mecanismos independientes de cap³⁴. Esto nos llevó a preguntarnos si

Dmp18 se mantuvo en una organización genética bicistrónica con Dmp8 para asegurar su traducción en situaciones de estrés. En estudios sobre la traducción de proteínas en moscas, se ha visto que el factor de inicio de la traducción, eIF4F, es inhibido en situaciones de estrés celular. La falta de este factor normalmente lleva a apoptosis celular. Sin embargo, cuando ocurre daño en las células, se traducen mensajeros de manera independiente del cap, es decir, que no requieren del factor eIF4F, mediante un mecanismo de traducción tipo IRES³⁴.

Para probar esta hipótesis, sometimos larvas de tercer instar a estrés por choque térmico y por radiación ultravioleta acoplados a ensayos tipo Western Blot para detectar diferencias en la cantidad de Dmp8 y Dmp18 en comparación con el nivel basal de cada proteína. Como se muestra en la Figura 22, es posible detectar los niveles de ambas proteínas en condiciones basales. Sin embargo, después de aplicar uno u otro tipo de estrés, la presencia de Dmp8 se reduce considerablemente, pero no ocurre así para la proteína Dmp18. Para ver si era posible que alguna situación de estrés estuviera provocando diferencias en la traducción de las proteínas de este bicistrón, se llevó a cabo un ensayo tipo Western Blot para analizar los cambios en el nivel de expresión de cada proteína al estresar a las larvas con luz UV o con choque térmico (HS) (Figura 22).

El ensayo parece mostrar que la proteína Dmp18 no sufre cambios en su nivel de expresión. Sin embargo, Dmp8 no es capaz de traducirse cuando se somete al tejido en situaciones de estrés. Esto indica una traducción diferencial entre ambas proteínas del bicistrón. Como el tejido de las glándulas salivales es alto en contenido de grasa, y esta parecía afectar la correcta detección de Dmp8, se hizo un ensayo tipo Western Blot con una extracción con TRIZOL, ya que de esta manera es más fácil eliminar todos los residuos de grasa y minimizar su impacto sobre la determinación de la expresión de Dmp8.

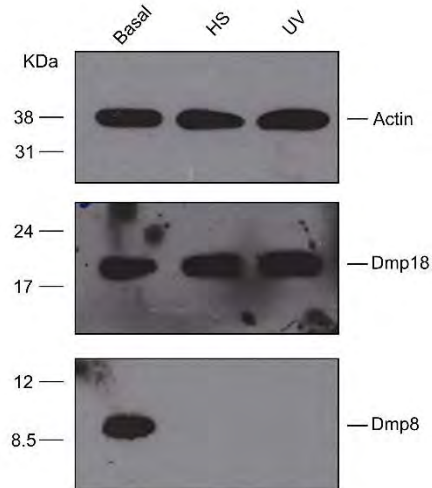


Figura 22. Detección de las proteínas del bicistrón en situaciones de estrés.

Ensayos tipos Western Blot de larvas de tercer instar en condiciones basales (primera columna), después de 1 hora de choque térmico a 37°C (segunda columna) y después de ser irradiadas con luz ultravioleta a 200 J/m². Se puede ver que en condiciones basales, ambas proteínas están presentes, sin embargo en presencia de ambos tipos de estrés, ya no es posible detectar la presencia de Dmp8. Se utilizó Actina como control de carga del experimento.

Los resultados con la extracción de TRIZOL, y ahora en larvas completas, nos indican que la traducción de la proteína Dmp18, no sólo es independiente de Dmp8, sino que está ocurriendo mediante mecanismos diferentes de traducción. Es posible que exista alguna secuencia que esté actuando como algún IRES.

6.7 Análisis del transcrito en situaciones de estrés

De las mismas larvas de las cuales se extrajo proteína para analizar la presencia de ambos miembros del bicistrón, hicimos extracción de RNA para analizar si había cambios en los niveles del transcrito mediante ensayos de RT-PCR. Como se observa en la figura 23, no hubo cambios en la presencia o cantidad del transcrito que codifica al bicistrón, lo cual nos da evidencia de que la falta de Dmp8 ocurre a nivel de traducción. Esto apoyaba nuestra hipótesis inicial de que Dmp18 podría beneficiarse de este arreglo bicistrónico.

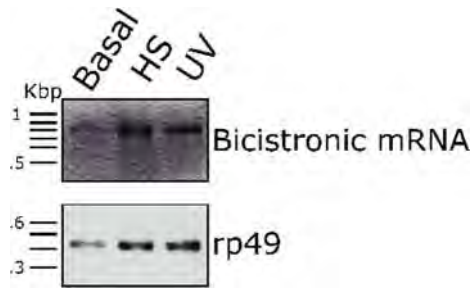


Figura 23. Análisis de transcripción del mensajero bicistrónico.

RT-PCR de las larvas sometidas a los diferentes tipos de estrés. Análisis del transcrito bicistrónico en condiciones basales (primera columna), después de 1 hora de choque térmico a 37°C (segunda columna) y después de ser irradiadas con luz ultravioleta a 200 J/m² (tercera columna). Si comparamos la cantidad del transcrito con aquella del control rp49 (gen ribosomal) se puede apreciar que no hay cambios en la cantidad de transcrito, lo que indica que la regulación no es a este nivel.

6.8 Análisis de traducción después de estrés celular

Para continuar con esta caracterización, decidimos hacer ensayos induciendo daño en las células previamente transfectadas con algunas de las construcciones disponibles. Para esto, se hicieron cinéticas de la expresión de Dmp18 en respuesta a los estímulos (Figura 24). La construcción que elegimos caracterizar fue la que tenía lo que representaba al bicistrón completo (Figura 24, ♦), que en ambos tipos de estrés resultó en un aumento de expresión posterior al estímulo. En el caso de la construcción que contaba solamente con las últimas 54 pares de bases de la región que codifica para Dmp8 (Figura 24, ▲) se comportó de manera similar al bicistrón completo en respuesta a ambos estímulos. Para corroborar que Dmp18 no necesitaba que existiera una traducción de Dmp8, evaluamos también la construcción con la estructura de tallo-asa en el extremo 5' (Figura 24, ◦). Sin embargo, la construcción que contenía a Dmp18 desde su secuencia Kozak, aunque fue capaz de traducirse, resultó incapaz de responder con un aumento de su traducción en respuesta a estímulos de estrés celular (Figura 24, ■). En conjunto, estos resultados apoyan la idea de que Dmp18 requiere de la presencia de, por lo menos, los últimos 54 nt de la región codificadora de Dmp8, pero no de la traducción de este para responder al estrés celular.

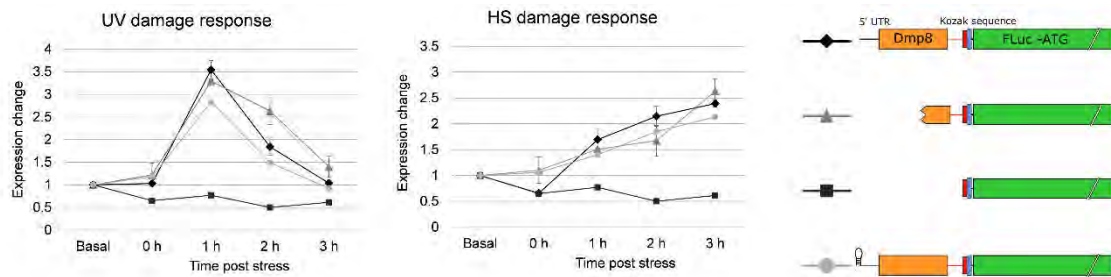


Figura 24. Cinéticas de luminiscencia en células sometidas a estrés.

Para analizar la respuesta de la traducción de Dmp18 en algunas construcciones, se sometió a estrés a las células después de ser transfectadas con estas construcciones. Los estímulos dados fueron radiación ultravioleta de 5 J/m^2 (primera gráfica) y choque térmico a 36° C . Las muestras se fueron midiendo a lo largo de 3 horas para evaluar la expresión de la luciferasa de luciérnaga. Las construcciones que contenían lo representativo al bicistrón completo (\blacklozenge), a la construcción con solo 54 pares de bases de Dmp8 (\blacktriangle) y la que tenía la secuencia que resulta en la estructura tallo-asa en el extremo 5' (\blacksquare) responden con un aumento de expresión después de recibir el estímulo, pero no así la construcción en la que sólo se incluye a Dmp18 desde su secuencia Kozak (\bullet). Se puede apreciar que cada estímulo genera una respuesta con una cinética diferente.

Esto nos pareció interesante, y nos llevó a tratar de acotar cuál era la región mínima requerida de estos 54 necesaria para llevar a cabo este efecto. Por ello, hicimos 3 construcciones más. De estas se midieron también los niveles de expresión de Dmp18 acoplado a la luciferasa de luciérnaga en condiciones basales. La primera con tan sólo los últimos 27 nt de la región codificante de Dmp8 (Figura 25, carril 5), la segunda con tan solo los 54 nt de la región intercistrónica (Figura 25, carril 6) y una última con solo la mitad de esta región (Figura 25, carril 7).

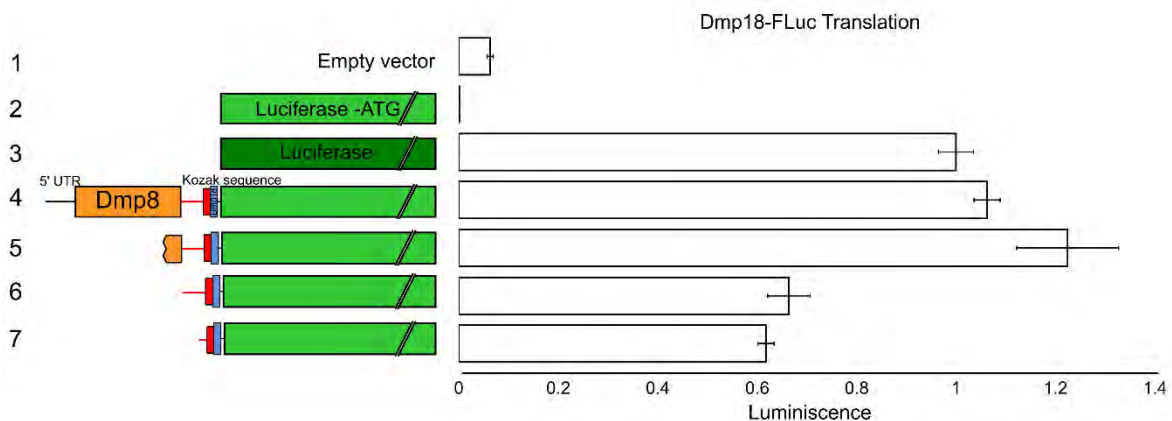


Figura 25. Análisis de traducción en regiones acotadas río arriba del sitio de inicio de traducción de Dmp18.

En un intento de acotar la región mínima necesaria para la respuesta de Dmp18 al estímulo de estrés, se hicieron construcciones con solamente 27 pares de bases de la región de Dmp8 (carril 5), con la región intercistrónica del bicistrón (carril 6) y con la mitad de esta secuencia intercistrónica (carril 7). Todas son capaces de traducir correctamente a Dmp18, como se esperaba porque en todas tiene su secuencia Kozak intacta.

Como era de esperarse porque ya habíamos visto que Dmp18 sólo requería de su secuencia Kozak para traducirse correctamente, todas estas construcciones fueron capaces de permitir la traducción de Dmp18. Ahora quisimos evaluar su respuesta a estrés celular, por lo que hicimos otras cinéticas de respuesta a estrés por choque térmico y radiación ultravioleta. Encontramos que la construcción con los últimos 27 pares de bases de la región codificante de Dmp8 respondía a estímulos de estrés (Figura 26, \blacklozenge), pero no así las construcciones que sólo tenían la región intercistrónica o parte de ella (Figura 26, \blacktriangle y \blacksquare respectivamente), corroborando una vez más que Dmp18 requiere estar pegado a Dmp8, pero no necesariamente que esta se traduzca para poder traducirse en respuesta a estrés.

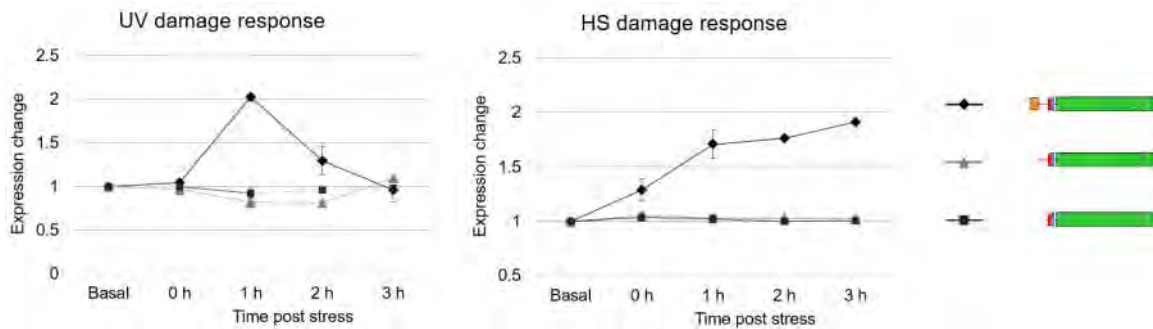


Figura 26. Cinética de respuesta a estrés por luz UV o choque térmico en regiones acotadas.

Medición de la respuesta a estrés de las construcciones con tan solo 27 nucleótidos de la región codificante de Dmp8 (\blacklozenge), la región intercistrónica completa (\blacktriangle), y la mitad de la región intercistrónica (\blacksquare). Como podemos ver, únicamente la construcción que tenía parte de la región codificante de Dmp8 fue capaz de

responder a los estímulos de estrés, pero no así las que tenían parte de la secuencia intercistrónica.

Finalmente, como nos surgió la hipótesis de que en las construcciones en las que había una región intercistrónica muy pequeña podía haber impedimento estérico para los ribosomas, a la construcción que tenía la región intercistrónica de *P. humanus* le añadimos la secuencia de la estructura de tallo-asa en su extremo 5' UTR. Lo que encontramos es que esta construcción no se traduce eficientemente al inicio (Figura 27, •), pero cuando existe estrés celular y se inhibe la traducción dependiente de cap, Dmp18 es capaz de traducirse eficientemente.

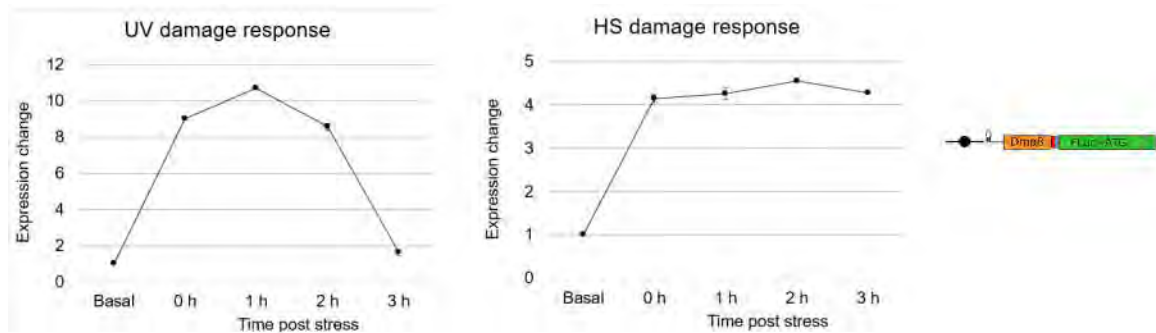


Figura 27. Cinética de traducción con región intercistrónica de *P. humanus*.

Por último evaluamos otra construcción, la cual contenía la región intercistrónica de *P. humanus*, en la que sospechábamos ocurría una inhibición de la expresión de Dmp18 por impedimento estérico, se le añadió la secuencia que resulta en tallo-asa en el 5'UTR (•). Nuestros resultados muestran que cuando se inhibe la traducción de Dmp8, hay un aumento importante en la expresión de Dmp18.

6.9 Análisis de bicistrones en *Drosophila melanogaster*

Se llevó a cabo un análisis de los bicistrones reportados en la página de FlyBase y de los reportados en el artículo de Irwin Jurngreis²⁷. Se seleccionaron las proteínas que corresponden al segundo cistrón, o en algunos casos también el tercero. De esta manera se obtuvieron 441 proteínas que se encuentran codificadas en bicistrones en *Drosophila melanogaster*. La tabla obtenida se incluye en el Anexo 1.

De estas proteínas, sólo 166 están caracterizadas y se conoce su función. Con esta información hicimos un análisis de ontología utilizando en el sistema de clasificación Panther⁴¹. Los resultados se muestran en la Figura 28.

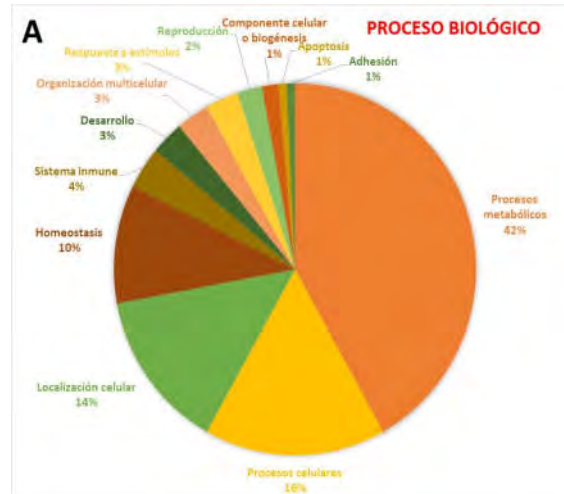


Figura 28. Posibles funciones de los segundos cistrones.

Representación de la ontología de los segundos cistrones en *Drosophila melanogaster* de las proteínas en arreglo policistrónico que están caracterizadas.

6.10 Búsqueda *in silico* de candidatos con una regulación similar

Los resultados mostrados previamente hicieron que nos preguntáramos si este fenómeno ocurría únicamente en el bicistrón Dmp8-Dmp18. Para contestar esta pregunta, seleccionamos las proteínas que correspondían al segundo cistrón, o en algunos casos también el tercero. De esta manera obtuvimos 437 proteínas que se encuentran codificadas en policistrones en *Drosophila melanogaster*. De estas proteínas, 283 correspondían a segundos cistrones. Este análisis nos llevó a seleccionar a aquellas proteínas que estuvieran relacionadas a estrés. De estas candidatas, intentamos identificar a las proteínas cuya función no dependiera del mismo mecanismo celular.

Posteriormente, intentamos identificar a las proteínas reportadas en el párrafo anterior en otros artrópodos para tratar de ver si conservaban este arreglo policistrónico. Esta búsqueda nos permitió identificar un bicistrón cuyas proteínas no están completamente caracterizadas aún, y se identifican con los símbolos CG42374 y CG9666. Un análisis de homología demostró que CG42374 corresponde a una

proteína miembro del complejo SOSS, que está involucrado en estabilidad genómica, y reparación en rotura de doble cadena⁴². Mientras que la proteína CG9666, el segundo cistrón, parece codificar a una N-6 metil-transferasa específica de adeninas, que en el caso de *Drosophila* se han visto involucradas en control de virulencia⁴³. Al igual que el bicistrón Dmp8-Dmp18, estas proteínas participan en mecanismos de respuesta a estrés, pero no participan en la misma vía. Además, los genes CG42374 y CG9666 están presentes en todos los organismos, pero solo en los artrópodos se encuentran con una organización bicistrónica. De la misma manera, están muy conservadas las proteínas, pero no así su región intercistrónica (Figura 29).

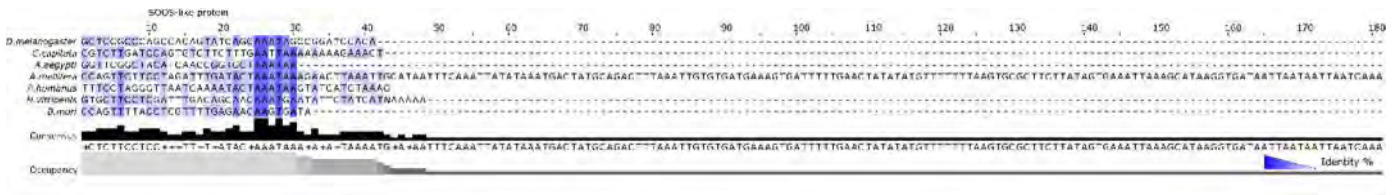


Figura 29. Alineamiento de la región intercistrónica del bicistrón CG 42374-CG9666. Al realizar un alineamiento de la región intercistrónica del bicistrón CG 42374-CG9666, se puede ver que una vez terminada la región que corresponde a la proteína, no existen regiones conservadas en su secuencia intercistrónica y no parece haber regiones en común.

VII. Discusión

Desde que se ha tenido acceso a los genomas completos secuenciados se ha demostrado que la presencia de genes en arreglos policistrónicos es más común de lo que anteriormente se creía. En la mayoría de los casos estudiados, estos requieren de elementos IRES para asegurar su correcta traducción^{27,44}. En este trabajo encontramos que el mensajero bicistrónico Dmp8-Dmp18, que se encuentra conservado en esta configuración en los artrópodos, es poco común no sólo porque la traducción de ambos parece ser independiente, sino que además parecen no requerir de la secuencia intercistrónica para su correcta traducción. Esto podría explicar por qué la distancia entre ambos marcos de lectura no se ha mantenido conservada. Estas observaciones implican que cada organismo regula de manera particular la traducción de la proteína Dmp18.

Nuestros resultados demuestran que, al menos en *Drosophila melanogaster*, es necesaria la presencia de Dmp8, ya que Dmp18 requiere de la presencia de los últimos 27 nt del marco abierto de lectura de Dmp8 antes de su inicio de traducción para permitir su correcta traducción cuando las células están sometidas a estrés. Esto tiene sentido, ya que como fue mencionado con anterioridad, ambas proteínas están involucradas en mecanismos de respuesta a estrés. La posible diferencia de regulación entre ellas puede deberse a que Dmp8 además actúa en iniciar la transcripción, proceso que es frenado cuando hay ciertos tipos de estrés^{4,6,45}. Es posible que Dmp8 y Dmp18 se mantuvieron en este arreglo porque Dmp8 es muy importante para la transcripción, y la célula cuenta con el transcrito siempre presente lo cual resultaba ventajoso a Dmp18 para permitir su traducción en situaciones de estrés, ya que sus homólogos en levadura, plantas, y humano se han visto implicados en mecanismos de respuesta a varios tipos de daño⁴⁶, como lo es su papel en el reclutamiento de p53 a sus genes blanco en respuesta a daño a DNA¹³.

