



BENEMERITA UNIVERSIDAD AUTÓNOMA DE PUEBLA



INSTITUTO DE CIENCIAS
CENTRO DE INVESTIGACIONES EN CIENCIAS
MICROBIOLÓGICAS

POSGRADO EN MICROBIOLOGÍA

***“IDENTIFICACIÓN Y CARACTERIZACIÓN DE LOS FACTORES
INVOLUCRADOS EN DECAIMIENTO DE ARN MENSAJERO
SIN SENTIDO (NMD) EN USTILAGO MAYDIS”***

TESIS

QUE PARA OBTENER EL GRADO DE:

DOCTORA EN CIENCIAS (MICROBIOLOGÍA)

PRESENTA:

M.C. NANCY MARTÍNEZ MONTIEL

ASESORA DE TESIS:

D.C. REBECA D. MARTÍNEZ CONTRERAS



BUAP

Puebla, Pue. a 20 de enero 2016.

**A LA ACADEMIA DEL POSGRADO
EN MICROBIOLOGÍA
CICM-ICUAP
PRESENTE.**

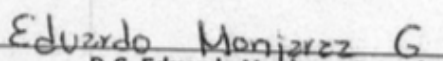
Por este conducto los abajo firmantes, integrantes del Comité revisor de Tesis de Doctorado de la alumna **Nancy Martínez Montiel**, les informamos que hemos revisado el escrito titulado:


"Identificación y caracterización de los factores involucrados en decaimiento del ARN mensajero sin sentido (NMD) en *Ustilago maydis*"

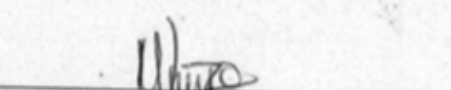
A nuestro juicio, la alumna Martínez Montiel puede proceder a la impresión de la tesis y a la presentación del examen de grado correspondiente.

Sin más que agregar, nos despedimos afectuosamente de ustedes.

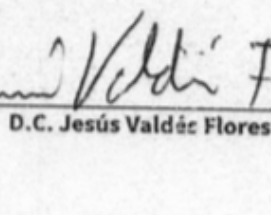
Atentamente.
"Pensar Bien Para Vivir Mejor"


D.C. Eduardo Monjaraz Guzmán


D.C. Miguel Castañeda Lucio


D.C. Mónica Ascención de Nova Ocampo


D.C. Luis Ernesto Fuentes Ramírez


D.C. Jesús Valdéz Flores

Posgrado en Ciencias Microbiológicas
Instituto de Ciencias (ICUAP)

Edificio IC11,
Ciudad Universitaria
Col. San Manuel, Puebla, Pue. C.P. 72570
01 (222) 229 55 00 Ext. 2522
posgrado.microbiologia@correo.buap.mx

AGRADECIMIENTOS

Este trabajo de investigación fue realizado en el laboratorio de Ecología Molecular Microbiana dentro del grupo de investigación Biología Molecular y Celular de Hongos liderado por la Dr. Rebeca D. Martínez Contreras Profesora e Investigadora Titular del Centro de Investigaciones en Ciencias Microbiológicas del Instituto de Ciencias de la Benemérita Universidad Autónoma de Puebla (CICM-ICUAP-BUAP).

Este trabajo fue financiado con los recursos provenientes del proyectos CB-2009-133317 otorgado por el Consejo Nacional de Ciencia y Tecnología (CONACyT).

La M.C. Nancy Martínez Montiel fue beneficiada con la beca 264278 otorgada por el Consejo Nacional de Ciencia y Tecnología (CONACyT) y la beca Becas-Tesis 2015 del Consejo de Ciencia y Tecnología del Estado de Puebla (CONCyTEP) para concluir sus estudios.

A mi familia

CONTENIDO

AGRADECIMIENTOS

DEDICATORIAS

TABLA DE CONTENIDOS

LISTA DE TABLAS

LISTA DE FIGURAS

RESUMEN

ABSTRACT

I. INTRODUCCIÓN

- I.I. Procesamiento y regulación de la expresión del ARNm
- I.II. Acople de los procesos de *splicing* y NMD
- I.III. Factores de NMD
 - I.III.I. UPF1 y el Núcleo de NMD
 - I.III.II. Otros factores relevantes para el mecanismo de NMD
- I.IV. Mecanismo de NMD
 - I.IV.I. El mecanismo de NMD en diferentes especies
 - I.IV.II. Sustratos de NMD
- I.V. NMD y enfermedades humanas
- I.VI. Terapias para enfermedades relacionadas a NMD
- I.VII. *Ustilago maydis* como modelo de estudio

II. JUSTIFICACIÓN

III. HIPÓTESIS

IV. OBJETIVOS

- IV.I. Objetivo general
- IV.II. Objetivos particulares

V. MATERIALES Y MÉTODOS

- V.I. Análisis bioinformático
 - V.I.I. Análisis porteómico y evolutivo de la ruta de vigilancia del ARN mensajero
 - V.I.II. Análisis de la estructura secundaria y terciaria de las proteínas del núcleo del NMD
- V.II. Material biológico
 - V.II.I. Microorganismos y condiciones de crecimiento
 - V.II.II. Plásmidos
- V.III. Técnicas generales de biología molecular
 - V.III.I. Técnicas experimentales con ácidos nucleicos
 - V.III.I.I. Extracción de ADN genómico
 - V.III.I.II. Extracción de ADN plasmídico mediante lisis alcalina
 - V.III.I.III. Extracción de ARN con Trizol
 - V.III.I.IV. Cuantificación de ADN ARN
 - V.III.I.V. Electroforesis de ADN y ARN en geles de agarosa
 - V.III.I.VI. Retrotranscripción
 - V.III.I.VII. Amplificación de fragmentos de ADN mediante (PCR)
 - V.III.I.VIII. Uso de enzimas de restricción
 - V.III.I.IX. Reparación de fragmentos de ADN
 - V.III.I.X. Ligación de moléculas de ADN
 - V.III.I.XI. Purificación y concentración de ADN
 - V.III.I.XII. Secuenciación
 - V.III.II. Técnicas de mutagénesis
 - V.III.II.I. Interrupción del gen *umUpf1* con un cassette de resistencia a higromicina B
 - V.III.III. Técnicas de transferencia genética
 - V.III.III.I. Transformación de células de *E. coli* DH5 α
 - V.III.III.II. Preparación de células quimiopetentes de *E. coli* DH5 α
 - V.III.III.II.I. Transformación de *E. coli* DH5 α
 - V.III.III.III. Electroporación de células de *U. maydis* FB1

V.III.III.IV.Obtención de células electrocompetentes de *U. maydis* FB1

V.III.III.IV.I.Electroporación de *U. maydis* FB1

V.III.IV.Técnicas de expresión inhibiendo el NMD

V.III.V. Técnicas experimentales con proteínas

V.III.V.I.Inducción de proteínas recombinantes

V.III.V.II.Electroforesis de proteínas

V.III.V.III.Tinción de geles de poliacrilamida

V.III.V.III.I.Tinción de azul de coomasie

V.III.V.III.II.Tinción de plata

V.III.V.IV.*Western-blott*

V.III.V.V.Ensayos de tipo *pull-down* para la interacción de proteínas recombinantes

VI. RESULTADOS

VI.I.Organización estructural y funcional del homólogo de UPF1 en *U. maydis*

VI.II.Modificaciones postraduccionales en umUPF1

VI.II.I.Relación filogenética del UPF1

VI.II.II.Interacciones proteicas de umUPF1

VI.II.III.El interruptor para la actividad helicasa de umUPF1

VI.II.IV.Posible interacción de umUPF1 con umUPF2

VI.II.V.Conformación ATPasa

VI.II.VI.Interacción de umUPF1 con ARN

VI.II.VII.Simulaciones de dinámica molecular

VI.III.El núcleo de NMD en *U. maydis*

VI.III.I.umUPF1, umUPF2 y umUPF3 de gen a proteína

VI.III.II.Identificación de mensajeros blanco de la vía de NMD en *U. maydis* FB1

VI.III.III.Interrupción de umUpf1 en *U. maydis*

VI.III.IV.Análisis de la interacción entre los factores del núcleo de NMD

VI.IV.Análisis comparativo de la maquinaria de NMD en diferentes especies

VI.IV.I.El núcleo central de la vía de NMD

VI.IV.II.Las proteínas SMG

VI.IV.III.El complejo EJC

VI.IV.IV.Factores del procesamiento del ARNm involucrados en NMD

VI.IV.V.Vía de degradación

VI.IV.VI.La remoción del CAP

VII. DISCUSIÓN

VII.I.Disertación de la vía de vigilancia del ARNm

VII.II.El núcleo de la maquinaria de NMD en *U. maydis*

VII.III.UPF1 en la viabilidad de *U. maydis*

VII.IV.Interacción umUPF1-umUPF2

VII.V.Funcionalidad de la vía de NMD en *U. maydis*

VII.VI.El núcleo central del proceso de NMD

VII.VII.Las proteínas SMG

VII.VIII.El complejo EJC

VII.IX.Factores del procesamiento del ARNm involucrados en NMD

VII.X.Vía de degradación

VII.XI.La remoción del Cap

VII.XII.El caso de *U. maydis*

VIII. CONCLUSIONES

IX. PERSPECTIVAS

X. BIBLIOGRAFIA

XI. ANEXOS

LISTA DE FIGURAS

- Figura 1. *Splicing* alternativo, maquinaria y eventos.
- Figura 2. NMD en *H. sapiens* y en *S. cerevisiae*.
- Figura 3. Estructuras y consecuencias de las interacciones entre los factores involucrados en NMD.
- Figura 4. Estructura dinámica e interacciones de UPF1.
- Figura 5. Modelo del EJC.
- Figura 6. Mecanismo general del proceso de NMD. Después del *splicing* se deposita el complejo EJC en el ARNm.
- Figura 7. Mecanismo molecular de NMD en mamíferos.
- Figura 8. Sustratos de NMD.
- Figura 9. Estructura de los ARNm que pueden ser blanco del proceso de NMD.
- Figura 10. Efecto de la introducción de un PTC debido a una mutación.
- Figura 11. Mecanismo de lectura a través del codón de paro (*readthrough*).
- Figura 12. Comparación del proteoma de *U. maydis*, *S. cerevisiae* y *H. sapiens*.
- Figura 13. Esquema del procedimiento empleado para los ensayos tipo *pull-down*.
- Figura 14. La organización de UPF1 es muy similar en *Ustilago maydis* y en *Homo sapiens*.
- Figura 15. La secuencia para UPF1 se encuentra conservada en *H. sapiens* y *U. maydis*.
- Figura 16. Modificaciones postraduccionales predichas para umUPF1.
- Figura 17. Relación evolutiva y organización de dominios para UPF1 en diferentes especies.
- Figura 18. Interacción de UPF1 con diferentes proteínas.
- Figura 19. La conformación cerrada de umUPF1 se mantiene gracias a la interacción entre los dominios CH y RecA2.
- Figura 20. El dominio CH de umUPF1 media la interacción con UPF2.
- Figura 21. Alteración *in silico* de la asociación CH-UPF2.
- Figura 22. Arreglo conformacional de umUPF1 en presencia de ATP o de ADP.
- Figura 23. Interacción de umUPF1 con ARN.
- Figura 24. Dinámica molecular de las proteínas UPF1 de *H. sapiens* y *U. maydis*.
- Figura 25. Presencia de los genes UM00237.1, UM03549.1, UM00233.1 en el genoma de *Ustilago maydis* FB1.
- Figura 26. Expresión de los genes UM00237.1, UM03549.1, UM00233.1 en *Ustilago maydis* FB1.
- Figura 27. Identificación de las proteínas UPF1, UPF2 y UPF3 en extractos totales de *U. maydis* FB1.
- Figura 28. Esquema del Ensayo de Emetina.
- Figura 29. Curva de crecimiento y estandarización del Ensayo de Emetina en *Ustilago maydis* FB1.
- Figura 30. Transcritos regulados por NDM en *Ustilago maydis* FB1.
- Figura 31. Construcción de la interrupción de umUpf1.
- Figura 32. Interrupción de Upf1 en la cepa FB1 de *Ustilago maydis*.
- Figura 33. Análisis de la mutante *U. maydis* umUpf1::Hyg.
- Figura 34. Construcción de los vectores para las proteínas recombinantes umHisUPF1-CH, umHisUPF2 y umHisUPF3.
- Figura 35. Proteínas recombinantes homólogas a UPF1, UPF2 y UPF3 de *Ustilago maydis*.
- Figura 36. Análisis de la interacción entre umHisUPF1-CH y umHisUPF2.

LISTA DE TABLAS

- Tabla 1. Proteínas de NMD y EJC
- Tabla 2. Genes mutados con más frecuencia (mostrando 10 de 106 genes)
- Tabla 3. Mutaciones más frecuentes en factores de NMD (Mostrando 10 de 13,243)
- Tabla 4. Plásmidos utilizados en este estudio
- Tabla 5. Secuencias de los oligonucleótidos utilizados en este trabajo
- Tabla 6. Anticuerpos empleados para los ensayos de *Western Blott*
- Tabla 7. Factores de NMD putativos en *Ustilago maydis*
- Tabla 8. Tamaños de los factores del núcleo de NMD en *U. maydis* y *H. sapiens*
- Tabla 9. Resultados del ensayo de emetina

RESUMEN

Los sistemas de control de calidad responsables de la adecuada traducción de proteínas son esenciales para fisiología y sobrevivencia de la célula. En células eucariotes, algunas moléculas de ARNm dejan el núcleo después de sufrir un evento de *splicing* y pueden desencadenar un proceso de decaimiento el cual puede ser dependiente de EJC (*Exon Junction Complex*) o de DSE (*Down Stream Element*) en mamíferos y levaduras, respectivamente y cuya finalidad es arrestar la traducción de dicho mensajero. Muchos eventos de *splicing* alternativo afectan la secuencia codificante y cerca de la mitad alteran el marco de lectura mientras que un tercio de ellos aparentemente desencadenan decaimiento del ARNm sin sentido (NMD), el cual no sólo degrada transcritos anómalos, sino que también regula el nivel basal de diferentes ARNm fisiológicos involucrados en una variedad de procesos celulares. Sin embargo, aún se desconoce la relevancia fisiológica y patológica del proceso de NMD en las diferentes especies. Desde *S. cerevisiae* hasta *H. sapiens*, la maquinaria de NMD requiere de una serie de factores reguladores conservados que incluyen al núcleo central conformado por las proteínas UPF1, UPF2 y UPF3. Además de estas proteínas, se requiere de un grupo de proteínas también conservadas que ayudan a realizar y regular este proceso. Finalmente, existen ligeras diferencias tanto en la maquinaria responsable del NMD y su regulación dependiendo del organismo en cuestión.

En este trabajo se realizó un análisis bioinformático de la maquinaria de NMD para un total de 299 eucariotes cuyo genoma se ha secuenciado completamente y la información correspondiente se encuentra liberada, incluyendo 99 metazoarios, 48 plantas, 109 hongos y 43 protistas. Se analizaron 70 factores para NMD empleando la base de datos KEGG y se determinó la distribución para estas proteínas. Los factores identificados participan en NMD pero pueden estar involucrados también en el *splicing* y procesamiento del ARNm, así como en la vigilancia, exporte y degradación del ARN. Se encontró que en los organismos estudiados la vía general de NMD es similar pero se pueden observar ligeras diferencias dependiendo del grupo en cuestión. En el caso particular de *U. maydis*, en donde el proceso de NMD no ha sido descrito, se identificó a los homólogos putativos para UPF1, UPF2 y UPF3, así como a diferentes centrales y auxiliares del NMD. UPF1 muestra el 66% de identidad entre *H. sapiens* y *U. maydis*, además de que las proteínas UPF2 y UPF3 también se encuentran altamente conservadas. Se realizó un análisis *in silico* de las características estructurales y bioquímicas para el homólogo putativo de UPF1 en *U. maydis* en donde se encontró que dicha proteína posee una secuencia, estructura, estabilidad mecánica, propiedades fisicoquímicas, organización espacial y potencial catalítico similares a las descritas para el homólogo de *H. sapiens*. Estas observaciones sugieren fuertemente que ambas proteínas pueden realizar una función biológica equivalente, introduciendo el uso potencial de *U. maydis* como un modelo para estudiar el proceso de NMD. Además, se obtuvo evidencia experimental de que los homólogos para UPF1, UPF2 y UPF3 se expresan (RT-PCR, *Western-Blott* e inmunofluorescencia) y son funcionales en el basidiomiceto (*Pull-down* y ensayos en presencia de emetina). Por otro lado, UPF1 parece ser esencial para la sobrevivencia del hongo ya que la mutante en donde se interrumpió el gen para UPF1 fue sub-lethal. Empleando ensayos

en los cuales *U. maydis* creció en presencia de emetina se identificaron diferentes ARNm que son blancos para NMD en el hongo, indicando que esta vía de regulación es funcional en el hongo.

Al comparar la maquinaria completa de NMD en *U. maydis* con la de *H. sapiens* se encontró una identidad entre el 40% y 60% para la mayoría de los factores, en donde el mayor porcentaje fue para eRF1 (71%) y MAGOH (66%). De manera relevante, encontramos en *U. maydis* el homólogo para algunos factores de NMD que se han descrito para *H. sapiens* pero que están ausentes en *S. cerevisiae* (SMG1, SMG5-7, MAGOH y algunos componentes del EJC). En concordancia, algunos factores de NMD descritos para *S. cerevisiae* pero ausentes en *H. sapiens* también parecen estar ausentes en *U. maydis*. Estas observaciones sugieren que el mecanismo y la regulación de NMD podrían ocurrir de manera similar en *H. sapiens* y *U. maydis*. Sabiendo que aproximadamente el 40% de los genes en *U. maydis* se encuentran interrumpidos por intrones, en donde el evento de *splicing* alternativo más frecuente es la retención de intrón, resulta probable la generación de transcritos aberrantes que puedan ser sustratos de NMD en este organismo. En conjunto, los resultados obtenidos sugieren un alto grado de conservación para el proceso de NMD y su regulación entre *H. sapiens* y *U. maydis*.

ABSTRACT

The quality control of protein translation is essential for the appropriate cell physiology and survival. In eukaryotic cells, some mRNA molecules leave the nucleus after splicing but elicit decay in an EJC- or DSE-dependent manner for mammals and yeasts respectively, in order to arrest translation. Most alternative splicing events affect the coding sequence, with half of these altering the reading frame and a third apparently leading to nonsense-mediated mRNA decay (NMD), which not only degrades faulty transcripts but also regulates the steady state level of many physiological mRNAs involved in a variety of different cellular processes. However, physiological and pathological significance of NMD across species remains largely unclear. From *S. cerevisiae* to *H. sapiens*, NMD requires a set of conserved regulatory factors where the core components of this mechanism are the proteins UPF1, UPF2 and UPF3. In order of NMD to occur, a full set of specific and accessory factors besides the UPF proteins are required. Moreover, slight differences on the NMD machinery and its regulation have been observed for some organisms.

Results: We performed a bioinformatic analysis of the NMD machinery in a total of 299 eukaryotes for which the full genome is available, including 99 metazoan, 48 plants, 109 fungi and 43 protists. The homologs for 70 different NMD factors were identified in the KEGG database. Using this approach, the distribution for the different factors involved in NMD was analyzed. The factors identified regulate NMD but can also be involved in pathways like mRNA processing and splicing, RNA surveillance, RNA export and degradation. Over these organisms, we found that the pathways are significantly similar but slight differences can be observed in some groups. In the particular case of *U. maydis*, where NMD has not been described, we identified the putative UPF1, UPF2 and UPF3 factors as well as several core and auxiliary proteins involved in NMD regulation. Surprisingly, UPF1 shows 66% identity at the protein level between *U. maydis* and *H. sapiens*. UPF2 and UPF3 are extremely conserved between the two organisms as well. Structural, biochemical and mutational *in silico* analysis of the UPF1 homolog in *U. maydis* indicate that this protein possess a very similar sequence, structural organization, mechanic stability, physicochemical properties, spatial organization and catalytic potential in comparison to the NMD factor depicted for *Homo sapiens*. These observations strongly support the notion that human and fungal UPF1 could perform equivalent biological activities, introducing the use of *Ustilago maydis* as an alternative model to study NMD. In addition, we have experimental evidence that the UPF1, UPF2 and UPF3 homologs are expressed (RT-PCR, Western-Blott and immunofluorescence) and functional in the basidiomycete (pull-down and emetine assays) and UPF1 appeared to be essential for cell survival because the fungal mutant where *Upf1* was interrupted resulted sub-lethal. Moreover, NMD targets in *U. maydis* were also identified when the fungal culture is treated with emetine, indicating that the regulatory pathway is functional in this model. The analysis of the full NMD machinery was performed in *U. maydis* in comparison to *H. sapiens* and identity for most putative NMD factors ranges between 40% and 60% with the most striking similarities at the protein level corresponded to eRF1 (71%) and MAGOH (66%). NMD factors that have been described for human, but that are absent in the classic yeast model *S. cerevisiae* were also found in *U. maydis* (MAGOH, SMG1, SMG5-7, EJC-components). This high

homology suggests an important degree of conservation between human and fungal control of NMD. Accordingly, NMD factors described in *S. cerevisiae* but absent in humans are also absent in *U. maydis*. Almost 40% of the genes in *U. maydis* are interrupted with introns, being intron retention the prevalent alternative splicing event, leading to a high probability of aberrant transcript generation, which could direct to NMD. Taken together, our results suggest an important degree of conservation between the NMD machinery and mechanistic control for this process between *H. sapiens* and *U. maydis*.

I. INTRODUCCIÓN

El proceso de NMD (Non-sense-Mediated mRNA Decay) es muy importante para las células eucariotas y constituye el principal sistema de control de calidad para la producción de transcritos funcionales en las células de estos organismos, ya que no sólo detecta transcritos defectuosos sino que además controla la producción de diferentes productos proteicos para mantenerlos en concentraciones fisiológicas reguladas dependiendo de los requerimientos celulares.

Un proceso íntimamente relacionado con el NMD es el *splicing*. En la actualidad no se sabe con certeza si ambos mecanismos actúan de manera acoplada o independiente y aunque la estrecha relación entre NMD y *splicing* en eucariotes superiores ha sido documentada para diversos genes, dicha relación no es tan clara en *S. cerevisiae*, *C. elegans* y plantas. Para entender en qué punto convergen estas maquinarias proteicas es necesario conocer el funcionamiento de ambos mecanismos.