Existen otros ejemplos de traducción diferencial de proteínas de un mismo bicistrón, como es el caso de un transportador de arginina/lisina, proteína involucrada en estrés nutricional. En este caso, en el momento en el que hay estrés, ocurre una inactivación del factor de inicio de traducción eIF2a, que al liberarse del mensajero resulta en la inhibición de la formación de una estructura de tallo. Esto a su vez, permite que se descubra un nuevo sitio de unión de la maquinaria de traducción y se traduzca la otra proteína del bicistrón^{47,48}.

En nuestro trabajo encontramos una contradicción en nuestros resultados, por un lado parecía que en el contexto del bicistrón, la región intercistrónica no era importante para la traducción de Dmp18, y por otro lado no logramos ver una traducción eficiente cuando sustituimos su región intercistrónica por aquella de *P. humanus*, la cual consiste de solo 8 nucleótidos. Una posible explicación es que existía un impedimento estérico causado por la proximidad de ambos marcos de lectura. Otra posible explicación es que en cada organismo opere un mecanismo diferente para la traducción de Dmp18 independiente de cap, y puede ser que en *P. humanus* solo se traduzca en respuesta a estrés, pero no en condiciones basales. Lo que podemos asumir considerando que esta configuración está conservada en artrópodos, y que la secuencia del marco de lectura de Dmp8 está muy bien conservada en todos ellos, es que Dmp18 requiere de los últimos 27 nt de la secuencia de Dmp8 para traducirse en respuesta a estrés. Esto no sería algo nuevo, hay múltiples casos de proteínas que tienen secuencias que mejoran su traducción en respuesta a estímulos^{23,49}.

Varios componentes de la maquinaria de traducción se han visto involucrados en el reconocimiento inicial para permitir el inicio de traducción de manera no canónica, haciendo que estudiar esto se vuelva sumamente complejo. En algunos casos, una estructura de tallo-asa puede inhibir la traducción³⁹, y otras veces son necesarias estas estructuras para permitir el inicio de traducción⁵⁰. Estas estructuras y secuencias se encuentran más comúnmente en el extremo 5' UTR³⁰, pero no es el caso del bicistrón Dmp8-Dmp18, que parece estar dentro del ORF de la primera proteína del bicistrón, o por lo menos los últimos 27 nucleótidos, como demostramos en este trabajo.

En general, varios mecanismos de traducción en respuesta a estrés se han identificado, lo cual amplía las opciones de regulación existentes en distintas situaciones y organismos. Mientras más se encuentre sobre esto, se abre campo para poder combatir enfermedades.

VIII. Conclusiones

- No existe ningún elemento conservado en la región intercistónica del mRNA de Dmp8-Dmp18 en los artrópodos, sugiriendo que si hay una conservación en el modo de traducción de Dmp18, no es relevante esta región.
- La presencia de las proteínas Dmp8 y Dmp18 a lo largo del desarrollo de *Drosophila melanogaster* se deben únicamente a la presencia de un mRNA.
- Se corroboró que la traducción de Dmp18 es independiente de la de Dmp8.
- En condiciones basales, Dmp18 puede traducirse con sólo su secuencia Kozak presente río arriba de su sitio de inicio de la traducción.
- En condiciones de estrés celular a causa de choque térmico y radiación ultravioleta hay traducción de la segunda proteína del bicistrón, Dmp18, pero no así de Dmp8.
- Esta traducción de Dmp18 en situaciones de estrés sólo posible cuando los últimos 27 nt del ORF de Dmp8 están presentes en la secuencia río arriba del sitio de inicio de traducción de Dmp18, aún en presencia de su secuencia Kozak.
- Es posible que este fenómeno no sea exclusivo del bicistrón Dmp8-Dmp18 en *Drosophila*, ya que se encontró otro bicistrón que comparte sus características.

VIX. Perspectivas

- Análisis de la secuencia de los 27 nucleótidos, para intentar identificar al factor asociado al reconocimiento y que posiblemente esté iniciando la traducción de la segunda proteína del bicistrón
- Usar esta secuencia en otros elementos genéticos para evaluar si es la única región necesaria para permitir la respuesta a estrés
- Estudiar al bicistrón CG42374-CG9666, para evaluar si ocurre un mecanismo similar al encontrado, o incluso si el mismo factor está reconociendo esa secuencia

X. Bibliografía

- 1 Kornberg R. The molecular basis of eukaryotic transcription. *Angew Chemie - Int Ed* 2007;46:6956–65.
- 2 Alberts B, Johnson A, Lewis J, Raff M, Roberts K. *Alberts - Molecular Biology Of The Cell 4th Ed.* 202AD doi:10.1017/CBO9781107415324.004.
- 3 Lee TI, Young R a. T Ranscription of Eukaryotic Protein- Coding Genes. *Annu Rev Genet* 2000;77–137.
- 4 Compe E, Egly J-M. TFIIH: when transcription met DNA repair. *Nat Rev Mol Cell Biol* 2012;13:343–54.
- 5 Kim TK, Lagrange T, Wang YH, Griffith JD, Reinberg D, Ebricht RH. Trajectory of DNA in the RNA polymerase II transcription preinitiation complex. *Proc Natl Acad Sci U S A* 1997;94:12268–73.
- 6 Egly JM, Coin F. A history of TFIIH: Two decades of molecular biology on a pivotal transcription/repair factor. *DNA Repair* 2011;10:714–21.
- 7 Herrera-Cruz M, Cruz G, Valadez-Graham V, et al. Physical and functional interactions between Drosophila homologue of Swc6/p18hamlet subunit of the SWR1/SRCAP chromatin-remodeling complex with the DNA repair/transcription factor TFIIH. *J Biol Chem* 2012;287:33567–80.
- 8 Singh A, Compe E, Le May N, Egly JM. TFIIH subunit alterations causing xeroderma pigmentosum and trichothiodystrophy specifically disturb several steps during transcription. *Am J Hum Genet* 2015;96:194–207.
- 9 Theil AF, Nonnekens J, Steurer B, et al. Disruption of TTDA results in complete nucleotide excision repair deficiency and embryonic lethality. *PLoS Genet* 2013;9:e1003431.
- 10 March-Díaz R, García-Domínguez M, Lozano-Juste J, León J, Florencio FJ, Reyes JC. Histone H2A.Z and homologues of components of the SWR1 complex are required to control immunity in Arabidopsis. *Plant J* 2008;53:475–87.
- 11 Berriri S, Gangappa SN, Kumar SV. SWR1 Chromatin-Remodeling Complex Subunits and H2A.Z Have Non-overlapping Functions in Immunity and Gene Regulation in Arabidopsis. *Mol Plant* 2016;9:1051–65.
- 12 Cuadrado A, Lafarga V, Cheung PCF, et al. A new p38 MAP kinase-regulated transcriptional coactivator that stimulates p53-dependent apoptosis. *EMBO J* 2007;26:2115–26.
- 13 Lafarga V, Cuadrado A, Nebreda AR. p18Hamlet mediates different p53-dependent responses to DNA damage-inducing agents. *Cell Cycle* 2007;6:2319–22.
- 14 Jackson RJ, Hellen CUT, Pestova T V. The mechanism of eukaryotic translation initiation and principles of its regulation. *Nat Rev Mol Cell Biol* 2010;11:113–

- 27.
- 15 Kozak M. Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes. *Proc Natl Acad Sci* 1990;87:8301–5.
- 16 Walsh D, Mohr I. Viral subversion of the host protein synthesis machinery. *Nat Rev Microbiol* 2011;9:860–75.
- 17 Nanda JS, Cheung YN, Takacs JE, et al. eIF1 Controls Multiple Steps in Start Codon Recognition during Eukaryotic Translation Initiation. *J Mol Biol* 2009;394:268–85.
- 18 Ryabova LA, Pooggin MM, Hohn T. Translation reinitiation and leaky scanning in plant viruses. *Virus Res* 2006;119:52–62.
- 19 Karginov TA, Pastor DPH, Semler BL, Gomez CM. Mammalian Polycistronic mRNAs and Disease. *Trends Genet* 2017;33:129–42.
- 20 Blumenthal T. Operons in eukaryotes. *Briefings Funct Genomics Proteomics* 2004;3:199–211.
- 21 Pelletier J, Sonenberg N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature* 1988;334:320–25.
- 22 Sendoel A, Dunn JG, Rodriguez EH, et al. Translation from unconventional 5' start sites drives tumour initiation. *Nature* 2017;541:494–99.
- 23 Lacerda R, Menezes J, Romão L. More than just scanning: the importance of cap-independent mRNA translation initiation for cellular stress response and cancer. *Cell Mol Life Sci* 2017;74:1659–80.
- 24 Linz B, Koloteva N, Vasilescu S, McCarthy JEG. Disruption of ribosomal scanning on the 5'-untranslated region, and not restriction of translational initiation per se, modulates the stability of nonaberrant mRNAs in the yeast *Saccharomyces cerevisiae*. *J Biol Chem* 1997;272:9131–40.
- 25 Gebauer F, Hentze MW. IRES unplugged. *Science* 2016; 351,6270:228.
- 26 Terenin IM, Smirnova V V., Andreev DE, Dmitriev SE, Shatsky IN. A researcher's guide to the galaxy of IRESs. *Cell Mol Life Sci* 2017;74:1431–55.
- 27 Jungreis I, Lin MF, Spokony R, et al. Evidence of abundant stop codon readthrough in *Drosophila* and other metazoa. *Genome Res* 2011;21:2096–2113.
- 28 Reiter LT, Potocki L, Chien S, Gribskov M, Bier E. A Systematic Analysis of Human Disease-Associated Gene Sequences In *Drosophila melanogaster*. 2001;1114–25.
- 29 Flagg RO. Carolina *Drosophila* Manual. *Carolina Biol Supply Co* 1979;36.
- 30 Weingarten-Gabbay S, Elias-Kirma S, Nir R, et al. Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. *Science* 2016;351. doi:10.1126/science.aad4939.
- 31 Antonucci L, D'Amico D, Di Magno L, et al. CNBP regulates wing development

- in *Drosophila melanogaster* by promoting IRES-dependent translation of dMyc. *Cell Cycle* 2014;13:434–39.
- 32 Andrews J, Smith M, Merakovsky J, Coulson M, Hannan F, Kelly LE. The stoned locus of *Drosophila melanogaster* produces a dicistronic transcript and encodes two distinct polypeptides. *Genetics* 1996;143:1699–1711.
- 33 Wall AA, Phillips AM, Kelly LE. Effective translation of the second cistron in two *Drosophila* dicistronic transcripts is determined by the absence of in-frame AUG codons in the first cistron. *J Biol Chem* 2005;280:27670–78.
- 34 Hernández G, Vázquez-pianzola P, Sierra JM, Va P, Rivera-pomar R. Internal ribosome entry site drives cap-independent translation of reaper and heat shock protein 70 mRNAs in *Drosophila* embryos. *RNA* 2004;10:1783–97.
- 35 Bucio-Mendez A, Cruz G, Valadez-Graham V, Dinkova D T, Zurita M. The Dmp8-Dmp18 bicistron mRNA enables unusual translation during cellular stress. *J Cell Biochem* 2018; In press.
- 36 Invitrogen. Schneider 2 (S2) cells Protocol. Version F.
- 37 Edgar RC. MUSCLE: Multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res* 2004;32:1792–97.
- 38 Hoskins RA, Landolin JM, Brown JB, et al. Genome-wide analysis of promoter architecture in *Drosophila melanogaster*. *Genome Res* 2011;21:182–92.
- 39 Babendure JR, Babendure JL, Ding J, Tsien RY. Control of mammalian translation by mRNA structure near caps. *RNA* 2006;851–61.
- 40 Clemens MJ. Translational regulation in cell stress and apoptosis. Roles of the eIF4E binding proteins. *J Cell Mol Med* 2001;5:221–39.
- 41 Mi H, Muruganujan A, Casagrande JT, Thomas PD. Large-scale gene function analysis with the PANTHER classification system. *Nat Protoc* 2013;8:1551–66.
- 42 Jun Huang, Zihua Gong, Gargi Ghosal, Chen J. SOSS complexes participate in the maintenance of genomic stability. *Mol Cell* 2010;35:384–93.
- 43 Luo G-Z, Blanco MA, Greer EL, He C, Shi Y. DNA N6-methyladenine: a new epigenetic mark in eukaryotes? *Nat Rev Mol Cell Biol* 2015;16:705–10.
- 44 Col B, Oltean S, Banerjee R. Translational regulation of human methionine synthase by upstream open reading frames. *Biochim Biophys Acta* 2009;1769:532–40.
- 45 Dubois MF, Vincent M, Vigneron M, Adamczewski J, Egly JM, Bensaude O. Heat-shock inactivation of the TFIIF-associated kinase and change in the phosphorylation sites on the C-terminal domain of RNA polymerase II. *Nucleic Acids Res* 1997;25:694–700.
- 46 Rosa M, Von Harder M, Aiese Cigliano R, Schlogelhofer P, Mittelsten Scheid O. The Arabidopsis SWR1 Chromatin-Remodeling Complex Is Important for DNA Repair, Somatic Recombination, and Meiosis. *Plant Cell* 2013;25:1990–2001.

- 47 Fernandez J, Yaman I, Huang C, et al. Ribosome stalling regulates IRES-mediated translation in eukaryotes, a parallel to prokaryotic attenuation. *Mol Cell* 2005;17:405–16.
- 48 Yaman I, Fernandez J, Liu H, et al. The zipper model of translational control: A small upstream ORF is the switch that controls structural remodeling of an mRNA leader. *Cell* 2003;113:519–31.
- 49 Wellensiek BP, Larsen AC, Stephens B, et al. Genome-wide profiling of human cap-independent translation-enhancing elements. *Nat Methods* 2013;10:747–50.
- 50 Paek KY, Hong KY, Ryu I, et al. Translation initiation mediated by RNA looping. *Proc Natl Acad Sci* 2015;112:1041–46.
- 51 Gramates LS, Marygold SJ, Dos Santos G, et al. FlyBase at 25: Looking to the future. *Nucleic Acids Res* 2017;45:D663–71.

Anexo 1. Tabla de policistrones reportados en *Drosophila melanogaster*

En esta tabla se reporta la compilación de los genes que se encuentran en arquitectura policistrónica en la mosca *Drosophila melanogaster*.

La información fue tomada de la base de datos reportada en la página FlyBase⁵¹ y en el artículo de Jungreis et al. sobre los reconocidos con Stop Codon Readthrough²⁷.

Se muestra el número con el que está anotado el gen, el nombre con el que se le reconoce (si es el caso), si es que tiene una función conocida y los genes que se encuentran asociados al mismo elemento.

| ID Cistrón 1 | Nombre | Función | ID Cistrón 2 | Nombre | Función |
|--------------|----------------------------|--|--------------|-------------------------------|--|
| CG10097 | - | Reductasa de fati-acil CoA | CG10096 | - | Reductasa de fati-acil CoA |
| CG10332 | - | - | CG33706 | Immune induced mol 18 | Involucrado en respuesta inmune |
| CG11181 | cup | transportador de eI4E | CG34310 | - | - |
| CG11714 | - | Dominio de dedo de Zn | CG6128 | α -L-fucosidase | Alfa-L-fucosidasa |
| CG11768 | - | - | CG34135 | - | - |
| CG11889 | - | Dominio de fosfo-transferasa | CG11891 | - | Dominio de fosfo-transferasa |
| CG12500 | stoned A | Regulador de sinapsis | CG12473 | stoned B | Regulador de sinapsis |
| CG12665 | Odorant-binding protein 8a | Percepción feromonas | CG15368 | - | Dedo de Zinc |
| CG12912 | - | - | CG33183 | Hormone receptor-like in 46 | reacciona a Ecdisona |
| CG12913 | - | - | CG46321 | - | - |
| CG13175 | - | - | CG33964 | - | - |
| CG13380 | - | - | CG4174 | - | Oxidoreductasa en homeostasis |
| CG13510 | - | Factor α de necrosis por LPS | CG13511 | - | Factor α de necrosis inducido por LPS |
| CG13511 | - | Factor α de necrosis por LPS | CG42565 | - | Factor α de necrosis inducido por LPS |
| CG14042 | - | - | CG14041 | SP555 | Importante en señalización |
| CG1410 | waw | GTPasa en traducción | CG1414 | bobby sox | Activador de transcripción |
| CG14164 | - | Dominio de unión a anti-codón | CG6709 | - | Dominio de polimerización de tubulina |
| CG14208 | Tyler | Transporte transmembranal mitocondria | CG14209 | Shawn | Transporte transmembranal mitocondria |
| CG14933 | - | Dominio de inhibidor de proteasa (Ser) | CG42486 | - | Dominio de inhibidor de proteasa (Ser) |
| CG14987 | Gustatory receptor 64d | Percepción sensorial (azúcares) | CG32258 | Gustatory receptor 64e | Percepción sensorial (azúcares) |
| CG15465 | - | - | CG44835 | rugose | unión a proteína cinasa A |
| CG15717 | - | Oxidoreductasa | CG4330 | Major Facilitator Transporter | Transportador transmembranal |
| CG1669 | K β -Ras | GTPasa | CG30503 | - | Catabolismo de lípidos |
| CG16944 | stress-sensitive B | Muchas cosas... | CG1683 | Nucleotide translocase 2 | Antiportador ATP/ADP |
| CG17139 | - | Canal (porina) | CG17140 | - | Canal (porina) |
| CG17159 | - | - | CG17162 | - | - |
| CG17197 | - | Palmitoil-transferasa de cisteína | CG17198 | - | Palmitoil-transferasa de cisteína |
| CG17226 | Odorant receptor 59c | Familia de quimiorreceptores | CG3569 | Odorant receptor 59b | Familia de quimiorreceptores |
| CG17767 | Tim9b | Translocasa, membrana interna | CG12788 | - | Dominio de asociación a la cromatina |
| CG17770 | - | Unión de calcio | CG17769 | Androcam | De la familia de miosina y calmodulina |
| CG17867 | Odorant receptor 10a | Percepción sensorial (olfativo) | CG32664 | Gustatory receptor 10a | Percepción sensorial (gustativo) |

| | | | | | |
|---------|-----------------------------|--|---------|------------------------------|--|
| CG18065 | - | - | CG13430 | - | Endopeptidasa de serina |
| CG18066 | Cuticular protein 57A | Estructura de quitina | CG30150 | Odorant-binding protein 57d | Percepción sensorial (a feromonas) |
| CG18170 | - | - | CG33791 | - | Oxoglutarato deshidrogenasa mitocondria |
| CG2217 | - | - | CG42740 | - | - |
| CG30073 | Odorant-binding protein 50b | Percepción sensorial | CG30072 | Odorant-binding protein 50c | Percepción sensorial (a feromonas) |
| CG30080 | - | - | CG42662 | - | - |
| CG30273 | - | Factor de necrosis alfa en LPS | CG30269 | - | Factor de necrosis alfa en LPS |
| CG30363 | comas sola | Espermatogénesis | CG30366 | swift-tuttle | Espermatogénesis |
| CG30364 | hug-bell | Espermatogénesis | CG30365 | spacewatch | Espermatogénesis |
| CG30479 | - | - | CG30480 | - | - |
| CG30481 | mitochondrial ribosomal L53 | Proteína ribosoma mitocondria | CG33155 | - | Hematopoesis de linfa en larva |
| CG30497 | - | - | CG45093 | - | - |
| CG31358 | - | Familia de Estomatina | CG42504 | - | - |
| CG31358 | - | Familia de Estomatina | CG42505 | - | - |
| CG31381 | - | Dimetil-alil-trasferasa de tRNA | CG11089 | - | Involucrada en curación de heridas |
| CG31542 | - | - | CG34277 | - | - |
| CG31975 | - | Transferencia de fósforo | CG31976 | - | Posible dominio de unión a lípidos |
| CG32152 | - | Metil-transferasa de arginina | CG5151 | - | - |
| CG32257 | Gustatory receptor 64b | Percepción sensorial (gustativo, azúcares) | CG32256 | Gustatory receptor 64c | Percepción sensorial (gustativo, azúcares) |
| CG32270 | - | Actividad de serina endopeptidasa | CG32269 | - | Actividad de serina endopeptidasa |
| CG32298 | - | - | CG32299 | - | - |
| CG32388 | - | - | CG43439 | - | - |
| CG32736 | - | - | CG42308 | - | - |
| CG33057 | - | Splicing de tRNA | CG7163 | monkey king protein | Nucleotidil transferasa |
| CG33096 | - | α/β hidrolasa de membrana | CG33095 | - | - |
| CG33105 | p24-related-2 | Transportador de membrana | CG31332 | - | Componente de nucleoplasma |
| CG33182 | Kdm4B | Desmetilasa de histona | CG32904 | sequoia | Morfogénesis cerebral |
| CG33182 | Kdm4B | Desmetilasa de histona | CG17724 | - | - |
| CG33261 | Trithorax-like | Compensación de dosis | CG42507 | - | - |
| CG33348 | Chemosensory protein B 42a | Regulador de canales | CG33349 | pickpocket 25 | Canal de sodio, respuesta a feromonas |
| CG33649 | - | tRNA amidotransferasa | CG33650 | DNA polymerase γ 35kD | Replicación |
| CG33670 | Stargazing-like | regulador de canal neuronal | CG11566 | - | - |

| | | | | | |
|---------|------------------------|------------------------------------|---------|------------------------------|---|
| CG33672 | - | - | CG33671 | - | Metabolismo de mevalonato |
| CG33677 | Like Sm protein 4 | Spliceosoma mRNA alternativo | CG17768 | - | Unión a snRNA |
| CG33714 | SLIRP1 | Unión a miRNA | CG33713 | - | Unión a fati-acil CoA |
| CG33718 | Pmi | Esencial para mitocondria | CG33717 | Peptidoglycan recognition LD | N-acetylmuramoil-L-alanina amidasa |
| CG33722 | - | - | CG18749 | - | co-factor en reacciones de reducción |
| CG33747 | primo-2 | Fosfatasa de proteínas (tirosina) | CG33748 | primo-1 | Fosfatasa de proteínas (tirosina) |
| CG33756 | Gonadal | Oogénesis y espermatogénesis | CG33755 | gdl-ORF39 | Espermatogénesis |
| CG33785 | - | Dirige a la pol III | CG33786 | - | Regulación de traducción |
| CG33931 | Rpp20 | Procesamiento de tRNA | CG33932 | - | Transmembrane Fragile-X-F (enfermedad) |
| CG33933 | Indy2 | Transportador piruvato y citrato | CG33934 | - | Transportador de sodio |
| CG33935 | mei-217 | Recombinación meiotica | CG8923 | meiotic 218 | Recombinación meiotica |
| CG33981 | - | Regulación del ciclo celular | CG6355 | fab1 kinase | Fosfatasa en endocitosis y señalización |
| CG33995 | - | - | CG44000 | - | - |
| CG34058 | pickpocket 11 | Canal de Na, respuesta a feromonas | CG34059 | pickpocket 16 | Canal de Na, respuesta a feromonas |
| CG34106 | - | - | CG12621 | beat-IIIa | Involucrado en adhesión celular |
| CG34131 | Muted | Biogénesis de lisosoma | CG7071 | - | Factor de ensamblaje de ER, ATPasa |
| CG34140 | - | Síntesis de Pseudouridina | CG34099 | MAP kinase phosphatase | Fosfatasa de MAP cinasa |
| CG34177 | - | - | CG34178 | - | - |
| CG34179 | - | - | CG9131 | slowmo | Dif. oocitos, espermatogénesis, peristalsis |
| CG34235 | - | - | CG30058 | - | - |
| CG34235 | - | - | CG34312 | - | - |
| CG34313 | - | - | CG31743 | - | Sulfotransferasa de carbohidratos |
| CG34314 | SWIM domain containing | Recombinación homóloga DSB | CG3240 | Radiation insensitive 1 | Checkpoint de ciclo celular |
| CG34315 | - | - | CG4088 | latheo | Reconocimiento de origen en replicación |
| CG34317 | - | Ribonucleasa de tRNA | CG7950 | - | Dominio CHCH |
| CG34440 | lemming A | Regulador de ciclo celular | CG34441 | lemming B | Dominio similar a dedo de Zn |
| CG3481 | Adh | Alcohol-deshidrogenasa | CG3484 | Adh-related | Actúa como NAD |
| CG3849 | lasp | Ensamblaje de músculo | CG43954 | - | - |
| CG3874 | fringe connection | Transporte de azúcares | CG32174 | - | Biosíntesis de ubiquinona |
| CG42245 | - | - | CG1244 | MEP-1 | Reg negativa de transcripción |
| CG42303 | snup | Unión a snRNA | CG42304 | - | - |
| CG42305 | - | - | CG17325 | - | - |

| | | | | | |
|---------|----------------|------------------------------|---------|-----------------------------|---|
| CG42306 | - | - | CG5469 | GDI interacting protein 3 | Dominio de ubiquitina |
| CG42307 | - | - | CG8601 | mutagen-sensitive 312 | Reparación, recombinación, replicación |
| CG42350 | Modigliani | Capping del telomero | CG44890 | Trimethylguanosine synthase | Metil-transferasa en capping |
| CG42353 | - | - | CG42354 | - | - |
| CG42362 | - | - | CG42363 | - | - |
| CG42362 | - | - | CG42690 | - | - |
| CG42371 | - | - | CG15386 | - | - |
| CG42372 | - | NADH deshidrogenasa | CG30100 | - | Factor de liberación en traducción |
| CG42373 | Dmp8 | Componente de TFIIF | CG31917 | Dmp18 | ¿Componente de Tip60? |
| CG42374 | - | Subunidad de complejo SOSS-C | CG9666 | - | DNA metil-transferasa |
| CG42374 | - | Subunidad de complejo SOSS-C | CG45081 | - | - |
| CG42375 | - | Dominio de citocromo oxidasa | CG9288 | - | Dominio de proteínas H cora glicina |
| CG42381 | - | Dominio hairpin de cisteína | CG9865 | Phosphatidylinositol glycan | Manosil transferasa |
| CG42381 | - | Dominio hairpin de cisteína | CG42379 | - | - |
| CG42381 | - | Dominio hairpin de cisteína | CG42380 | - | Ubiquinona de NADH deshidrogenasa |
| CG42384 | tarsal-less 1A | Organizador de actina | CG42385 | - | Morfogénesis de epitelio, regula actina |
| CG42384 | tarsal-less 1A | Organizador de actina | CG42386 | - | Morfogénesis de epitelio, regula actina |
| CG42384 | tarsal-less 1A | Organizador de actina | CG42387 | - | Morfogénesis de epitelio, regula actina |
| CG4241 | att-ORFA | Transportador en defensa | CG33488 | att-ORFB | - |
| CG42455 | - | - | CG11989 | variable nurse cells | Alfa-N-Acetiltransferasa de péptidos |
| CG42456 | - | Biosíntesis de manosa | CG12016 | - | Ribosil nicotinamida (vit B3) cinasa |
| CG42487 | - | - | CG4884 | - | Dominio de unión a RNA |
| CG42488 | - | - | CG5116 | - | GTP trifosfato hidrolasa |
| CG42496 | - | Dominio de citocromo oxidasa | CG30290 | PPCDC | Flovoproteína (FAD) |
| CG42497 | - | - | CG9878 | Translocase of inner mem 10 | Importante en respuesta inmune |
| CG42498 | - | Familia CTAG/Pcc1 | CG14550 | - | Síntesis de GPI |
| CG42502 | - | - | CG10570 | - | Inhibición lateral (desarrollo) |
| CG42503 | - | - | CG10238 | Molybdenum cofactor 2 | Actividad de sintasa de molibdeno |
| CG42511 | - | Ensamblaje de cilio | CG8135 | - | Similar a proteína de membrana |
| CG42512 | - | - | CG32573 | - | - |
| CG42516 | - | Parte de TFIIC | CG42515 | PSEA-binding protein 49kD | Complejo de activación de snRNA |
| CG42517 | Med9 | Componente del mediador | CG42518 | - | Componente del mediador |

| | | | | | |
|---------|------------------------|--------------------------------|---------|-------------------------|---|
| CG42557 | - | - | CG42558 | - | - |
| CG42559 | - | - | CG42560 | - | - |
| CG42567 | DnaJ-like-60 | Espermatogénesis | CG42568 | - | - |
| CG42599 | PFTAIRE-interacting 1A | Unión a ADN en transcripción | CG33720 | PFTAIRE-interacting 1B | Unión a proteínas en transcripción |
| CG42630 | Ccdc56 | Esencial para mitocondria | CG42631 | Mitochondrial Factor B1 | Metil-transferasa de rRNA |
| CG42634 | - | - | CG42635 | - | - |
| CG42660 | - | - | CG16991 | Tetraspanin 66A | Proteína de membrana |
| CG42660 | - | - | CG42661 | - | - |
| CG42726 | - | Dedo de Zn | CG42727 | - | Metil-transferasa S-adenosil-L-metionina |
| CG42730 | B9d2 | Función en transición ciliar | CG42731 | tectonic | Vía de regulación de hedhehog |
| CG42734 | Ankyrin 2 | Citoesqueleto de cerebro | CG44195 | - | - |
| CG42837 | - | Regulador de transporte axonal | CG44243 | - | lipoil-transferasa |
| CG43194 | - | - | CG43210 | - | - |
| CG43195 | - | - | CG43195 | - | - |
| CG43200 | - | - | CG43172 | - | - |
| CG43200 | - | - | CG46305 | - | - |
| CG43201 | - | - | CG43171 | - | - |
| CG43201 | - | - | CG43178 | - | * Señal de localización nuclear* |
| CG43325 | - | - | CG43326 | - | - |
| CG43327 | - | - | CG43328 | - | - |
| CG43345 | - | Dominio de fosfolipasa | CG43346 | - | - |
| CG43355 | - | - | CG4922 | spalt-adjacent | Importante en la gastrulación |
| CG43397 | - | - | CG12902 | - | - |
| CG43675 | - | - | CG45050 | - | Dedo de Zn |
| CG43731 | - | - | CG17344 | - | - |
| CG43732 | - | - | CG42613 | - | Receptor de lipoproteína de baja densidad |
| CG43739 | - | - | CG9339 | skywalker | RAB GTPase proteína activadora |
| CG43775 | - | - | CG43776 | - | - |
| CG43775 | - | - | CG43777 | - | - |
| CG43780 | - | - | CG43781 | - | - |
| CG43816 | - | - | CG13245 | - | - |
| CG43999 | - | Familia de OPA3 | CG43998 | - | Familia de OPA3 |

| | | | | | |
|---------|---------------------|----------------------------------|---------|--------------------------------|--|
| CG44194 | - | Cadena respiratoria mitocondria | CG17327 | - | Aminoacil-tRNA sintetasa |
| CG44434 | - | - | CG44435 | - | - |
| CG4456 | Hsp67Bb | Respuesta a choque térmico | CG4460 | Heat shock protein 22 | Respuesta a choque térmico |
| CG4468 | Xport | Chaperona en foto-transducción | CG42508 | - | - |
| CG45064 | - | - | CG45065 | - | Glucosa-metanol-colina oxidoreductasa |
| CG45068 | - | GPI-manosil transferasa | CG45069 | - | - |
| CG45075 | - | - | CG45074 | Kruppel homolog 1 | regulador de transcripción, cerebral |
| CG45084 | - | - | CG4824 | Bicaudal C | Prot de unión a mRNA en desarrollo |
| CG45085 | - | - | CG8026 | - | Transportador de ac fólico |
| CG45086 | - | - | CG17759 | G protein α q subunit | Proteína G heterotrimérica |
| CG45087 | - | - | CG17725 | PEP carboxykinase | Gluconeogénesis |
| CG45088 | - | - | CG6191 | - | Regulación de ciclo celular, fagocitosis |
| CG45089 | - | - | CG18624 | NADH dehydrogenase MNLL | NADH a ubiquinona en mitocondria |
| CG45091 | Sarcolamban A | Regulación del músculo cardiaco | CG45091 | Sarcolamban B | Regulación del músculo cardiaco |
| CG45092 | - | - | CG3409 | - | Transportador de ac monocarboxílico |
| CG45099 | - | - | CG6015 | - | Splicing de mRNA |
| CG45100 | - | - | CG2128 | Histone deacetylase 3 | Desacetilasa de histonas |
| CG45101 | - | - | CG7869 | Supp of Under-Replication | Organiza cromatina, papel en replicación |
| CG4542 | xiantuan | Glucosil transferasa | CG4547 | Ataxin 1 | Asociado a gen en ataxia humana |
| CG46244 | - | - | CG10443 | Leukocyte-antigen-related-like | Formación de axones en embrión |
| CG4975 | - | - | CG4984 | - | Componente de membrana (claudina) |
| CG5189 | - | - | CG30120 | - | - |
| CG5317 | Rpl7-like | Proteína ribosomal | CG34164 | - | - |
| CG6294 | - | Zn metalopeptidasa | CG6299 | - | Transportador de glicolípidos |
| CG6715 | KP78a | Serina-Treonina cinasa | CG17216 | KP78b | Serina-Treonina cinasa |
| CG6726 | - | Aminoacilasa | CG17110 | - | Aminoacilasa |
| CG6783 | Fabp | Transportador unión a ác. graso | CG6782 | scheggia | Transportador de citrato en mitocondria |
| CG6813 | - | Factor de transcripción, dedo Zn | CG6808 | - | Factor de transcripción, dedo de Zn |
| CG7215 | - | Dominio de ubiquitina | CG7217 | Peroxiredoxin 5 | Oxidoreductasa en homeostasis |
| CG7938 | Serendipity β | Factor de transcripción, dedo Zn | CG17957 | Serendipity α | Unión de actina, formación actinmiosina |
| CG7964 | Menl-1 | Metabolismo de malato | CG7969 | Malic enzyme like-2 | Metabolismo de malato |
| CG8788 | - | - | CG44286 | - | - |

| | | | |
|--------------------------------|---------|------------------------------|--------------------------------------|
| Recombinación meiotica | CG33935 | meiotic 217 | Recombinación meiotica |
| Endopeptidasa espermatogénesis | CG30461 | - | - |
| - | CG9133 | - | - |
| Homeostasis/Sinapsis | CG9960 | - | Metil-transferasa |
| - | CG1220 | Kaz1-ORFB | Inhibidor de endopeptidasa de serina |
| | CG43736 | - | Motivos ricos en YVTN |
| | CG43743 | Glutamate receptor IB | Receptor de glutamato ionotrópico |
| | CG43745 | Myosuppressin receptor 2 | Receptor de Miosupresina |
| | CG43759 | - | Responsable de agresión en machos |
| | CG43770 | Sex lethal | Determinación sexual |
| | CG43934 | Hr4 | Responde a Ecdisona |
| | CG34360 | - | regulador de transcripción |
| | CG46149 | Fatty acid transport protein | Proteína de membrana |

Anexo 2. Artículos publicados durante el doctorado

1. Bucio-Mendez A, Cruz-Becerra G, Valadez-Graham V, Dinkova TD, Zurita M. The Dmp8-Dmp18 bicistron messenger RNA enables unusual translation during cellular stress. *Journal of Cellular Biochemistry*, 2018, Early View DOI: 10.1002/jcb.27670.
2. Cruz-Becerra G, Valerio-Cabrera S, Juárez M, Bucio-Mendez A, Zurita M. TFIIH localization is highly dynamic during zygotic genome activation in *Drosophila*, and its depletion causes catastrophic mitosis. *Journal of Cell Science*, 2018, 131, jcs211631 DOI:10.1242/jcs.21163

The Dmp8-Dmp18 bicistron messenger RNA enables unusual translation during cellular stress

Alyeri Bucio-Mendez¹ | Grisel Cruz-Becerra¹ | Viviana Valadez-Graham¹ |
Tzvetanka D. Dinkova² | Mario Zurita¹ 

¹Department of Developmental Genetics and Molecular Physiology, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Cuernavaca, Morelos, México

²Department of Biochemistry and Molecular Biology, Facultad de Química, Universidad Nacional Autónoma de México, Ciudad de México, México

Correspondence

Mario Zurita, Department of Developmental Genetics and Molecular Physiology, Instituto de Biotecnología, Universidad Nacional Autónoma de México, Av Universidad 2001 Cuernavaca 62250, Morelos, México.
Email: marioz@ibt.unam.mx

Funding information

Consejo Nacional de Ciencia y Tecnología, Grant/Award Numbers: 250588, 219673; Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica, Grant/Award Number: IN-200315

Abstract

Alternatives to the cap mechanism in translation are often used by viruses and cells to allow them to synthesize proteins in events of stress and viral infection. In *Drosophila* there are hundreds of polycistronic messenger RNA (mRNA), and various mechanisms are known to achieve this. However, proteins in a same mRNA often work in the same cellular mechanism, this is not the case for *Drosophila's* Swc6/p18^{Hamlet} homolog Dmp18, part of the SWR1 chromatin remodeling complex, who is encoded in a bicistronic mRNA next to Dmp8 (Dmp8-Dmp18 transcript), a structural component of transcription factor TFIIH. The organization of these two genes as a bicistron is conserved in all arthropods, however the length of the intercistronic sequence varies from more than 90 to 2 bases, suggesting an unusual translation mechanism for the second open reading frame. We found that even though translation of Dmp18 occurs independently from that of Dmp8, it is necessary for Dmp18 to be in that conformation to allow its correct translation during cellular stress caused by damage via heat-shock and UV radiation.

KEYWORDS

bicistron mRNA, *Drosophila*, stress, TFIIH, translation

1 | INTRODUCTION

Many advances have been made in recent years to understand translation mechanisms in eukaryotic cells. Since translation is the most energetically expensive process in the cell, having many ways to regulate the synthesis of a given protein helps to avoid a waste of energy.^{1,2} Failure to regulate translation often results in diseases such as cancer.^{2,3} These regulatory mechanisms vary from simple ones such as adjusting messenger stability through the size of the poly(A) tail^{4,5} and alternative codon usage to regulate transfer RNA affinity for the ribosome.^{6,7} It is known that viruses use polycistronic arrays to improve the efficiency of their proteins' translation at the expense of the host in which

they reside, and the mechanisms by which they achieve this are fairly well understood.⁸⁻¹⁰

Currently, the existence of several polycistronic messenger RNA (mRNA) has been identified in a series of eukaryotic organisms, including *Drosophila*, that has approximately 407 genes in this configuration.^{11,12} Several mechanisms have been described to achieve polycistronic translation in viruses, and now it is clear that the cellular machinery also takes advantage of these noncanonical methods of translation. For instance, various forms of leaky scanning mechanisms often involve bypassing of a first nonfavorable AUG to produce at least two proteins with a shared C-terminus.¹³ Another mechanism used by both viruses and eukaryotic cells is stop codon readthrough, which involves reinitiation of

the translation machinery resulting in a second protein with extra amino acids that are later removed by endopeptidases, encoded by viral or cellular genomes.¹¹ A well understood method to achieve polycistronic translation involves recruitment of the translation machinery internally to sites in the mRNA other than the 5'-capped end. Such mechanisms are designated internal ribosome entry sites (IRES) and are widely used by viruses to recruit ribosomes to their mRNAs.¹⁰ Intense studies on a variety of IRES behaviors revealed that alternative ribosome recruitment displays great versatility.¹⁴ For instance, it is known that several proteins of the translation initiation complex participate in achieving translation initiation at the 5'-untranslated region (UTR) of the mRNA,^{3,15} at the intercistronic region,¹⁶ or even at the 3'-UTR.¹⁷

One of the polycistronic arrays present in the fly genome is the Dmp8 and Dmp18 bicistronic mRNA.¹⁸ This bicistronic messenger comprises Dmp8, which is a structural part of the core complex of transcription factor II (TFIIH), that has an important role not only in DNA transcription, but also in the DNA damage response to UV radiation,¹⁹ whereas Dmp18 is a homolog of Swc6/p18^{Hamlet}, a component of the chromatin remodeling complex SWR1 responsible for histone H2A.Z variant deposition in nucleosomes that are involved in DNA repair, somatic recombination, and meiosis and was also found to participate in the p38 repair response to different stress conditions.²⁰⁻²² Even though both, Dmp8 and Dmp18, seem to partake in DNA repair and that a physical and genetic interaction has been proven between them, the molecular function of this interaction remains unknown. Furthermore, this bicistronic configuration is maintained in all of the Arthropoda phylum, but not in other organisms, although they have the genes that code for both proteins.¹⁸

Our previous work with this bicistronic messenger in *Drosophila* focused on demonstrating a functional interaction between its two products. We provided evidence that Dmp18 translation can occur independently from that of Dmp8, since we found that in a mutant fly with a transposon inserted in the Dmp8 coding region, there was a truncated nonfunctional transcript for Dmp8 but Dmp18 translation remained unaffected.¹⁸ This led us to the belief that both proteins can achieve translation independently from each other, especially because throughout the whole *Drosophila* development, only one transcript of this bicistronic messenger is present.¹⁸ This is fairly unusual for proteins encoded within the same mRNA.

In this study, we aimed at understanding why Dmp8 and Dmp18 are conserved in a bicistronic configuration in the Arthropoda phylum. Are there any advantages for

either protein to maintain this genomic configuration? We demonstrated that the translation of both is indeed independent from each other, and we shed some light as to why this configuration succeeded in Arthropoda, since without it, Dmp18 would not be able to be translated in response to cellular stress.

2 | MATERIALS AND METHODS

2.1 | Fly strains

Fly stocks were maintained at 25°C on standard food. Oregon was used as the wild-type strain.

2.2 | Antibodies

Antibodies against Dmp8 and Dmp18 reported in Herrera-Cruz et al¹⁸ were used. Monoclonal antibodies against β -tubulin (E7) and actin (JLA20) were obtained from the Developmental Studies Hybridoma Bank.

2.3 | Transcription/translation assays

Complete Dmp8-Dmp18 cDNA and variant constructions were cloned in the pBluescript II KS(+) vector with the selected oligonucleotides with the *NotI*, *Sall* sites (see Supporting Information Table 1). Each construction was used in the one-step in vitro transcription and translation TNT system (Promega, Fitchburg, WI) in the presence of [³⁵S]-methionine, according to the product protocol and later separated in an sodium dodecyl sulfate-polyacrylamide gel electrophoresis followed by autoradiography detection.

2.4 | Protein detection

Proteins from the selected developmental stage were obtained with the TRIZOL reagent (Invitrogen, Merelbeke, Belgium) according to the product protocol, separated in an sodium dodecyl sulfate-polyacrylamide gel electrophoresis and analyzed by immunoblotting using standard procedures. Immunodetection with the corresponding antibodies: rabbit polyclonal anti-Dmp8 (1:1000), rabbit polyclonal anti-Dmp18 (1:1000), mouse monoclonal anti- β -tubulin (1:3000), and mouse monoclonal anti-actin (1:3000). Secondary antibodies with horseradish peroxidase allowed autoradiography.

2.5 | Cell culture and transfection assays

Drosophila S2R+ cell were maintained at 25°C in Schneider culture medium supplemented with antibiotics, and 10%

fetal bovine serum. A modification to the pAc5.1/V5-His A vector was performed with oligonucleotides to change its polylinker. Reporter vectors were performed by adding firefly luciferase with and without its start codon, and *Renilla* luciferase (Supporting Information Figure 2). Different constructions were added between *Sac*II and *Eco*RI sites in the vector with firefly luciferase without its ATG (see Supporting Information Table 1); S2R+ cells were independently transfected with each of the constructions for transient expression by means of calcium phosphate as indicated in the protocol of the cells (Invitrogen catalog no. R690-07 version F; Invitrogen). After 48 hours, the cells were harvested and treated with the Dual-Glo Luciferase Assay System according to the product protocol (Promega), and the activity was measured in a Monolight 3010 Luminometer (BD Biosciences, San Jose, CA).

2.6 | Stress response experiments

Wild-type larvae were exposed to heat-shock (HS) at 42°C for 1 hour, 200 J/m² of UV radiation in a UV Stratalinker 2400 (LA JOLLA, CA), or with no stimuli, and protein extracts were obtained after 1 hour of the stimuli.

In the cell transfections, the selected constructions were individually transfected in 35 × 10 mm cell culture plates, and individually subjected to either 1 hour HS at 36°C, or 5 J/m² of UV radiation. Samples were taken prior the stimuli, immediately after and then samples taken each hour for 3 hours.

3 | RESULTS

3.1 | The intergenic sequence present in the Dmp8/Dmp18 bicistron in the Arthropoda phylum is not conserved between species

Since the bicistronic Dmp8/Dmp18 gene organization is present in all of the sequenced genomes of the Arthropoda phylum and both CDS are very well conserved,¹⁸ we first decided to analyze if its intergenic region contains a signature or conserved motif, in an attempt to identify a possible regulatory element in it. We performed a sequence alignment of these intergenic regions using multiple sequence alignment with MUSCLE (MUltiple Sequence Comparison by Log-Expectation)²³ (Figure 1) in which we found that not only is the difference in size more clearly visualized (varying from 2 to 92 nucleotide [nt]) but there is also no sign of a motif or conserved sequence in them. An interesting aspect to take into account is that, in general, the intergenic sequence is A/T rich, as is the final coding region of Dmp8. If we assume that in all these arthropods both proteins are translated from this bicistronic messenger, then this result suggests that the translation of the second open reading frame (ORF) does not require the intergenic region, even though the bicistronic configuration has been maintained during the arthropod evolution. However, species-specific regulation is possible.