I.I. PROCESAMIENTO Y REGULACIÓN DE LA EXPRESIÓN DEL ARNm

Los genes de organismos eucariotes se encuentran frecuentemente interrumpidos por secuencias no codificantes que deben ser eliminadas durante la expresión, las cuales reciben el nombre de intrones. El *splicing* es el proceso mediante el cual los intrones son removidos para producir moléculas de ARN mensajero (ARNm) funcionales al ligarse las secuencias codificantes o exones. Para que este mecanismo se lleve a cabo, se asocian al mensajero las ribonucleoproteínas pequeñas nucleares (snRNPs) U1, U2, U4, U5 y U6 además de proteínas accesorias para regular el proceso, como la familia de SR y hnRNP (Martínez-Contreras R. *et al.*, 2006) (Fig. 1A). Todas estas proteínas son importantes para el *splicing* pero también para otros procesos co- y post-transcripcionales, ya que de éstas dependerá la calidad y cantidad de los mensajeros. El *splicing* alternativo ocurre cuando algunos exones, intrones o porciones de los mismos son incluidos de manera diferencial para producir diversas moléculas de ARNm a partir de un mismo precursor inmaduro (pre-ARNm) como se muestra en la figura 1B. El *splicing* alternativo es empleado por el 75% de los genes humanos y constituye la principal fuente de diversidad proteica en eucariotes (Johnson *et al.*, 2003). Aunque se estima que entre el 15 y el 50% de las enfermedades genéticas humanas tienen su origen en algún defecto a nivel de *splicing* (Tazzi *et al.*, 2009), la extraordinaria contribución del *splicing* alternativo en diferentes procesos biológicos y patológicos en plantas y mamíferos apenas empieza a dilucidarse. Como ejemplo, se sabe que diferentes eventos de *splicing* alternativo son responsables de determinar los caracteres sexuales en *Drosophila* (Förch & Valcárcel, 2003), de generar hormonas con actividades biológicas distintas o de producir proteínas de membrana o citoplásmicas con funciones antagónicas (Martínez-Contreras *et al.*, 2003). La regulación del *splicing* ha sido estudiada principalmente en células humanas en cultivo, así como en *S. cerevisiae*, organismo en el que sólo el 5% de los genes se encuentran interrumpidos. Además del evento de *splicing*, el pre-ARNm es modificado mediante la adición del CAP en el extremo 5' en donde quedan unidos los factores CBP20/CBP80 y de la cola de poliA en el extremo 3', en donde se unen las proteínas PABP1 y PABP2. Una vez que el ARNm cuenta con estas modificaciones, entonces

presenta su forma madura que puede salir del núcleo hacia el citoplasma para ser traducido. Ya en el citoplasma, existen sistemas de vigilancia que se aseguran de que la traducción de dichos mensajeros se complete de manera regulada, precisa y eficaz, como la vía de decaimiento del ARNm sin sentido o NMD. Este mecanismo constituye un sistema de control de calidad que contribuye a la generación adecuada de proteínas funcionales (Maquat, 1995).

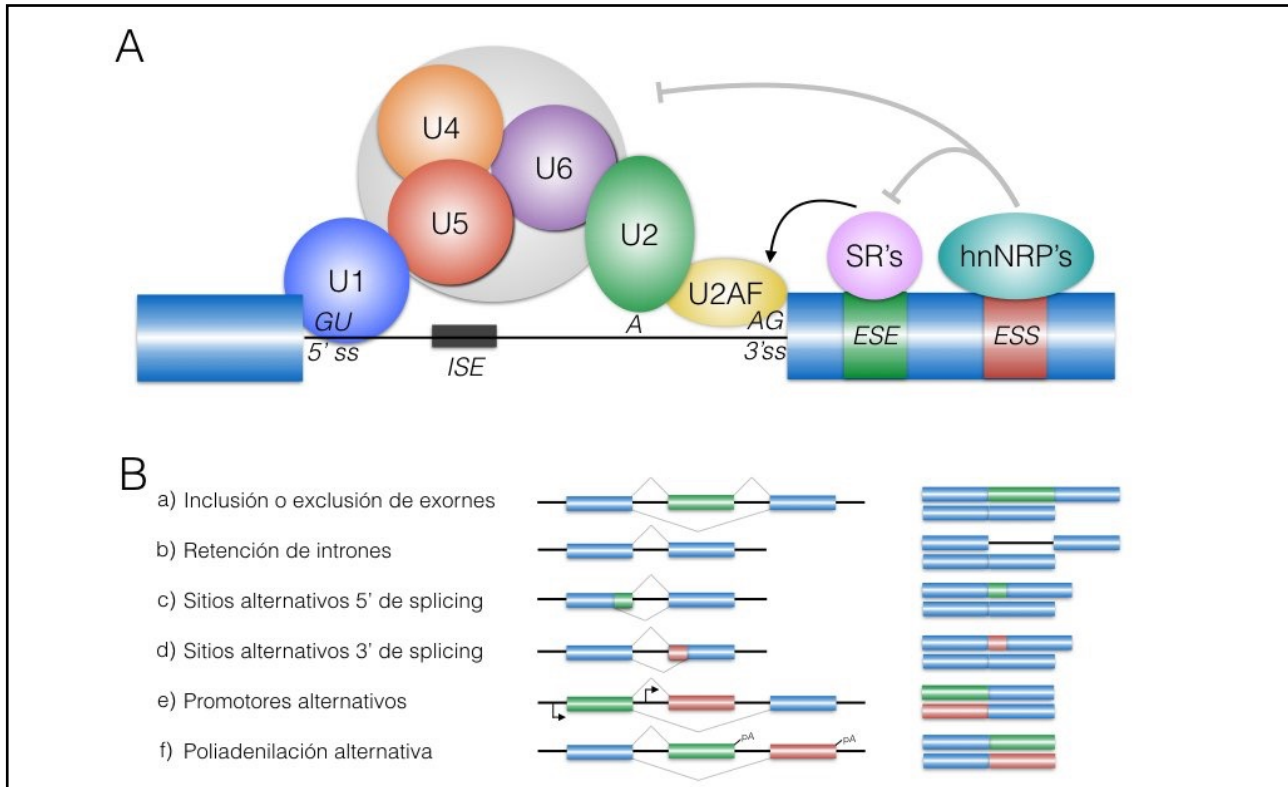


Figura 1. Splicing alternativo, maquinaria y eventos. A) Los factores de *splicing* reconocen secuencias a lo largo del transcrito incluyendo el sitio 5' donador de *splicing* y el 3' aceptor de *splicing* (5' ss y 3' ss) los cuales interactúan con el *spliceosoma* compuesto principalmente por las snRNP (U1, U2, U4, U5 y U6). Otros elementos reguladores pueden ubicarse en los exones (ESE y ESS) o en los intrones (ISE) y pueden ser reconocidos por la familia de proteínas SR o hnRNP. De manera general las SR facilitan el *splicing* mientras que las hnRNP lo regulan negativamente. B) Posibles eventos de *splicing*: las barras en azul, verde o rojo indican secuencias codificantes, en líneas negras están las secuencias no codificantes; el nombre del evento representado se encuentra a la izquierda (Martínez Montiel *et al.*, 2015).

I.II. ACOUPLE DE LOS PROCESOS DE SPLICING Y NMD

La expresión genética en eucariotes comprende una serie de pasos interconectados de modulación fina. En este sentido existen diferentes mecanismos involucrados en procesos de regulación a nivel co- y post-transcripcional, incluido el *splicing*, el silenciamiento de genes y el *editing* por mencionar algunos. Una vez en el citoplasma existen otros mecanismos de regulación que se encargan del control de calidad en la expresión de proteínas, los cuales son importantes para asegurar el funcionamiento óptimo de la célula; tal es el caso del NMD. Este último ocurre de manera post-transcripcional y es un mecanismo de vigilancia que detecta mensajeros con codones de paro prematuros o PTC (Premature Termination Codon) o "sin sentido", llevándolos a degradación para evitar la expresión de proteínas truncas o erróneas (Maquat, 1995). Desde el descubrimiento del fenómeno de decaimiento de ARNm asociado a mutaciones sin sentido por Losson y Lacroute (1979) y posteriormente llamado sistema NMD, se han hecho diversos estudios de este mecanismo.

Actualmente se sabe que en mamíferos existen tres características determinantes para llevar a un ARNm a degradación por NMD: la presencia del CAP unido a CBP20 y CBP80 en el extremo 5', la unión de proteínas de poliadenilación en el extremo 3' y un complejo denominado EJC (*exon junction complex*, en español complejo de unión exón-exón) que se deposita en el ARNm durante el *splicing* a una distancia de alrededor de 20 nucleótidos río arriba de la unión entre dos exones. Para que un EJC active la vía de NMD, debe encontrarse de 50 a 55 nt río abajo del PTC (Lejeune & Maquat, 2005) como se muestra en la figura 2 de lado izquierdo. En otros organismos como en *S. cerevisiae* la vía de NMD se desencadena por un mecanismo al parecer diferente, en donde el elemento regulador es una secuencia del ARNm denominado DSE (*DownStream Element*), el cual se localiza río abajo de los codones de paro prematuros y es reconocida por la proteína Hrp1 (González *et al.*, 2001) como se muestra a la derecha en la figura 2. En *S. cerevisiae* no es extraño que se utilice otro mecanismo ya que únicamente el 5% de los genes en este organismo están interrumpidos por intrones, es por ello que no presentan *splicing* ni la formación de complejos EJC.

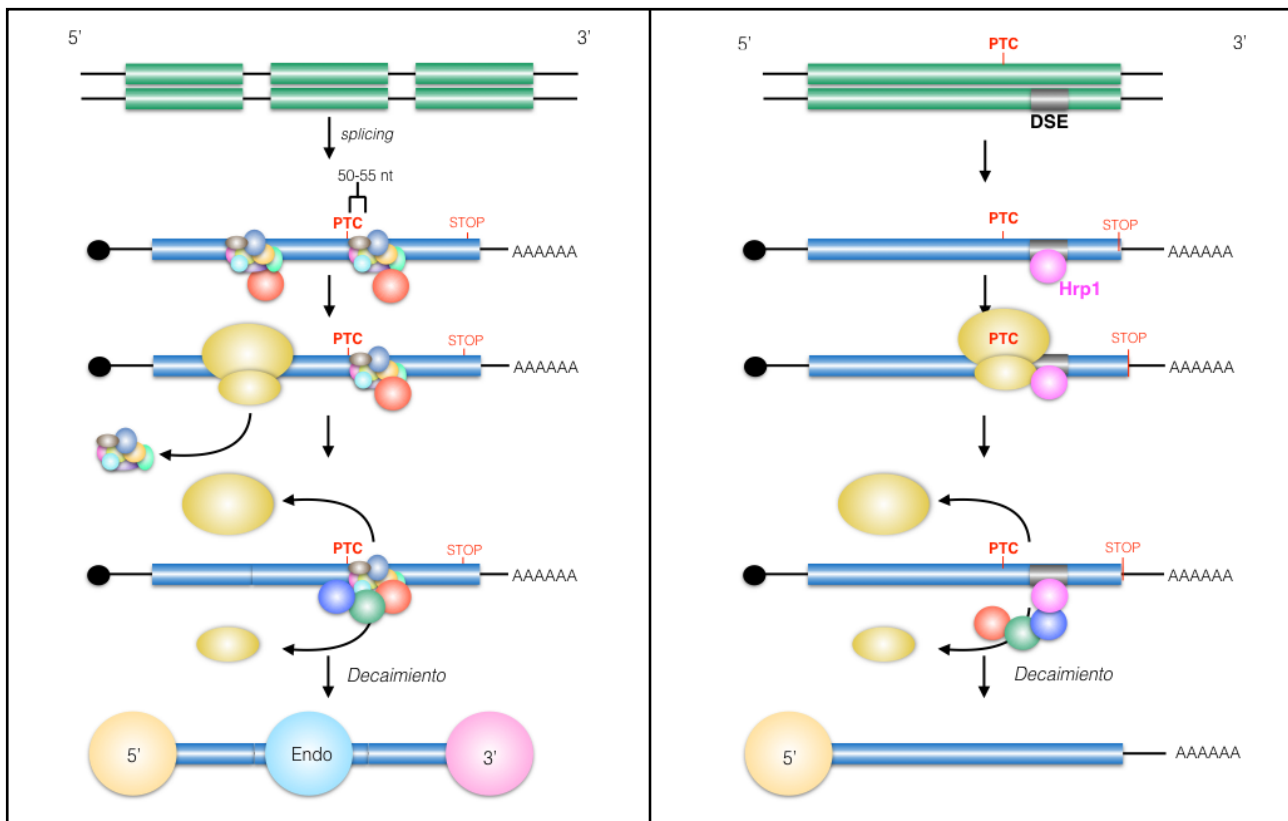


Figura 2. NMD en *H. sapiens* y en *S. cerevisiae*. El mecanismo de NMD en *H. sapiens* (izquierda) involucra la participación del EJC, el cual es depositado durante el *splicing* y se localiza a más de 50-55 nt río abajo del PTC para promover NMD. El transcrito que presenta estas características puede ser detectado por la maquinaria de NMD la cual detiene la traducción, desensambla al ribosoma y arresta al mensajero llevándolo a degradación por el extremo 3', 5' o por endonucleasas al interior del transcrito. Algo similar ocurre en *S. cerevisiae* (derecha). En este modelo, se encuentra una señal consenso (DSE) río abajo del PTC que es reconocida por una proteína llamada Hrp1. La otra diferencia para este modelo radica en que el transcrito únicamente se degrada por el extremo 5'. (Modificado de Martínez-Montiel & Martínez-Contreras, 2015).

El mecanismo de NMD ocurre en todos los organismos eucariotes estudiados a la fecha. Sin embargo, principalmente en mamíferos se ha podido corroborar el acople entre los procesos de *splicing* y NMD, ya que de manera general el complejo EJC que se deposita durante el *splicing* permite diferenciar aquellos codones “sin sentido” que desencadenan el proceso de NMD de los codones de paro

funcionales. Otra característica de los eucariotes superiores es que presentan un número importante de genes que sufren eventos de *splicing* (Blencowe, 2006, Johnson et al., 2003). En consecuencia, el *splicing* del pre-ARNm es una de las principales fuentes de inserción de PTCs en el ARNm, ya sea como un medio para regular la expresión genética o bien como resultado de un proceso ineficiente o aberrante de *splicing* (Lejeune & Maquat, 2005).

A la fecha no se sabe con certeza si los procesos de *splicing* y NMD actúan de manera coordinada o independiente y aunque la estrecha relación entre NMD y *splicing* en eucariotes superiores ha sido ampliamente documentada para diversos genes (McGlincy & Smith 2008), dicha relación no es tan clara en *S. cerevisiae*, *C. elegans* y plantas (Le Jeune & Maquat, 2005). Sin embargo, otra línea de evidencia sugiere que en algunos eucariotes inferiores como en el hemiascomiceto *Yarrowia lipolytica*, los eventos de *splicing* que consisten en la retención de intrones facilitan el NMD y regulan así la expresión genética (Mekouar et al., 2010). Existen algunos factores auxiliares de *splicing* cuya participación en NMD ha sido establecida (Zhang & Krainer, 2004); sin embargo, el único factor central de *splicing* cuya interacción directa con UPF1 ha sido demostrada (empleando el método doble híbrido), es la snRNP U1-70K (He & Jacobson, 1995).

I.III. FACTORES DE NMD

El centro de la maquinaria del NMD comprende tres factores en trans, llamados proteínas UPF (*up-frameshift*), que fueron inicialmente descubiertas en *Saccharomyces cerevisiae* y después identificadas en diversas especies (Leeds et al., 1992, Cui et al., 1995). UPF1, UPFP2 y UPF3, han sido descritos en todos los organismos eucariotes estudiados a la fecha, desde *S. cerevisiae* y *C. elegans* hasta *H. Sapiens* (Serin et al., 2001). Todas las proteínas y sus características principales las podemos encontrar en la figura 3.

Durante el proceso de NMD, la UPF1 es reclutada a los ARNm cuando la maquinaria de traducción reconoce algún PTC (Czaplinski et al., 1998,

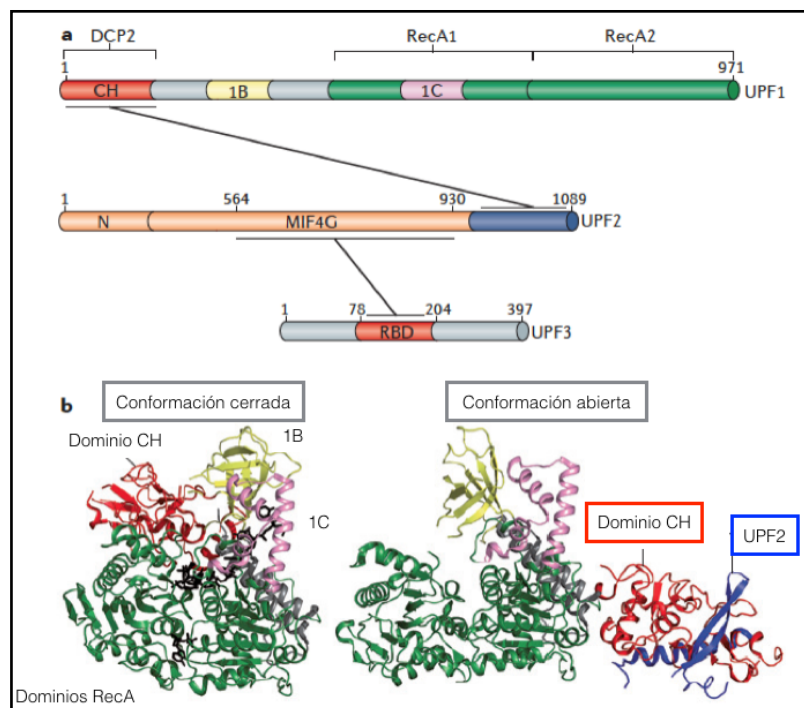


Figura 3. Estructuras y consecuencias de las interacciones entre los factores involucrados en NMD. a) Proteínas UPF y sus interacciones en *S. cerevisiae*. Los números refieren el tamaño de la proteína o la región en donde ocurren las interacciones de interacción. En UPF1 se muestran los dominios CH (dominio rico en cisteínas e histidinas) y 1B (incluido el conector, en color gris) y 1C (rosa) que son dominios regulatorios adicionales, MIF4G (naranja) se refiere a la porción media del factor eIF4G y RBD (Dominio de unión a RNA). b) La interacción de UPF2 sirve de “switch” a UPF1 para llevarla de una conformación cerrada a una abierta. El cristal de UPF1 de *S. cerevisiae* a la izquierda, mostrando en colores dominios de interés y la unión a RNA y ADP (estructuras en negro) y a la derecha el heterodímero UPF1-UPF2 de humano (Modificado de Kervestin & Jacobson 2012).

Kashima *et al.*, 2006). El rápido decaimiento del ARNm que contiene PTCs se activa gracias a la interacción de UPF1 con UPF2 y UPF3. En células de mamífero, UPF3 forma parte del gran complejo de proteínas del EJC (Chang *et al.*, 2007).

I.III.I. UPF1 Y EL NÚCLEO DE NMD

La proteína UPF1 es el factor más conservado entre los organismos estudiados y por tanto la más estudiada por su papel central en NMD (Culbertson & Leeds, 2003).

UPF1 es una proteína compleja que presenta actividad ARN helicasa 5'-3' dependiente de ATP y por tanto un dominio con actividad ATPasa (Czaplinski *et al.*, 1995; Chamieh *et al.*, 2008). Estas actividades catalíticas son necesarias para que se lleve a cabo el proceso de NMD tanto en *S. cerevisiae* como en mamíferos (Weng *et al.*, 1996; Mendell *et al.*, 2002; Franks *et al.*, 2010; Kashima *et al.*, 2006). UPF1 presenta un dominio denominado CH en el extremo N-terminal, una zona flexible (que permite la conexión de los dominios en la estructura secundaria) que incluye el dominio regulador 1B, seguida de dos dominios RecA helicasa donde se incluye el dominio regulador 1C como se muestra en la figura 3 (Czaplinski *et al.*, 1995; Chamieh *et al.*, 2008). Un análisis estructural mostró que el dominio CH interactúa con el dominio de helicasa, regulando así su propia actividad en una conformación “cerrada” de la proteína; este mismo dominio es responsable de la unión entre UPF1 y UPF2 con lo que la proteína adquiere una conformación “abierta” (Clerici *et al.*, 2009; Chakrabarti *et al.*, 2011). Cuando el dímero UPF2/UPF3 se une al dominio CH de UPF1 genera un cambio conformacional importante que activa la función de helicasa y ATPasa de UPF1 (Chamieh *et al.*, 2008; Chakrabarti *et al.*, 2011). UPF1 interactúa también de manera directa con eRF1 y eRF3 a través del dominio CH (Ivanov *et al.*, 2008). Los dos factores de liberación eRF1 y eRF3 pueden inhibir la actividad ATPasa de UPF1 *in vitro* (Czaplinski *et al.*, 1998).

Los extremos N- y C-terminal de la proteína tienen varios residuos de serina que son de importancia central en NMD porque están sujetos a un ciclo de fosforilación y desfosforilación por proteínas SMG (Yamashita *et al.*, 2001; Ohnishi *et al.*, 2003). La proteína SMG1 es una cinasa que fosforila residuos de serina y treonina (Yamashita *et al.*, 2001), promoviendo el estado fosforilado de UPF1 (Wilkinson, 2003). En contraste, las proteínas SMG5, SMG6 y SMG7 promueven la desfosforilación de UPF1 por medio del reclutamiento de PP2A (Anders *et al.*, 2003). El descubrimiento de que NMD requiere a las proteínas SMG para fosforilar y desfosforilar a UPF1, sugiere que este ciclo dirige el proceso de NMD en mamíferos y en *C. elegans* (Wilkinson, 2003).

UPF1 presenta una estructura dinámica que depende de la interacción de UPF1 con UPF2, ATP y RNA, con lo que se activa un rearrreglo en el espacio que sirve como una especie de interruptor responsable de la regulación de la actividad de la proteína. Este cambio conformacional origina que UPF1 pase de una conformación “cerrada” en donde la actividad helicasa se encuentra bloqueada, hacia una conformación “abierta” que favorece la actividad de helicasa gracias a la interacción con UPF2.

En la conformación “cerrada” (Fig. 4B, izquierda), UPF1 interactúa con una molécula de ATP a través de una cavidad formada por los dominios RecA, los cuales se acercan en cuanto se lleva a cabo dicha unión, mientras que la interacción de UPF1 con ADP ocasiona un ligero arreglo espacial en el sentido opuesto. En esta misma conformación, UPF1 se encuentra en contacto con el ARN (en orientación 5'-3') a través de un canal que se forma al centro de la proteína por los dominios RecA, 1B, 1C y la zona flexible. Se ha observado que en este mismo arreglo espacial pero en ausencia del dominio CH, el dominio 1B rota alejándose del extremo 3' del ARN y favoreciendo la actividad ATPasa y helicasa de la proteína (Chakrabarti *et al.*, 2011). En la parte inferior derecha de la figura 4B podemos observar que debido a la interacción del dominio CH (UPF1) con UPF2 ocurre un cambio conformacional significativo en donde el dominio CH se desplaza (Fig. 4) originando la conformación “abierta” de la proteína en la cual la actividad helicasa se ve potenciada (Clerici *et al.*, 2009; Chakrabarti *et al.*, 2011).