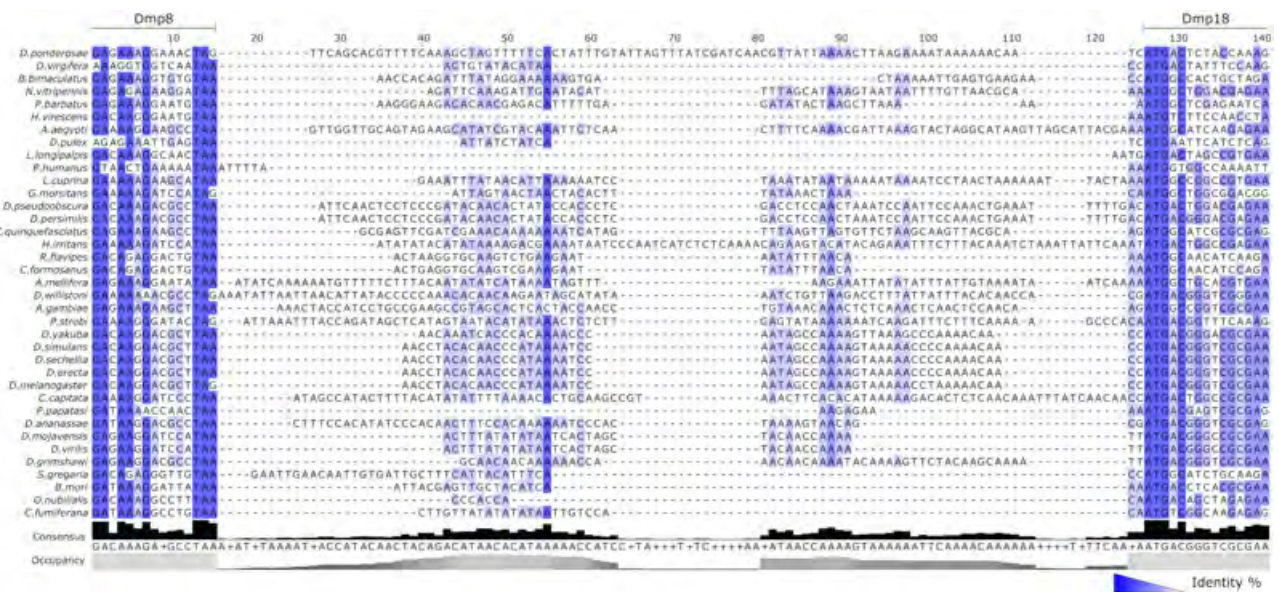


FIGURE 1 Dmp8-Dmp18 intergenic regions do not show a defined length, consensus sequence, or motif. A sequence alignment of the intergenic regions of all sequenced members of the Arthropoda phylum, that goes from 2 to 92 nucleotide in length, shows no consensus sequence or a specific motif. MUSCLE²³ was used to perform the alignment. Shown in a blue scale is the sequence identity conservation, indicating that the only conserved sequences belong to the portions corresponding to Dmp8 and Dmp18 open reading frames

3.2 | Dmp8/Dmp18 bicistronic mRNA allows both proteins to be translated

To understand both proteins' translation mechanism, first we needed to assess if both of them were indeed translated in the fly throughout its whole lifespan. To do that, we analyzed their presence with whole protein extracts of each of the developmental stages of the *Drosophila* fly, spanning from: early embryo (0 to 3 hours), late embryo (3 to 21 hours), the three stages of larval development, then pupae and pharate, to culminate with the adult form (Figure 2A). We found that even though the amount of expressed protein apparently varies, both proteins are present in each of the developmental stages analyzed, supporting the idea that both proteins are simultaneously translated from the same transcript. The presence of both products during the fly development was not surprising since both proteins are part of canonical homeostatic mechanisms.

Based on this observation and since previous work showed there is only one mRNA spanning the whole developmental process,¹⁸ we decided to determine if this bicistronic mRNA was enough for the cell to produce both protein products. To do that, we cloned the bicistronic coding region spanning from the complete 5'-UTR to the end of the Dmp18 3'-UTR into pBluescript II KS(+), to perform in vitro coupled transcription and translation

assays. We checked that only one transcript was being produced in the system (Supporting Information Figure 1). As a positive control, we used plasmid constructions in which separated Dmp8 and Dmp18 mRNAs were cloned in the pGADT7 AD vector that adds a tag and therefore makes the protein heavier (data not shown) and allows efficient in vitro translation using the vector elements for translation (Figure 2B, constructions 2 and 3). We found that both ORFs were able to translate on their own and that the bicistronic mRNA was capable of producing both proteins (Figure 2B, construction 4). It is interesting to notice that in the in vitro conditions, Dmp18 apparently is less efficiently translated than Dmp8. However, we have to take into account that part of this difference in the detection of the two proteins could be related to the fact that the proteins are detected by the incorporation of ³⁵S-methionine and Dmp18 has two methionine residues compared with Dmp8 that has four. Then we performed a time kinetics experiment that showed that both proteins could be detected as early as 20 minutes in the translation reaction, and earlier times only show the presence of Dmp8 (Figure 2C), as far as the sensitivity that this method allows us to visualize. Altogether, these results show that in vitro, Dmp8 and Dmp18 are translated from the same transcript and that all of the sequences necessary for their translation are present in their bicistronic mRNA.

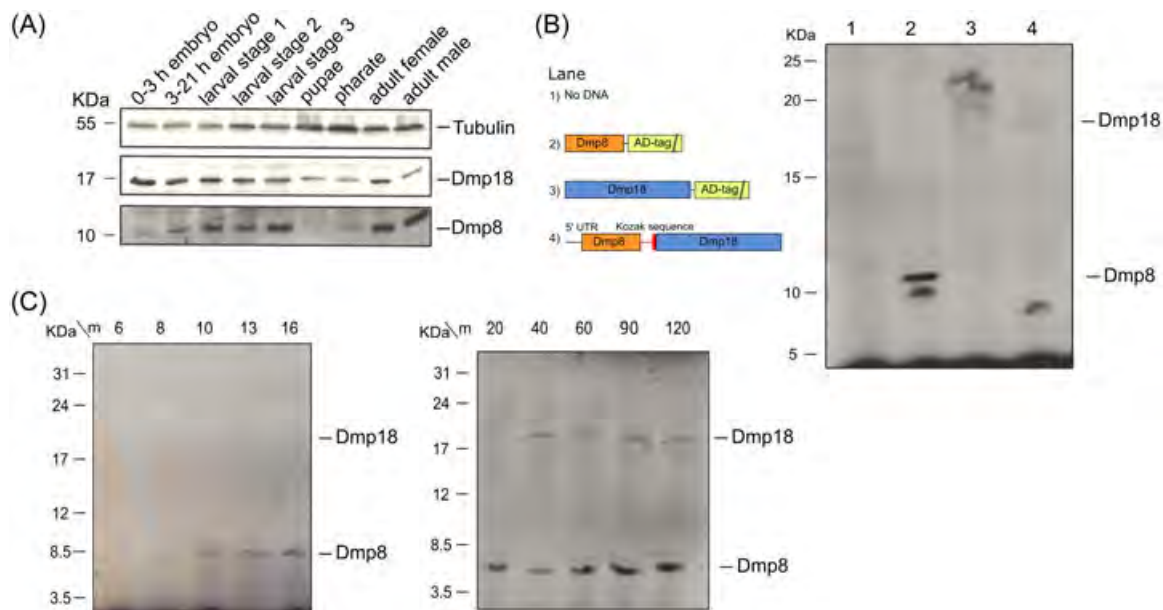


FIGURE 2 Dmp8 and Dmp18 are translated from the same messenger RNA (mRNA) and are present throughout *Drosophila* development. A, Western blot analysis shows that Dmp8 and Dmp18 are present through all fly developmental stages, from embryo to adult. β -Tubulin was used as loading control for the total protein extracts. B, In vitro coupled transcription and translation of constructs of both proteins with a tag revealed their correct translation (lanes 2 and 3). Also shown is a construction spanning the whole bicistronic mRNA including its 5'-UTR, the intergenic region is indicated in red, and Dmp18 Kozak sequence represented by a red rectangle (lane 4). All mRNAs include 21 additional nucleotide from the expression vector used. C, Time kinetics depicting the apparition of both proteins of the bicistronic messenger; at early times only Dmp8 is visible, but both proteins appear at the 20 minutes mark

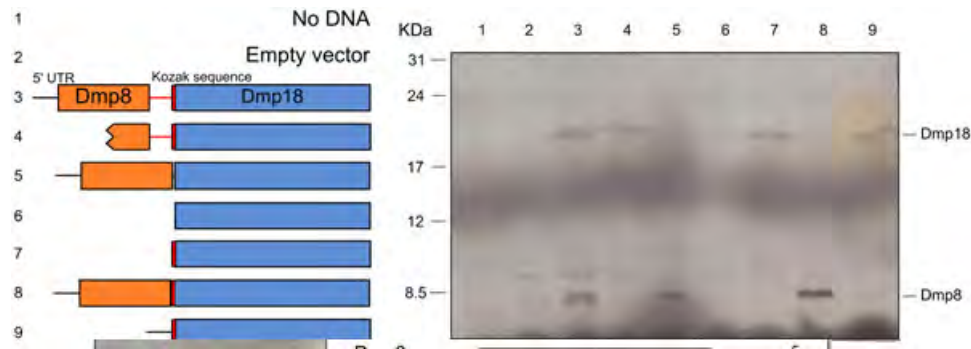


FIGURE 3 The distance between Dmp8 and Dmp18 affects translation of the second cistron. In vitro coupled transcription/translation was performed to analyze Dmp18 translation efficiency when its 5'-upstream sequences were changed. Lane 3 shows both proteins of the bicistronic array translated from the bicistronic messenger with Dmp8 (orange rectangle) and Dmp18 (blue rectangle), including its 5'-untranslated region (UTR), its intergenic region (red line) with Dmp18 Kozak sequence (red rectangle). Lane 4 shows that Dmp18 is able to translate when only 110 nucleotide (nt) of Dmp8 are present, but when the intergenic region is not present, Dmp18 translation is ablated. In lane 6 we can see that Dmp18 is able to translate with only the 5'-UTR region of the expression vector, and that this translation appears to be enhanced when its Kozak sequence is reconstituted (lane 7). In lane 8 we can see that when only the Kozak sequence is present between the two open reading frames, translation of Dmp18 is unable to proceed. All messenger RNAs also include 21 nt added from the expression vector

3.3 | The context of what lies upstream of Dmp18 dictates its translation

Since the Dmp8-Dmp18 bicistronic mRNA was able to translate the two ORFs in vitro, we decided to search for elements that could be important for Dmp18 translation within the same mRNA. To achieve this we generated a series of constructions that may alter Dmp18 translation (Figure 3) and compared them with the construction that spans the whole bicistronic region. Since a mutant fly generated by the insertion of a transposon in the Dmp8 coding region resulted in a truncated Dmp8 coding region, but still translated Dmp18,¹⁸ we generated a construction emulating the Dmp8 part that would remain in the transcript after the insertion of the transposon (Figure 3, construction 4). We found that indeed a transcript that contains the final 110 nt of the Dmp8 coding region was sufficient to allow Dmp18 translation (Figure 3, construction 4).

Since there is no sequence homology in the intergenic region of all sequenced Arthropoda (Figure 1), and some organisms have a region as short as 2 nt, we assumed that this region was not relevant for Dmp18 translation, so we made a construction that lacked *Drosophila*'s intergenic region. Note that this means that its Kozak sequence was obliterated, leaving only 6 nt of the restriction enzyme used between Dmp8's stop codon and Dmp18's ATG (Figure 3, construction 5). Interestingly, in this configuration, Dmp18 translation was clearly reduced, may be because of some kind of steric impediment of termination vs initiation complexes, since both coding regions were too close to each

other and earlier experiments show that Dmp8 is translated first, or it may be due to Dmp18's Kozak sequence absence despite that it maintains its ATG codon. Intriguingly, a construction in which we removed the intergenic sequence still affected Dmp18 translation, even when the Dmp18 Kozak region was restored (Figure 3, construction 8). This result suggests that the proximity of the first ORF affects the translation efficiency of Dmp18 in vitro, and this is in agreement with the fact that Dmp8 gets translated first. This was unexpected since some arthropods have a very short intergenic sequence between both ORFs.

Then, we tested if Dmp18 was able to be translated on its own with only the 21 nt in its 5'-UTR transcribed next to it in the vector used, but nothing of the bicistronic array (Figure 3, construction 6), and in this case, we could still observe translation at low levels. As expected, its translation was substantially improved when its Kozak region was added (Figure 3, construction 7). Finally, we tested whether adding the 5'-UTR region of Dmp8 and the Dmp18 Kozak region (Figure 3, construction 9) were also sufficient for Dmp18 translation, and it seems that as long as it has its Kozak sequence, Dmp18 is capable of being translated independently of the 5'-UTR sequence.

The in vitro translation experiments performed up to this point allowed us to determine that both ORFs can be translated from the same bicistronic transcript and gave us an initial approximation of the sequence requirements for the second ORF translation. However, these experiments were not quantitative and were conducted in a heterologous system. Therefore, for our next experiments, we performed

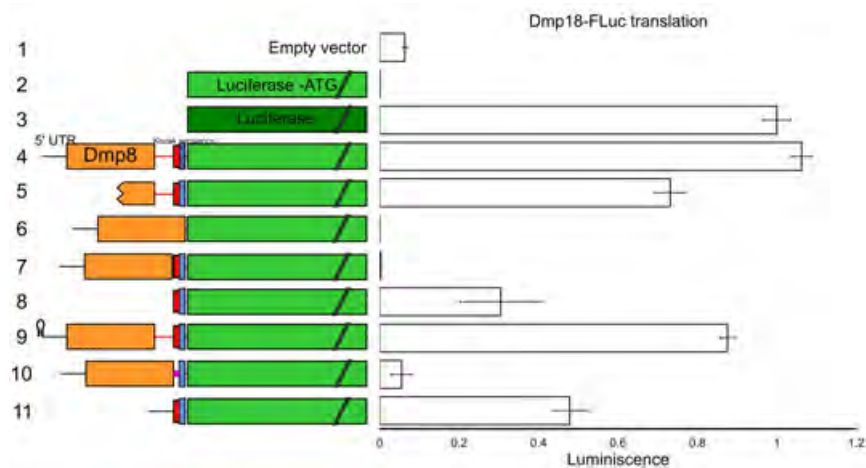


FIGURE 4 In vivo translation efficiency of luciferase reporter constructs. A fusion between firefly luciferase (FLuc, green rectangle) and the first four amino acids of Dmp18 was tested in S2R⁺ *Drosophila* cell line transfection assays. The activity of the FLuc reporter revealed that translation of the fusion construct bearing Dmp18 (blue rectangle) with its Kozak sequence (red rectangle) next to Dmp8 (orange rectangle), and its whole intergenic region (red line) is comparable to that of FLuc alone (lanes 4 and 2, respectively). When only 110 nucleotide (nt) of Dmp8 are present, translation of the second cistron is also achieved (lane 5), but it is impaired when the whole intergenic sequence is removed (lane 6), and is not restored by the presence of the Kozak sequence (lane 7), even though the Kozak sequence alone drives translation of a monocistronic construction (lane 8). The addition of a sequence forming a stem-loop near the 5'-end of the messenger (lane 9) does not affect translation of the fusion reporter. The intergenic sequence corresponding to another arthropod (purple box), *Pediculus humanus*, was not able to support translation of the fusion reporter (lane 10). Finally, the 5'-untranslated region (UTR) sequence next to the Dmp18 Kozak sequence does not affect its translation as a monocistron (lane 11). All messenger RNAs include 85 nt transcribed from the expression vector

transient expression in the *Drosophila* cell line S2R⁺, to test the bicistronic behavior in an endogenous system. We did this by coupling our constructions to a firefly luciferase (FLuc) reporter, using *Renilla* luciferase (RLuc) as a reporter of transfection efficiency (Figure 4, Supporting Information Figure 2). For the constructions, we removed the FLuc translation start site and inserted Dmp18's first four amino acids including its Kozak element. We found that in the Dmp18-FLuc fusion coupled to Dmp8 beginning at its 5'-UTR and including the intergenic region construction, which is similar to the bicistronic configuration previously used, the translation of the luciferase was equivalent to that of the complete FLuc (Figure 4, construction 4 compared with construction 3). In this system, the construction with the last 110 nt of the Dmp8 coding sequence was translated, although the luciferase activity was slightly reduced, if compared with the wild-type configuration (Figure 3, construction 5 compared with construction 4).

Like the in vitro assays, when the intergenic sequence was deleted and the Dmp8 ORF was closer to the Dmp18 translation initiation codon, the hybrid Dmp18-FLuc translation was ablated, even in the presence of the Kozak region (Figure 4, constructions 6 and 7). The presence of only the Dmp18 Kozak sequence plus the first four codons preceded by the 85 nt 5'-UTR sequence transcribed from the vector used (Figure 4, construction 8) was enough to allow translation, albeit at

lower levels than the construct containing the intergenic region and 110 nt of Dmp8 (compare Figure 3, constructions 8 with 5).

Since Dmp18 is able to be translated without Dmp8 translation, we can rule out leaky scanning and stop codon readthrough as possible Dmp18 translation mechanisms. Therefore, the alternative is that Dmp18 translation occurs due to the presence of some region that promotes an alternative translation initiation mechanism. To test this idea, we added a stem-loop structure, which is known to prevent ribosome attachment to the 5'-UTR in a cap-dependent context, next to the 5'-end of the construction, and therefore we expected to ablate Dmp8 translation in this bicistronic context, and if there is indeed a separate attachment site for the translation machinery somewhere before the Dmp18 translation start site, its translation should not be affected. The result shown for construction 9 in Figure 4 demonstrates indeed that the addition of a stem-loop that ablates 5'-end cap-dependent translation²⁴ (Supporting Information Figure 3) did not affect FLuc translation.

To revise our idea that the intergenic region is not important for Dmp18 translation, we changed the intergenic region to that of the lice *Pediculus humanus* (Figure 4, construction 10) with a length of 8 nt, one of the shortest reported for this region in arthropods (Figure 1). Surprisingly, we found that this construction

was unable to translate the FLuc reporter. This result further supports the idea that Dmp8 translation interferes with Dmp18 production when both ORFs are placed close to each other, in this case 8 nt apart. Finally, when we used other unrelated sequences, such as the 5'-UTR of the bicistronic messenger with the Dmp18 Kozak sequence (Figure 4, construction 11), translation was the same as that with only the Kozak sequence (compare with Figure 4, construction 8), reinforcing the idea that Dmp18 requires its Kozak sequence to translate correctly. Remarkably, the results obtained in the $S2^+$ cells were similar to the ones obtained using the *in vitro* translation system. Cell protein extracts were performed to establish the correct expression of Dmp8 in cells transfected with constructions were Dmp8 was complete (Supporting Information Figure 3).

3.4 | Bicistronic architecture allows Dmp18 translation in response to cellular stress

Since it appears that Dmp18 is capable of translating correctly on its own, the question arose as to why it has maintained its bicistronic array next to Dmp8. It is known that when cells undergo stress from different

sources, cap-dependent translation is completely ablated.²⁵ For that reason, proteins that are necessary for damage repair, among other types of proteins, depend on cap-independent mechanisms to allow their translation.² Therefore, the next question was whether Dmp18 is maintained in this bicistronic configuration to allow for its translation in response to stress, since its plant and human homologs are active during stress response mechanisms, such as UV radiation and HS.^{21,26}

To answer this question, we analyzed both proteins in whole protein extracts from *Drosophila* wild-type larvae subjected to 200 J/m² UV radiation and 42°C HS treatments (Figure 5A). We found that in response to both types of stress, after 45 minutes, Dmp8 was not detected, but Dmp18 was still present. Since transcription of the bicistronic messenger was unaffected (Supporting Information Figure 4) this result suggests that part of the evolutionary selection to maintain Dmp8 and Dmp18 as a bicistron in arthropods is due to the role of Dmp18 in stress responses.