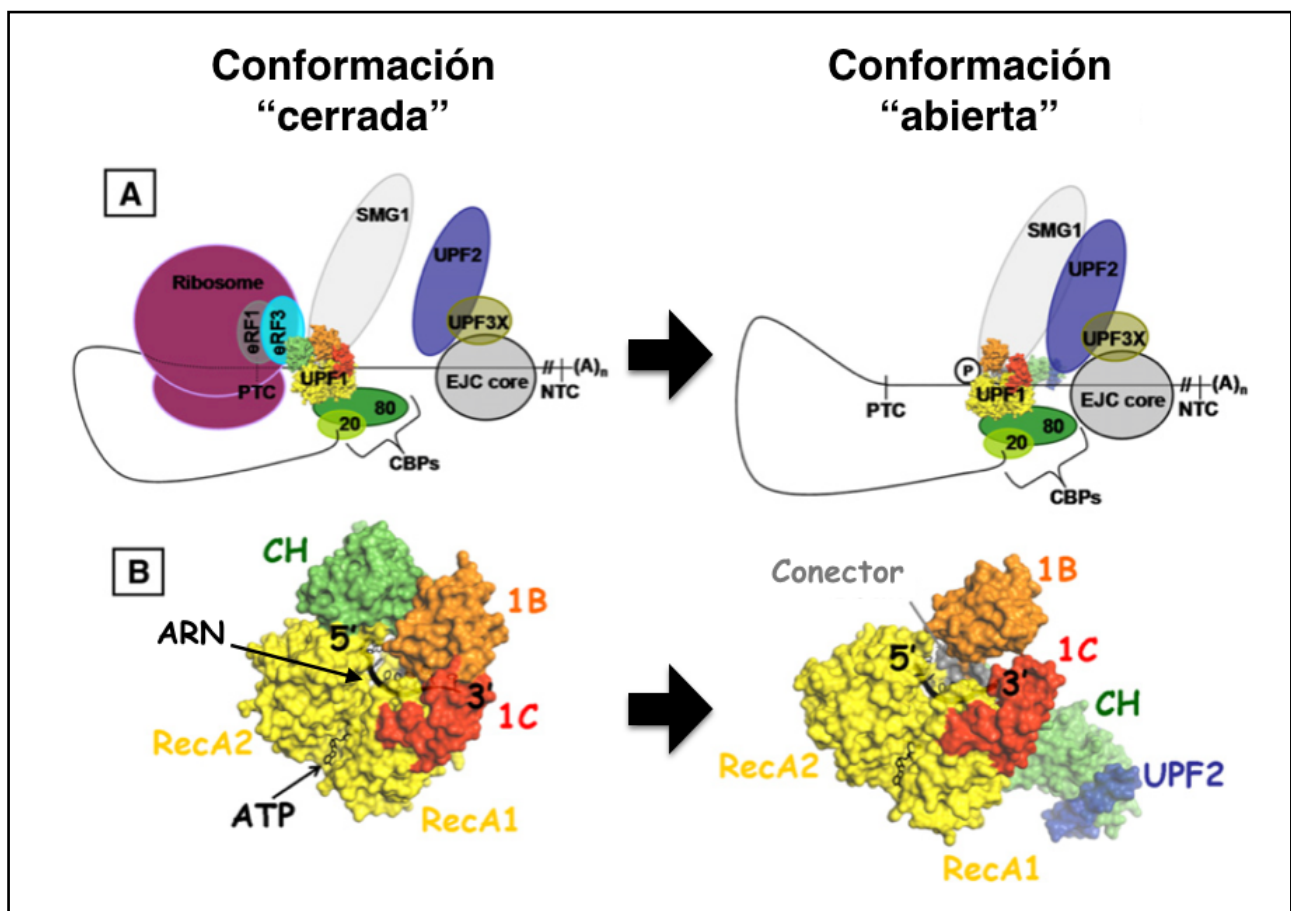


Figura 4. Estructura dinámica e interacciones de UPF1. A) A la izquierda, interacciones y conformación de UPF1 después de que el complejo SURF (SMG-1 kinase, Upf1, eRF1 y eRF3) se forma en el PTC. Cuando UPF1 se encuentra en la conformación “cerrada” CBP80 puede interactuar por medio de una región expuesta de RecA; también en esta conformación el dominio CH interactúa con eRF3 (Ivanov *et al.*, 2008). La proteína SMG1 se encuentra posicionada cerca del extremo C-terminal de UPF1 donde se han demostrado algunos blancos de fosforilación, aunque el sitio exacto de unión es aun desconocido. A la derecha, tras la interacción SMG1-UPF1, el heterodímero UPF2-UPF3 (parte del EJC) se une a UPF1 por medio de UPF2. Esta interacción libera al ARN y aumenta la actividad helicasa de UPF1. Estas acciones probablemente sean simultáneas a la separación de los factores de liberación (eRF) de UPF1, ya que eRF3 compite por el mismo sitio de unión que UPF2 en la proteína (Ivanov *et al.*, 2008). El impedimento estérico desensambla al ribosoma, lo que resulta en la represión traduccional por parte de UPF1 y posteriormente a la degradación del mensajero (Isken, *et al.*, 2008; Franks *et al.*, 2010). B) Representaciones estructurales de UPF1 en la conformación “cerrada” (izquierda) y “abierta” (derecha). En la conformación “cerrada”, el ARN (línea gruesa en negro) se encuentra al centro de la proteína mientras que el ATP (línea delgada en negro) se ancla en una cavidad formada por los dominios helicasa RecA. En la conformación “abierta”, la interacción con UPF2 desplaza al dominio CH en dirección opuesta, los dominios 1B y 1C junto con el conector son ligeramente reordenados en la proteína; esta información se obtuvo de los cristales generados por Chakrabarti en 2011 (Gleghorn & Maquat, 2011).

La proteína UPF2 presenta tres dominios MIF4G y la capacidad de interactuar con ARN (Kadlec *et al.*, 2004). Inicialmente se le consideró únicamente una molécula adaptadora ya que sirve de puente entre UPF1 y UPF3 (Czaplinski *et al.*, 1995; Chamieh *et al.*, 2008; Melero *et al.*, 2012). Sin embargo, en estudios posteriores se demostró que UPF2 funciona como activador, regulando la función de UPF1 (Chakrabarti *et al.*, 2011). En *S. cerevisiae*, UPF2 se une directamente a eRF3 a través del dominio intermedio entre los sitios de unión a UPF1 y a UPF3 (Wang, 2001). UPF2 también interactúa de manera directa con SMG1 y con UPF3 en mamíferos, independientemente de la presencia de UPF1 (Kashima *et al.*, 2006). El dominio de unión a UPF3 (Fig. 3) tiene una secuencia idéntica al motivo de interacción proteína-proteína MIF4G inicialmente descubierto en el factor de inicio de la traducción eIF4G (Kadlec *et al.*, 2004). Esta proteína se localiza principalmente en el citoplasma, sin embargo, se ha propuesto que forma parte del EJC y que es reclutada al ARNm durante el proceso de *splicing* en el núcleo (Mendell *et al.*, 2000).

UPF3 es codificada por dos genes en mamíferos: *Upf3a* y *Upf3b*, ambos genes tienen una función redundante y tienen la capacidad de unirse a UPF2 y al EJC (Kunz *et al.*, 2006; Chamieh *et al.*, 2008; Melero *et al.*, 2012), mientras que *C. elegans*, *D. melanogaster* y *S. cerevisiae* solo tienen una copia (He *et al.*, 1997; Serin *et al.*, 2001; Gatfield *et al.*, 2003). Existe una variante generada por *splicing* alternativo denominada *Upf3aS*, la cual carece de la habilidad para unirse a UPF2, pero puede interactuar con UPF1, así como con el complejo SMG5/SMG7 (Ohnishi *et al.*, 2003; Kunz *et al.*, 2006). UPF3 es predominantemente nuclear pero como parte del EJC es exportada con el mensajero y con frecuencia se encuentra en el citoplasma (Lykke-Andersen *et al.*, 2000).

I.III.II OTROS FACTORES RELEVANTES PARA EL MECANISMO DE NMD

Dentro de las proteínas con más relevancia para el mecanismo de NMD se encuentran una serie de factores que están relacionados con procesos co-transcripcionales y traduccionales como: *splicing*, protección de los extremos 3' y 5', el núcleo del EJC y proteínas accesorias del EJC, además de las proteínas SMG. Algunos de estos factores se encuentran resumidos en la Tabla 1.

Las proteínas SMG1, SMG5, SMG6 y SMG7 (*Suppressors with Morphological effect on Genitalia*) son importantes para NMD ya que regulan el estado de fosforilación de la proteína UPF1 y por tanto a este mecanismo (Wilkinson, 2003). En *C. elegans* las proteínas SMG son necesarias para NMD, ya que mutaciones en cualquiera de estas cuatro proteínas eliminan el NMD en el nemátodo (Pulak & Anderson, 1993), mientras que en células de humano la inhibición o atenuación de éstas no causan efectos severos (Yamashita *et al.*, 2001; Denning *et al.*, 2001; Usuki *et al.*, 2006). SMG1 es una cinasa que fosforila residuos de serina y treonina (Page *et al.*, 1999; Yamashita *et al.*, 2001; Denning *et al.*, 2001) y se ha descrito que regiones discretas en el N- y C-terminal son indispensables para su actividad como cinasa (Morita *et al.*, 2007). Se ha reportado que SMG1 fosforila *in vitro* al menos siete motivos STQ presentes en UPF1 (Yamashita *et al.*, 2001) y al menos cuatro de estos sitios se fosforilan también *in vitro* (Okada-Katsuhata *et al.*, 2012). De todos estos sitios, aquellos que han sido implicados

directamente en la regulación de la función de UPF1 son T28 y S1096, ya que cuando T28 se encuentra fosforilada, UPF1 es capaz de reclutar a SMG6 hacia una región conservada en el extremo N-terminal, mientras que el heterodímero SMG5/SMG7 se une a la región rica en SQ del extremo C-terminal de UPF1 cuando la posición S1096 se encuentra fosforilada (Okada-Katsuhata *et al.*, 2012). Tanto el complejo SMG5/SMG7 como SMG6 son capaces de unirse de manera simultánea con UPF1 cuando ésta se encuentra fosforilada en las posiciones mencionadas y esta unión es requerida para la defosforilación de UPF1 mediada por PP2A, lo que desencadena la disociación del complejo proteico que se encuentra unido al ARNm antes de su degradación (Ohnishi *et al.*, 2003; Okada-Katsuhata *et al.*, 2012). Además, se ha observado que la inactivación de SMG5, SMG6, SMG7 o PPA2 reprime el proceso de NMD mediante la inhibición de la actividad de UPF1 (Ivanov *et al.*, 2007; Okada-Katsuhata *et al.*, 2012). El complejo SMG5/SMG7 es indispensable para la disociación del ribosoma, de los eRFs, de UPF2 y del EJC. Esta función se debe a que la inactivación de SMG5 correlaciona con la acumulación de ribonucleopartículas que contienen además de las proteínas mencionadas, a UPF1 fosforilada, junto con las proteínas CBP y PABP. Por otro lado, SMG7 es capaz de activar a la enzima Dcp2, la cual remueve al CAP y a la exonucleasa Xrn1 (Unterholzner & Izaurralde, 2004; Okada-Katsuhata *et al.*, 2012). Por su parte, SMG6 es una endonucleasa requerida para la degradación del ARNm y se propone que su actividad empieza en el PTC (Huntzinger *et al.*, 2008; Eberle *et al.*, 2009).

Las proteínas SMG tienen motivos tipo 14-3-3, conocidos por unir residuos de serina fosforilados; mutaciones en este dominio para SMG7 tienen como consecuencia que no se reclute a UPF1 a los cuerpos de procesamiento (cuerpos-P) que son compartimientos citoplásmicos en donde se lleva a cabo la degradación de mensajeros y donde se pueden encontrar UPF1, UPF2 y UPF3 (Fukuhara *et al.*, 2005). En *S. cerevisiae* no existe ningún ortólogo reportado para estas proteínas. Sin embargo, la proteína Est1 parece tener una función similar a la de SMG5, SMG6 y SMG7 en este modelo. Est1 es una proteína que actúa en la elongación de los telómeros y en humano sus ortólogos EST1A, EST1B y EST1C presentan secuencias idénticas a las de SMG5, SMG6 y SMG7 respectivamente (Luke *et al.*, 2007). Además, todas estas proteínas conservan los dominios funcionales Est1 (proteína activadora de telomerasa) y Est1 de unión a ADN/ARN; aunque las SMG adicionalmente tienen un dominio PIN asociado.

En mamíferos, los codones de paro prematuros se reconocen principalmente gracias a la ayuda del EJC, el cual es de vital importancia para el mecanismo de NMD. En la figura 5 podemos observar la formación de un complejo proteico estable que rodea al ARN y que incluye en la parte central a MAGOH, Y14, eIF4A3 (eIF4AIII) y MLN51 (Le Hir *et al.*, 2000; Ballut *et al.*, 2005). En la periferia el complejo se asocia a UAP56, REF/ALY, TAP, p15, PYM, SRm160, PININ, RNPS1, ACINUS, UPF2 y UPF3. Este complejo protege al ARN además y participa en el transporte del núcleo al citoplasma; la calidad y cantidad de estas proteínas servirá para la correcta traducción del transcrito (Le Hir *et al.*, 2000; Le Hir *et al.*, 2001; Tange *et al.*, 2000). Los factores del EJC también están involucrados en otros procesos, como RNPS1 y SRm160 que participan en *splicing* (Le Hir *et al.*, 2001, Luo & Reed, 1999; Zhou *et al.*, 2000).

El complejo CBC (CAP Binding Complex) o complejo de unión al CAP es un heterodímero formado por CBP20 y CBP80 que se une al CAP durante la transcripción (Izaurre et al., 1995; Gorlich et al., 1996; Visa et al., 1996; Lewis & Izaurre, 1997) y tiene un papel relevante en NMD pues CBP80 interactúa con UPF1. Esta interacción potencia la unión con UPF2 (Hosoda et al., 2005) y hace más eficiente la interacción de UPF1 con mensajeros que tienen PTCs (Hwang et al., 2010). Las PABP (Poly-A-Binding Protein) o proteínas de unión a la cola de poli-A también participan en NMD. De manera particular, PABP1 interactúa con UPF1 compitiendo por el mismo sitio que eRF3 y de esta forma inhibe el NMD (Singh et al., 2008).

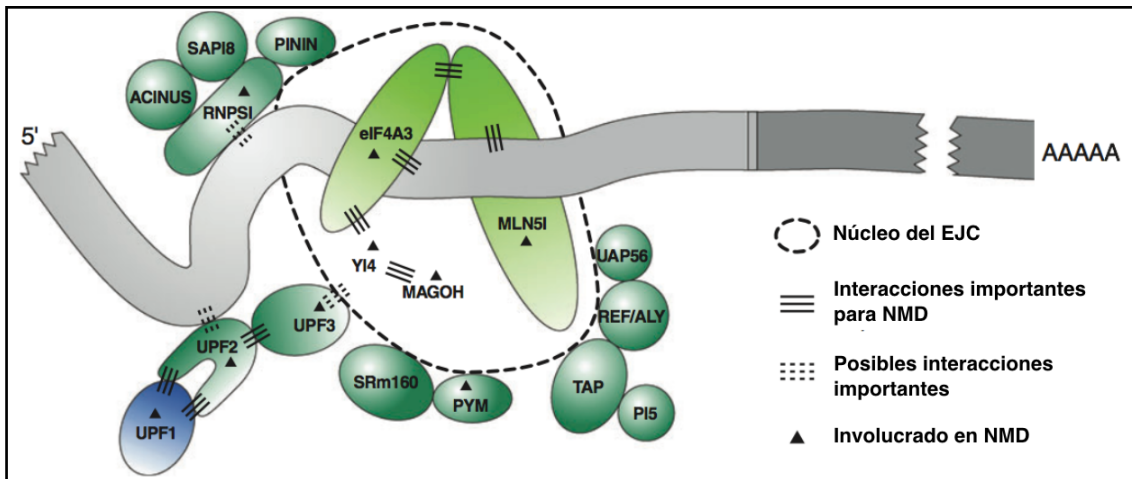


Fig 5. Esquema del EJC. Cuatro proteínas conforman el centro del complejo proteico que se une a ARN (eIF4AIII, MAGOH, MLN51 y Y14); éstas interactúan con algunos otros factores asociados al mismo complejo y participan en el transporte desde el núcleo al citoplasma (Modificado de Chang et al., 2007).

Tabla 1. Características y funciones de los principales factores del EJC y NMD

FACTOR	Características y funciones
UPF1	NMD; promueve la traducción; decaimiento del mRNA de hispanos; decaimiento de mRNA mediado por Stau1; ATPasa; fosfoproteína sustrato de SMG1; recluta a factores eRF hacia los codones de paro
UPF2	NMD; promueve la traducción; adaptador para el EJC que une UPF1 y UPF3; se une al ARN <i>in vitro</i>
UPF3a (UPF3)	NMD (débil); promueve la traducción (débil); proteína del EJC que presenta isoformas variables las cuales se asocian con UPF1 formando diferentes complejos proteicos los cuales se distribuyen de manera diferencial en el citoplasma
UPF3b (UPF3X)	NMD; promueve la traducción; proteína del EJC que interactúa directamente con UPF2; puede interactuar también con Y14
SMG1	NMD; miembro de la familia de proteínas relacionadas a la fosfoinositido-3-cinasa; fosforila UPF1
SMG5	NMD; interactúa con PP2A y promueve la defosforilación de UPF1
SMG6	NMD; interactúa con PP2A y promueve la defosforilación de UPF1
SMG7	NMD; interactúa con PP2A y promueve la defosforilación de UPF1; recluta a UPF1 a los cuerpos P si se sobreexpresa
Y14	NMD; promueve la traducción; componente central del EJC que forma un heterodímero estable con MAGOH
MAGOH	NMD; promueve la traducción; componente central del EJC que forma un heterodímero estable con Y14; se une a TAP
eIF4AIII	NMD; RNA helicasa; componente central del EJC que posiblemente ancla a otras proteínas del EJC al ARNm que es sustrato de NMD
MLN51 (BTZ, CASC3)	NMD; proteína del núcleo del EJC que interactúa directamente con eIF4AIII; se une al ARN <i>in vitro</i>
RNPS1	NMD; promueve la traducción; coactivador del splicing; factor del EJC que interactúa con PININ, SAPI8 y ACINUS
PYM	NMD; factor del EJC que interactúa directamente con Y14/MAGOH; proteína de unión al ARN
UAP56	Splicing del ARN; exporte del ARNm; factor del EJC que recluta a REF/ALY
REF/ALY	Exporte del ARNm; factor del EJC que recluta a TAP/p15; interactúa con UAP56 y Y14
TAP (NXF1)	Exporte del ARNm; levemente asociada con el EJC; interactúa con REF/ALY y con componentes del complejo del poro nuclear
p15	Exporte del ARNm; levemente asociada con el EJC; interactúa con REF/ALY y con componentes del complejo del poro nuclear
SRm160	Coactivador del splicing; factor del EJC
PININ	Exporte del ARNm; factor del EJC que se une a RNPS1
ACINUS	Apoptosis; factor del EJC que forma un heterodímero estable con RNPS1
SAPI8	Función desconocida; componente del complejo represor de la transcripción SIN3-HDAC; factor del EJC que se une a RNPS1
PP2A	fosfatasa que al unirse con SMGs desfosforila a UPF1

I.IV. MECANISMO DE NMD

En células eucariotas el proceso de NMD constituye el principal sistema de control de calidad para la producción de transcritos funcionales en las células de estos organismos, ya que detecta transcritos defectuosos pero además controla la producción de diferentes productos proteicos para mantenerlos en concentraciones fisiológicas reguladas dependiendo de los requerimientos celulares (Conti & Izaurralde, 2005).

De manera general, el modelo de ensamble de la maquinaria de NMD se activa como el resultado del reconocimiento de codones de paro prematuros durante una primera ronda de traducción, como se observa en la figura 5 (Maquat, 1995). Para recapitular, en *Homo sapiens* el NMD es activado

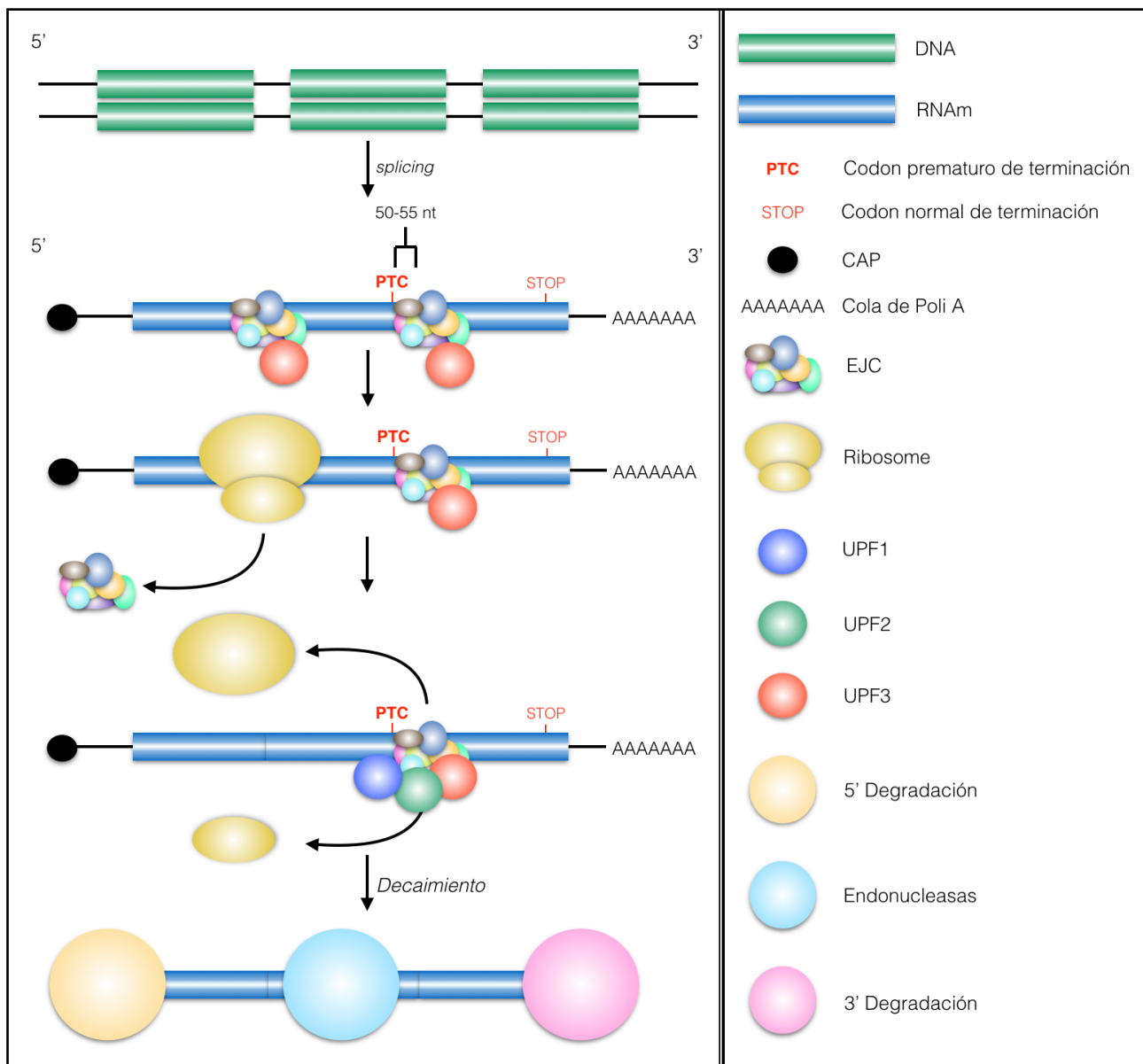


Figura 6. Mecanismo general del proceso de NMD. Después del *splicing* se deposita el complejo EJC en el ARNm. Al detectar un PTC, la maquinaria de NMD detiene la traducción y arresta al mensajero. Después de este reconocimiento inicial, UPF2 y UPF3 se unen al EJC y el ribosoma se desensambla del ARNm. Finalmente, la degradación del ARNm aberrante involucra diferentes procesos de degradación y actividades nucleolíticas (Martínez-Montiel & Martínez-Contreras, 2016).

específicamente cuando la proteína UPF1 interactúa con UPF2 (Reed, 2003) que se une a su vez a UPF3 (Kashima *et al.*, 2006). UPF2 y UPF3 forman parte del EJC (Le Hir *et al.*, 2001) que es reclutado en las uniones exón-exón durante el *splicing* del ARNm (Le Hir *et al.*, 2000). En una transcripción aberrante, existe por lo menos un EJC que se localiza río abajo del PTC, permitiendo la interacción de UPF1 con el CBC (CAP Binding Complex, formado por CBP20 y CBP80) y los factores de liberación eRF1 y eRF3 (Kashima *et al.*, 2006, Hosoda *et al.*, 2005). La interacción entre UPF1 y UPF2 se ve reforzada por CBC (Hosoda *et al.*, 2005), que se une al 5' CAP del ARNm en la primera ronda de la traducción (Ishigaki *et al.*, 2001). Por el contrario, en una transcripción normal se evita el NMD porque todos los EJC están río arriba del codón de paro y por tanto son desplazados por el ribosoma antes de que UPF1 sea reclutada (Dostie & Dreyfuss, 2002). Después de este paso, los transcritos "normales" intercambian proteínas a los extremos 5' y 3', procediendo a la traducción en masa (Ishigaki *et al.*, 2001). Por el contrario, en los productos aberrantes se arresta la traducción, eRF1 y eRF3 se liberan con lo que el transcrito es degradado por alguno de los complejos de degradación, como el mediado por DCP1-DCP2, XRN1, el exosoma o por endonucleasas (Fig. 6D).