We next performed time kinetics of the change of translation of the luciferase reporter in response to 5 J/m² UV damage (Figure 5B) and 36°C HS in $S2R^+$ cells transfected with some of the FLuc constructions shown in Figures 4 and 5C. To analyze the results of

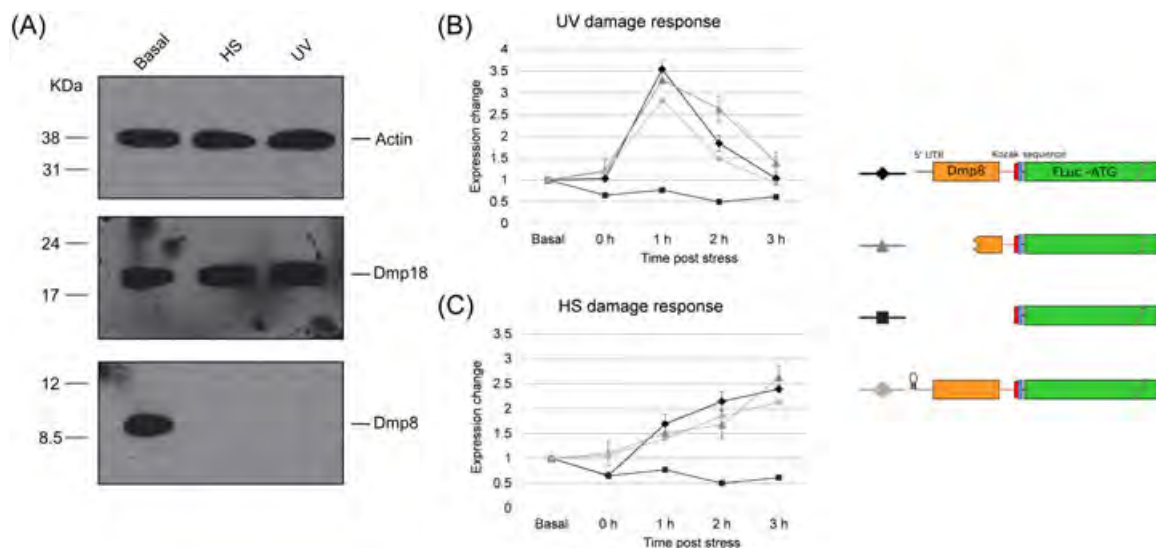


FIGURE 5 The bicistronic configuration is necessary to enable Dmp18 translation under stress response. A, Western blot analysis shows that following UV and heat-shock (HS) damage to wild-type *Drosophila* larvae causes Dmp8 translation arrest, but translation of Dmp18 is still possible. Kinetics of FLuc fusion constructs in comparison to control RLuc following UV (B) and HS (C) damage to transfected cells reveal that Dmp18 has a spike of translation following the damage stimuli when it is located next to Dmp8 and its intercistronic region, but not in the presence of its Kozak region alone (black square line). An increase in FLuc activity under the UV stress was observed at 1 hour of the stimulus for the whole bicistronic construct, both in the presence or absence of a stem-loop at the 5'-end. The increment was time-dependent since it dropped for both constructions at 2 hours after the stress and reached basal levels at 3 hours. The construction with the last 110 nucleotide (nt) of Dmp8 showed similar behavior, whereas that with only the Kozak sequence did not show any change upon the UV stress. The kinetics observed for HS stress was different from the UV response, since FLuc activity gradually increased with time after the stimulus for the same three constructs, but overall stress response was similar. All messenger RNAs include 85 nt transcribed from the expression vector. FLuc, firefly luciferase; RLuc, *Renilla* luciferase

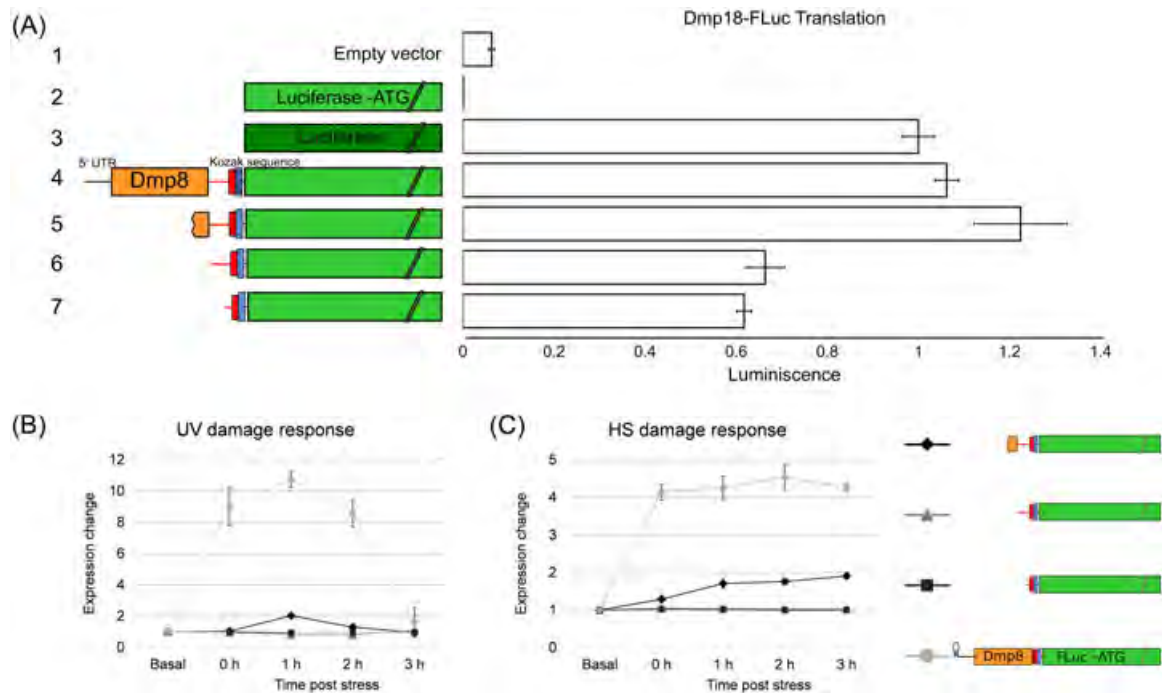


FIGURE 6 The last portion of Dmp8 open reading frame (ORF) is required for the second cistron translation in response to stress. A, The fusion reporter is able to translate very efficiently in the presence of the last 27 nucleotide (nt) of Dmp8 ORF (lane 5), the intercistronic sequence (lane 6), or half of the intercistronic sequence (lane 7). B, In response to UV damage, and (C) heat-shock (HS) stress, translation stimulation was achieved only in the presence of at least the last 27 nt of Dmp8 ORF. Stimulation was observed also in the absence of intercistronic region (gray circle line), resulting in a more dramatic change since this construction is unable to translate in basal conditions. All messenger RNAs include 85 nt transcribed from the expression vector

these experiments, we measured the ratio between FLuc and the RLuc control activities, considering its initial translation as 1. The constructions used in this experiment were the ones that emulate the whole bicistronic mRNA or only the last 110 nt from Dmp8 (Figure 4, constructions 4 and 5, respectively), the one with the stem-loop structure placed at the 5'-end (Figure 4, construction 9), and the one that only has the Kozak sequence upstream of the Dmp18 translation start site (Figure 4, construction 8). We chose these constructions because all of them are able to translate in the absence of stress conditions. We found that constructions with at least the 110 nt of Dmp8 had an increase in translation in comparison with the *Renilla* luciferase control, due to the fact that it depends on cap translation. This apparent rise happens in response to both types of stress, but the same cannot be said for the one with only the Kozak sequence, since its translation seemed to be unchanged and it followed the dynamics of the cap-dependent *Renilla* luciferase (Figure 5B and 5C).

The results presented suggest that Dmp18 kept the bicistronic architecture because some upstream element allows its increased translation under stress conditions. Because the construction spanning from 110 nt of Dmp8 is able to result in translation of the Dmp18 (four amino

acids) FLuc fusion protein in response to stress, we decided to determine the minimal sequence from the 110 nt to the Dmp18 Kozak sequence required for translation. Therefore, we generated three additional constructions (Figure 6A), one that included 27 nt of Dmp8 and the intercistronic region (Figure 6A, construction 5), another one with only the Dmp8/Dmp18 bicistron intercistronic region (Figure 6A, construction 6), and one with half the intercistronic region, the last 27 nt (Figure 6A, construction 7), and all of them kept the Dmp18 Kozak sequence. As expected, we found that all of the constructions were able to be translated under basal conditions, although the constructions lacking any Dmp8 sequence translated at slightly lower levels. Next, we subjected the cells bearing each of the three new constructions to UV and HS stress treatments and measured their response (Figure 6B and 6C). In both types of stimuli, we could see that while they were all able to translate as under basal conditions, the only one whose translation response was higher than the control in response to the damage stimuli was the one still bearing the last 27 nt of the Dmp8 ORF.

These results suggest that the last 27 nt from the Dmp8 sequence are necessary for Dmp18 translation in response to damage. To test this, we used under the same

conditions the construction with only the Dmp18 Kozak sequence in the intercistronic region (Figure 4, construction 7). As shown in Figure 4, this construction was not able to be translated in basal conditions, perhaps because of the closeness of the translation machinery involved in Dmp8 termination and the Dmp18 initiation start site. However, in the presence of UV or HS stress to the cells, where the cap-dependent translation initiation of Dmp8 is hindered (Figure 6A), Dmp18 translation resulted in an apparent exacerbated increase in response to the stress compared to the control (Figure 6B and 6C). This result confirmed that when the Dmp8 ORF is too close to the Dmp18 start codon, it interferes with its correct translation under normal conditions. It also supports a role for the last part of Dmp8's CDS in the enhancement of Dmp18 translation under UV or HS stress.

4 | DISCUSSION

Since the annotation of several animal genomes, the discovery of genes encoded in different multicistronic configurations has increased. In most cases, these genes are organized as bicistrons, and in general the translation of the second ORF requires the presence of IRES or IRES-like elements.^{11,27} We have found that the Dmp8/Dmp18 bicistronic messenger, conserved in the Arthropoda phylum, is unusual since although there is a sequential translation of both proteins, the intercistronic region is dispensable for the translation of the second ORF. This might explain why the distance and sequence of the intercistronic region is not conserved in the Arthropoda phylum. However, our analysis of different reporter constructs suggests that, at least in *Drosophila*, an arrangement of Dmp8 to Dmp18 is required to allow for Dmp18 translation when the cell is under stress conditions. This makes sense, because as we mentioned earlier, both proteins are part of complexes that act in DNA repair mechanisms, and Dmp8 is very important for general transcription mechanisms as well as repair.^{28,29} This arrangement ensures that both proteins are always present in the cell, but they need to respond differently to stress. When damage arises in the cell, cap-dependent translation is ablated²⁵ to allow for the induction of stress response mechanisms, and Dmp18 homologs have been identified as active during various types of DNA damages.²⁶ In addition, p18^{Hamlet} has been found to help in the recruitment of p53 to target genes in response to DNA damage.²¹ On the other hand, although Dmp8 participates in the nucleotide excision repair mechanism, it also has a role as a component of TFIIH in RNA polymerase II transcription, which is mostly arrested during stress,^{19,30} which might be explained by its

complete lack of translation immediately after damage by HS or UV radiation (Figure 5A).

Such a difference in the translation of two proteins in the same genomic unit is rare, but not unique. In previous reports, a protein involved in nutritional stress, the arginine/lysine transporter (cat-1), is only able to translate in response to starvation that is accompanied by eIF2a activity ablation, allowing for the translation of an upstream open reading frame and making it possible to inhibit the formation of a loop structure.^{31,32}

On the one hand, our results show that in the context of the bicistron, the intercistronic sequence does not seem to be important for Dmp18; on the other hand, we found no translation when *Pediculus humanus* intercistronic region was present. This intercistronic sequence only consists of 8 nt between the Dmp8 stop and Dmp18 initiation codons, suggesting that this proximity interferes with ribosome binding, hindering Dmp18 translation (construction 10 in Figure 4). One possibility to explain this result is that different mechanisms for the translation of Dmp18 exist, which may explain why in other organisms this bicistronic configuration is not present. Another possibility could be that in *P. humanus*, Dmp18 is only translated in response to stress. In the Arthropoda phylum, the presence of a second ORF that allows its translation in a cap-independent manner may be important, and indeed this is nothing new, since elements that promote translation initiation in a cap-independent manner are more common than previously thought.³³ In addition, many genes have been found to possess sequences that enhance their translation.^{2,34} Diverse components of the translational initiation machinery have been found to be involved in this initial recognition that allows for a noncanonical translation initiation, making it more complex, since there is no typical behavior of translation regulation. Sometimes, a stem-loop structure can hinder translation,²⁴ and other times, such secondary structure is needed for its recognition by translation factors to allow efficient translation.³⁵ This structure is generally found in the 5'-UTR¹⁷; however, this seems not to be the case for the Dmp8-Dmp18 bicistron in *Drosophila*. Our results show that at least the last 27 nt coding for the C-terminal portion of Dmp8 are necessary for Dmp18 translation under stress conditions. This suggests that this sequence may be recognized by a factor that allows Dmp18 translation during stress. The Dmp8 amino acids sequence is highly conserved in all eukaryotes to maintain its functions, this guarantees the conservation of its DNA sequence to be the target of a factor that modulates Dmp18 translation under stress conditions. Furthermore, even when the intercistronic region is not conserved throughout arthropods it is intriguing that

the Dmp18 Kozak element seems to be very well conserved, and its sequence is identical to the reported consensus sequence.^{36,37}

In a search through all of *Drosophila*'s bicistronic messengers (data not shown), we found another protein array with similar characteristics to the Dmp8-Dmp18 bicistronic messenger. This contained two proteins that did not appear to have a role in the same mechanism, and both are part of repair mechanisms. These proteins are identified with the annotation symbols CG42374 and CG9666, the first of which corresponds to a member of the SOSS complex involved in genomic stability, and double-strand break repair.³⁸ The second ORF encodes for an N-6 adenine-specific DNA methyltransferase that has been implicated in regulation during virulence control.³⁹ This bicistronic array is conserved in arthropods, as is the case with the Dmp8-Dmp18 mRNA, and similarly, even though both proteins are significantly conserved, its intercistronic sequence is not (Supporting Information Figure 5). Both proteins are present in other organisms, but not in a bicistronic array, like Dmp8 and Dmp18. This is interesting because it reveals that the mechanism we found in Dmp8-Dmp18 bicistronic messenger might not be an isolated case.

In general, several translation mechanisms that only arise in the presence of stress have been identified. Some of them involve a decrease in particular translation factors activity to allow for translation initiation at a site different from the canonical AUG.⁴⁰ There is also the case of the recently discovered translation regulation through methylation of adenines in the mRNA, which allows cells to translate their proteins involved in stress response pathways.⁴¹ As more research is done in this area, additional alternative translational mechanism will be understood, and we may be able to use this in fighting diseases with an etiology involving translation failure.

ACKNOWLEDGMENTS

We would like to thank E. Lopez, S. Becerra, J. Yañez, and P. Gaytan in the Synthesis and Sequencing Unit for the production of the different oligonucleotides and constructions sequencing. This study was supported by Consejo Nacional de Ciencia y Tecnología (219673 and 250588 to MZ, 177393 to VV-G); Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (IN-200315 for MZ); and AB-M received a scholarship from the Consejo Nacional de Ciencia y Tecnología (334762). The authors declare no competing financial interests.

AUTHOR CONTRIBUTIONS

AB-M performed leader investigation, data curation, and writing of the original draft. GC-B executed supporting investigation. VV-G did supporting funding acquisition and supporting review and editing. TDD and MZ equally did the conceptualization and leader review and editing. Leader funding acquisition was performed by MZ.

ORCID

Mario Zurita  <http://orcid.org/0000-0002-8404-2173>

REFERENCES

1. Sonenberg N, Hinnebusch AG. Regulation of translation initiation in eukaryotes: mechanisms and biological targets. *Cell*. 2009;136:731-745.
2. Lacerda R, Menezes J, Romão L. More than just scanning: the importance of cap-independent mRNA translation initiation for cellular stress response and cancer. *Cell Mol Life Sci*. 2017; 74:1659-1680.
3. Sendoel A, Dunn JG, Rodriguez EH, et al. Translation from unconventional 5' start sites drives tumour initiation. *Nature*. 2017;541:494-499.
4. Burgess HM, Gray NK. mRNA-specific regulation of translation by poly(A)-binding proteins. *Biochem Soc Trans*. 2010;38: 1517-1522.
5. Park JE, Yi H, Kim Y, Chang H, Kim VN. Regulation of poly(A) tail and translation during the somatic cell cycle. *Mol Cell*. 2016;62:462-471.
6. Yu CH, Dang Y, Zhou Z, et al. Codon usage influences the local rate of translation elongation to regulate co-translational protein folding. *Mol Cell*. 2015;59:744-754.
7. Hinnebusch AG. Molecular mechanism of scanning and start codon selection in eukaryotes. *Microbiol Mol Biol Rev*. 2011;75:434-467.
8. Terenin IM, Smirnova VV, Andreev DE, Dmitriev SE, Shatsky IN. A researcher's guide to the galaxy of IRESS. *Cell Mol Life Sci*. 2017;74:1431-1455.
9. Elfakess R, Sinvani H, Haimov O, Svitkin Y, Sonenberg N, Dikstein R. Unique translation initiation of mRNAs-containing TISU element. *Nucleic Acids Res*. 2011;39:7598-7609.
10. Pelletier J, Sonenberg N. Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA. *Nature*. 1988;334:320-325.
11. Jungreis I, Lin MF, Spokony R, et al. Evidence of abundant stop codon readthrough in *Drosophila* and other metazoa. *Genome Res*. 2011;21:2096-2113.
12. Gramates LS, Marygold SJ, Santos G, et al. FlyBase at 25: looking to the future. *Nucleic Acids Res*. 2017;45:D663-D671.
13. Ryabova LA, Pooggin MM, Hohn T. Translation reinitiation and leaky scanning in plant viruses. *Virus Res*. 2006;119: 52-62.
14. Yamamoto H, Unbehaun A, Spahn CMT. Ribosomal chamber music: toward an understanding of IRES mechanisms. *Trends Biochem Sci*. 2017;42:655-668.

15. Murray J, Savva CG, Shin BS, Dever TE, Ramakrishnan V, Fernández IS. Structural characterization of ribosome recruitment and translocation by type IV IRES. *eLife*. 2016;5:1-24.
16. Wang QS, Jan E. Switch from cap- to factorless IRES-dependent 0 and +1 frame translation during cellular stress and dicistrovirus infection. *PLoS One*. 2014;9:17-20.
17. Weingarten-Gabbay S, Elias-Kirma S, Nir R, et al. Comparative genetics. Systematic discovery of cap-independent translation sequences in human and viral genomes. *Science*. 2016;351:240-251. <https://doi.org/10.1126/science.aad4939>
18. Herrera-Cruz M, Cruz G, Valadez-Graham V, et al. Physical and functional interactions between *Drosophila* homologue of Swc6/p18hamlet subunit of the SWR1/SRCAP chromatin-remodeling complex with the DNA repair/transcription factor TFIIF. *J Biol Chem*. 2012;287:33567-33580.
19. Singh A, Compe E, Le May N, Egly JM. TFIIF subunit alterations causing xeroderma pigmentosum and trichothiodystrophy specifically disturb several steps during transcription. *Am J Hum Genet*. 2015;96:194-207.
20. Berriri S, Gangappa SN, Kumar SV. SWR1 chromatin-remodeling complex subunits and H2A.Z have non-overlapping functions in immunity and gene regulation in *Arabidopsis*. *Mol Plant*. 2016;9:1051-1065.
21. Lafarga V, Cuadrado A, Nebreda AR. p18Hamlet mediates different p53-dependent responses to DNA damage-inducing agents. *Cell Cycle*. 2007;6:2319-2322.
22. Cuadrado A, Lafarga V, Cheung PCF, et al. A new p38 MAP kinase-regulated transcriptional coactivator that stimulates p53-dependent apoptosis. *EMBO J*. 2007;26:2115-2126.
23. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res*. 2004;32:1792-1797.
24. Babendure JR, Babendure JL, Ding J, Tsien RY. Control of mammalian translation by mRNA structure near caps. *RNA*. 2006;12:851-861.
25. Hernández G, Vázquez-pianzola P, Sierra JM, Va P, Riverapomar R. Internal ribosome entry site drives cap-independent translation of reaper and heat shock protein 70 mRNAs in *Drosophila* embryos. *RNA*. 2004;10:1783-1797.
26. Rosa M, Von Harder M, Aiese Cigliano R, Schlogelhofer P, Mittelsten Scheid O. The *Arabidopsis* SWR1 chromatin-remodeling complex is important for DNA repair, somatic recombination, and meiosis. *Plant Cell*. 2013;25:1990-2001.
27. Col B, Oltean S, Banerjee R. Translational regulation of human methionine synthase by upstream open reading frames. *Biochim Biophys Acta*. 2009;1769:532-540.
28. Egly JM, Coin F. A history of TFIIF: two decades of molecular biology on a pivotal transcription/repair factor. *DNA Repair*. 2011;10:714-721.
29. Compe E, Egly J-M. TFIIF: when transcription met DNA repair. *Nat Rev Mol Cell Biol*. 2012;13:343-354.
30. Theil AF, Nonnekens J, Steurer B, et al. Disruption of TTDA results in complete nucleotide excision repair deficiency and embryonic lethality. *PLoS Genet*. 2013;9:e1003431.
31. Fernandez J, Yaman I, Huang C, et al. Ribosome stalling regulates IRES-mediated translation in eukaryotes, a parallel to prokaryotic attenuation. *Mol Cell*. 2005;17:405-416.
32. Yaman I, Fernandez J, Liu H, et al. The zipper model of translational control: a small upstream ORF is the switch that controls structural remodeling of an mRNA leader. *Cell*. 2003;113:519-531.
33. Hultmark D, Klemenz R, Gehring WJ. Translational and transcriptional control elements in the untranslated leader of the heat-shock gene hsp22. *Cell*. 1986;44:429-438.
34. Wellensiek BP, Larsen AC, Stephens B, et al. Genome-wide profiling of human cap-independent translation-enhancing elements. *Nat Methods*. 2013;10:747-750.
35. Paek KY, Hong KY, Ryu I, et al. Translation initiation mediated by RNA looping. *Proc Natl Acad Sci*. 2015;112:1041-1046.
36. Kozak M. Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Res*. 1984;12:857-872.
37. Cavener DR. Comparison of the consensus sequence flanking start sites in *Drosophila* and vertebrates. *Nucleic Acids Res*. 1987;15:1353-1361.
38. Huang J, Gong Z, Ghosal G, Chen J. SOSS complexes participate in the maintenance of genomic stability. *Mol Cell*. 2010;35:384-393.
39. Luo G-Z, Blanco MA, Greer EL, He C, Shi Y. DNA N6-methyladenine: a new epigenetic mark in eukaryotes? *Nat Rev Mol Cell Biol*. 2015;16:705-710.
40. Nanda JS, Cheung YN, Takacs JE, et al. eIF1 controls multiple steps in start codon recognition during eukaryotic translation initiation. *J Mol Biol*. 2009;394:268-285.
41. Meyer KD, Patil DP, Zhou J, et al. 5' UTR m6A promotes cap-independent translation. *Cell*. 2015;163:999-1010.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Bucio-Mendez A, Cruz-Becerra G, Valadez-Graham V, Dinkova TD, Zurita M. The Dmp8-Dmp18 bicistron messenger RNA enables unusual translation during cellular stress. *J Cell Biochem*. 2018;1-11. <https://doi.org/10.1002/jcb.27670>

RESEARCH ARTICLE

TFIIH localization is highly dynamic during zygotic genome activation in *Drosophila*, and its depletion causes catastrophic mitosis

Grisel Cruz-Becerra, Sarai Valerio-Cabrera, Mandy Juárez, Alyeri Bucio-Mendez and Mario Zurita*

ABSTRACT

In *Drosophila*, zygotic genome activation occurs in pre-blastoderm embryos during rapid mitotic divisions. How the transcription machinery is coordinated to achieve this goal in a very brief time span is still poorly understood. Transcription factor II H (TFIIH) is fundamental for transcription initiation by RNA polymerase II (RNAPII). Herein, we show the *in vivo* dynamics of TFIIH at the onset of transcription in *Drosophila* embryos. TFIIH shows an oscillatory behaviour between the nucleus and cytoplasm. TFIIH foci are observed from interphase to metaphase, and colocalize with those for RNAPII phosphorylated at serine 5 (RNAPIIS5P) at prophase, suggesting that transcription occurs during the first mitotic phases. Furthermore, embryos with defects in subunits of either the CAK or the core subcomplexes of TFIIH show catastrophic mitosis. Although, transcriptome analyses show altered expression of several maternal genes that participate in mitosis, the global level of RNAPIIS5P in TFIIH mutant embryos is similar to that in the wild type, therefore, a direct role for TFIIH in mitosis cannot be ruled out. These results provide important insights regarding the role of a basal transcription machinery component when the zygotic genome is activated.

KEY WORDS: Zygotic genome activation, ZGA, TFIIH, *Drosophila*, Transcription

INTRODUCTION

In many organisms, the initial stages of embryonic development are directed by the action of maternal products deposited in eggs during oogenesis. Maternal mRNAs encoding factors that are required for the first developmental stages are eventually degraded; simultaneously, the expression of zygotic genes occurs in transcriptional waves (Zurita et al., 2008; Lee et al., 2014). In *Drosophila*, after fertilization, the embryo is a syncytium in which the nucleus undergoes a series of rapid, synchronized mitotic divisions that consist of only S and M phases (Mazumdar and Mazumdar, 2002). It has been proposed that the first wave of zygotic transcription begins at nuclear cycle eight (NC8) (Pritchard and Schubiger, 1996; De Renzis et al., 2007); however, transcription of a few zygotic genes prior to NC7 was recently reported (Ali-Murthy et al., 2013). Therefore, the exact moment of

zygotic transcription onset at pre-midblastula transition (pre-MBT) embryos is currently controversial. After NC10, mitotic cycles slow down, and by NC14, expression of most of the zygotic genome is activated at the midblastula transition (MBT; Farrell and O'Farrell, 2014).

In *Drosophila*, zygotic genome activation (ZGA) is mediated by the transcription factors *vielfältig/Zelda* and the GAGA factor (Staudt et al., 2006; Li et al., 2014; Blythe and Wieschaus, 2016) in coordination with components of the basal transcription machinery, that must be recruited to the promoters of genes expressed either at the pre-MBT or MBT (Satija and Bradley, 2012). At the pre-MBT, ~100 genes are reportedly expressed (Lott et al., 2011), and a recent study detected unpaused RNA polymerase II (RNAPII) on the promoters of 77 genes during pre-MBT, probably because these genes must be transcribed in a very short time span (Chen et al., 2013). Therefore, analysis of the dynamic factors that activate transcription is fundamental to understanding the mechanism of ZGA.

A central component of basal transcription machinery is TFIIH, a multi-protein complex composed of core [XPB (Hay), XPD, p62 (Tfb1), p52 (Mrn), p44 (Ssl1), p34 (Tfb4) and p8 (Tfb5) subunits] and CAK (Cdk7, CycH and MAT1 subunits) subcomplexes. The core and CAK subcomplexes form the holo-TFIIH complex, which participates in transcription mediated by RNAPII (Zurita and Merino, 2003; Compe and Egly, 2012). The XPB and XPD subunits are DNA helicases/ATPases, while the XPB subunit is also a DNA translocase required for transcription initiation (Egly and Coin, 2011; Fishburn et al., 2015). Additionally, phosphorylation of serine residues 5 and 7 of the C-terminal domain (CTD) of the largest RNAPII subunit by Cdk7 is required for RNAPII promoter escape and recruitment of mRNA modification and processing factors during transcription elongation (Hsin and Manley, 2012). In addition, the TFIIH core subcomplex participates in the nucleotide excision DNA repair mechanism (Coin et al., 2007), while CAK participates in controlling the cell cycle as an activating kinase for cyclin-dependent kinase complexes (reviewed in Schachter and Fisher, 2013).

Furthermore, a transitory complex between the CAK and XPD subunits has also been identified (Aguilar-Fuentes et al., 2006; Abdulrahman et al., 2013). In accordance with this interaction, XPD reportedly regulates Cdk7 localization and negatively modulates its function during cell cycle regulation in *Drosophila* embryos (Chen et al., 2003; Li et al., 2010). However, XPD has been demonstrated to be a component of the MMXD complex, which is involved in chromosome segregation in human cells; thus, XPD depletion produces catastrophic mitosis (Ito et al., 2010). Furthermore, in *Drosophila*, XPD physically and genetically interacts with *Galla*, the fly homologue of a component of the human MMXD complex, which when defective also causes catastrophic mitosis during early embryonic nuclear division (Yeom et al., 2014).

Departamento de Genética del Desarrollo y Fisiología Molecular, Instituto de Biotecnología. Universidad Nacional Autónoma de México, Av Universidad 2001, Cuernavaca Morelos 62250, México.

*Author for correspondence (marioz@ibt.unam.mx)

© S.V., 0000-0001-6676-8440; M.Z., 0000-0002-8404-2173

Received 3 October 2017; Accepted 3 April 2018

Therefore, the general transcription factor TFIID is involved in at least three fundamental cellular processes: transcription, DNA repair and cell cycle regulation. Although several studies have concentrated their efforts on the analysis of TFIID functions *in vitro* and *in vivo*, fundamental questions remain to be answered, including whether mutations in other core subunits of TFIID also affect mitosis, whether all TFIID subunits have the same dynamics during early embryonic development, and how TFIID subunits participate in zygotic gene activation. In this work, we have addressed these questions by analysing the *in vivo* dynamics of TFIID subunits during early mitotic division cycles in fly embryos as well as the effect of the absence of TFIID components on the syncytial blastoderm stage.

RESULTS

TFIID shows highly dynamic behaviour at transcription onset in the syncytial blastoderm

We previously reported the localization of the XPB, XPD, Cdk7 and MAT1 TFIID subunits during early *Drosophila* embryonic development (Aguilar-Fuentes et al., 2006). Using immunostaining analysis, we showed that during the early pre-blastoderm stage, XPB and XPD were preferentially cytoplasmic. At NC10 some XPB and XPD signals were detected inside the nucleus, while most Cdk7 and MAT1 remained cytoplasmic. Furthermore, chromatin

immunoprecipitation (ChIP) experiments indicated that both core and CAK subunits could be found at some promoters at pre-MBT embryos (Aguilar-Fuentes et al., 2006). In contrast, a recent report using a different protocol for embryo immunostaining preparation revealed Cdk7 in the nucleus at the NC13 and NC14 interphases (Li et al., 2010).

Based on this information, we decided to refine our previous analysis of TFIID dynamics during early embryonic development and determine the *in vivo* dynamics of TFIID during fast mitotic cycles at the pre-MBT by performing immunostaining and via tagged fluorescent proteins. p8 is the smallest TFIID subunit, and although it does not have enzymatic activity, it has been proposed to play an important role in the maintenance of steady-state TFIID levels in human fibroblasts (Giglia-Mari et al., 2004). Initially, we studied the dynamics of this protein at pre-MBT embryos when ZGA takes place. Using immunostaining analysis and confocal microscopy, we found that at NC10 interphase, p8 was mostly nuclear. Furthermore, overlapping signals for p8 and chromosomes were observed (Fig. 1A). Intriguingly, at metaphase, most of the p8 signal was observed surrounding the chromosomes (Fig. 1B,C), and only slightly overlapped with the DNA and the mitotic spindle (Fig. 1D). In contrast, the p8 signal did not overlap with the nuclear envelope (which is not completely disassembled during syncytial blastoderm mitosis; Foe and Alberts, 1983; Harel et al., 1989), as

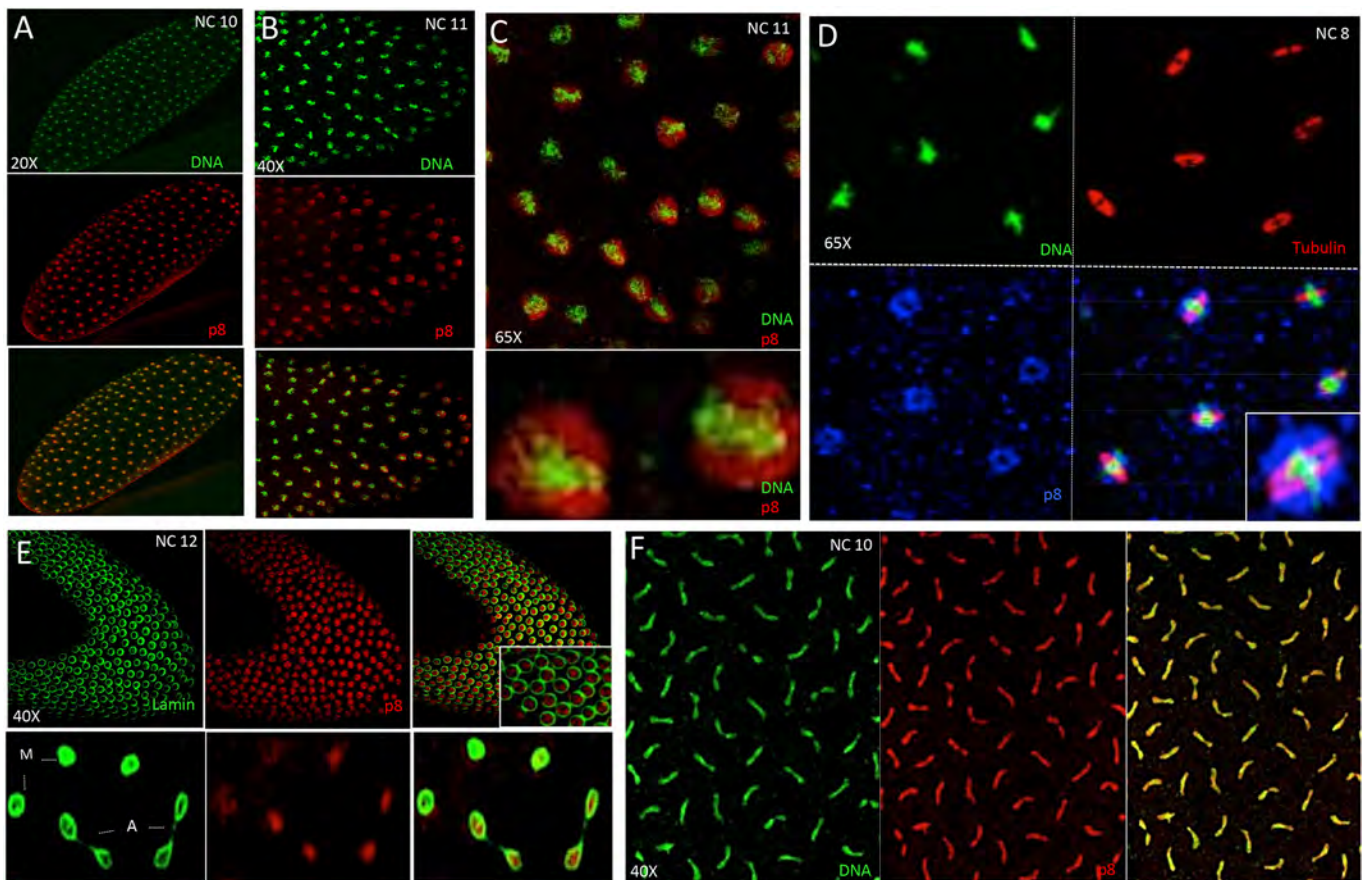


Fig. 1. p8 is preferentially nuclear during the mitotic cycles in the syncytial blastoderm embryo. (A) p8 is distributed in the nucleoplasm during interphase. (B,C) Most p8 (red) surrounds the chromosomes (green), although some p8 colocalizes with the chromosomes at metaphase. (D) At metaphase, p8 surrounds the mitotic structures, and also colocalizes with chromosomes and the mitotic spindle. p8, blue; β -tubulin, red; DNA, green. (E) Immunolocalization of lamin (green) and p8 (red) in syncytial blastoderm embryos at interphase (left) and mitosis (right) shows that p8 does not overlap with the nuclear membrane during the nuclear division cycle. M, metaphase; A, anaphase. (F) p8 (red) overlaps with chromosomes (green) at anaphase. All images correspond to projections of z-stacks acquired by confocal microscopy and the NC for each embryo is indicated in the figure. Magnifications relate to objective lens use.

determined by anti-p8 and anti-Lamin co-immunostaining (Fig. 1E), indicating that most p8 at metaphase is maintained inside the nuclear lumen. In addition, the p8 signal still clearly overlapped with chromosomes at anaphase (Fig. 1F). This localization of p8 during mitosis has not been reported for any TFIID subunits, suggesting that p8 could have different dynamics than those of other TFIID components.

In the context of TFIID, p8 physically interacts with p52, and this interaction has been reported to be important for TFIID functions (Kainov et al., 2008; Cruz-Becerra et al., 2016). We therefore decided to analyse whether the dynamics of p52 were similar to those of p8 during syncytial blastoderm mitotic cycles. Although several very specific anti-p52 antibodies exist for western blot analysis (Villicaña et al., 2013), these do not work well for

immunostaining. Therefore, we generated transgenic flies expressing recombinant p52 tagged with enhanced yellow fluorescent protein (EYFP). We previously reported that this EYFP-p52 construct could rescue several lethal *p52* alleles (Cruz-Becerra et al., 2016), suggesting that this fusion protein is functional. Furthermore, this result provides confidence that the dynamics of this protein observed *in vivo* are similar to those of the wild-type protein. To visualize the different mitotic cycle phases through chromatin dynamics, we took advantage of the previously reported transgenic line in which histone H2Av is fused to red fluorescent protein (H2Av-RFP) (Schuh et al., 2007). Fig. 2A (also see Movie 1 and Fig. S1) shows *z*-stacks projections of confocal images of an embryo expressing EYFP-p52 and H2Av-RFP. Although some signal was observed in the cytoplasm at the NC11 interphase, most EYFP-p52

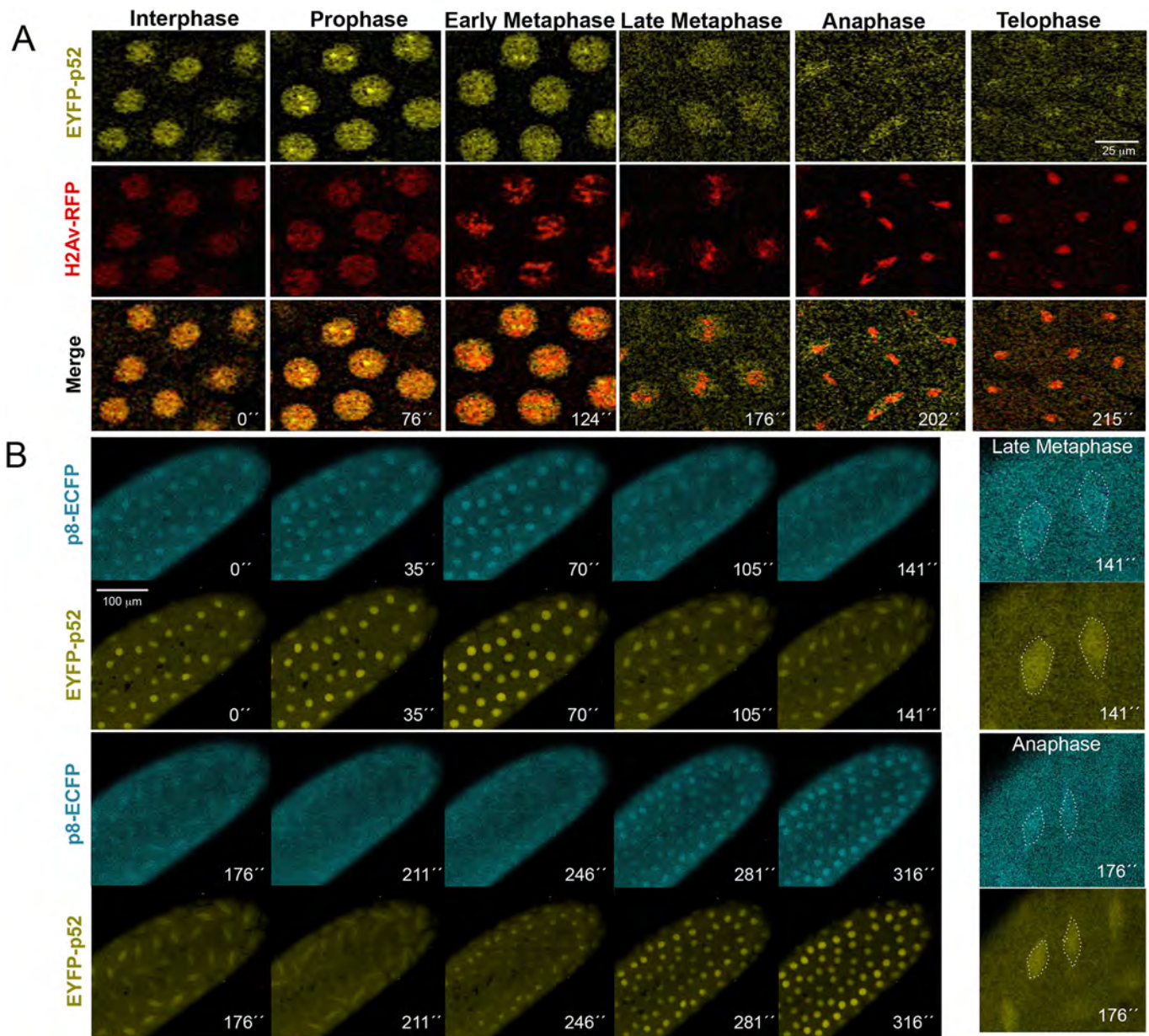


Fig. 2. The core and CAK TFIID subunits follow similar dynamics during the syncytial blastoderm stage. (A) Visualization of an *EYFP-p52/H2Av-RFP* (yellow/red) transgenic pre-MBT embryo during a nuclear division cycle (the time in seconds is shown). The different stages of the mitotic cycle were identified using H2Av-RFP. Also see Movie 1. (B) Time-lapse analysis of a pre-MBT embryo expressing EYFP-p52 (yellow) and p8-ECFP (cyan) (also see Movies 2 and 3, and Fig. S1).

signal overlapped with chromosomes, and intriguingly, EYFP-p52 nuclear foci were clearly observed (Fig. 2A; Movie 1). During early prophase, the EYFP-p52 signal was reduced in the cytoplasm and enriched in the nucleus. Furthermore, the number of EYFP-p52 nuclear foci was increased (in general, two to four foci per nucleus could be observed), and they were more evident at late prophase (Fig. 2A; Movie 1). The nuclear EYFP-p52 signal was reduced, but some EYFP-p52 surrounded chromosomes during metaphase and overlapped with the chromosomes at anaphase. While nuclear EYFP-p52 was significantly decreased at telophase, some EYFP-p52 overlap with the chromosomes was always detectable (Fig. 2A; Movie 1; Fig. S1).

Although the p8 chromosomal localization detected by immunostaining is clearest (Fig. 1), probably due to the tissue being fixed, the dynamic behaviour observed *in vivo* for EYFP-p52 was similar during NC8 to NC12 (Fig. 2), suggesting that these proteins follow the same dynamics in pre-MBT embryos. To confirm these results, we next analysed the dynamics of p8 *in vivo* using transgenic flies expressing recombinant p8 tagged with enhanced cyan fluorescent protein (ECFP). We have previously successfully demonstrated the functionality of the p8-ECFP recombinant protein by performing rescue experiments in a p8-null background (Cruz-Becerra et al., 2016). To directly compare the *in vivo* dynamics of p8 and p52, we generated flies that simultaneously expressed EYFP-p52 and p8-ECFP. Fig. 2B (see also Movie 2) shows the time-lapse sequence of one mitotic cycle in a pre-MBT embryo expressing both transgenes. We observed overlap in the signal of these proteins at all nuclear division stages, indicating highly similar dynamic oscillations between the nucleus and cytoplasm for these two components of the TFIH core subcomplex. Specifically, the nuclear localization of p52 and p8 during mitosis, which has never been described for any TFIH subunits in any organism, was intriguing.

The p52 subunit of TFIH is essential for anchoring XPB to the complex (Jawhari et al., 2002); in addition, its role in modulating the ATPase activity of XPB has been extensively documented (Coin et al., 2007; Fregoso et al., 2007). Thus, we reasoned that if the TFIH core is assembled during mitosis, XPB would behave similarly to p8 and p52. In our previous studies, we observed the exclusion of XPB from mitotic chromosomes in gastrulated embryos using an antibody against XPB (Aguilar-Fuentes et al., 2006). We therefore analysed the *in vivo* dynamics of fluorescently tagged XPB in syncytial blastoderm embryos, as was done for p8 and p52. XPB-EGFP or XPB-mCherry recombinant proteins follow dynamics along the mitotic cycle similar to those of p8 and p52 (Movie 3 and Fig. S2). Therefore, the *in vivo* results indicate that these core TFIH subunits are retained in the nuclear lumen and suggest colocalization with the chromosomes during mitosis.

As mentioned previously, the holo-TFIH complex (composed of the core and CAK subcomplexes) is required for transcription by RNAPII. Considering that ZGA occurs during the pre-MBT, we expected to observe core and CAK colocalization in transcriptionally active nuclei in pre-MBT embryos (i.e. at the interphase of NC8 and onwards). However, in addition to the nuclear localization of TFIH core subunits during interphase, in this study, we noted some nuclear localization during mitosis in pre-MBT embryos (Figs 1 and 2). Thus, we considered the possibility of nuclear CAK localization during mitosis. To answer this question, we tested several antibodies against Cdk7, MAT1 and CycH for immunostaining in wild-type embryos (data not shown, Fig. S3). Intriguingly, CycH, a CAK component whose localization has not been analysed in pre-MBT embryos, showed

similar dynamics to the core components of TFIH described in this study. The CycH signal overlapped with chromosomes during mitosis, although a significant signal was detected in the cytoplasm, as expected (Fig. S3). Furthermore, we constructed flies that expressed a recombinant Cdk7-EGFP protein that showed the same dynamics as CycH and the core TFIH subunits (Fig. S1; Movie 4). This indicates that at least part of both the core and CAK subunits remain in the nucleus during mitosis, and partially overlap with the mitotic chromosomes in pre-MBT embryos.

TFIH foci are present throughout most of the syncytial blastoderm mitotic cycle and colocalize with active transcriptional foci

As indicated previously, EYFP-p52 was observed in granular structures in the nuclei of syncytial blastoderm embryos. These nuclear foci could also be visualized in transgenic embryos expressing XPB-mCherry and XPB-EGFP (Movie 3). The foci appeared simultaneously in nuclei at the anterior and posterior poles during interphase and were maintained until early metaphase (Fig. 3A,C). In addition, the foci located randomly in the nuclear lumen, with no specific apical or basal positioning. Interestingly, these EYFP-p52 foci did not overlap with the areas showing highest chromatin compaction (indicated by the low intensity of the H2Av-RFP nuclear signal in these regions) (Fig. 3A,C), suggesting that transcription could take place in these foci. Furthermore, it is known that in syncytial blastoderm nuclei, the form of RNAPII that is phosphorylated at serine 5 (RNAPIIS5P) is enriched in two foci (Seydoux and Dunn, 1997; Chen et al., 2013) that correspond to histone locus bodies (HLBs; these contain the histone genes that are organized in several tandem repeats, as well as factors required for histone mRNA synthesis; Salzler et al., 2013). Histone genes are transcribed at the S-phase during the pre-MBT (Nizami et al., 2010; Salzler et al., 2013); thus, these foci could be sites of TFIH enrichment. Therefore, by assessing the dynamics of TFIH foci in time-lapse experiments, we could potentially visualize spatial and temporal gene expression in pre-MBT stages. Immunostaining of EYFP-p52 early syncytial blastoderm embryos using an antibody against the RNAPIIS5P revealed that the signal of active RNAPII overlapped with some EYFP-p52 foci (Fig. 3B), suggesting that these nuclear structures are sites of basal transcription machinery recruitment for the zygotic transcription of the histone genes.