I. IV.1 EL MECANISMO DE NMD EN DIFERENTES ESPECIES

Se sabe que el proceso de NMD en *S. cerevisiae* es desencadenado por el complejo ribonucleoproteico que se forma como parte del proceso de terminación de la traducción. Se han planteado modelos alternativos que explican la forma en la que un codón de terminación es reconocido como prematuro. Este modelo establece que la terminación de la traducción prematura es un evento ineficiente debido a que el complejo proteico que se forma no es el apropiado para finalizar la transcripción. De manera particular, existe evidencia de que en *S. cerevisiae* el ARNm que funciona como blanco de NMD puede ser reconocido debido a la ausencia de factores río abajo del codón de terminación (Amrani *et al.*, 2004). Este mecanismo se conoce como el modelo del "3'UTR falso" (Fig. 6B), en éste el ribosoma detecta el sitio del fin de la traducción en un codón que se encuentra muy alejado de la cola de poli A y por lo tanto la traducción termina de manera ineficiente. Este modelo propone que el mecanismo de NMD se activa debido a que factores de unión a la cola de Poli A, como PABP1, no se encuentran en la proximidad del ribosoma (Amrani *et al.*, 2006). En ausencia de la señal oportuna para la terminación de la traducción, las proteínas UPF interactúan con los factores de liberación eRF1 y eRF3 (Czaplinski *et al.*, 1998), dando como resultado un evento aberrante de terminación de la traducción y la activación de NMD (Kim *et al.*, 2005; Yepiskoposyan *et al.*, 2011; Eberle *et al.*, 2008; Singh *et al.*, 2008). Esta activación de NMD lleva a la remoción del CAP mediado por el complejo DCP1-DCP2, seguida de la degradación del mRNA a partir del extremo 5' por la exoribonucleasa XRN1 (Cao & Parker, 2003; Meaux *et al.*, 2008); la actividad de DCP1-DCP2 es potenciada por el complejo Lsm1-7 (Fillman & Lykke-Andersen, 2005).

Por otro lado, en *S. cerevisiae* menos del 5% de los genes contienen al menos un intrón. A pesar de ello, la mayoría de los mRNAs sin sentido que contienen PTCs son sustratos para NMD. Ya que las proteínas para el EJC no se han descrito en *S. cerevisiae*, resulta evidente la necesidad de un

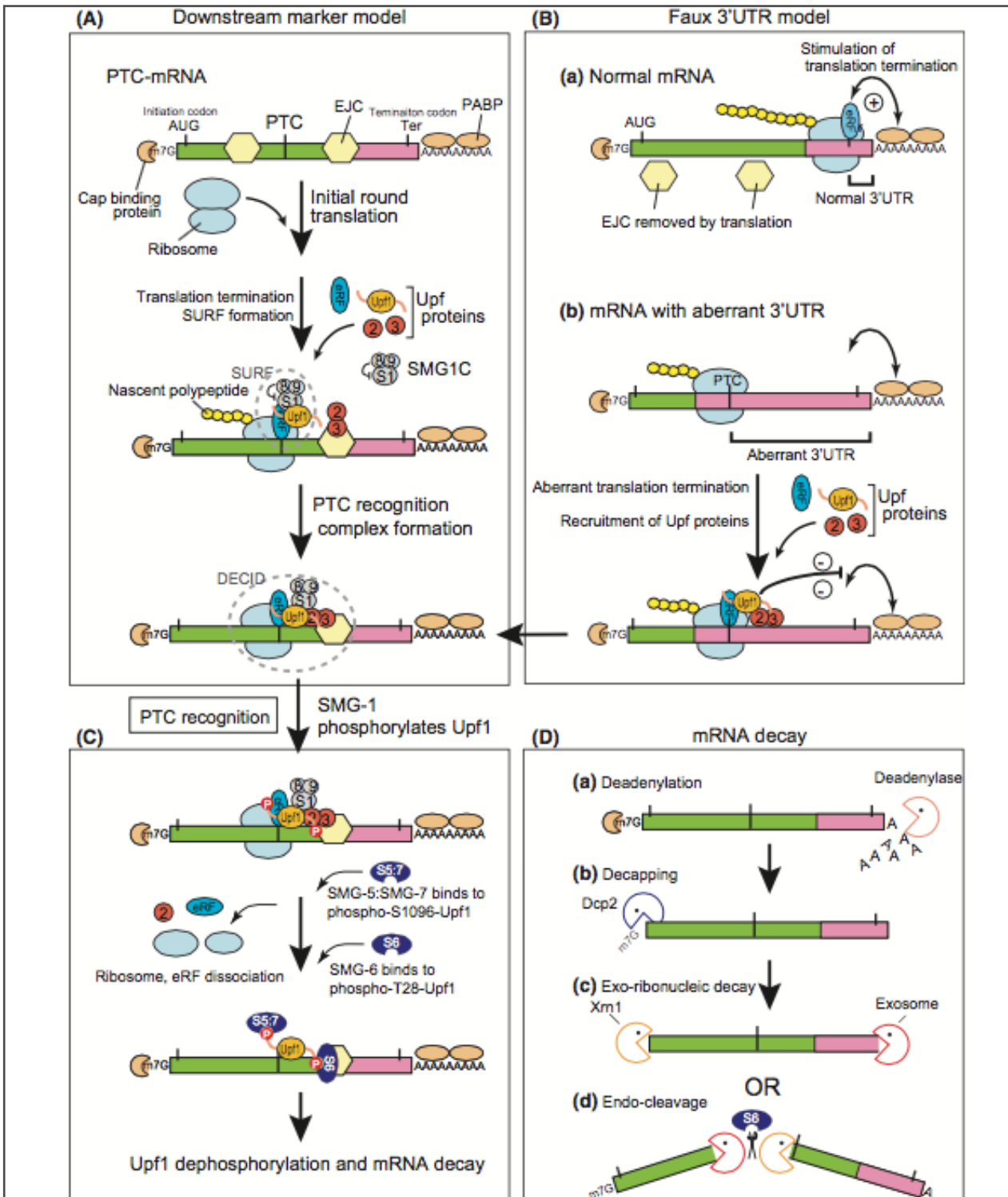


Fig. 7. Mecanismo molecular de NMD en mamíferos. A) Modelo del reconocimiento del PTC por un marcador “río abajo”. En este modelo, el reconocimiento del PTC se establece debido a la formación de los complejos de vigilancia denominados “SURF” y “DECID”, los cuales se forman en los sitios en los cuales el ribosoma y el EJC han identificado a un PTC durante la ronda primaria de la traducción. Al final de la traducción, el complejo SURF que incluye a SMG1C (SMG-1/SMG-8/SMG-9 complex), UPF1 y eRF, detecta al EJC que se encuentra río abajo y forma al complejo DECID, el cual induce la fosforilación de UPF1 por SMG1. El complejo SMG5/SMG7 se une a UPF1 fosforilada en S1096 para disociar al ribosoma y al factor de liberación de UPF1. SMG6 se une a UPF1 fosforilada en la posición T28 para inducir la disociación de UPF1 del ARNm. B) Modelo de reconocimiento del PTC por un 3’UTR falso. La terminación de la traducción se induce cuando el ARNm presenta un 3’UTR aberrante. El ribosoma que detecta este sitio aberrante puede reclutar a las proteínas UPF para prevenir la unión de eRF3 y PABP, las cuales se reclutan a un ARNm normal. C) Degradación del ARNm por NMD. La degradación se inicia por una desadenilación bifásica en mamíferos. Después de la primera fase de desadenilación de aproximadamente 110nt, el ARNm es degradado por ambos extremos. La endonucleasa SMG6 es responsable de la actividad endonucleolítica que ocurre al degradar el ARNm. AUG, codón de inicio; Ter, codón de terminación; eRF, complejo eRF1/eRF3; 2, UPF2; 3, UPF3; S8, SMG8; S9, SMG9; eRF, complejo eRF1/eRF3; P, grupo fosfato; S5:7, complejo SMG5/SMG7; S6, SMG6 (Modificado de Yamashita, 2013).

mecanismo alternativo que pueda desencadenar NMD en ausencia del EJC. En este sentido, existen en la levadura dos proteínas de exporte del núcleo que participan en NMD: UPF3 y Hrp1 (Gonzalez *et al.*, 2000). La mayoría de los mRNAs contienen secuencias ricas en AU que funcionan como DSE (Peltz *et al.*, 1993). Hrp1 tiene la capacidad de unirse a la secuencia DSE y en ausencia de dicho gen, el proceso de NMD se encuentra afectado, por lo que se propone que esta proteína podría realizar funciones semejantes a las proteínas del EJC, RNPS1 o Y14 (Fig. 2), a pesar de que no presentan homología a nivel de secuencia (Culbertson & Leeds, 2003).

En metazoarios (Figuras 6A y 6C), se requieren factores adicionales para que el proceso de NMD se complete de manera exitosa. Estos factores incluyen a las proteínas SMG, descritas inicialmente en *C. elegans*. Las proteínas SMG se han reportado en diversos organismos y llevan a cabo diferentes funciones. SMG1, SMG5 y SMG7 regulan el estado de fosforilación de UPF1 (SMG2 en *C. elegans*) como ya se ha mencionado, mientras que SMG6 también funciona como una endonucleasa. Proteínas adicionales necesarias para que ocurra el proceso de NMD en metazoarios son las que conforman el núcleo del complejo multiproteico EJC (eIF4AIII, Y14, MAGOH, MNLS51 y UPF3).

Inicialmente se propuso que durante la primera ronda de la traducción del ARNm se determina si dicha molécula es blanco para NMD. En los ARNm recién sintetizados, la estructura del CAP en el 5' presenta un complejo que incluye a las proteínas CBP20 y CBP80 y la presencia de este complejo es requisito para entrar a la primera ronda de la traducción en la que entran los ARNm que son blanco para NMD (Ishigaki *et al.*, 2001). Estos ARNm entran a NMD cuando el ribosoma encuentra el fin de la traducción a una distancia de 50-55 nt río arriba del EJC. Para la mayoría de los ARNm, el codón de paro funcional se encuentra en el último exón, de tal manera que los EJC son desplazados del ARNm antes de encontrar el codón de paro funcional. Por el contrario, el ribosoma se arresta cuando encuentra un PTC al asociarse con el complejo SURF (SMG-1 kinase, Upf1, eRF1 y eRF3). En este contexto, recordemos que UPF1 interactúa con el EJC que permanece unido al ARNm río abajo. Esta asociación ocurre debido a la interacción entre UPF2 y UPF3 que se encuentra formando parte del EJC, enviando al ARNm a NMD. Por otro lado, reportes recientes han mostrado que el ARNm unido a CBP y a eIF4E también pueden ser blancos para NMD, sugiriendo que moléculas que han entrado a la ronda de traducción en masa también pueden ser blanco para NMD (Maderazo *et al.*, 2003; Rufener & Mühlemann, 2013; Durand & Lykke-Andersen, 2013). Además de la participación del EJC, otra diferencia central entre los procesos de NMD en *S. cerevisiae* y en mamíferos, radica en que la degradación del ARNm ocurre a partir del extremo 5' luego de la remoción del CAP, mientras que el proceso de NMD en mamíferos activa el corte endonucleolítico, seguido de la degradación de los fragmentos de ARN generados por ambos extremos (Mühlemann & Lykke-Andersen, 2010).

En *C. elegans*, los ARNm que presentan una mutación sin sentido o una región 3'UTR extendida son blancos para NMD (Pulak & Anderson, 1993), así como transcritos fisiológicos que contienen algún PTC (Mitrovich & Anderson, 2000; Ramani *et al.*, 2009; Barberan-Soler *et al.*, 2009). Igual que para *Drosophila* y mamíferos, en *C. elegans* la funcionalidad de los factores centrales de NMD es regulada

por las proteínas SMG. Por el contrario, los componentes del EJC no son necesarios para el NMD y los PTC son detectados en función a la posición relativa con respecto al 3'UTR (Longman *et al.*, 2007; Mango, 2001). Al igual que en *S. cerevisiae*, el mecanismo de NMD no es esencial para la viabilidad de *C. elegans* (Anderson, 2006).

En *D. melanogaster*, existen diversos componentes del EJC que se encuentran conservados, los cuales participan en el metabolismo del ARNm; sin embargo, no parecen ser determinantes para definir un PTC o para desencadenar el proceso de NMD (Wittkopp *et al.*, 2009). En este sentido, la definición del PTC en *Drosophila* ocurre de manera independiente al *splicing* y a la definición de las uniones entre exones (Lejeune & Maquat, 2005). Por otro lado, la degradación de transcritos sin sentido en *Drosophila* se inicia por el corte endonucleolítico cerca del PTC. El fragmento 5' resultante es degradado rápidamente a partir del extremo 3' por el exosoma en un proceso que involucra a los integrantes del complejo Ski, mientras que el fragmento 3' es degradado a partir del extremo 5' por XRN1 (Juneau *et al.*, 2007). De esta manera, los fragmentos de ARN son degradados sin necesidad de que se remuevan el CAP ni la cola de poliA, procesos que son necesarios tanto en *S. cerevisiae* como en mamíferos. De esta manera, en *Drosophila* no se requiere el complejo Lsm1-7 responsable de la remoción del CAP (Juneau *et al.*, 2007).

Aunque el proceso de NMD es esencial también en plantas, todavía se desconoce una buena parte de los detalles de regulación para este mecanismo en dicho modelo. Las plantas corresponden al único grupo de eucariotes en donde ambos tipos de elementos en cis: los 3'-UTR largos y los intrones ubicados en los 3'-UTR son capaces de inducir NMD de manera eficiente (Kurihara *et al.*, 2009; Kertesz *et al.*, 2006). A la fecha se ha demostrado que los factores de NMD UPF1, UPF2 y SMG7 son esenciales para ambos sustratos de NMD, mientras que UPF3 juega un papel esencial en la vía de NMD desencadenada por 3'-UTR largos (Hori *et al.*, 2005; Kerényi *et al.*, 2008). Esta ruta de NMD es similar en *S. cerevisiae* y en invertebrados, en donde UPF1 interactúa con eRF3 impidiendo su interacción con PABP (Kerényi *et al.*, 2006). Por otro lado, los intrones incluidos en el 3'UTR pueden desencadenar eventos de NMD dependiendo de la posición de dicho intrón y se han identificado los homólogos para componentes del EJC, como Y14 y MAGOH, que se requieren para la ruta de NMD en plantas (Kerényi *et al.*, 2006), principalmente en la vía que reconoce intrones presentes en el 3'UTR (Nyikó *et al.*, 2013). En cuanto a la regulación por fosforilación, se ha demostrado que el ARNm para SMG7 es blanco de NMD en plantas (Rayson *et al.*, 2012; Nyikó *et al.*, 2013).

I.IV.II. SUSTRATOS DE NMD

Para asegurar el correcto funcionamiento de la célula, el NMD ocurre desde el núcleo y es activo desde la primera ronda de la traducción (Carter *et al.*, 1996; Belgrader *et al.*, 1994), esto sugiere que todos los mensajeros son escaneados antes de llegar al citoplasma y tan pronto empiezan a salir del núcleo, todavía anclados a la membrana, son reconocidos por la maquinaria de NMD (Izaurralde *et al.*, 1996).

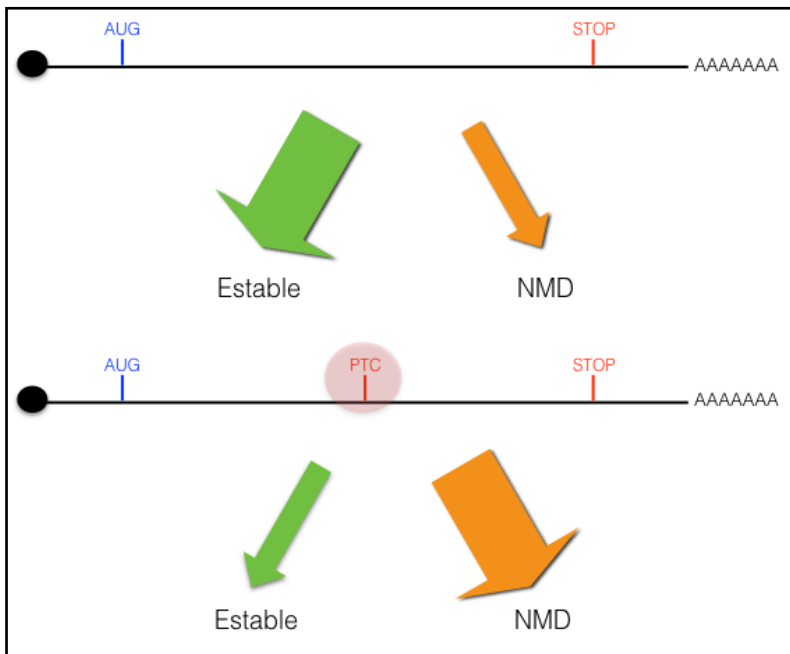


Fig. 8. Sustratos de NMD. El NMD puede reconocer mensajeros con y sin codones de paro prematuros (PTCs). El NMD reconoce principalmente ARNm que contienen PTCs, pero algunos mensajeros que contienen PTCs se escapan del sistema (figura inferior) pues son estables. También pueden reclutarse mensajeros que no contienen PTCs para controlar la abundancia y estabilidad de transcritos funcionales (Rebbapragada & Lykke-Andersen, 2009).

La mayoría de los transcritos en una célula son escaseados por la maquinaria de NMD (Figura 8), sin embargo, la maquinaria los reconoce como estables y los libera pero si encuentra alguna anomalía como en el caso de transcritos que contienen PTCs la maquinaria es capaz de reconocerlos a tiempo para degradarlos. En algunos casos estos transcritos con codones de paro prematuros son isoformas alternativas pero funcionales de *splicing* o mutaciones estables que son liberadas por el sistema (Rebbapragada & Lykke-Andersen, 2009).

Además de los ARNm que contienen PTCs generados por mutaciones,

otros blancos de NMD corresponden a transcritos que han sufrido eventos anómalos de *splicing* y muchos transcritos que han sido generados por *splicing* alternativo pueden presentar al menos una isoforma que incluye un PTC (Green *et al.*, 2003). La degradación de estas isoformas que presentan PTCs mediante NMD ha sido confirmada experimentalmente en líneas celulares de mamífero y en tejidos de ratón (Pan *et al.*, 2006; Weischenfeldt *et al.*, 2012).

A pesar de que inicialmente se propuso que solamente aquellos transcritos que presentaban al menos un PTC eran blancos de NMD, ahora se sabe que esta vía de decaimiento también regula la abundancia y la estabilidad de diversos ARNm funcionales en vertebrados (Fig. 8) (Mendell *et al.*, 2004; Thoren *et al.*, 2010; Tani *et al.*, 2012).

Aquellos ARNm que también han sido descritos como sustratos de NMD son los siguientes (Fig. 9):

- Un ARNm que presenta un PTC cuya inclusión se asocia a un proceso de retención de intrón.
- Las moléculas de ARNm en donde se hayan incluido exones que presenten un PTC, ya sea de manera natural o debido a la inserción de una mutación.
- El ARNm para ciertos ARNs no codificantes, como los productos de los genes para pequeños ARN nucleolares (snoRNAs) SNGH1 y SNGH2 (Ideue *et al.*, 2007; Yamashita *et al.*, 2009).
- Los ARNm que incluyen un marco de lectura abierta adyacente río arriba, como los transcritos para ATF3 y NAT9 (Mendell *et al.*, 2004; Viegas *et al.*, 2007).

- e) Un ARNm que presenta un 3'UTR largo sensible a NMD, como es el caso de SMG5 (Mendell *et al.*, 2004; Singh *et al.*, 2008).
- f) Productos de genes como GADD45B y TBL2, los cuales contienen un intrón embebido en la región 3'UTR (Viegas *et al.*, 2007).
- g) Proteínas que incluyen PTCs para regular su propia expresión, como ocurre en algunos factores reguladores de *splicing* alternativo como proteínas SR y para ciertas proteínas ribosomales (Cuccurese *et al.*, 2005; Lareau *et al.*, 2007).
- h) El codón UAG para selenocisteínas puede ser reconocido como PTC en condiciones de baja concentración de selenio para transcritos como el de la glutatión peroxidasa 1 GPx1 (Sun *et al.*, 2000; Usuki *et al.*, 2011).

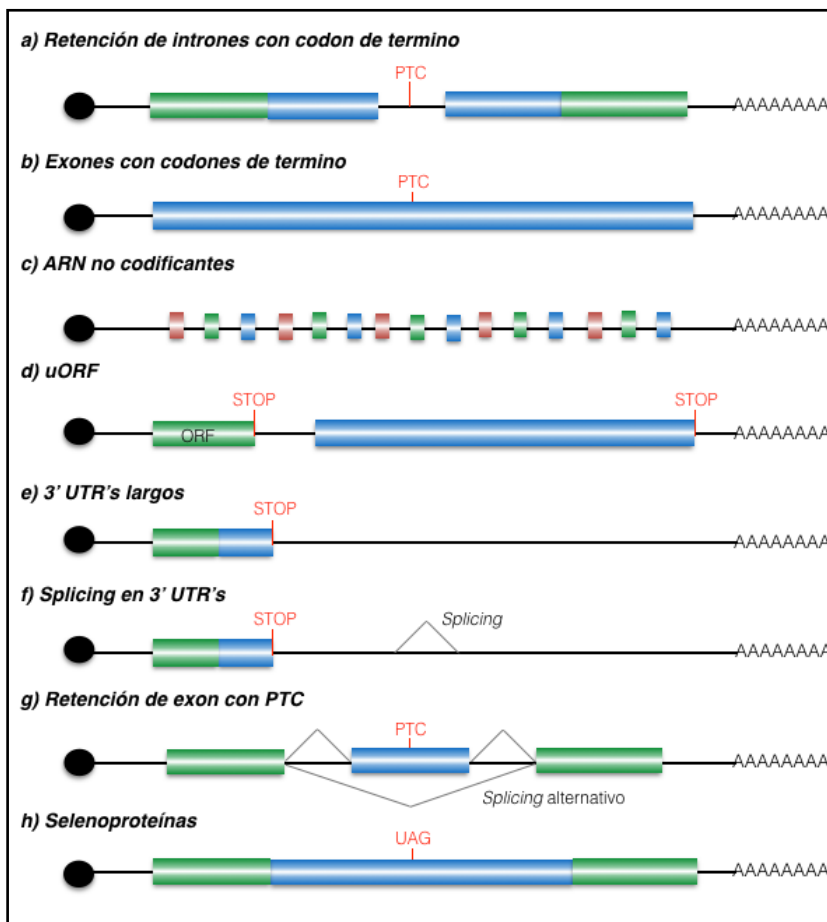


Figura 9. Estructura de los ARNm que pueden ser blanco del proceso de NMD. En todos los casos, se observa que el ARNm presenta Cap (círculo negro) en el extremo 5' y cola de poliA en el extremo 3'. Los rectángulos azules o verdes indican regiones exónicas. Se muestra la posición relativa del codón de paro prematuro (PTC), el codón de paro (STOP), el codón que se puede confundir con PTC en las selenoproteínas (UAG), así como los sitios en donde puede ocurrir un evento de *splicing*. El tipo de ARNm se indica en cada caso.

I.V. NMD Y ENFERMEDADES HUMANAS

Hasta finales del siglo pasado, en muchos casos no había sido posible establecer una función clara para ciertas mutaciones puntuales que habían sido descritas en algunas enfermedades genéticas. Ahora se ha podido demostrar que en algunos casos dichas mutaciones insertan PTCs, los cuales dependiendo de su posición y contexto celular pueden desencadenar una patología o pueden presentar diversos grados de afectación en un individuo (Fig. 10). La primera evidencia que indicaba que un PTC podría estar relacionado con la disminución en la traducción de un ARNm particular provino de estudios del transcrito del gen de β -globina provenientes de pacientes con β -talasemia (Cui *et al.*, 1995; He *et al.*, 1997). En condiciones normales, la hemoglobina está constituida por un tetrámero conformado por dos subunidades de α -globina y dos de β -globina. En la forma recesiva común de la β -talasemia, se ha

evidenciado la presencia de PTCs en la porción 5' para ambas copias del ARNm de la β -globina. En esta situación, la α -globina es degradada debido a la falta de β -globina funcional, lo que conduce a una pobre producción de hemoglobina activa, desencadenando el desarrollo de anemia severa en el paciente. Por otro lado, cuando los PTCs se localizan en el último exón del mRNA de β -globina, estas formas anómalas tienen la capacidad de escapar del proceso de NMD y se producen isoformas truncas de la proteína, lo que genera el fenotipo dominante de anemia heredada (Danckward *et al.*, 2002). El caso de la β -talasemia es uno de los pocos ejemplos de enfermedades genéticas para los cuales se ha estudiado de manera exhaustiva el efecto que tiene el proceso de NMD. Sin embargo, se ha propuesto que un mecanismo similar al que se ha observado en la β -talasemia puede también explicar las

siguientes enfermedades: la susceptibilidad a infecciones por micobacterias relacionadas con el gen *IFNGR1* (Jouanguy *et al.*, 1999), la bradidactilia tipo B o el síndrome de Robinson, ambos causados por mutaciones en el gen *ROR2* (Schwabe *et al.*, 2000), la enfermedad de von Willebrand (Scneppeinheim *et al.*, 2001), así como la deficiencia del factor X (Millar *et al.*, 2000) o la retinitis pigmentosa, éstos dos últimos asociados a mutaciones en el gen *CRX* (Rivolta *et al.*, 2001).

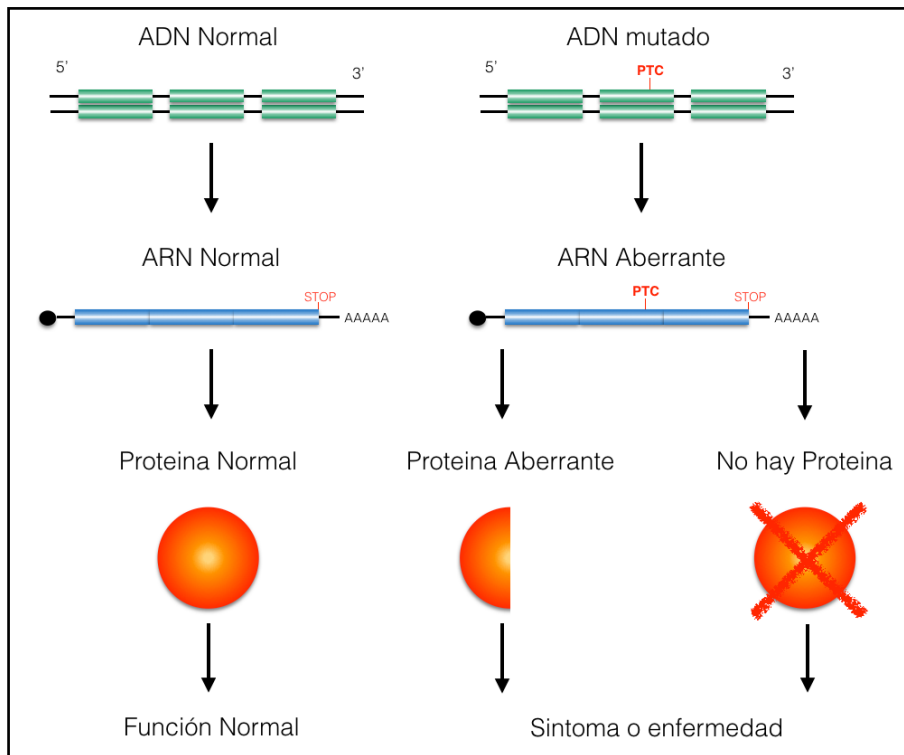


Fig 10. Efecto de la introducción de un PTC debido a una mutación. En condiciones normales, un ADN silvestre genera un ARNm y una proteína acorde a la función asociada. Cuando se presenta una mutación en el ADN que introduce un codón de paro prematuro (PTC) se puede generar un ARN aberrante que puede dar origen a una proteína anómala o puede ocasionar que no se produzca proteína. En ambos casos, es posible que se presente algún síntoma o que se desarrolle cierta enfermedad.

Otro ejemplo interesante de una enfermedad genética que se ha relacionado con NMD es la osteogénesis imperfecta, también conocida como enfermedad de los huesos de cristal. La principal causa de esta enfermedad es la variabilidad observada en la secuencia para los genes *col1a1* y *col1a2*, los cuales codifican para la cadena alfa1 y alfa2 de I colágeno tipo 1. Variaciones en la secuencia de las uniones entre los exones pueden generar la eliminación de un exón alternativo o bien la inserción de PTCs, los cuales a su vez pueden desencadenar el proceso de NMD. La osteogénesis imperfecta causada por PTCs que se insertan debido a sustituciones de un solo nucleótido se ha relacionado con un total de 39 mutaciones detectadas en el gen *col1a1*, lo que sugiere que los PTCs insertados en el gen *col1a2* podrían correlacionar con un fenotipo recesivo. También se han reportado

mutaciones en los exones 50 y 51, las cuales generan PTCs en el último exón, el cual no desencadena NMD pero genera un efecto dominante negativo de la enfermedad (Dagleish, 2014).

La principal evidencia que sustenta el papel de factores de NMD en las enfermedades genéticas humanas deriva de pacientes que presentan diversas patologías relacionadas con el desarrollo neuronal y que se deben a mutaciones observadas en el gen *Upf3b*. En este sentido, se han realizado estudios genéticos en familias que presentan múltiples individuos afectados, en donde los pacientes con mutaciones en dicho gen presentan problemas intelectuales, los cuales se acompañan en ciertos casos de problemas psiquiátricos que van de la esquizofrenia al autismo (Addington *et al.*, 2011; Laumonier *et al.*, 2010; Lynch *et al.*, 2012; Tarpey *et al.*, 2007). Además de las mutaciones en el gen *Upf3b*, *Upf2* también se ha asociado con defectos neuronales y del desarrollo (Singh *et al.*, 2012). De manera particular, deleciones heterocigotas de regiones genómicas que incluyen al gen *Upf2* se han relacionado a fenotipos clínicos de esquizofrenia e incluso una mutación sin sentido fue detectada en un paciente con esquizofrenia (Gulsuner *et al.*, 2013).

Por otro lado, se realizó un estudio en el cual se evaluó la variabilidad en el número de copias para 18 genes implicados en NMD que se sabe que se relacionan con desordenes neuronales, en donde *UPF2* y *RBM8A* se asociaron con diversos defectos neuronales, los cuales se acompañaban en algunos casos de anormalidades congénitas (Nguyen *et al.*, 2013). También se han asociado ciertas irregularidades genómicas en *upf3a*, *smg6*, *eif4a3* o *rnp1* con enfermedades neuronales o con problemas psiquiátricos. De acuerdo con la información depositada en la base de datos de proyectos de cáncer (International Cancer Genome Projects Database CGPD), los genes de NMD que muestran una mayor frecuencia de mutación corresponden a diferentes genes de SMG (Tabla 2). En total, se han reportado 13,242 mutaciones para genes de NMD en individuos afectados por diferentes tipos de cáncer, en donde aquellos con mayor número de mutaciones corresponden a *UPF3A*, *SMG7*, *SMG1* y *UPF2* (Tabla 2).

Tabla 2. Genes de NMD mutados con más frecuencia en pacientes con cáncer (10 de 106 genes)

Símbolo	Nombre	Ubicación	No. Donadores Afectados	No. Mutaciones
SMG1	cinasa SMG1 relacionada a la fosfatidilinositol 3 cinasa	chr16:18816175-18937776	468 / 3,239 (14.45%)	613
SMG6	Factor de NMD SMG6	chr17:1963133-2207065	464 / 3,239 (14.33%)	787
SMG7	Factor de NMD SMG7	chr1:183441351-183567381	456 / 3,239 (14.08%)	601
RPS9	Proteína ribosomal S9	chr19:54704610-54752862	434 / 3,239 (13.40%)	506
UPF2	UPF2 homólogo del regulador de transcritos sin sentido (Sce)	chr10:11962021-12085169	382 / 3,239 (11.79%)	486
SMG5	Factor de NMD SMG5	chr1:156219015-156252620	248 / 3,239 (7.66%)	290
RPL37A	Proteína ribosomal L37a	chr2:217362912-217443903	241 / 3,239 (7.44%)	330
SyPABPC1	Proteína citoplásmica 1 de unión a Poli(A)	chr8:101698044-101735037	233 / 3,239 (7.19%)	241
EIF4G1	Factor eucariote de inicio de la traducción 4, gamma 1	chr3:184032283-184053146	229 / 3,239 (7.07%)	261
PPP2R1A	Proteín-fosfatasa 2, subunidad reguladora A, alfa	chr19:52693292-52730687	222 / 3,239 (6.85%)	238

Tabla 3. Mutaciones más frecuentes que ocurren en genes de NMD reportadas para pacientes de diferentes tipos de cáncer

ID	Cambio en el ADN	Tipo	Consecuencia	No. Donadores afectados
MU80428	chr1:g. 6257785T>-	delección de <=200bp	<i>Frameshift</i> : RPL22 K15 5 UTR: RPL22 Exón: RPL22	41 / 8,038 (0.51%)
MU4473544	chr1:g. 153963239C>T	Sustitución de una sola base	5 UTR: RPS27 Río arriba: RAB13 - RP11-422P24.9 Exón: RPS27 Río abajo: NUP210L	25 / 8,038 (0.31%)
MU612590	chr13:g. 115047559C>T	Sustitución de una sola base	Río arriba: UPF3A Sinónimo: UPF3A L91L Exón: UPF3A Intrón: UPF3A	25 / 8,038 (0.31%)
MU3892625	chr19:g. 54754843A>G	Sustitución de una sola base	<i>Missense</i> : LILRB5 S598P Río arriba: AC010492.4 - CTD-2337J16.1 Exón: LILRB5 Río abajo: RPS9 Intrón: LILRB5	10 / 8,038 (0.12%)
MU4307851	chr1:g. 183496786G>-	delección de <=200bp	Exón: SMG7 Río abajo: SMG7 Intrón: SMG7	10 / 8,038 (0.12%)
MU35402822	chr8:g. 101713496CAA>-	delección de <=200bp	Río abajo: PABPC1 Intrón: PABPC1	9 / 8,038 (0.11%)
MU1800908	chr19:g. 52715971C>G	Sustitución de una sola base	<i>Missense</i> : PPP2R1A P219R, P179R, P124R 5 UTR: PPP2R1A Río arriba: PPP2R1A Exón: PPP2R1A Río abajo: PPP2R1A	8 / 8,038 (0.10%)
MU715959	chr16:g. 18875513T>C	Sustitución de una sola base	Río arriba: SMG1 Río abajo: SMG1 Intrón: SMG1	8 / 8,038 (0.10%)
MU64881	chr10:g. 11990443T>A	Sustitución de una sola base	<i>Missense</i> : UPF2 E1033D	7 / 8,038 (0.09%)
MU4555399	chr19:g. 54726833T>C	Sustitución de una sola base	<i>Missense</i> : LILRB3 T6A Río arriba: LILRB3 - CTB-83J4.1 Exón: LILRB3 Intrón: RPS9 - LILRB3 - LILRA6	7 / 8,038 (0.09%)

Otro factor de NMD que se ha asociado fuertemente con diversas enfermedades es Y14, el cual es parte central del complejo EJC y participa en NMD así como en otros eventos del metabolismo del ARNm, además de que facilita la traducción. Como parte del complejo EJC, Y14 forma un dímero con MAGOH, el cual interactúa de manera directa con eIF4AIII inhibiendo su actividad ATPasa para facilitar la interacción entre el EJC y el ARNm (Ballut *et al.*, 2005). Esta asociación se mantiene aún después del proceso de exporte al citoplasma (Tange *et al.*, 2005). El papel central de Y14 en el proceso de NMD se confirmó al observar que la depleción de este factor inhibía la degradación de transcritos que presentaban PTCs (Gehring *et al.*, 2003). Además del papel en la regulación de la vía de NMD; Y14 es capaz de aumentar la traducción (Diem *et al.*, 2007; Lee *et al.*, 2009), aumentar la metilación de proteínas SM mediada por PRMT5 (Chuang *et al.*, 2011) e inhibir la remoción del CAP (Chuang *et al.*, 2013). Otras funciones adicionales de Y14 incluyen la producción de formas anti-apoptóticas de genes de sufren *splicing* alternativo, como Bcl-x, Bim y Mcl1 (Michelle *et al.*, 2012), así como de otros genes que regulan la progresión del ciclo celular (Rehwinkel *et al.*, 2005). Teniendo en cuenta todas estas importantes funciones de Y14 en la biogénesis y maduración del ARNm, así como en la traducción, resulta lógico proponer que este factor pueda estar relacionado en diversas patologías. En este sentido,

Y14 ha sido implicado en el síndrome TAR (*thrombocytopenia-absent radius*), el cual se presenta cuando un alelo de RBM8A que codifica para Y14 se encuentra deletado al mismo tiempo que se encuentra mutado un SNP regulador en el otro alelo (Albers *et al.*, 2012; Klopocki *et al.*, 2007). Los pacientes que padecen este síndrome presentan un bajo número de precursores de plaquetas en médula ósea, además de retraso mental (Silver *et al.*, 2010; Brunetti-Pierri *et al.*, 2008; Skorka *et al.*, 2005). De esta manera, Y14 parece estar relacionado con la generación de células hematopoyéticas y con el desarrollo cerebral (Chuang *et al.*, 2015).

En la actualidad, existe cada vez más evidencia que indica que el proceso de NMD es un mecanismo clave en la regulación de la expresión genética. En este sentido, se ha descubierto que diversas enfermedades como la fibrosis quística, diferentes tipos de distrofias y algunos tipos de cáncer se generan debido a la presencia de PTCs y se ha estimado que el 12% de todas las mutaciones reportadas en la base de datos para genes humanos (Human Gene Mutation Database) generan PTCs (Krawczak *et al.*, 2000). Si se consideran todo tipo de mutaciones, incluyendo delecciones, inserciones y mutaciones de *splicing*, la cantidad de eventos capaces de inducir la generación de un producto aberrante que puede desencadenar el proceso de NMD se eleva al 30% de todas las enfermedades genéticas (Frischmeyer & Dietz, 1999).

En cuanto a la relación entre el mecanismo de NMD y el proceso de cáncer, se considera que algunos carcinomas pueden verse reforzados debido a anomalías en el proceso de NMD. Por lo general, las células cancerosas muestran defectos en eventos como la reparación del ADN, la progresión del ciclo celular y el mecanismo de apoptosis, lo que permite proponer que diferentes tipos de mutaciones pueden acumularse en tumores, contribuyendo a eventos de proliferación y malignidad. Sin embargo, el proceso de NMD puede contribuir también al mantenimiento de la homeostasis celular gracias a que puede degradar productos aberrantes que se generan a partir de genes esenciales que de otra manera resultan tóxicos. Existe también la posibilidad de que el proceso de NMD afecte la expresión de genes asociados a cáncer debido a que pueda existir un cambio en la cantidad de variantes de *splicing* producidas para dichos genes. En este sentido, se ha demostrado que existen variantes de *splicing* que contienen PTCs y que pueden entrar en NMD para genes supresores tumorales como BRCA1, TP53 y WT1 (Martínez-Montiel & Martínez-Contreras, 2016).

Aún cuando solamente en pocos casos se ha establecido una clara conexión entre un factor de NMD y una enfermedad particular, resulta lógico proponer que exista una relación real más compleja entre la maquinaria de NMD y diferentes procesos celulares o patologías humanas.

I.VI. TERAPIAS PARA ENFERMEDADES RELACIONADAS A NMD

Una posibilidad para el tratamiento de enfermedades causadas por mutaciones que generan PTCs consiste en reducir la eficiencia de la terminación de la traducción, de manera que sea posible restaurar la producción de proteínas funcionales que no sean truncas, sino que presenten el marco de

lectura completo. En eucariotes, la terminación de la traducción ocurre cuando un codón de paro interactúa con el sitio A del ribosoma. Este reconocimiento del codón de paro es mediado por el factor de liberación eRF1, que suple la interacción codón-anticodón que ocurre de manera natural durante la traducción. Después del reconocimiento inicial, eRF1 genera la liberación del polipéptido naciente del peptidil-tRNA localizado en el sitio P del ribosoma (Kisselev *et al.*, 2003). Por lo general, eRF1 compete con el casi-cognado aminoacil-tRNA, el cual puede unirse al sitio A del ribosoma aún cuando contenga solo dos de los tres nucleótidos correspondientes a la secuencia del codón de paro. Bajo condiciones diferentes, la incorporación del casi-cognado aminoacil t-RNA puede ser estimulada para convertir los codones de paro en codones que puedan ser traducidos (Keeling *et al.*, 2011). Este proceso se denomina “supresión de la terminación”. Cuando esto ocurre sobre un PTC que se encuentre en fase, la síntesis de la proteína anómala se restablece, con lo que se obtiene nuevamente la expresión de la proteína funcional mediante un evento denominado “supresión del codón de paro”. Este proceso ocurre de manera natural y puede implementarse como herramienta terapéutica beneficiando a pacientes que presenten ciertas enfermedades genéticas (Fig. 11). Se han empleado diferentes moléculas que tienen la capacidad de inducir este tipo de supresión tanto *in vitro* como *in vivo*; dentro de dichas moléculas encontramos aminoglicósidos (gentamicina, amikacina, tobramicina), aminoglicósidos modificados (NB30, NB54, NB84) y compuestos como el ataluren (PTC124) o el RTC13. Dentro de los blancos para estas moléculas existe una diversidad de enfermedades entre las que se incluyen las siguientes: fibrosis quística, distrofias musculares de Becker o de Duchenne, atrofia muscular espinal, ataxia telangiectasia, síndrome de Rett, síndrome de Usher tipo I, síndrome de Hurler, síndrome de Maroteaux-Lamy, lipofuscinosis infantil neuronal cerosa, cistinosis, diabetes insipidus nefrogénica ligada al cromosoma X, hemofilia, obesidad y cáncer (Miller & Pearce, 2014). Los aminoglicósidos también han mostrado su efecto como supresores de mutaciones sin sentido en los genes p53 y ATM (Keeling & Bedwell, 2011). Algunas enfermedades se encuentran ya en la etapa clínica empleando la terapia de supresión, dentro de las que se encuentran la fibrosis quística, las distrofias musculares de Duchenne o de Becker, la deficiencia del factor VII, la enfermedad de Hailey-Hailey, hemofilia A y B, deficiencia de adhesión leucocitaria 1 y la enfermedad de McArdle (Lee & Dougherty, 2012; Keeling *et al.*, 2012).

Desafortunadamente, varios aminoglicósidos empleados para la terapia de supresión han mostrado ciertos riesgos, como nefrotoxicidad reversible y ototoxicidad permanente, problemas que han sido desarrollados por 2-25% de los pacientes que han recibido estos tratamientos (Keeling *et al.*, 2012). Debido a estos inconvenientes, se han desarrollado esfuerzos adicionales para disminuir los efectos adversos de los aminoglicósidos, entre los que se incluye la administración simultánea con antioxidantes o polianiones, con lo que se ha disminuido de manera importante la toxicidad (Reiter *et al.*, 2011; Thibault *et al.*, 1995). Otra alternativa exitosa ha sido encapsular los aminoglicósidos en liposomas (Meers *et al.*, 2008), pero su aplicación en terapia de supresión requiere todavía ajustes adicionales (Yukihara *et al.*, 2011). Finalmente, un monitoreo cuidadoso y la regulación de la dosis también han sido efectivos para reducir la toxicidad sin sacrificar el efecto terapéutico (Smyth & Bhatt, 2012).

Además del desarrollo de estrategias alternativas para reducir la toxicidad de los aminoglucósidos, el diseño de nuevos análogos efectivos para la supresión pero que sean más seguros y menos agresivos para el paciente ha mostrado resultados esperanzadores, siendo Ataluren, RT13 y RT14 las moléculas más eficaces. Por ejemplo, Ataluren ha demostrado poseer una baja toxicidad y ha sido aprobada por la FDA y su equivalente europeo (EMA) para el tratamiento de la fibrosis quística y la distrofia muscular de Duchenne originadas por mutaciones sin sentido (Peltz et al., 2013), así como para diferentes enfermedades genéticas de diversa índole (Martínez-Montiel & Martínez-Contreras, 2016).