Interestingly, by following the same foci from interphase to late metaphase in the mitotic nuclei of syncytial blastoderm embryos (arrows in Fig. 3C) revealed that although histone genes are highly transcribed at S phase (Guglielmi et al., 2013), the TFIH foci were maintained when mitosis was initiated, suggesting that transcription might overlap with the first stages of mitosis at the pre-MBT. Furthermore, for the p52 subunit of TFIH, we observed three or four foci in some nuclei (Fig. 3A; Movies 1 and 3); however, the functions of the additional foci remain to be determined.

Essential role of TFIH for proper mitosis in pre-MBT stages

The overlapping of the TFIH signal with chromosomes at prophase, metaphase and anaphase suggests a possible role for TFIH in mitosis in the pre-MBT embryo. The CAK TFIH subcomplex plays an important role in controlling the cell cycle (reviewed in Schachter and Fisher, 2013); mutations in Cdk7 disrupt mitosis (Larochelle et al., 2007). Furthermore, depletion of XPD in *Drosophila* embryos reportedly causes mitotic defects by inducing Cdk7 localization on mitotic chromosomes (Lee et al., 2010). Nonetheless, a TFIH-independent role for the XPD core subunit in

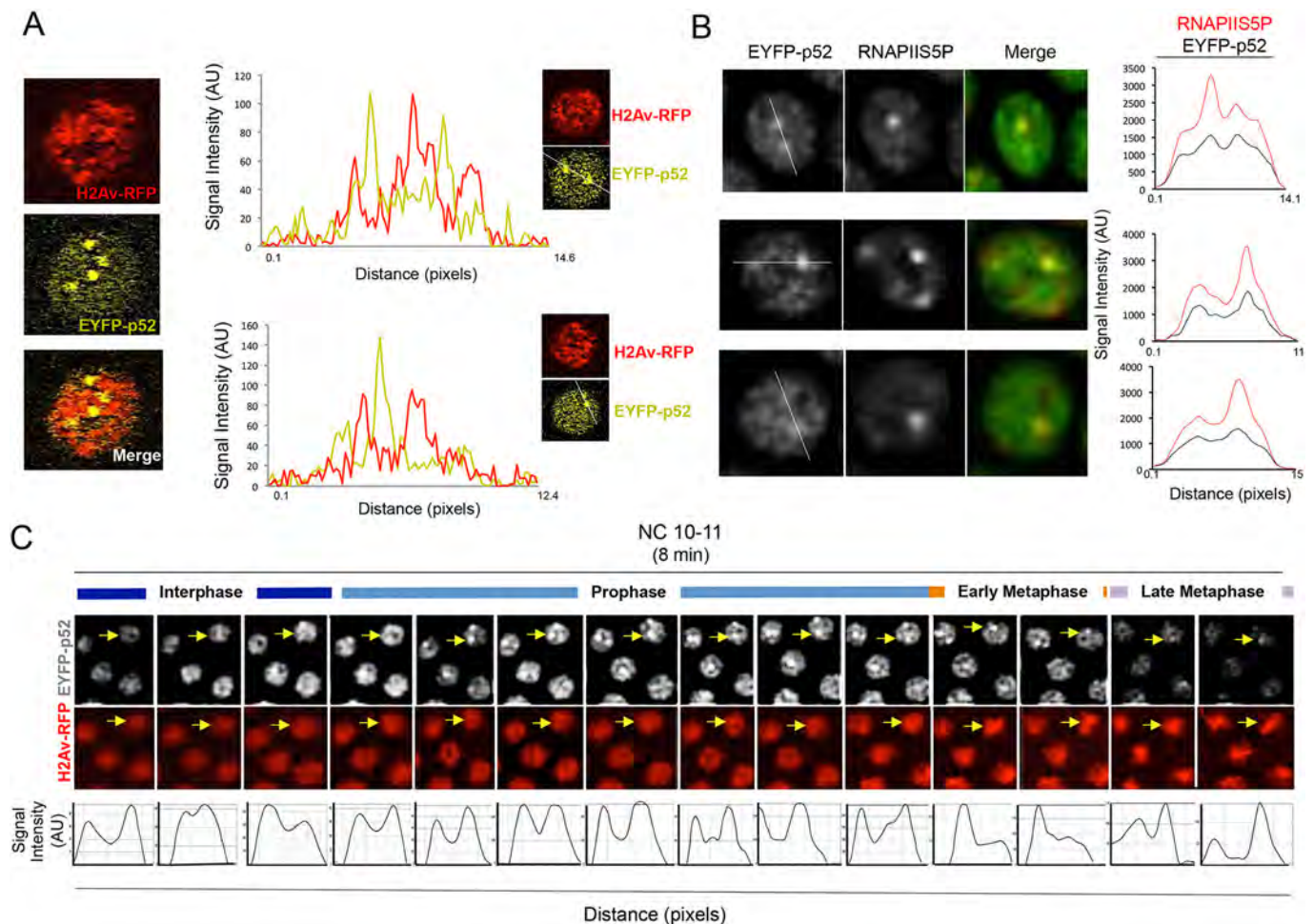


Fig. 3. TFIIF is enriched at transcribed chromatin in most of the mitotic cycle in the syncytial blastoderm. (A) Nuclei at interphase expressing EYFP-p52 (yellow) and H2Av-RFP (red) recombinant proteins in which TFIIF foci are observed. The plots are scans of both signals in a single confocal plane showing that regions of TFIIF enrichment are reduced in the H2Av-RFP signal. The lines on the nuclei indicate the scanning track. (B) RNAPII-enriched regions colocalize with TFIIF foci at the syncytial blastoderm. The plots show the location in a single plane of both signals for the line shown in the left panel. (C) Time-lapse analysis of TFIIF foci during a replicative cycle. TFIIF (EYFP-p52) foci (indicated by the arrows) were followed during a replicative cycle using the H2Av-RFP recombinant protein as a reference for stage. The time-lapse shows nuclei from the beginning of the interphase to metaphase in a syncytial blastoderm embryo between the NC10 and NC11 replicative cycles. The plots show the location in a single plane of the TFIIF foci. AU, arbitrary units.

chromosome segregation during mitosis has been suggested (Ito et al., 2010). However, whether core TFIIF subunits other than XPD participate in mitosis remains unknown. Thus, we analysed the effects of the absence of core TFIIF subunits on syncytial blastoderm embryos.

We recently characterized a *p8*-null allele that shows semi-lethality, as ~25% of the expected homozygous organisms develop into adults. Of these adults, the males are sterile, and the females lay fewer eggs than wild-type flies (Herrera-Cruz et al., 2012). Importantly, the embryos laid by homozygous *p8*-null females do not have maternally deposited *p8* and are thus an excellent tool to analyse the phenotypes produced by the absence of *p8* in the early embryos. As shown in Fig. 4A, in homozygous *p8*-null females, although oogenesis could be completed, ~90% of the egg chambers and laid eggs were smaller than wild type (Fig. 4C); this suggests transcriptional deficiencies during oogenesis, as we have previously observed in other organs with TFIIF loss of function (Fregoso et al., 2007; Villicaña et al., 2013). We next analysed the effect of the absence of *p8* in early syncytial blastoderm embryos, which is when embryos are committed to nuclear division cycles. Intriguingly ~20% of the embryos escape lethality and develop into adults

without apparent phenotype. By staining microtubules and chromosomes in *p8*-null embryos, we observed failure to exit mitosis and mitotic catastrophes with different levels of severity in ~60% of the embryos after NC8 (Fig. 4C–F; Fig. S4A). In *p8*-null nuclei at metaphase, the spindle was misaligned and longer than that in wild-type embryos; furthermore, some metaphase chromosomes were under-compacted (Fig. 4D), and isolated centrosomes and chromosomes were observed (Fig. 4D,E). During anaphase, the localization of histone H3 phosphorylated at serine 10 (H3S10P) was restricted to the telomeric region of the chromosomes in wild-type embryos, while in *p8*-null embryos, the localization of this mitotic marker was maintained throughout the chromosome bodies, even at telophase (Fig. 4F). All these defects in mitosis were suppressed in *p8*-null embryos expressing the *p8*-ECFP recombinant protein (Cruz-Becerra et al., 2016).

Importantly, human cells derived from TTDA patients, affected in the human *p8* gene (*GTF2H5*) have been shown to exhibit reduced levels of all other TFIIF subunits (Vermeulen et al., 2000; Giglia-Mari et al., 2004). However, we recently reported that the absence of *p8* in *Drosophila* testes does not affect the levels of the other TFIIF subunits (Cruz-Becerra et al., 2016). To determine

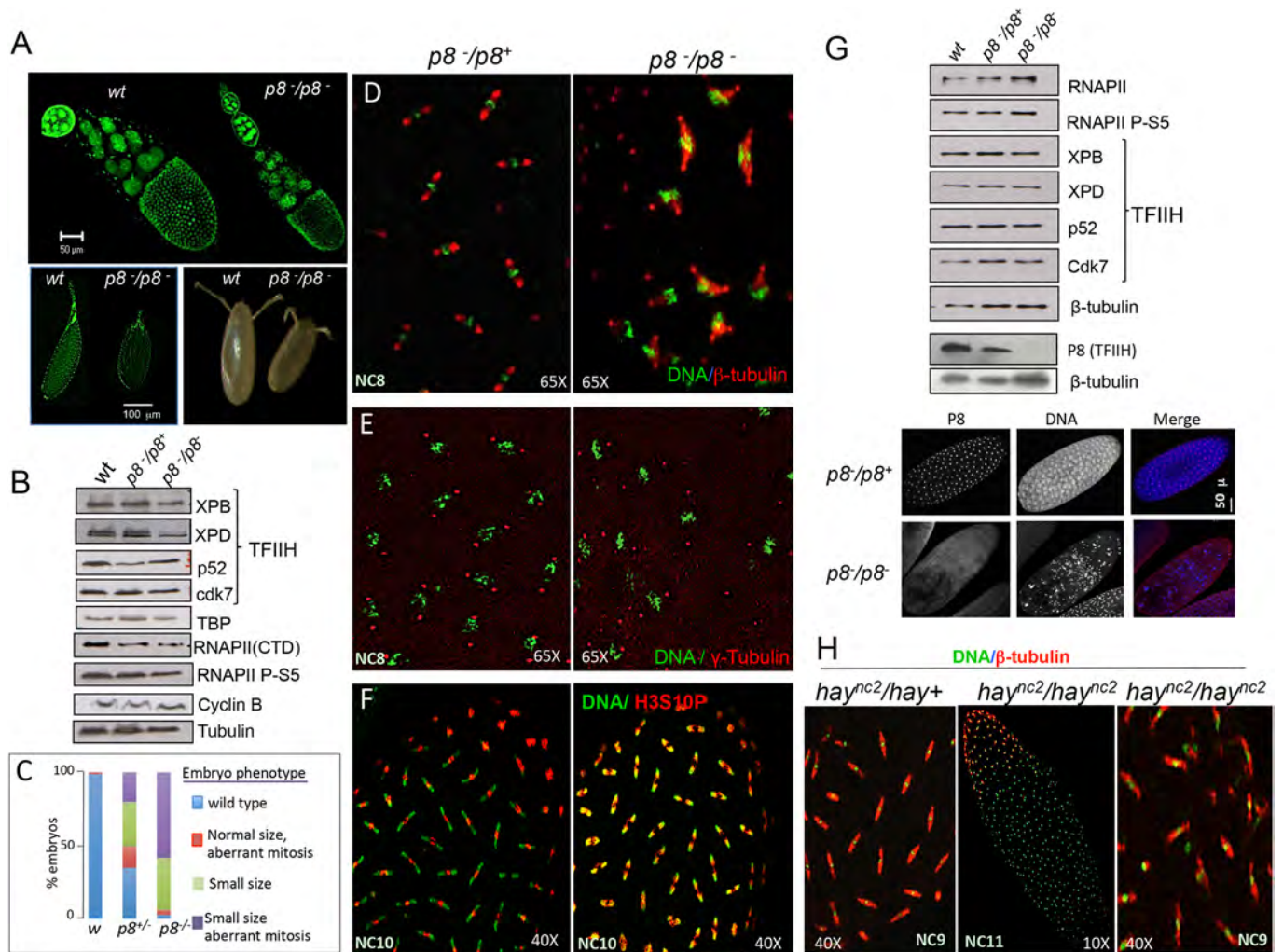


Fig. 4. The absence of functional TFIIH causes catastrophic mitosis in the syncytial blastoderm embryo. (A) Wild-type (*wt*) and *p8*-null ovaries, and the embryos laid, stained with SYTOX Green to visualize the DNA, from wild-type and *p8*-null mothers. (B) Western blot from total ovaries stained for proteins for different subunits of TFIIH, components of the basal transcription machinery and β -tubulin as loading control. (C) Penetrance of the small embryo and catastrophic mitosis phenotypes. (D) Comparison of mitosis between wild-type and *p8*-null embryos. The mitotic spindle was detected with an anti- β -tubulin antibody (red); DNA (green) was detected with SYTOX Green. (E) Detection of centrosomes in wild-type and *p8*-null embryos using a γ -tubulin antibody (red) and DNA (green) with SYTOX Green. (F) Alterations in mitosis cycle timing in *p8*-null embryos as compared to wild-type organisms as determined by visualizing the presence of H3 phosphorylated at serine 10 (H3S10P). (G) Upper panel: western blot from *p8*-null or wild-type embryos to determine the levels of different TFIIH subunits (including *p8*) and the RNAPII CTD phosphorylated at serine 5 (RNAPII P-S5). Lower panel: immunostainings of homozygous *p8*-null or *p8*-heterozygous embryos using the anti-*p8* antibody. (H) Syncytial blastoderm embryos from heterozygous *hay*^{nc2} females and embryos from homozygous *hay*^{nc2} females stained with the anti- β -tubulin antibody and with SYTOX Green. A desynchronized embryo is presented in the middle, and a magnification showing catastrophic mitosis in a homozygous *hay*^{nc2} embryo is presented on the right. All the genotypes and antibodies used are indicated in the figure.

whether the catastrophic mitosis observed in syncytial blastoderm embryos lacking *p8* could be an indirect effect of the deficiency of other TFIIH subunits, we analysed the levels of some core and CAK subunits in egg chambers and embryos from *p8*-null flies. Intriguingly, we did not find a significant reduction in the levels of the TFIIH components analysed, nor did we find changes in the levels of other transcription factors or RNAPIIS5P (Fig. 4G). Therefore, the mitotic catastrophes observed in *p8*-null embryos were only due to the absence of *p8*, suggesting that *p8* might have a role in mitosis.

The question of whether mutations in TFIIH subunits other than *p8* and XPD also affect mitosis in early embryos emerged. Thus, we analysed the XPB (*haywire* in fly) mutant allele *hay*^{nc2} (Mounkes et al., 1992; Merino et al., 2002). This allele is also semi-lethal, and homozygous females having this allele lay few fertilized eggs.

Intriguingly, ~60% of these embryos showed defects in mitosis, including free centrosomes, incompletely compacted chromosomes at metaphase and isolated chromosomes not attached to the mitotic spindle (Fig. 4F). On the other hand, less than 10% of the embryos from heterozygous *hay*^{nc2} females presented defects in mitosis. Because both XPB and *p8* are core TFIIH subunits, analysing whether CAK loss of function also generates embryos with defects in mitosis was relevant. To achieve this, we constructed flies that expressed RNAi against the Cdk7 transcript during oogenesis, thereby generating embryos with reduced levels of this CAK component (Fig. S4B). These embryos showed mitotic defects, including aberrant mitotic spindles and isolated chromosomes (Fig. S4C,D). In summary, altering the XPB and Cdk7 functions in pre-MBT embryos causes defects in mitosis similar to those resulting from depleting XPD (Li et al., 2010) or *p8* (this study).

Therefore, deficiency of core and CAK TFIIF subunits in the early fly embryo catastrophically affects mitosis.

Misregulation of the expression of maternal transcripts caused by *p8* depletion is correlated with the penetrance of mitotic defects in the early embryo

In addition to the fundamental activity of TFIIF in serine 5 phosphorylation of the CTD of RNAPII, TFIIF is also necessary for opening the DNA at the transcription initiation site (Kouzine et al., 2013; Alekseev et al., 2017). Therefore, although wild-type levels of RNAPIIS5P were observed in *p8*-null ovaries and early embryos (Fig. 4B,G), we were intrigued by the possibility that the mitotic phenotypes observed were caused by transcriptional defects during oogenesis or in pre-MBT stages. To determine the effect of the absence of *p8* on global transcription, we performed RNA-seq analysis of individual *p8*-null ($p8^{-}/p8^{-}$) and *p8*-heterozygous ($p8^{+}/p8^{-}$) embryos, which show mitotic defects and wild-type phenotypes, respectively (Fig. 5A). Using this strategy, we could determine whether the mitotic defects in *p8*-depleted embryos were associated with the deregulation of maternal transcripts or the

deregulation in genes expressed pre-MBT. *p8*-null or *p8*-heterozygous females expressing H2Av-RFP were used to collect embryos between NC9 and NC10. Then, total RNA from two heterozygous (Fig. 5A, embryos G and H) and three *p8*-null (Fig. 5A, embryos A, B and E) organisms were subjected to single-embryo transcriptome analyses by following the single-cell protocol for RNA-seq. These analyses showed genome map rates higher than 90% and identified nearly 8000 different transcripts per embryo and almost 4000 genes that were differentially regulated between the wild-type and the mutant embryos (Tables S1–S7).

The correlation analysis of the two control embryo (*p8* heterozygous) samples indicated similar levels of most of the detected transcripts; although, as expected, some variability was observed for the biological samples (Fig. 5B, embryo H versus embryo G). In contrast, comparison of the controls (embryos G or H) and *p8*-null embryos (embryos A, B and E) showed a large variation in the levels of many transcripts (Fig. 5B; Fig. S6). The *p8*-null embryos A and B clearly showed more severe mitotic phenotypes than embryo E, which showed less dramatic defects (Fig. 5A). Interestingly, the global gene expression in embryo E was

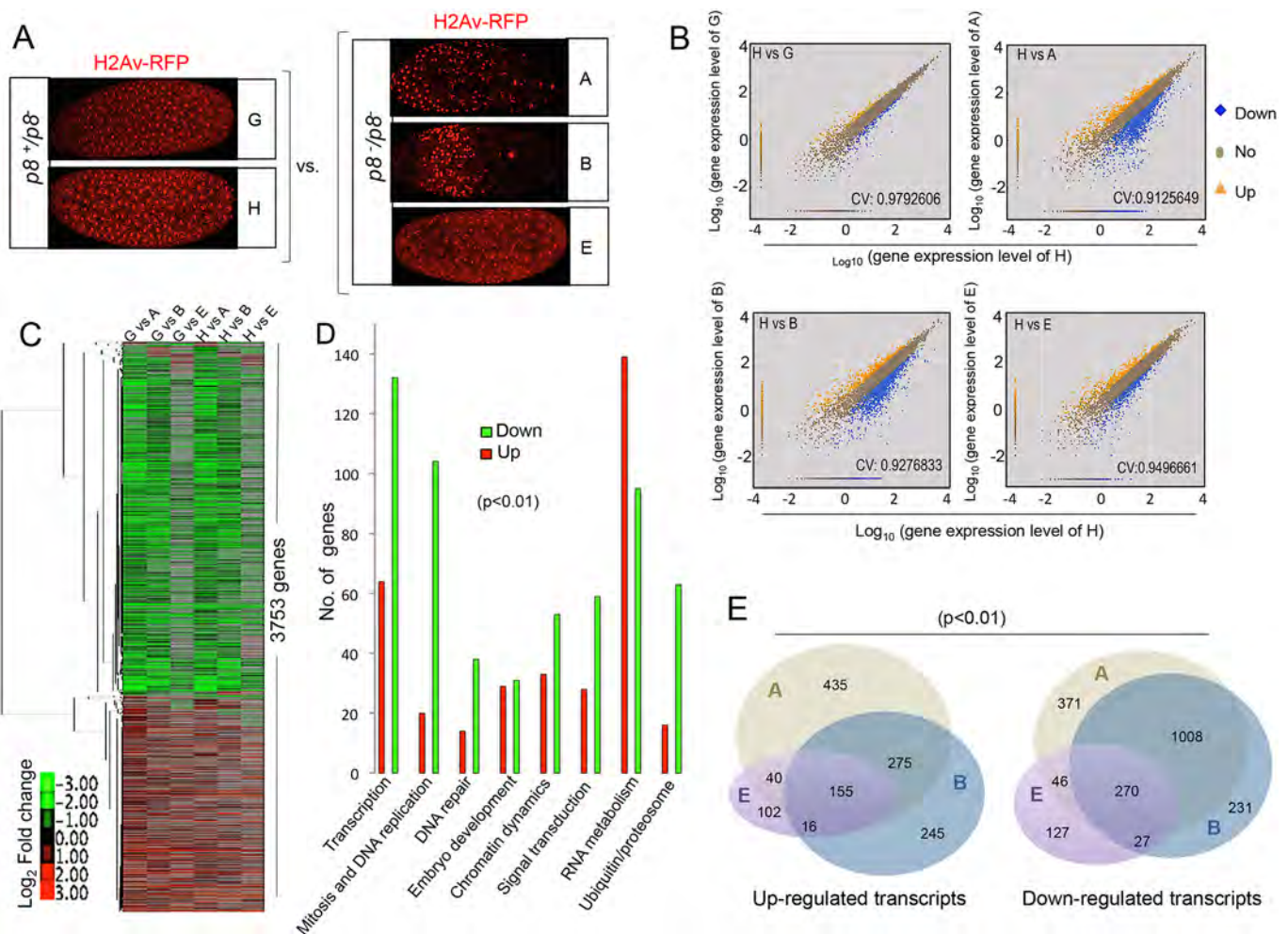


Fig. 5. Transcriptomic analysis of *p8*-null or *p8*-heterozygote embryos. (A) Heterozygous *p8*-null ($p8^{+}/p8^{-}$; *H2Av-RFP/+*) embryos (embryos G and H) and homozygous *p8*-null ($p8^{-}/p8^{-}$; *H2Av-RFP/+*) embryos, with defects in mitosis (embryos A, B and E) at division NC9 and NC10, used for single embryo RNA-seq analysis. (B) Correlation plots indicating the differences in gene expression between *p8*-null homozygous (A, B and E) and control (G and H) embryos. (C) Heat map showing the variations in gene expression between the two control embryos (G and H) and the three mutants (A, B and E). Note the low variation in embryo E compared with embryos A and B with more severe mutant phenotype. (D) Ontology analysis of affected genes identified by RNA-seq, whose function may be related to the catastrophic mitosis phenotype. (E) Venn diagrams that indicate the number of down- and upregulated genes affected in all the *p8*-depleted embryos.

less affected (Fig. 5B,C), indicating a correlation between the severity of the mitotic defect phenotype and the degree of influence on global gene expression. We identified a group of 270 downregulated genes shared among all of our p8-depleted samples, suggesting that these genes are particularly sensitive to the transcriptional deregulation caused by the absence of p8 (Fig. 5E). The RNA-seq data showed that most of the deregulated transcripts in the p8-null embryos came from the maternal contribution (Table S7), suggesting aberrant transcription during oogenesis in p8-null females. In agreement with the p8-null embryo phenotypes, gene ontology analysis based on cellular functions that may be related to the observed phenotypes showed the downregulation of 104 genes encoding factors that participate in DNA replication or mitosis (Fig. 5D; Table S8). Intriguingly, 139 genes that participate in RNA metabolism are upregulated (Fig. 5D; Table S9). Taken together, these observations suggest that deregulated gene expression during oogenesis caused by mutations in TFIIH leads to catastrophic mitosis in early syncytial blastoderm embryos.