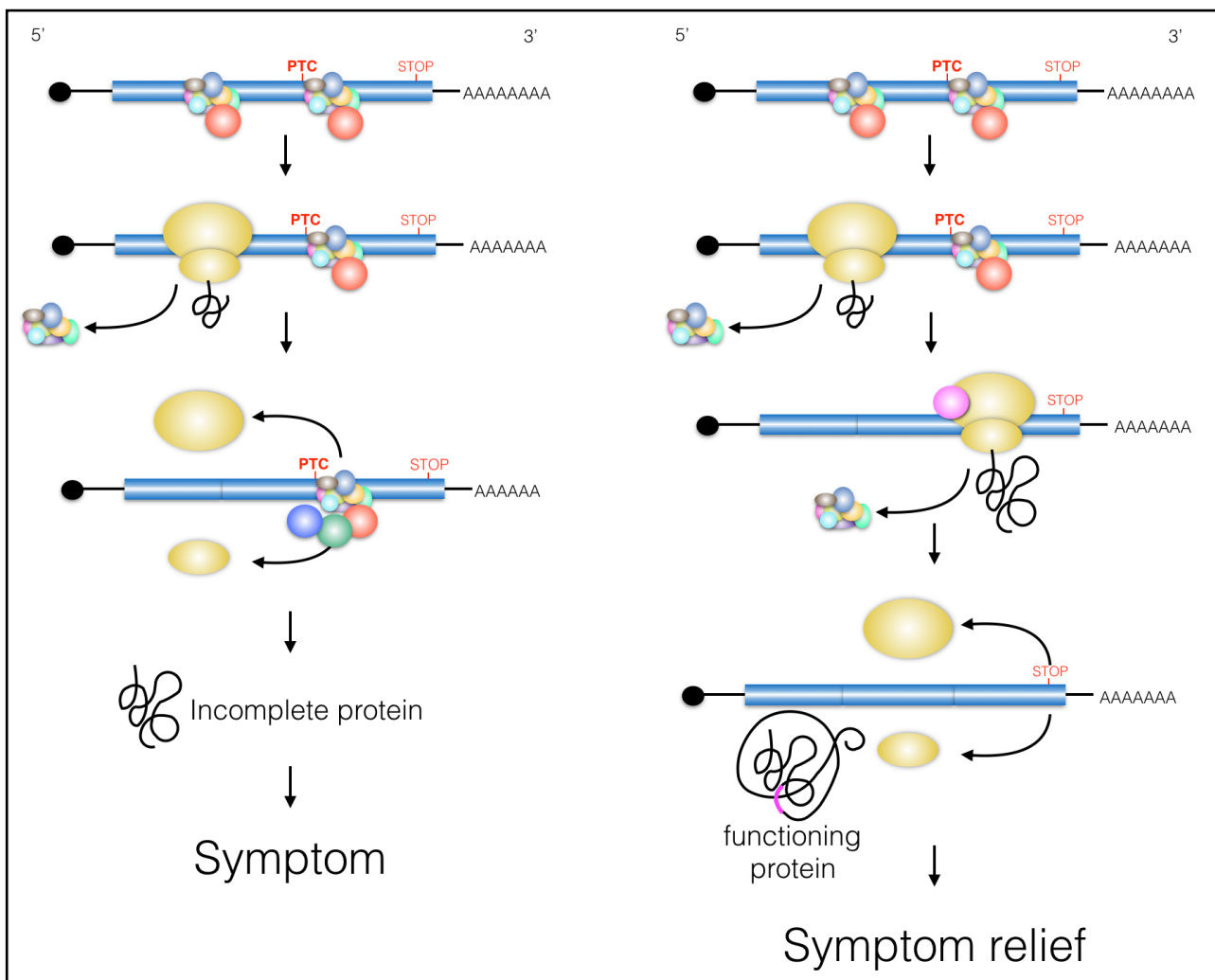


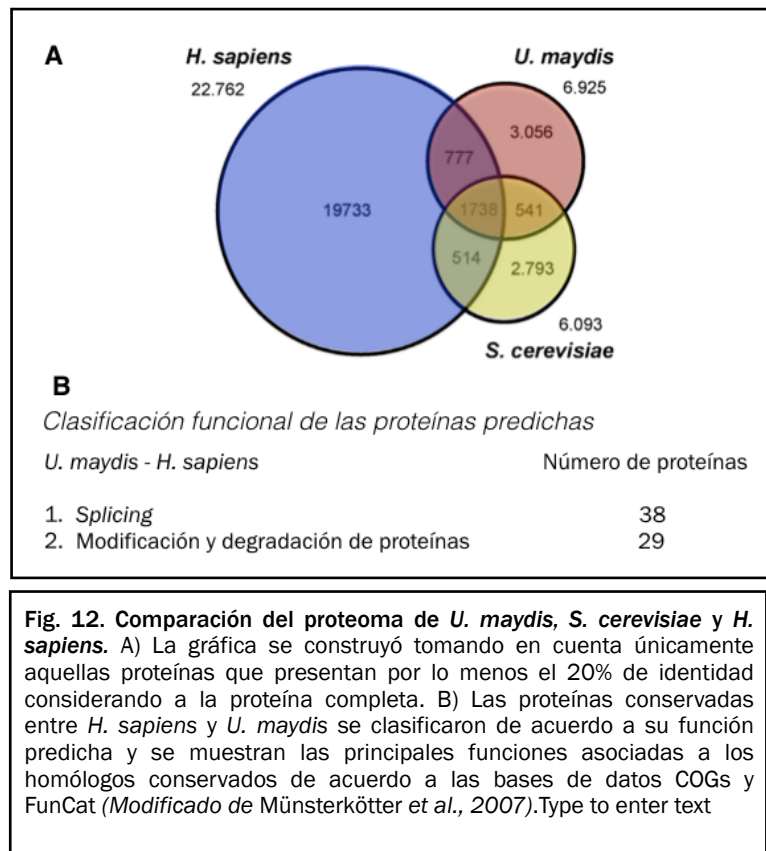
Fig. 11. Mecanismo de lectura a través del codón de paro (readthrough). En condiciones normales, cuando se detecta un PTC, la maquinaria de NMD es reclutada hacia el ARNm y lo degrada. En algunos casos, cuando esto ocurre se produce la versión trunca de una proteína y se puede observar un síntoma o fenotipo asociado (izquierda). Cuando un aminoglucósido o una molécula similar (círculo rosa) interactúa con el ribosoma, el codón de paro anómalo puede ser ignorado por lo que se continúa la traducción y se puede sintetizar una proteína funcional aliviando el síntoma causado por la mutación que insertó el PTC (derecha) (Martínez-Montiel & Martínez-Contreras, 2016).

I.VII. *Ustilago maydis* COMO MODELO DE ESTUDIO

Existen varios sistemas que han sido adoptados para el estudio de la organización genómica y de la regulación de la expresión genética, como el basidiomiceto *Ustilago maydis*. Dicho microorganismo presenta varias de las ventajas que se conocen en las levaduras para este tipo de estudios, como el tamaño del genoma (similar en ambos microorganismos), así como su rápido crecimiento; además de

la capacidad de sus esporas para germinar y formar productos por meiosis haploide, los cuales pueden ser propagados formando colonias compactas como lo hacen algunas levaduras. Es por ello que este modelo ha sido empleado para estudiar procesos como la recombinación y reparación del ADN en la célula (Holliday, 1974). Las conclusiones obtenidas en estos estudios se han extrapolado a eucariotes superiores y son vigentes en la actualidad. Por otro lado, el acceso al genoma completo de *U. maydis* lo perfila como un importante modelo de estudio para diferentes procesos en sistemas eucariotes (http://www.broad.mit.edu/annotation/fungi/ustilago_maydis/).

Por mucho tiempo, *Ustilago maydis* ha sido un modelo bien establecido para el estudio de la fitopatología molecular. Sin embargo, recientemente se volvió evidente que *U. maydis* y *H. sapiens* comparten proteínas y procesos celulares que no se encuentran en el modelo fúngico *S. cerevisiae*. En este sentido, se comparó y analizó *in silico* el proteoma de *U. maydis*, *S. cerevisiae* y *H. sapiens*, encontrando sorprendentemente que el proteoma de *U. maydis* es más semejante al de humano que al de *S. cerevisiae*. Además, *U. maydis* comparte con el humano 777 proteínas, las cuales están ausentes en *S. cerevisiae*. Estas proteínas adicionales participan en importantes procesos celulares y moleculares en mamíferos, incluyendo la señalización mediada por proteínas G, la oxidación de ácidos grasos y el *splicing* del ARNm, entre otros (Münsterkötter *et al.*, 2007).



II. JUSTIFICACIÓN

En eucariotes, la regulación post-transcripcional de la expresión de diferentes ARNm, incluye el proceso conocido como NMD (*Nonsense-Mediated mRNA Decay*). Debido a que existen dificultades técnicas para estudiar la regulación de la expresión genética en humanos, proponemos al basidiomiceto *Ustilago maydis* como un modelo alternativo para el estudio de este mecanismo.

La identificación de los homólogos para los principales factores proteicos que catalizan el proceso de NMD en el basidiomiceto, así como el análisis *in silico* de los dominios conservados en dichos factores, podrán orientar las pruebas de interacción para validar funcionalmente a los homólogos identificados. Más aún, el análisis global de los factores de NMD en diferentes especies podría sugerir la forma en la que ocurre este mecanismo en *U. maydis*.

Con este abordaje global pretendemos validar a *Ustilago maydis* como un modelo experimental para el estudio en eucariotes del decaimiento del ARNm.

III. HIPÓTESIS

El hongo dimórfico *Ustilago maydis* posee la maquinaria proteica funcional responsable del mecanismo de decaimiento del ARN mensajero sin sentido (NMD).

IV. OBJETIVOS

IV.I. Objetivo General

Identificar y caracterizar los principales factores proteicos responsables del decaimiento de ARN mensajero sin sentido (NMD) del en *Ustilago maydis*.

IV.II. Objetivos Particulares

- Analizar *in silico* algunos componentes de la maquinaria de NMD en *Ustilago maydis* así como en otras especies importantes.
- Comprobar la presencia en el genoma y la expresión de los factores principales de NMD, UPF1, UPF2 y UPF3 en la cepa de *Ustilago maydis* FB1.
- Identificar posibles mensajeros blanco de la vía de NMD en *Ustilago maydis* FB1.
- Analizar mediante ensayos de interrupción las posibles repercusiones funcionales de Upf1 en el desarrollo del hongo, incluyendo las posibles alteraciones en la supervivencia del basidiomiceto.
- Demostrar experimentalmente la interacción proteína-proteína entre los factores principales de NDM, UPF1 y UPF2 de *Ustilago maydis*.

V. MATERIALES Y MÉTODOS

V.I. Análisis Bioinformático

La bioinformática es una herramienta muy útil para analizar datos biológicos, con la cual es posible examinar secuencias de diferente índole. En el presente trabajo se utilizaron algunas de estas herramientas para la identificación de secuencias dentro de los bancos de datos (NCBI y KEGG), diseño de oligonucleótidos, análisis de los resultados de secuenciación, construcción de matrices y árboles filogenéticos para determinar el grado de relación evolutiva, estudio de la estructura secundaria y terciaria de las proteínas, modelado de proteínas a partir de cristales, dinámica molecular para el comportamiento de nuestras proteínas, etc.

V.I.I. Análisis proteómico y evolutivo de la ruta de vigilancia del ARNm

Se estudiaron las proteínas de la ruta de vigilancia del ARNm en *U. maydis* y otras 299 especies de eucariotes. Además de las proteínas de vigilancia se analizaron algunas otras de *splicing*, transporte y degradación de mensajeros todo esto con ayuda del KEGG (*Kyoto Encyclopedia of Genes and Genomes*) y su herramienta *KEGG Pathway*, además de algunas otras como KOALA, GFIT, DBsearch y tablas de ortólogos y parálogos. Los alineamientos se hicieron utilizando Clustal, MAFFT, MUSCLE o PRRN, los árboles filogenéticos se hicieron con el programa MEGA 5 y CLC *Sequence Viewer 7*.

V.I.II. Análisis de la estructura secundaria y terciaria de las proteínas del núcleo del NMD

El análisis de la estructura secundaria de las proteínas se hizo utilizando la base de datos InterPro (*protein sequence analysis & classification*) del EMBL-EBI (*European Molecular Biology Laboratory*). La información de los cristales de humano y *Saccharomyces* se obtuvo de la base de datos PDB (*Protein Data Bank*), a partir de estos cristales los modelados se obtuvieron con SWISS-MODEL, las interacciones con STRING. Empleando *Phosphosite plus* obtuvimos la información del estado de fosforilación, ubiquitinación y otros arreglos de las proteínas.

V.II. Material Biológico

V.II.I. Microorganismos y condiciones de crecimiento

El organismo principal de este estudio fue *U. maydis* FB1 (cepa de referencia, *Broad Institute*); para el mantenimiento de esta cepa se utilizaron diferentes medios de cultivo líquidos y sólidos, cuya composición y preparación se detallan a continuación:

YEPS (Yeast Extract, Peptone, Sugar): este medio se compone (por cada litro) de Peptona de Caseína 10-g, Extracto de Levadura 5 g, Cloruro de Sodio (NaCl) 10 g, para el medio líquido. En algunos experimentos se utilizó medio sólido al que se le agregaron 15 g por litro de agar.

En algunos casos el YEPS se suplementó con HEPES 1 M o Sorbitol a una concentración final de 1 M en el medio.

Para los experimentos de biología molecular se utilizó la bacteria *E. coli* DH5 α . Para el mantenimiento de esta cepa se utilizó un medio de cultivo líquido y sólido, cuya composición y preparación se detallan a continuación:

LB (Luria-Bertani; Sambrook *et al.*, 1989): medio de cultivo líquido utilizado de forma habitual para el crecimiento bacteriano, en este trabajo se utilizó de manera regular para crecer *E. coli* DH5 α . El medio se compone (por cada litro) de Peptona de Caseína 10 g, Extracto de Levadura 5 g, Cloruro de Sodio (NaCl) 10 g, para el medio líquido; en algunos experimentos se utilizó medio sólido al que se le agregan 15 g por litro de agar.

En caso de ser necesario los medios de cultivo fueron suplementados con antibióticos los cuales se añadieron al medio estéril a partir de una solución concentrada. La preparación se hizo siguiendo las instrucciones reportadas (Sambrook *et al.*, 1989), empleando las siguientes concentraciones:

- Ampicilina (Ap) (sal sódica, Roche): se preparó una solución concentrada a 100 mg/ml en agua bidestilada para obtener en el medio una concentración de 100 μ g/ml, se esterilizó por filtración.
- Higromicina B (Hyg) (Sigma): la solución madre se preparó a 250 mg/ml en agua bidestilada, se esterilizó por filtración y se guardó protegida de la luz, en el medio de cultivo se encuentra a una concentración final de 250 μ g/ml.

V.II.II. Plásmidos.

Los plásmidos utilizados en este trabajo y sus características se detallan en la tabla 4

Tabla 4. Plásmidos utilizados en este estudio

Plásmido	Uso	Características	Referencia
pCDNA3.1/Hygro	Cassette de resistencia a Higromicina	rep (pUC), bla (Amp), promotor CMV, Sitio múltiple de clonación (SMC), hph (HygB), BGH pA	Invitrogen
pJet1.2	Reservorio para productos de PCR	rep (pMB1), bla (Ap), eco 47IR, PlacUV5, promotor T7, Sitio múltiple de clonación (SMC)	ThermoFisher
pPROExHT (a, b y c)	Expresión de proteínas recobinantes	rep (pBR322), His(6) N-terminal, bla (Ap), lacI, promotor Trc, Sitio múltiple de clonación (SMC)	ThermoFisher

V.III. Técnicas Generales De Biología Molecular

V.III.I. Técnicas Experimentales Con Ácidos Nucleicos

V.III.I.I. Extracción de ADN genómico (*Kit Wizard Genomic DNA Purification, Promega*)

Para la extracción del ADN genómico con el Kit Wizard Genomic DNA Purification, el hongo se cultivó en 5 ml de medio YPD o YPD-Hyg líquidos (ver apartado 5.1.2) durante toda la noche a 30 °C en agitación constante. El cultivo completo se centrifugó por 2 min a 13,000 rpm decantando el sobrenadante, se lavó la pastilla con 1 ml de MgSO₄ 10 mM y se decantó el exceso con la punta. Para la lisis celular se añadieron 600 µl de *Nucleic lysis solution* mezclando suavemente con una punta. Se incubó 5 min a 80 °C y se enfrió a temperatura ambiente. Se añadieron 3 µl de *RNase solution* y se mezcló con la punta. Se incubó a 37 °C unos 30 min, enfriándose a temperatura ambiente. Para la precipitación de proteínas se añadieron 200 µl de *Protein Precipitation Solution*, se mezcló con vortex e incubó en hielo. Tras centrifugar a 13,000 rpm durante 3 min se llevó a cabo la precipitación y rehidratación de ADN; para ésta se transfirió el sobrenadante a un microtubo nuevo y se añadieron 600 µl de isopropanol, se mezcló suavemente y se centrifugó por 2 min a 13000 rpm, desechando el sobrenadante; luego se añadieron 600 µl de etanol al 70% y se mezcló. Enseguida se centrifugó por 2 min a 13000 rpm aspirando el etanol, se secó el *pellet* al aire unos 15 min. Por último, para la rehidratación se añadió al *pellet* de ADN 100µl de *Rehydration Solution* incubando por 1 h a 65 °C.

V.III.I.II. Extracción de ADN plasmídico mediante lisis alcalina (*Green & Sanbrook, 2012*)

Para el aislamiento del ADN plasmídico se cultivaron las células en los medios LB para *E. coli* durante toda la noche a 37 °C, en agitación. Una vez crecidas las bacterias, se centrifugaron 2 ml de cultivo durante 3 minutos a 13000 rpm y se resuspendieron en 100 µl de la solución I (50 mM de glucosa, 25 mM de Tris HCl pH 8, 10 mM de EDTA), que se mezcló con las células invirtiendo suavemente el tubo o utilizando una punta de micropipeta. La mezcla se mantuvo durante 5 minutos a temperatura ambiente, añadiendo a continuación 200 µl de la solución II (0,2 N de NaOH y 1% de SDS), preparada al momento. Se homogenizó la mezcla mediante agitación suave hasta que la solución se volvió transparente debido a la lisis celular. A continuación, se incubó durante 5 min en hielo. A la mezcla se le añadieron 150 µl de la solución III (Acetato potásico 3 M). Los lisados se centrifugaron durante 10 min a 13000 rpm. A los sobrenadantes obtenidos se les añadió 1 volumen de fenol:cloroformo:alcohol isoamílico (24:24:1). Tras centrifugar durante 5 minutos se recuperó la fase acuosa. Para precipitar el ADN, a ésta fase se le añadió 1 ml de etanol absoluto frío y se mezcló con vortex (ligeramente), se centrifugó nuevamente y a continuación se añadió 1 ml de etanol al 70% para lavar la pastilla. Finalmente, se eliminaron los restos de alcohol, dejando secar en una campana de extracción y se resuspendieron los precipitados en 50 µl de TE o H₂O MQ con 0.01% de RNAsa.

V.III.I.III. Extracción de ARN con Trizol (Trizol LS, Invitrogen)

El TRIZOL es un reactivo que sirve para aislar ARN total de diferentes tipos de muestra. El reactivo consiste en una solución monofásica de Fenol e Isotiocianato de Guanidina, el cual corresponde a una modificación del método original desarrollado por Chomczynski y Sacchi (1987). Para la extracción del ARN se tomaron 3 ml de la suspensión celular de un cultivo de toda la noche con una DO600 de 1.0 (ajustando el volumen si no se tenía la densidad adecuada), se centrifugó durante 3 min a 13000 rpm decantando el sobrenadante; a continuación las células se lavaron con 1ml de PBS y se repitió la centrifugación para después resuspender en 750 μ l de TRIZOL LS (Invitrogen) y se mezcló. Para la fase de separación, la muestra homogenizada se incubó por 5 min a temperatura ambiente (para permitir la completa disociación de los complejos nucleoprotéicos) y se adicionaron 200 μ l de cloroformo, se agitó vigorosamente con la mano por 15 segundos y posteriormente se incubó 15 min a temperatura ambiente (15 a 30°C), centrifugando luego la muestra a 12000g por 10 min a 4°C. Para eliminar cualquier residuo de ADN, la fase acuosa se incubó con 2 μ l de DNasa por 2 horas. En la fase de precipitación se transfirió la fase acuosa a un tubo limpio mezclándola con 500 μ l de alcohol isopropílico, se incubó la muestra por 10 min a temperatura ambiente y se centrifugó a 12000g por 10 min a 4°C. A continuación se removió el sobrenadante con mucho cuidado y se lavó el ARN una vez con 1 ml de etanol al 75% para después mezclar la muestra vortexeando y volver a centrifugar a 7500g por 5 min a 4°C. Por último, para resuspender el ARN se secó rápidamente la pastilla y se resuspendió en agua libre de RNAsas pasando la solución varias veces por la pipeta e incubando por 10 min a 55°C.

V.III.I.IV. Cuantificación de ADN y ARN (Green & Sanbrook, 2012)

La cuantificación de ADN y ARN se realizó mediante espectrofotometría. Se tomaron dos medidas, una a 280 nm y otra a 260 nm. La relación 260/280 proporciona una idea del grado de pureza del ADN. Si este valor es próximo a 1.8, la preparación se considera pura. El valor de la absorbancia a 260 nm sirvió para estimar la cantidad de ADN, utilizando la siguiente relación: [1 unidad de absorbancia a 260 nm = 50 μ g/ml de ADN de doble cadena] y para ARN [1 unidad de absorbancia a 260 nm = 40 μ g/ml de ARN de cadena sencilla].

V.III.I.V. Electroforesis de ADN y ARN en geles de agarosa (Green & Sanbrook, 2012)

En las distintas electroforesis se usaron geles de agarosa al 1% (p/v) para ADN y 1.5% (p/v) para ARN, ambos en tampón TAE (40 mM de Tris-acetato pH 8 y 1 mM de EDTA). Este mismo tampón fue usado para el corrimiento. Las muestras de ADN, mezcladas con el tampón de carga (0,25% de azul de bromofenol (p/v) 0,25% xilencianol (p/v) y 15% de Tris-Ficoll (p/v) en Agua, pH 8), se depositaron en los pocillos de los geles solidificados y se corrieron a 100 V. Como patrón de referencia se utilizó el marcador de peso molecular de 1kb (Fermentas). Tras la electroforesis, los geles se tiñeron con bromuro de etidio (1 μ g/ml) y se visualizaron con un transluminador de luz ultravioleta (300 nm de longitud de onda).

V.III.I.VI. Retrotranscripción (M-MuLV Reverse Transcriptase, Fermentas)

Para la obtención de transcritos a partir de ARN se emplea la técnica llamada retrotranscripción. En ésta, a través de una transcriptasa reversa se sintetizan fragmentos de ADN complementario (ADNc), teniendo como templado el mRNA. Para este procedimiento se agregó a un tubo lo siguiente: H₂O-DEPC 10 µl, Oligo reverso 1 µl y RNA 1 µl (500ng-1µg) y se mezcló; a continuación se incubó a 65 °C durante 5 min para después agregar, Buffer 5X 4 µl, Ribo Lock RNase Inhibitor 1 µl, mezcla de dNTPs 10 mM 2 µl y M-MuLV Reverse Transcriptase 1 µl (Fermentas). Todo se mezcló y se incubó a 37 °C una hora, la reacción se detiene calentando a 70 °C durante 5 min.

V.III.I.VII. Amplificación de Fragmentos de ADN mediante PCR (New England BioLabs)

Para amplificar los diferentes fragmentos descritos en el trabajo, las reacciones de PCR se llevaron a cabo en un volumen final de 25 µl, el cual contenía 2.5 µl de Buffer de PCR (10X), 1.5 µl de MgCl₂ (25 mM), 1 µl de dNTP's (10 mM), 0.5 µl de cada oligonucleótido (10 mM) y 0.2 µl de Taq ADN polimerasa (5 U), aproximadamente 10-100 ng de templado y c.b.p. (cuanto baste para) 25µl de H₂O MQ. El programa de amplificación general se describe a continuación: pre-desnaturalización de 95 °C por 5 min, seguida de 30 a 35 ciclos de las siguientes condiciones: 94 °C de 30 a 45 segundos para desnaturalizar, la temperatura de alineamiento correspondiente por 30 a 45 segundos y una extensión a 72 °C por el tiempo necesario para amplificar el fragmento esperado, tomando 1 minuto por cada mil pares de bases a amplificar. Al finalizar los ciclos se programó una extensión final de 72 °C por 7 minutos. La temperatura de alineamiento se calculó a partir de la media de las T_m (temperatura media) del par de oligonucleótidos usados en cada reacción.

La lista de los oligonucleótido y las temperaturas medias se describe en la siguiente tabla:

Tabla 5. Secuencias de los oligonucleótidos utilizados en este trabajo.