DISCUSSION

ZGA in the *Drosophila* embryo occurs in two waves, with the first involving the transcription of a small group of genes at the pre-MBT, and the second including the transcription of an important fraction of the zygotic genome at the MBT (Lee et al., 2014). At these stages, the dynamics of basal transcription machinery have been poorly studied. In *Drosophila*, only a few reports by our group and others on some TFIIH subunits (Aguilar-Fuentes et al., 2006; Li et al., 2013) and RNAPII and TATA-binding protein (TBP) dynamics (Seydoux and Dunn, 1997; Wang and Lindquist, 1998) exist. In this work, we have extended our analysis of the dynamics of TFIIH by immunostaining and *in vivo* studies using recombinant fluorescent proteins, and found that some subunits of TFIIH overlap with mitotic chromosomes during nuclear division cycles in pre-MBT embryos. In addition, we found that mutations in some TFIIH core subunits cause catastrophic mitosis during early embryonic development, at least partially resulting from the deregulation of global transcription during oogenesis.

Dynamic behaviour of TFIIH in *Drosophila* early syncytial blastoderm nuclear division cycles

Several studies have been performed to observe the subcellular and molecular dynamics of TFIIH during transcription, and the DNA damage response in cells during interphase (Giglia-Mari et al., 2006; Nonnekens et al., 2013). Nonetheless, the dynamics of TFIIH components throughout the cell cycle progression have been less extensively studied. Importantly, although a previous study showed the dynamics of the XPB core subunit *in vivo* in a mouse XPB-YFP knock-in model, that report focused on comparing the mobility of fluorescently tagged XPB in interphase cells that had the capacity to proliferate and in post-mitotic non-proliferative cells. By using FRAP, Giglia-Mari and colleagues showed that XPB mobility was lower in post-mitotic differentiated cells (Giglia-Mari et al., 2009). Thus, to our knowledge, our work is a pioneer study of the *in vivo* dynamics of TFIIH throughout the cell cycle progression in the context of the whole organism. Here, we showed that TFIIH is highly dynamic during fast nuclear division cycles in the early fly embryo. Our *in vivo* analyses showed that, although TFIIH overlaps with chromosomes, the intensity of the nuclear signal changed with cell cycle progression. During interphase, cytoplasmic localization was detected, but most of the TFIIH signal was observed in the nucleus. Interestingly, during chromatin condensation at prophase,

the signal further increased inside the nucleus, while decreasing in the cytoplasm, suggesting translocation of TFIIH into the nucleus at this stage. Additionally, TFIIH foci were clearly observed at this stage, and these data suggest that some transcription at the pre-MBT occurs at this point.

We have demonstrated that TFIIH oscillated between the cytoplasm and the nucleus during cell cycle progression in pre-MBT embryos. A previous study described the localization of general transcription factors in HeLa cells during the transition from mitosis to G1. Unlike TBP, which was found on chromosomes throughout mitosis, other general transcription factors, including RNAPII, XPB and XPD, were localized in daughter nuclei only after mid-telophase (Prasanth et al., 2003). Hence, this intriguing behaviour observed for TFIIH in fly embryos could reflect a mechanism involved in the regulation of transcription during very short nuclear division cycles, such as those that occur at pre-MBT stages, where incomplete nuclear membrane breakdown occurs.

Importantly, the observation that the CycH and Cdk7 signals overlap with chromosomes during mitosis in wild-type embryos differs from a previous study suggesting that CAK is nuclear during interphase, but is excluded from mitotic chromosomes through an XPD-dependent mechanism. Specifically, in syncytial blastoderm embryos, XPD-depletion causes defects in mitosis and aberrant Cdk7 overlapping with chromosomes at anaphase (Li et al., 2010). Despite our efforts to analyse the Cdk7 localization pattern by immunostaining with several available antibodies, we could not observe clear nuclear signals during interphase in wild-type embryos. However, we did observe identical localization using the Cdk7-EGFP protein to that observed for other TFIIH subunits during replicative cycles in the syncytial blastoderm, including overlap between the CAK and mitotic chromosomes (Fig. S4; Movie 4). Thus, this discrepancy could be a consequence of the antibodies used to recognize Cdk7 not generating a good signal in embryonic immunostaining. Indeed, in agreement with our observations for other TFIIH subunits in pre-MBT embryos, Ito and colleagues reported that the XPD signal spreads in the nucleoplasm, but is enriched at the mitotic spindle at metaphase in human cells (Ito et al., 2010). Interestingly, an increasing number of studies have recently highlighted the existence of 'mitotic bookmarking' by several transcription factors (reviewed in Kadauke and Blobel, 2013). Our observations on TFIIH dynamics suggest that TFIIH might also be involved in bookmarking during early embryonic development, and this is currently under investigation.

TFIIH is maintained on mitotic chromosomes in the early fly embryo

An interesting feature of TFIIH dynamics is its enrichment in nuclear foci that overlap with active RNAPII foci that could correspond to HLBs. Transcriptional activation of the histone gene cluster during the pre-MBT in *Drosophila* has been well documented (Günesdogan et al., 2014). Except for histone H1, the transcription of histone genes depends on TATA box promoters and TBP binding. Indeed, HLBs are sites of TBP and RNAPII enrichment (Isogai et al., 2007; Guglielmi et al., 2013). Importantly, HLBs have been observed from interphase through metaphase (Isogai et al., 2007), as we also observed for TFIIH foci. As previously reported for HLBs (Salzler et al., 2013), TFIIH foci are maintained throughout the division cycle and disappear at late metaphase; a similar behaviour is observed for active RNAPII (Blythe and Wieschaus, 2016). Despite the generally accepted model of transcription being shut off in mitotic nuclei (Gottesfeld and Forbes, 1997), some transcription can reportedly occur during

mitosis; for instance, transcription of the *CycB* gene in HeLa cells (Sciortino et al., 2001) and α -satellite DNA in centromeres (Chan et al., 2012). In the pre-MBT embryo, the time period for transcription is very short because the nuclear cycle has only a brief interphase and mostly consists of the S and M phases (Farrell and O'Farrell, 2014). Histone genes are transcribed during S phase (Guglielmi et al., 2013); however, for the pre-MBT stages, it is possible that transcription could be extended during the fast nuclear division cycle to achieve the amount of pre-MBT mRNA required during this developmental stage. The overlap of some TFIID subunits with chromosomes in all nuclei division cycle stages in early embryos suggests a mechanism to retain some TFIID for fast transcriptional re-activation after mitosis.

Early embryonic catastrophic mitosis and transcriptional defects caused by mutations in the TFIID core and CAK subunits

Previous reports have suggested a role for the TFIID XPD subunit in controlling Cdk7 functionality in cell cycle modulation during early fly embryogenesis (Chen et al., 2003; Li et al., 2010). In addition, XPD is a component of the MMXD complex involved in chromosome segregation in human cells (Ito et al., 2010). Furthermore, the ATPase activity of XPD is well known to be important for the formation of the open complex in RNAPII-transcribed genes, and it acts as a scaffold to maintain the integrity of TFIID during transcription (Lainé et al., 2006; Kuper et al., 2014). Thus, further analyses to clarify the role of XPD in cell cycle regulation are required. Herein, we showed that mutations in the p8 and XPB core subunits of TFIID, as well as the depletion of Cdk7, cause defects in mitosis, suggesting that this phenotype is not exclusive to the Cdk7 delocalization seen upon XPD depletion in fly embryos. In accordance, Matsuno and colleagues reported that mutations in p52, XPB and Cdk7 produced similar defects in cell cycle progression during fly eye imaginal disc development, suggesting that CAK activity in cell cycle regulation is performed by the holo-TFIID complex (Matsuno et al., 2007). However, although a direct role for the holo-TFIID complex in mitosis cannot be discarded, the evidence presented in this work indicates that the absence of p8 leads to deregulated gene expression during oogenesis, thus affecting the expression of many genes that participate in coordinated mitosis at the pre-blastoderm and syncytial blastoderm stages. On the other hand, our results that show the presence of the CAK subunits on mitotic chromosomes, challenging the model that proposes that only in the absence of XPD is Cdk7 maintained on the chromosomes during mitosis, causing catastrophic mitosis (Li et al., 2010).

Interestingly, embryos from females with decreased abundance of the transcriptional activator *vielfältig/Zelda*, which activates the onset of zygotic transcription, display mitotic defects with some similarities to those reported for XPD-deficient embryo phenotypes, and those observed here for p8-null, homozygous *hay^{nc2}* and Cdk7-depleted embryos (Staudt et al., 2006). It is thus difficult to distinguish between possible roles of TFIID subunits in mitosis and indirect effects due to defective gene expression, particularly because some of the genes expressed at the pre-MBT stage may also participate in mitosis.

Despite the essential role for p8 in maintaining a steady-state level of cellular TFIID proposed in human cells (Giglia-Mari et al., 2004), we did not observe a significant reduction in other TFIID subunits in p8-depleted ovaries or embryos (this work) or in spermatogenesis (Cruz-Becerra et al., 2016). Furthermore, our transcriptome analyses of p8-depleted embryos highlight the importance of this protein in

global transcription during development, as was recently observed in a p8-knockout mouse model, where absence of p8 caused embryonic lethality (Theil et al., 2013).

In conclusion, our analysis of TFIID at the pre-MBT shows how this complex is distributed at chromosomes in waves during mitotic cycles, and that depletion of TFIID subunits affects transcription during oogenesis and causes catastrophic mitosis in syncytial blastoderm embryos. This knowledge provides important insights into the functions of TFIID during early embryonic development.

MATERIALS AND METHODS

Fly stocks

Oregon R was used as the wild-type strain, except when indicated. The *His2Av-RFP* transgenic lines (BL23650 and BL23651) and the GAL4 driver line (BL7063) were obtained from the Bloomington *Drosophila* Stock Center. The *Cdk7i* line (v10424) was obtained from the Vienna *Drosophila* Resource Center. The EYFP-p52, p8-ECFP and XPB-EGFP transgenic lines, as well as the p8-null and *hay^{nc2}* alleles were previously described (Cruz-Becerra et al., 2016; Herrera-Cruz et al., 2012; Regan and Fuller, 1988).

Transgenic flies

Recombinant DNA was generated by tagging the 5' or 3' end of the corresponding full-length cDNA coding sequence with the DNA sequence of the different fluorescent proteins used in this work. The recombinant DNA was cloned into the pCaSper-Hsp83 vector and used by BestGene Inc. to generate transgenic flies through standard microinjection protocols.

RNA interference assays

Line *v10442* (<http://flybase.org/reports/FSbf0000074319.html>) expressing a dsRNA against the *cdk7* transcript under the control of the *UAS-Gal4* system was crossed with the *P{mata4-GAL-VP16}* driver that expressed Gal4 from the α -*Tubulina67c* promoter in the ovaries. The resulting females flies with the genotype *v10442/+; P{mata4-GAL-VP16}/+* were crossed with wild-type males and the ovaries and the embryos analysed in western blot experiments and embryo immunostainings.

Immunostaining

Embryos were fixed in formaldehyde and heptane as previously described (Reyraud et al., 1999; Aguilar-Fuentes et al., 2006; Chen et al., 2013), except that methanol was used as fixative for tubulin staining. Immunostainings were performed by standard methods with the antibodies against p8 (1:100; Herrera-Cruz et al., 2012), H₃S₁₀P (1:1500; cat. no SC-8656, Santa Cruz Biotechnology), AA4.3 (1:800; cat. no AB-579793, DSHB), ADL67.1 (1:600; cat. no AB-528336, DSHB), H14 (1:500; cat. no MMS-134-500, Covance), TBP (1:100; cat. no SC-273, Santa Cruz Biotechnology), CTD4H8 (1:100; cat. no SC-7495, Santa Cruz Biotechnology), GTU-88 (1:500; cat. no SC-11316, Abcam), CycH (1:500; cat. no SC-1662, Santa Cruz Biotechnology), and Alexa Fluor-conjugated secondary antibodies (Invitrogen). DNA was visualized by staining with either SYTOX Green (Invitrogen) or DAPI (Roche). To avoid flattening, the embryos were supported in handcrafted mounting chambers.

Live imaging

Embryos collected on apple juice plates were dechorionated with 50% bleach, extensively washed with water, placed on heptane glue-treated coverslips, immediately covered with Halocarbon oil 700 (Sigma) and analysed by confocal microscopy. Time-lapse movies from live embryos during at least two consecutive nuclear division cycles were acquired with an Olympus FV1000 confocal system coupled to an inverted microscope. Several embryos from at least two independent transgenic lines of each fluorescently labelled protein were analysed.

Western blot

Total protein extracts were analysed by immunoblotting with standard procedures. The following antibodies were used: 8WG16 (1:1500; cat. no

- Jawhari, A., Lainé, J.-P., Dubaele, S., Lamour, V., Poterszman, A., Coin, F., Moras, D. and Egly, J.-M. (2002). p52 Mediates XPB function within the transcription/repair factor TFIIH. *J. Biol. Chem.* **277**, 31761-31767.
- Kadauke, S. and Blobel, G. A. (2013). Mitotic bookmarking by transcription factors. *Epigenet. Chromatin.* **6**, 6.
- Kainov, D. E., Vitorino, M., Cavarelli, J., Poterszman, A. and Egly, J.-M. (2008). Structural basis for group A trichothiodystrophy. *Nat. Struct. Mol. Biol.* **15**, 980-984.
- Kouzine, F., Wojtowitz, D., Yamane, A., Resch, W., Kieffer-Kwon, K.-R., Bandle, R., Nelson, S., Nakahashi, H., Awasthi, P., Feigenbaum, L. et al. (2013). Global regulation of promoter melting in naive lymphocytes. *Cell* **153**, 988-999.
- Kuper, J., Braun, C., Elias, A., Michels, G., Sauer, F., Schmitt, D. R., Poterszman, A., Egly, J.-M. and Kisker, C. (2014). In TFIIH, XPD helicase is exclusively devoted to DNA repair. *PLoS Biol.* **12**, e1001954.
- Lainé, J. P., Mocquet, V. and Egly, J. M. (2006). TFIIH enzymatic activities in transcription and nucleotide excision repair. *Methods Enzymol.* **408**, 246-263.
- Larochelle, S., Merrick, K. A., Terret, M.-E., Wohlbold, L., Barboza, N. M., Zhang, C., Shokat, K. M., Jallepalli, P. V. and Fisher, R. P. (2007). Requirements for Cdk7 in the assembly of Cdk1/cyclin B and activation of Cdk2 revealed by chemical genetics in human cells. *Mol. Cell* **25**, 839-850.
- Lee, M. T., Bonneau, A. R. and Giraldez, A. J. (2014). Zygotic genome activation during the maternal-to-zygotic transition. *Annu. Rev. Cell Dev. Biol.* **30**, 581-613.
- Li, X., Urwyler, O. and Suter, B. (2010). *Drosophila* Xpd regulates Cdk7 localization, mitotic kinase activity, spindle dynamics, and chromosome segregation. *PLoS Genet.* **6**, e1000876.
- Li, X.-Y., Harrison, M. M., Villalta, J. E., Kaplan, T. and Eisen, M. B. (2014). Establishment of regions of genomic activity during the *Drosophila* maternal to zygotic transition. *eLife* **3**, e03737.
- Lott, S. E., Villalta, J. E., Schroth, G. P., Luo, S., Tonkin, L. A. and Eisen, M. B. (2011). Noncanonical compensation of zygotic X transcription in early *Drosophila* melanogaster development revealed through single-embryo RNA-seq. *PLoS Biol.* **9**, e1000590.
- Matsuno, M., Kose, H., Okabe, M. and Hiromi, Y. (2007). TFIIH controls developmentally-regulated cell cycle progression as a holo-complex. *Genes Cells* **12**, 1289-1300.
- Mazumdar, A. and Mazumdar, M. (2002). How one becomes many: blastoderm cellularization in *Drosophila melanogaster*. *BioEssays* **24**, 1012-1022.
- Merino, C., Reynaud, E., Vázquez, M. and Zurita, M. (2002). DNA repair and transcriptional effects of mutations in TFIIH in *Drosophila* development. *Mol. Biol. Cell* **13**, 3246-3256.
- Mounkes, L. C., Jones, R. S., Liang, B.-C., Gelbart, W. and Fuller, M. T. (1992). A *Drosophila* model for xeroderma pigmentosum and Cockayne's syndrome: haywire encodes the fly homolog of ERCC3, a human excision repair gene. *Cell* **71**, 925-937.
- Nizami, Z. F., Deryusheva, S. and Gall, J. G. (2010). Cajal bodies and histone locus bodies in *Drosophila* and *Xenopus*. *Cold Spring Harb. Symp. Quant. Biol.* **75**, 313-320.
- Nonnekens, J., Cabantous, S., Slingerland, J., Mari, P.-O. and Giglia-Mari, G. (2013). In vivo interactions of TTDA mutant proteins within TFIIH. *J. Cell Sci.* **126**, 3278-3283.
- Prasanth, K. V., Sacco-Bubulya, P. A., Prasanth, S. G. and Spector, D. L. (2003). Sequential entry of components of gene expression machinery into daughter nuclei. *Mol. Biol. Cell* **14**, 1043-1057.
- Pritchard, D. K. and Schubiger, G. (1996). Activation of transcription in *Drosophila* embryos is a gradual process mediated by the nucleocytoplasmic ratio. *Genes Dev.* **10**, 1131-1142.
- Regan, C. L. and Fuller, M. T. (1988). Interacting genes that affect microtubule function: the nc2 allele of the haywire locus fails to complement mutations in the testis-specific beta-tubulin gene of *Drosophila*. *Genes Dev.* **2**, 82-92.
- Reynaud, E., Lomelí, H., Vázquez, M. and Zurita, M. (1999). The *Drosophila* melanogaster homologue of the Xeroderma pigmentosum D gene product is located in euchromatic regions and has a dynamic response to UV light-induced lesions in polytene chromosomes. *Mol. Biol. Cell* **10**, 1191-1203.
- Salzler, H. R., Tatomer, D. C., Malek, P. Y., McDaniel, S. L., Orlando, A. N., Marzluff, W. F. and Duronio, R. J. (2013). A sequence in the *Drosophila* H3-H4 Promoter triggers histone locus body assembly and biosynthesis of replication-coupled histone mRNAs. *Dev. Cell* **24**, 623-634.
- Satija, R. and Bradley, R. K. (2012). The TAGteam motif facilitates binding of 21 sequence-specific transcription factors in the *Drosophila* embryo. *Genome Res.* **22**, 656-665.
- Schachter, M. M. and Fisher, R. P. (2013). The CDK-activating kinase Cdk7: taking yes for an answer. *Cell Cycle* **12**, 3239-3240.
- Schuh, M., Lehner, C. F. and Heidmann, S. (2007). Incorporation of *Drosophila* CID/CENP-A and CENP-C into centromeres during early embryonic anaphase. *Curr. Biol.* **17**, 237-243.
- Sciortino, S., Gurtner, A., Manni, I., Fontemaggi, G., Dey, A., Sacchi, A., Ozato, K. and Piaggio, G. (2001). The cyclin B1 gene is actively transcribed during mitosis in HeLa cells. *EMBO Rep.* **2**, 1018-1023.
- Seydoux, G. and Dunn, M. A. (1997). Transcriptionally repressed germ cells lack a subpopulation of phosphorylated RNA polymerase II in early embryos of *Caenorhabditis elegans* and *Drosophila melanogaster*. *Development* **124**, 2191-2201.
- Staudt, N., Fellert, S., Chung, H. R., Jäckle, H. and Vorbrüggen, G. (2006). Mutations of the *Drosophila* zinc finger-encoding gene vielfältig impair mitotic cell divisions and cause improper chromosome segregation. *Mol. Biol. Cell* **17**, 2356-2365.
- Theil, A. F., Nonnekens, J., Steurer, B., Mari, P.-O., de Wit, J., Lemaitre, C., Marteijn, J. A., Raams, A., Maas, A., Vermeij, M. et al. (2013). Disruption of TTDA results in complete nucleotide excision repair deficiency and embryonic lethality. *PLoS Genet.* **9**, e1003431.
- Villicaña, C., Cruz, G. and Zurita, M. (2013). The genetic depletion or the triptolide inhibition of TFIIH in p53-deficient cells induces a JNK-dependent cell death in *Drosophila*. *J. Cell Sci.* **126**, 2502-2515.
- Wang, Z. and Lindquist, S. (1998). Developmentally regulated nuclear transport of transcription factors in *Drosophila* embryos enable the heat shock response. *Development* **125**, 4841-4850.
- Yeom, E., Hong, S.-T. and Cho, K.-W. (2014). Crumbs interacts with Xpd for nuclear division control in *Drosophila*. *Oncogene* **34**, 2777-2789.
- Zurita, M. and Merino, C. (2003). The transcriptional complexity of the TFIIH Complex. *Trends Genet.* **10**, 578-584.
- Zurita, M., Reynaud, E. and Aguilar-Fuentes, J. (2008). From the beginning: the basal transcription machinery and onset of transcription in the early animal embryo. *Cell. Mol. Life Sci.* **65**, 212-227.