Nombre del Oligo	Secuencia	Tm	Referencia
Upf1Fwd	ATG TCC GAG CTC TAC CAG TTT	62	Este trabajo
Upf1RevLong	TCA AAA CGT CGT AAA TGA AGA C	60	Este trabajo
Upf1RevCH	CGG CTT GTC GAG ATC GTC AA	62	Este trabajo
Upf1RevRT	GGC GTG CTT GGA GTG TTT CT	62	Este trabajo
Upf2Fwd	ATG GCA GCT TCC GCT AGC T	60	Este trabajo
Upf2RevLong	TTA GAA GCC GGG ACT GTT CC	62	Este trabajo
Upf2RevRT	GCG CTC TTT GTA AAT TGC CTG	62	Este trabajo
Upf3Fwd	ATG TCG CAG TCA CAA AGA GG	60	Este trabajo
Upf3RevLong	CTA GAG GAT GTG CAG CTT GG	62	Este trabajo
Upf3RevRT	CGC CCT GTC GGT TTG TTG C	62	Este trabajo
CHY2Fwd	TGG TCA ACC TGG ACA AGC TC	62	Este trabajo
CHY2Rev	AAC CTT ACC GTA ACC GGC AG	62	Este trabajo

Nombre del Oligo	Secuencia	Tm	Referencia
ppiFwd2	GCC AAG TTC GAA GAC GAG AA	60	Este trabajo
ppiRev2	ACG GTG GTG ATG AAG AAC TG	60	Este trabajo
PARFwdRT	AGA AGC TCG CTT CGG TTC	56	Ho et al., 2007
PARRevRT	AGA AGC TCG CTT CGG TTC	56	Ho et al., 2007
CRMPFWD	TCA CTT TGC TCG ATG CAG TT	58	Ho et al., 2007
CRMPREV	GGA TCC GAC TCG ATT GAA GA	60	Ho et al., 2007
PARFWD1	CAC GAC ACT TTT CTC AAA GTA CCT A	70	Ho et al., 2007
PARREV	CGT CAT CGG ACT CTT CCT TC	62	Ho et al., 2007
FraFwd	CAC GGA CAG CCA ATA TCA CA	60	Ho et al., 2007
FraRev	TCG AAA TCG AAT CGC TTG GG	60	Ho et al., 2007
ZFPFwd	CAA AAG CTC AGT CGA GCG TG	62	Ho et al., 2007
ZFPRev	GCC GGA TTC AGT GTA GTA AC	60	Ho et al., 2007
GSKFwd	CTC AAC GGC GTC AAG CTC A	60	Ho et al., 2007
GSKRev	ACC ACG TTA GGG TGC TTT AC	60	Ho et al., 2007
HygFwd	GGT GAT GGT TCA CGT AGT GG	62	Este trabajo
HygRev	AGC TCA CTC ATT AGG CAC CC	62	Este trabajo
OligodT	TTT TTT TTT TTT TTT T	32	Este trabajo
RNPS1FwdRT	CAC AAT CAA GAC GTG CCA AAC	62	Este trabajo
RNSP1Rev	TCA TCG TGT GTG TTG GCG G	60	Este trabajo
HBS1Fwd	ATG TCA CGA CAT CGC GCC G	62	Este trabajo
HBS1RevRT	TTG CTT CTC TTC GGC TCT TT	58	Este trabajo
SMG1Fwd	ATG ATT GCT GCC AGT ATC GTG	62	Este trabajo
SMG1RevRT	CTC TCC TCG TTC AAC AAC GA	60	Este trabajo
SMGsFwd	ATG TCA CAG AGG GAC CAG C	60	Este trabajo
SMGsRevRT	GTG GTG TTG CCA GAT GAT GA	60	Este trabajo
MAGOHFwd	ATG TCT GAC TTT TAC GTC CGA	60	Este trabajo
MAGOHRevRT	TCC ACA TCT CTT TGC GGA TTA G	64	Este trabajo
MusashiFwd	ATG GCA AGT CAG GAA GAC AC	60	Este trabajo
MusashiRevRT	CTG CTC ATC GCC AGC ATG G	62	Este trabajo

V.III.I.VIII. Uso de enzimas de restricción

Las endonucleasas o enzimas de restricción son aquellas que pueden reconocer una secuencia específica de nucleótidos dentro de una molécula de ADN y causar una ruptura de los enlaces fosfodiéster en la doble hebra; este corte origina dos extremos que pueden ser romos o cohesivos. Las enzimas utilizadas en este trabajo fueron suministradas por la casa comercial Roche o Fermentas y se utilizaron siguiendo las instrucciones del fabricante. Generalmente las reacciones se incubaron de 1 a 16 horas a 37 °C excepto las FastDigest (Fermentas) que se incubaron 5 min y se inactivaron incubando

a 65 u 80°C de 5 a 20 min. La concentración de Buffer y de la enzima fue 10% del volumen de reacción total siendo la única constante. La eficiencia de la restricción se comprobó mediante electroforesis en gel de agarosa.

V.III.I.IX. Reparación de fragmentos de ADN.

Cuando los extremos de ADN después de las restricciones enzimáticas son cohesivos dejan un extremo dispar en la doble hebra, si este tipo de extremos se quisieran ligar deben coincidir; de no ser así la unión no se lleva a cabo, para que esto no sea un impedimento al hacer una construcción entre vector y fragmento, éstos se reparan. La reacción se lleva a cabo con una enzima que sea capaz de rellenar los extremos insertando los nucleótidos que faltan en la hebra. La enzima utilizada fue el fragmento Klenow (Thermo Scientific), la cual corresponde al fragmento largo de la polimerasa I y se usó al 2% en una reacción final de 20 µl en los que se incluyeron 2 µl de Buffer, 1 µl de mezcla de dNTPs (10 mM), 0.1 a 4 µg de ADN y el resto de H₂O MQ; la mezcla se incubó a 37°C por 10 min y se termina la reacción incubando 10 min a 75°C.

V.III.I.X. Ligación de moléculas de ADN (T4 DNA ligase, Fermentas)

Para clonar un fragmento de ADN en un vector, los fragmentos y vectores linearizados y reparados de ser necesario, fueron purificados y concentrados para posteriormente poder ser ligados. Con esta reacción se buscaba unir los extremos romos del fragmento con los del vector; para esto, se mezcló en un volumen final de 20 µl lo siguiente: 2 µl de Buffer 10X, 2 µl (1 U/µl) de ligasa del fago T4, de 25 a 50ng de ADN lineal en una relación de 3:1 inserto:vector, adicionando el resto de H₂O MQ. La mezcla se incubó durante toda la noche a temperatura ambiente.

V.III.I.XI. Purificación y concentración de ADN (High Pure PCR Cleanup, Roche)

El kit de Roche *High Pure PCR Cleanup* permite purificar y concentrar productos de PCR pero también recuperar ADN a partir de bandas de geles de agarosa. Esta técnica se empleó para purificar y concentrar productos únicos de PCR, vectores linearizados a partir de geles de agarosa, preparar productos para secuenciar, etc. Para productos en fase acuosa el volumen fue de 100 µl, de no ser así se ajustó con H₂O MQ para después adicionar 400 µl de *Binding Buffer*, se mezcló bien y centrifugó brevemente. Para el caso del ADN proveniente de un gel de agarosa es necesario correr el ADN en un gel de agarosa, teñirlo y cortar la banda deseada depositándola en un microtubo y pesándola (el peso resulta de la diferencia entre el tubo vacío y el tubo con la banda); por cada 100 mg de agarosa se adicionaron 300 µl de *Binding Buffer*. Para liberar el ADN de la agarosa se agitó con vortex la mezcla y se centrifugó brevemente para después calentar a 56°C 10 min, se mezcló fuertemente de 2 a 3 veces durante la incubación y una vez que la mezcla se disolvió completamente se adicionaron 100 µl de *Binding Enhancer* por cada 100 mg de agarosa, se mezcló bien con vortex y se centrifugó brevemente. A partir de este punto se unen los protocolos; se colectó la mezcla de los pasos anteriores y se hizo

pasar por una columna insertada en un tubo colector, se centrifugó de 30-60 segundos a 8000 rpm y se desechó el sobrenadante; a continuación se adicionaron a la columna 400 µl de *Wash Buffer* y se centrifugó de 30-60 segundos a 8000 rpm, se volvió a lavar la columna ahora con 300 µl de *Wash Buffer*, con centrifugación en las mismas condiciones desechando el sobrenadante. Para eliminar los restos del *Wash Buffer* se centrifugó 1 min a máxima velocidad; para terminar se reemplazó el tubo colector por un microtubo limpio y se adicionaron de 10 a 20 µl en la matriz de la columna centrifugando 1 min a 8000 rpm.

V.III.I.XII. Secuenciación

El ADN destinado a la secuenciación fue aislado mediante el kit descrito en la sección anterior de este apartado y preparado según las consideraciones de la unidad de secuenciación de MacroGen. El producto de la secuenciación fue posteriormente analizado como se describe en el apartado de análisis bioinformático.

V.III.II Técnicas de mutagénesis

V.III.II.I Interrupción del gen de *umUpf1* en *U. maydis* con un cassette de resistencia a Higromicina B

Para la interrupción del gen *Upf1* en el genoma de *U. maydis* FB1 se siguió el siguiente procedimiento. Inicialmente se amplificó por PCR y a partir de ADN genómico el marco de lectura abierto (MAL) completo para el homólogo de *hUpf1* en la cepa FB1 (UM00237.1). Los oligonucleótidos empleados fueron *Upf1Fwd* y *Upf1RevLong* descritos en la tabla 5; una vez obtenido el producto de PCR de aproximadamente 3.2 Kb, se clonó en el vector reservorio *pJet2.1* (Thermo Scientific). Habiendo verificado la correcta clonación y obtención del plásmido *pJet2.1/Upf1* con un tamaño aproximado de 6.3 Kb, este se linearizó con la enzima *HpaI* la cual tenía un solo sitio de reconocimiento en el plásmido, al centro de MAL de UM00237.1. Por otro lado pero de manera simultánea se amplificó también por PCR el cassette de resistencia a Higromicina B (*hph*) a partir del plásmido *pcDNA3.1/Hygro* (Invitrogene), el cual tiene un tamaño aproximado de 2 Kb, ambas amplificaciones se hicieron con una enzima de alta fidelidad. Tanto el casete como el plásmido *pJet2.1/Upf1* linearizado se obtuvieron con extremos romos y se purificaron para después ser ligados con la finalidad de crear un plásmido que contuviera el gen de *Upf1* interrumpido por el cassette de resistencia a Higromicina B el cual incluía su propio promotor; una vez que se obtuvo esta clonación, el nuevo plásmido denominado *pJet2.1Upf1::Hyg* fue linearizado con la enzima *EcoRV*, la cual presenta un sitio único de corte en este vector (fuera de la constricción de interés). El plásmido lineal se purificó para poder ser electroporado en células de *U. maydis* FB1 como se detalla en el apartado V.III.III.III.

La selección se realizó recuperando las células resistentes a Higromicina B, que por medio de recombinación homóloga, intercambiaron el gen viable del genoma por nuestra construcción con la versión interrumpida del gen.

V.III.III. Técnicas de transferencia genética

V.III.III.I. Transformación de células quimiocompetentes de *E. coli* DH5 α

V.III.III.I.I. Preparación de células quimiocompetentes de *E. coli* DH5 α

La preparación de células competentes se realizó mediante tratamiento con CaCl₂ (Cohen *et al.*, 1972). A partir de un inóculo crecido toda la noche en LB se hizo una dilución 1:100 en medio de cultivo fresco y se incubó a 37°C en agitación hasta fase exponencial temprana, es decir, hasta una absorbancia aproximada de 0.5 a 600 nm. El cultivo ya crecido se incubó en tubos de propileno 10 minutos en hielo y se centrifugó a 3000 rpm durante 10 minutos a una temperatura de 4°C. A continuación se resuspendieron las células en la mitad de volumen de CaCl₂ 100 mM frío y estéril (el cloruro de sodio facilita la permeabilidad de la pared celular bacteriana), la mezcla se incubó en hielo por 30 min y se centrifugaron de nuevo en las mismas condiciones ahora durante 5 min. Finalmente, las células se resuspendieron en un volumen de CaCl₂ que representaba 1/20 del volumen inicial con 15% de glicerol para hacer alícuotas de 100 μ l cada una y poder almacenarlo posteriormente a -70°C. Las células se encuentran viables al menos 6 meses bajo estas condiciones.

V.III.III.I.II. Transformación de células quimiocompetentes de *E. coli* DH5 α

Para la transformación de las células quimiocompetentes de *E. coli* DH5 α , las alícuotas de 100 μ l se mezclaron con 10-100 ng de ADN, en algunos casos plasmídico y en otros casos productos de PCR. Esta mezcla se incubó 30 min en hielo y a continuación se sometieron a un choque térmico a 42°C durante 40 segundos. Se añadió medio de cultivo LB fresco hasta un volumen de 1 ml y se dejó incubar a 37°C durante una hora o dos para permitir la expresión de los marcadores fenotípicos transferidos (mayoritariamente genes de resistencia a antibióticos), que permitieron seleccionar las células que adquirieron el ADN exógeno al sembrar en medio sólido de selección. Después del tiempo de incubación, las células se sembraron en medio LB sólido suplementado con los antibióticos adecuados para la selección de los transformantes y se incubaron a 37°C, al menos 16 horas.

V.III.III.II. Electroporación de células de *U. maydis* FB1 (*MicroPulserElectroporation Apparatus*)

Este método permite la captación de ADN mediante la permeabilización de las membranas provocada por una descarga eléctrica (Dower *et al.*, 1988). Con este método la eficiencia de la transformación aumenta considerablemente, obteniendo entre 10⁹-10¹⁰ transformantes por cada μ g de ADN.

V.III.III.II.I. Obtención de células electrocompetentes de *U. maydis* FB1 (Cregg and Russell, 1998).

Se inocularon 250ml de YEPS (con una alícuota fresca de *Ustilago maydis* FB1 incubada toda la noche) y se incubó durante toda la noche a una temperatura de 30°C en agitación constante (300 rpm) hasta

llegar a una densidad de población de aproximadamente $5-7 \times 10^7$ células/ml. Una vez crecido el cultivo se centrifugó durante 5 minutos a 3000g a una temperatura de 4°C (a partir de este paso todo se realiza a una temperatura cercana a 0°C) para posteriormente decantar el sobrenadante con mucho cuidado. Al paquete celular se le adicionaron 50 ml de YEPS/HEPES 1M (100 ml de YEPS, 20 ml de 1 M HEPES, pH 8.0) estéril adicionando con 1.25 ml de 1M de DTT (1.55 g ditionitrol, disolver en 8 ml de agua y llevar el volumen a 10 ml. esterilizar por filtración) y esto se incubó a 30°C durante 15 minutos. Posteriormente se adicionaron 200 ml de una solución 1M de Sorbitol fría y estéril, se mezcló y se centrifugó a 3000g durante 5 min a 4 °C; con mucho cuidado se descartó el sobrenadante. A la pastilla celular se le adicionó ~50 ml con la misma solución de Sorbitol, se agitó con vortex para resuspenderla y después llevarla a un volumen final de 250 ml de esta misma solución de Sorbitol siempre conservando la temperatura y la esterilidad. Posteriormente con una solución 1M de Sorbitol fría y estéril, se realizaron algunos lavados (con 200 ml, 50 ml llevándolo después a 250 ml y 30 ml) y centrifugaciones sucesivas (a 3000 rpm durante 5 min a 4°C) teniendo mucho cuidado siempre al descartar el sobrenadante. Después se llevó a un volumen final de 250 ml de esta misma solución de Sorbitol. Para finalizar se resuspendió la pastilla celular con 0.5 ml de la solución de Sorbitol 1M (el volumen final debería ser ~1.3 ml y la concentración final debe ser $\sim 1 \times 10^9$ células/ml), con esta suspensión se hicieron alícuotas de 40 μ l que se mantuvieron en hielo y se electroporaron lo más pronto posible después de finalizar este protocolo.

V.III.III.II.II. Electrotransformación de células competentes de *U. maydis* FB1

Para este método se mezclaron 40 μ l de células con hasta 10 ng de ADN y mientras los tubos permanecían en hielo se procedió a mezclar gentilmente la mezcla, a continuación se transfirió esta a cubetas de 0.2 cm estériles y previamente enfriadas. Se ajustó el electroporador (*MicroPulser™ Electroporation Apparatus*, BIO-RAD) en el programa "Pic". La electroporación se realizó bajo las siguientes condiciones: $E = 12.5$ kV/cm y $\tau = 5$ ms (resistencia de 400 Ω y capacitancia de 25 μ F). Se colocó la cubeta en la cámara y se deslizó hasta hacer contacto con el equipo y una vez en la posición correcta se dio el pulso. Inmediatamente después de la descarga eléctrica se removió la cubeta y se adicionó 1 ml de YEPS/1M de Sorbitol frío. La suspensión de células se transfirió con mucho cuidado a un microtubo de 1.5 ml y se incubó durante 1 hora a 30°C con agitación a 225 rpm. Las células transformadas se seleccionaron en medio YEPS/1M de Sorbitol líquido adicionado con una concentración final de 250 μ g/ml de Higromicina B y se incubó a 30°C por 2 hrs sin agitación; por último se hizo crecer la células en placas de YEPS/1M de Sorbitol sólido adicionado con una concentración final de 250 μ g/ml de Higromicina B y se incubaron a 30°C de 72 a 96 horas.

V.III.IV. Técnicas de expresión inhibiendo el NMD

Para observar el NMD en *Ustilago maydis* se ha utilizado un fármaco llamado Emetina. Anteriormente usado como anti-helmíntico, la emetina es capaz de inhibir el NMD sin ser letal para las células (Noensie & Dietz, 2001). En este ensayo el crecimiento de *Ustilago maydis* FB1 se realizó en medio

YEPS y el tiempo total de incubación fueron 12 horas. A partir de un preinóculo de una $DO_{620} = 0.6$ se tomó 1 ml y se inocularon 100 ml de medio de cultivo, donde se hicieron crecer a las levaduras incubando a 30 °C en agitación constante. La Emetina (Dihidrocloruro de emetina hidratado, Sigma) se utilizó en una concentración final de 100 µg/ml y fue adicionado al medio de cultivo 6 hrs antes de la extracción de ARN. Se obtuvieron tres muestras con tratamiento y tres sin tratamiento. Para concluir con el objetivo de la técnica y saber si la expresión de algunos genes en *Ustilago maydis* FB1 era diferencial en presencia y ausencia de Emetina, se midieron algunos transcritos (oligonucleótidos descritos en el apartado de PCR) como Hrp1 y CYH2 por RT-PCR de punto final.

V.III.V. Técnicas experimentales con proteínas

V.III.V.I. Inducción de proteínas recombinantes (pPROEX HT Prokaryotic Expression System)

Para una producción de proteínas a pequeña escala en este sistema, la inducción de proteínas recombinantes se realizó el siguiente procedimiento: se inocularon 2 ml de medio LB con ampicilina a una concentración final de 100 µg/ml (LB Amp100) con una UFC de la cepa transformada y se incubó el cultivo toda la noche a 37 °C con agitación constante. Al día siguiente, con estos 2 ml se inocularon 10 ml de LB Amp100 y se incubó a 37 °C con agitación. Cuando el cultivo alcanzó una DO_{590} entre 0.5, se removió 1 ml del cultivo (muestra sin inducir). Al cultivo restante, se adicionó IPTG a una concentración final de 0.6 mM y se continuó con la incubación a 37 °C. A partir de aquí y de manera constitutiva se tomaron muestras las siguientes 4 horas de la inducción. Todas las muestras se prepararon de la misma forma lavando con 100 ml de PBS 1X una vez.

V.III.V.II. Electroforesis de proteínas (Miniprotean4 BioRad)

Para la preparación de las muestras de proteínas para la electroforesis en gel de poliacrilamida, de manera posterior a la obtención y lavado de la muestra se adicionaron 60 µl de buffer Laemli (0.5 M Tris-HCl pH 6.8, 10 % SDS, 1 % Azul de bromofenol, glicerol, 50 µl de β-mercapto-etanol a un volumen final de 1 ml) y todas las muestras fueron sonicadas a una amplitud de 3 kHz durante 10 seg, repitiendo esta operación 3 veces con intervalos de 10 seg. Finalmente las muestras se calentaron a 94 °C por 10 minutos.

Para verificar la integridad del extracto proteico e identificar la proteína de interés se realizó una electroforesis en gel de poliacrilamida (SDS-PAGE), para mejorar la nitidez de las bandas se utilizan dos geles: el *Resolving gel* y el *Stacking gel*, la concentración del gel varió dependiendo del peso de las proteínas. De manera general para UPF1 y UPF3 se preparó un gel de poliacrilamida al 12 % (para un volumen final de 10 ml se debe agregar 3.4 ml Agua DDI, 4.0 ml de Acrilamida 30 %, 2.5 ml Buffer: 1.5 M Tris-HCl pH 8.8 para el *Resolving gel* y 0.5 M Tris-HCl pH 6.8 para el *Stacking gel*, 0.1 ml de SDS 10 %, 100 µl APS y 5 µl TEMED). Para UPF2 el gel tuvo una concentración del 8 % (para un volumen final de 10 ml se debe agregar 4.7 mL Agua DDI, 2.7 mL de Acrilamida 30 %, 2.5 mL Buffer: 1.5 M Tris-HCl

pH 8.8 para el *Resolving gel* y 0.5 M Tris-HCl pH 6.8 para el *Stacking gel*, 0.1 ml de SDS 10 %, 100 μ l APS y 5 μ l TEMED). Se cargaron 30 μ l de la muestra (aproximadamente 20 μ g de proteína) en los pocillos y como patrón de referencia se utilizó un marcador de proteínas con un rango de 15 - 170 KDa (*Page Ruler*, Thermo Scientific). La electroforesis se corrió en Buffer Tris - Glicina (500 mM Tris, 500 mM Glicina) a 80 mV por un tiempo de 3 horas.

V.III.V.III. Tinción de geles de poliacrilamida

V.III.V.III.I. Tinción con azul de coomasie

Para visualizar las proteínas, éstas se tiñeron con Azul de Coomassie (0.1 % Azul brillante de Coomassie, 45 % metanol, 10 % ácido acético) durante 15 min en agitación, posteriormente se retiró el exceso de colorante y el gel se lavó con agua. Para desteñir el gel se lavó con una solución 10 % metanol, 5 % de ácido acético, dejando en agitación por 18 horas, después de este tiempo el gel se documentó.

V.III.V.III.II. Tinción de proteínas con nitrato de plata

La tinción de proteínas con nitrato de plata es muy útil cuando la concentración que se tiene de éstas llega a ser muy baja ya que el método llega a detectar hasta 10 ng, mejorando la capacidad de detección de la tinción con azul de Coomasie (1 mg).

Para este protocolo se necesitan recipientes de vidrio y que todas las soluciones sean frescas. En este caso primero se fijaron las proteínas al gel de poliacrilamida incubando con la solución de fijado (40% Etanol, 10% Acido acético, 0.018% Formaldehído) durante al menos 2 horas en agitación constante. Lavar 2 veces el gel con etanol al 30% durante 20 minutos. Posterior a esto se trató el gen con la solución de sensibilización (Tiosulfato de Sodio $\text{Na}_2\text{S}_2\text{O}_3$ 0.02%) únicamente 2 minutos y se hicieron 3 lavados consecutivos de 30 segundos cada uno con agua destilada, se decantó el agua. Se adiciona la solución de tinción (AgNO_3 0.02%, Formaldehído 0.028%) y se incubó en agitación constante, se hicieron 3 lavados consecutivos de 15 segundos cada uno con agua destilada y decantar al final el líquido; por último se incubó el gel con la solución de revelado (Carbonato de Sodio 6%, Tiosulfato de Sodio $\text{Na}_2\text{S}_2\text{O}_3$ 0.0004%, Formaldehído 0.028%) en agitación constante el tiempo necesario hasta encontrar el contraste (en nuestro caso con 15 minutos fue suficiente). Por último se detiene la reacción con una solución de 1% de Ácido acético incubando 5 minutos, se lavó el gel tres ocasiones con agua destilada para quitar residuos y se documentó.

V.III.V.IV. Western-Blot

Para el análisis tipo western-blot se prepararon extractos proteicos totales (tanto de *Ustilago maydis* FB1 como de *E. coli* DH5 α en caso de ser necesario) usando aproximadamente 20 μ g (la concentración

de proteínas se cuantificó con el Reactivo de Bradford-BioRad), los cuales se corrieron en un gel al 8% o al 12% (según fuera necesario) de SDS-PAGE. Las proteínas se transfirieron a una membrana de PVDF (Millipore) bloqueando con Casec (Nestle) al 5% durante 2 horas en agitación constante y se incubó con el anticuerpo primario (normalmente a una dilución de 1:750) durante toda la noche para posteriormente al siguiente día lavar con PBS-Tween 0.05% e incubando nuevamente en agitación constante con el anticuerpo secundario (normalmente en una dilución 1:2000). Posterior a esto y antes de revelar la membrana se lavó muy bien 3 veces con la solución de PBS-Tween 0.05%. Los anticuerpos secundarios están acoplados a peroxidasa, por lo que se detectó con luminol (*Santa cruz Biotechnology*) y la placa radiográfica se reveló empleando las soluciones de revelado Kodak.

La lista de los anticuerpos utilizados se detalla en la tabla siguiente:

Tabla 6. Anticuerpos empleados para los ensayos de WB

NOMBRE DEL PRODUCTO	NÚMERO DE CATÁLOGO:	ISOTIPO	EPÍTOPO	APLICACIONES	MARCA COMERCIAL
Rent1 (H-300)	sc-48802	conejo IgG	71-370 (h)	WB, IP, IF, ChIP, Gel shift	Santa Cruz Biotechnology
Rent2 (H-300)	sc-48801	conejo IgG	121-420 (h)	WB, IP, IF, ChIP, Gel shift	Santa Cruz Biotechnology
Rent3 (H-190)	sc-48800	conejo IgG	51-240 (h)	WB, IP, IF, ChIP, Gel shift	Santa Cruz Biotechnology
His-probe (H-3)	sc-8036	raton IgG	Histidinas	WB, IP, ELISA	Santa Cruz Biotechnology
goat anti-rabbit IgG-HRP	sc-2030	-	-	-	Santa Cruz Biotechnology
goat anti-rabbit IgG-HRP	sc-2005	-	-	-	Santa Cruz Biotechnology

V.3.5.3. Ensayos de tipo pull-down para la interacción de proteínas recombinantes (*ProFound Pull-Down PolyHis Protein:Protein Interacción, Thermo Scientific*)

Elucidar la función de un gen involucra la determinación de su función, para esto se debe analizar el producto proteico. En la célula, las proteínas participan en muchos procesos que requieren la interacción entre diversas proteínas, las cuales pueden guiarnos para entender la estructura de la proteína y su función, identificando así su relevancia biológica. Para este protocolo se siguieron las especificaciones generales que incluyen los siguientes pasos esquematizados en la figura 13 y que se detallan más adelante.

Preparación de la proteína anzuelo. Tomando 5 ml de cultivo inducido (apartado V.3.3.1.) de UPF1 se centrifugó a 5000g durante 5 minutos y se decantó el sobrenadante, se resuspendió el *pellet* en 1 ml de TBS y se transfirió a un microtubo de 1.5 ml. Nuevamente se centrifugó a 5000g durante 5 minutos y se decantó el sobrenadante; se resuspendió el *pellet* en 200 µl de TBS frío, se añadieron 200 µl de *ProFound Lysis Buffer* y se sonicó la muestra. Posteriormente se incubó durante 30 minutos invirtiendo los tubos periódicamente (durante esta incubación se procedió a comenzar con el paso dos), por último se centrifugó a 5000g durante 5 minutos y se transfirió el sobrenadante a un microtubo nuevo de 1.5 ml y se conservó en frío (esta muestra es el anzuelo). Para la *Inmovilización de la proteína anzuelo*

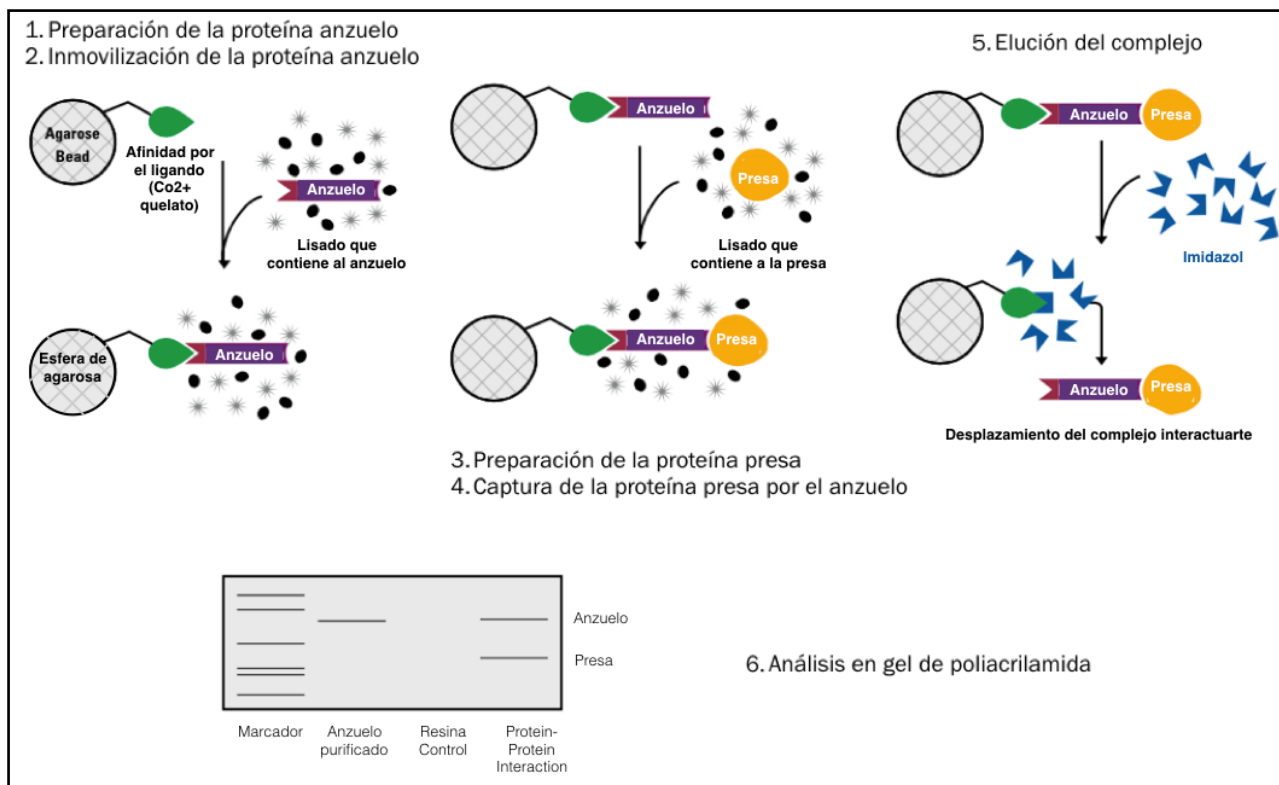


Figura 13. Esquema del procedimiento empleado para los ensayos tipo pull-down. Inicialmente se prepara la proteína que se empleará como anzuelo y se inmoviliza sobre esferas de agarosa. Por otro lado, se prepara la proteína que fungirá como presa y se permite su interacción con el anzuelo. El complejo formado luego de esta interacción se desplaza al agregar un exceso de imidazol. Finalmente, el resultado de la interacción se analiza mediante una electroforesis desnaturizante en gel de poliacrilamida, el cual incluye los controles correspondientes.

primero se tiene que equilibrar la resina de cobalto, para esto se hicieron 5 lavados en las columnas con la solución *TBS:Profoun Lysis Buffer* a la cual se le adicionó Imidazol a una concentración final de 40 mM empleando la Stock Solution y después se adicionó la muestra anzuelo a la columna y se incubó durante 2 horas, en frío y en agitación constante. Para la *Preparación de la proteína presa*, por otro lado se tomaron 5 ml de cultivo inducido (apartado V.3.3.1.) de UPF2 y se trató de igual forma que la proteína UPF1 al inicio del protocolo con la diferencia de que a esta muestra se le adiciona una concentración final de Imidazol de 40 mM (esta muestra es la presa). *Captura de la proteína presa por el anzuelo.* Se adiciona a la columna, en donde previamente se inmovilizó el anzuelo, 800 µl de la proteína presa y se incubó a 4°C durante 2 horas con agitación constante pero suave, después de la incubación se centrifugó la columna a 1250 g durante 30 segundos y se lavó la columna 5 veces añadiendo 400 µl de *wash solution* y centrifugando a 1250 g por 30 segundos. *Elución del complejo.* Lavar la columna una vez más con 1 ml de Solución de Elución (290 mM Imidazol en solución de lavado) se incubó 5 minutos en agitación y se centrifugó a 1250g por 1 minuto, esta es la muestra de la interacción. Posterior a esto se hace la electroforesis en gel de poliacrilamida como se indica en el apartado V.3.5.2., incluyendo las muestras correspondientes para analizar los resultados.

VI. RESULTADOS

VI.I. ORGANIZACIÓN ESTRUCTURAL Y FUNCIONAL DEL HOMÓLOGO DE UPF1 EN *U. maydis*.

Inicialmente se realizó el análisis de la estructura primaria del homólogo putativo para el factor de NMD UPF1 identificado en *U. maydis* (umUPF1), el cual corresponde al locus UM00237.1 y para el cual se identificaron los dominios estructurales y funcionales que se han reportado para UPF1 en otras especies (Fig. 1). Mediante este análisis, se identificaron en umUPF1 los mismos dominios reportados previamente para el homólogo en *H. sapiens* (hUPF1). En la Figura 1A se muestra el esquema para hUPF1 y umUPF1 resaltando los principales dominios funcionales que incluyen el dominio CH, RecA1, RecA2, 1B y 1C (Cheng *et al.*, 2007; Kadlec *et al.*, 2006; Chakrabarti *et al.*, 2001; Yamashita, 2013). Las principales características para cada dominio se resumen en la Fig. 1B, en donde se puede observar que el tamaño y la posición relativa de los diferentes dominios se encuentra altamente conservada para ambas proteínas.

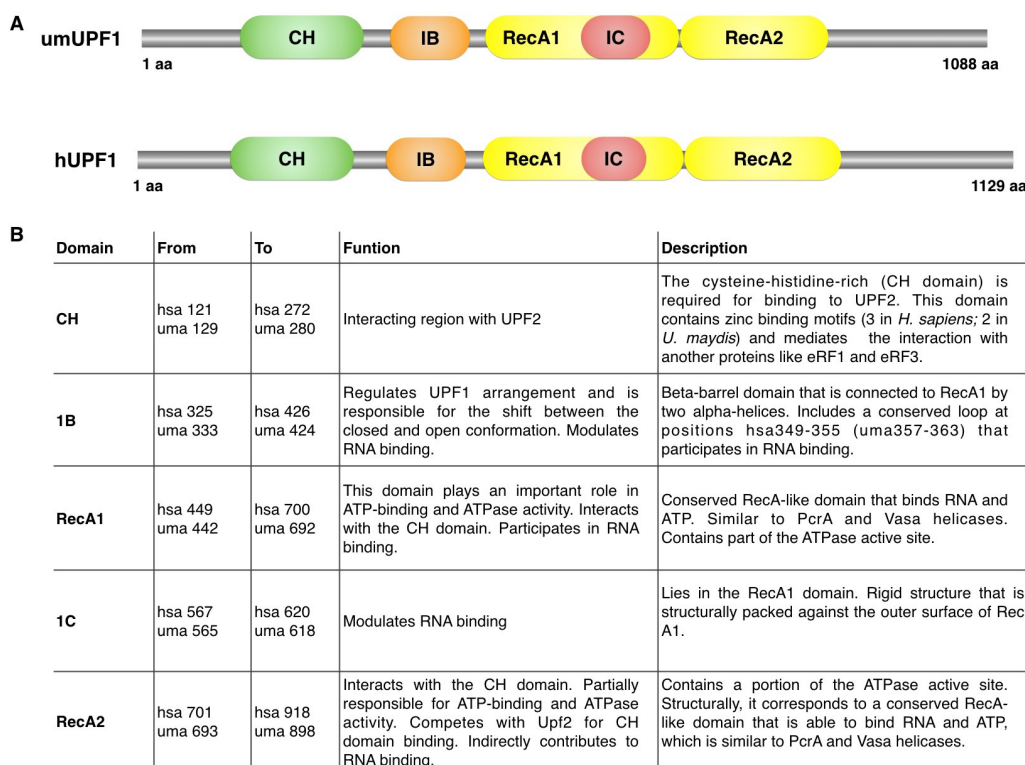


Figura 14. La organización de UPF1 es muy similar en *Ustilago maydis* y en *Homo sapiens*. A) Representación esquemática de los dominios presentes en UPF1 de *Ustilago maydis* (umUPF1) y de *Homo sapiens* (hUPF1). El dominio CH (verde) es responsable de la interacción con UPF2, eRF1 y eRF3. La región de helicasa incluye dos dominios RecA (amarillo). Otros regiones reguladoras incluyen el dominio 1B (naranja) y el dominio 1C (rojo). Se indica el aminoácido correspondiente al inicio y al final de cada péptido. B) Posición relativa para cada dominio en *H. sapiens* y en *U. maydis*, junto con el resumen de las principales características reportadas para el factor en humano. Las posiciones mostradas para la proteína de humano corresponden a la isoforma 1 (Q92900-1).

Para analizar con mayor detalle las principales características de la proteína putativa umUPF1, se realizó el alineamiento de las secuencias aminoacídicas para UPF1 de *H. sapiens* y *U. maydis* (Fig. 2) con lo que se encontró una identidad del 66% entre los dos organismos. Además de los dominios estructurales ya identificados, al realizar este alineamiento también resaltaron varios motivos que han

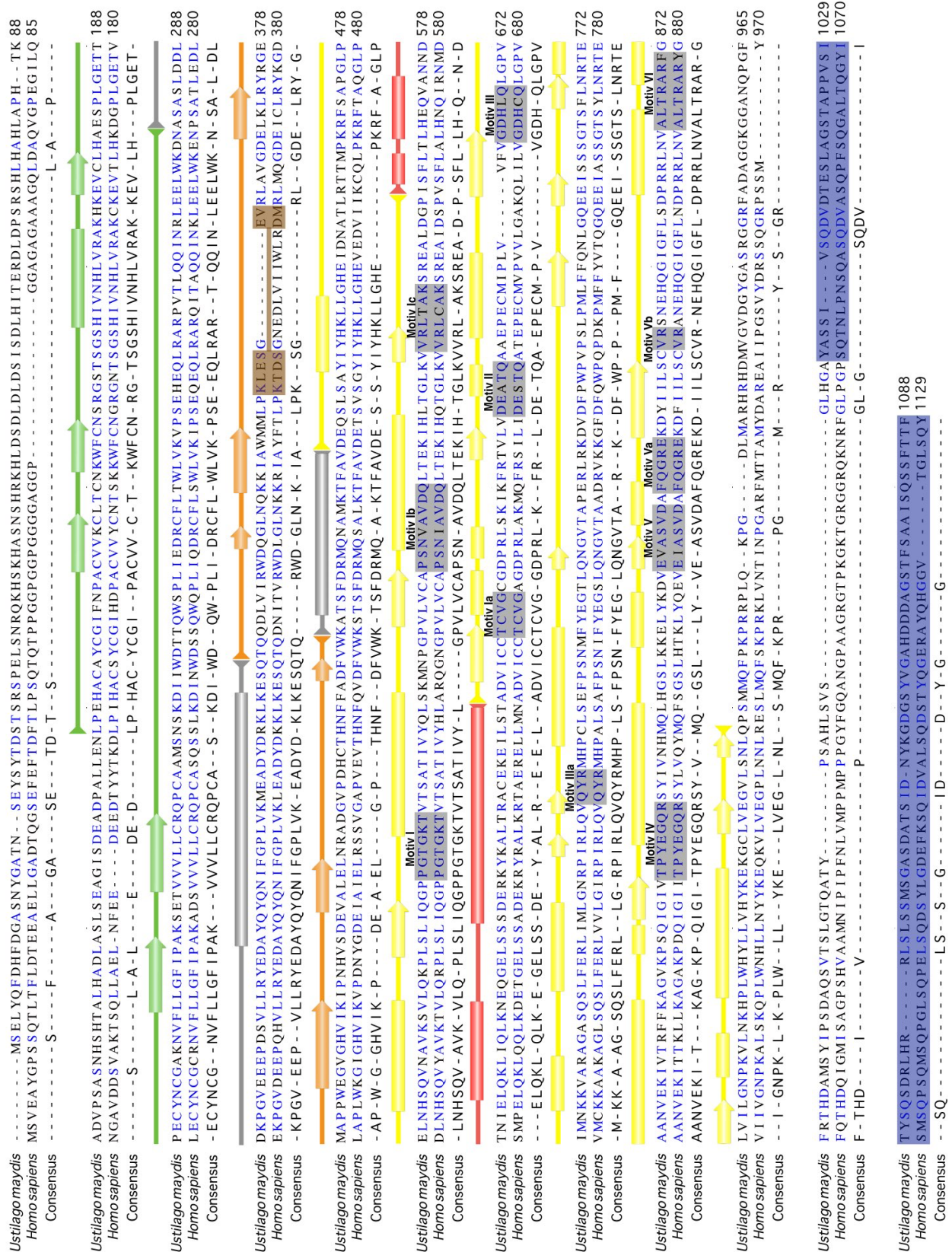


Figura 15. La secuencia para UPF1 se encuentra conservada en *H. sapiens* y *U. maydis*. Se muestra el alineamiento de la secuencia aminoacídica completa para umUPF1 y hUPF1 y los residuos conservados se muestran en azul. Cada dominio se indica en la parte superior de la secuencia empleando el mismo código de colores de la figura 1. También se incluyen los elementos de estructura secundaria en donde los rectángulos representan hélices y las flechas ilustran láminas. Los dominios de helicasa conservados (I, II, III, IV, V y VI) se muestran como cajas grises. El asa 349-355, la cual se encuentra interrumpida en la isoforma 1 debido a una secuencia intrónica, se resalta en café. El motivo SQ corresponde a la caja azul oscura.

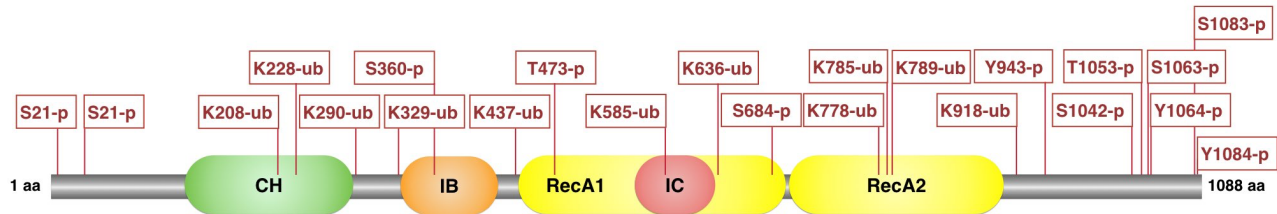
sido reportados para UPF1 (Cheng *et al.*, 2007), incluyendo el asa 349-355, los motivos helicasa I-VI conservados en las helicasas SF1/2, así como el motivo rico en Gln/Ser (QS) presente en el dominio C-terminal. En el alineamiento se puede observar que los aminoácidos que se localizan en los dominios estructurales y funcionales se encuentran conservados entre los dos organismos, siendo los más conservados el dominio CH y los dos dominios RecA2. Para el caso del dominio CH, se determinó una homología del 75% entre las secuencias de *U. maydis* y *H. sapiens*, sugiriendo la relevancia funcional de esta región para la proteína umUPF1 identificada. Por otro lado, se identificaron dos dominios del tipo RecA en umUPF1, los cuales son similares en posición y secuencia a los que se han descrito para *H. sapiens* (Fig. 2) y muestran una identidad del 37% y 71% respectivamente entre los dos organismos. En este análisis también se identificaron a los dominios 1B y 1C (Fig. 1), los cuales mostraron 50% y 58% de identidad entre los dos organismos comparados (Fig. 2). Estos dominios son exclusivos de UPF1 y participan en las actividades catalíticas de este factor. El dominio 1B de umUPF1 está conformado por seis láminas beta, mientras que el dominio 1C presenta cuatro hélices alfa. Esta organización espacial concuerda con la estructura reportada para hUPF1 (Cheng *et al.*, 2007). umUPF1 muestra los motivos característicos de las helicasas pertenecientes a la familia SF1/2 (Applequist *et al.*, 1997) los cuales se localizan en los dos dominios RecA (Fig. 2). Adicionalmente se identificó en umUPF1 la secuencia correspondiente al asa 349-355 (así denominada de acuerdo a la posición de esta secuencia en hUPF1; Cheng, 2007). Aún cuando la secuencia primaria de esta región no se encuentra completamente conservada, el arreglo estructural propuesto para esta porción exhibe también un arreglo en forma de asa, la cual posiblemente pueda realizar la misma función en *U. maydis*.

Finalmente, se identificaron tres sitios SQ conservados dentro del motivo rico en residuos de Ser/Gln en umUPF1. En *H. sapiens*, el extremo C-terminal de UPF1 contiene varios repetidos SQ y varios residuos de serina son fosforilados por la cinasa SMG1 (Kojic *et al.*, 2002). Una vez fosforilados, los extremos N- y C-terminal de UPF1 reclutan a SMG6 y SMG5/7, respectivamente (Okada-Katsuhata *et al.*, 2012). Porciones del N- y del C-terminal de UPF1 se encuentran conservadas en metazoarios y plantas pero no están en *S. cerevisiae*. En este estudio se identificaron diversos residuos de serina conservados en umUPF1, además de repetidos SQ en los extremos de la proteína, los cuales podrían ser funcionales en esta proteína putativa.

VI.I.I MODIFICACIONES POSTRADUCCIONALES EN umUPF1

Para identificar algunos sitios putativos que pudieran ser blancos de modificaciones postraduccionales en umUPF1, se empleó la herramienta PhosphositePlus (Hornbeck *et al.*, 2014), con la cual se encontraron algunos blancos posibles de fosforilación y ubiquitinación en hUPF1, incluyendo algunos sitios que han sido validados de manera experimental. De los datos obtenidos, se seleccionaron solamente aquellos sitios que se encuentran conservados tanto en *U. maydis* como en *H. sapiens*. De esta manera se evidenciaron once sitios putativos de ubiquitinación en umUPF1, los cuales se encuentran conservados en *H. sapiens* (Fig. 3). Se ha reportado que el dominio CH de UPF1 en *S.*

cerevisiae es sustrato para auto-ubiquitinación luego de la asociación de dicho factor con UPF3, por lo que se ha sugerido que el proceso de ubiquitinación podría ser relevante para la regulación de NMD (Takahashi *et al.*, 2008). A pesar de que este es el único trabajo en el que se reporta la ubiquitinación de UPF1, el número de blancos putativos de ubiquitinación conservados para UPF1 encontrados en este trabajo permiten suponer que esta modificación podría jugar un papel importante en la función de UPF1 y en la regulación de la vía de NMD.



HTP	aa	<i>Homo sapiens</i>	aa	<i>Ustilago maydis</i>
1	S10-p*	VEAYGFPSSQTLTFLD	S2-p	_____MSELYQFDH
1	S31-p	LGADTQGSEFEFTDF	S21-p	SNYGATNSEYSYTD
8	K200-ub	LLGFIPAKADSVVVL	K208-ub	LLGFIPAKSETVVVL
3	K220-ub	CASQSSSLKIDINWDSS	K228-ub	CASQSSSLKIDINWDSS
2	K282-ub	ATLEDLEKPGVDEEP	K290-ub	ASLDDLDKPGVEEPP
1	K321-ub	ADYDKKLEKESQTQDN	K329-ub	ADYDRKLEKESQTQDD
1	S352-p	FTLPKTDSGNEDLVI	S360-p	MMLPKLES
43	K439-ub	DRMQSALKTFVADET	K437-ub	DRMQNAMKTFVADEQ
1	T475-p	CQLPKRFTAQGLPDL	S473-p°	TTPMKRFSAPGLPEL
3	K587-ub	DSMPLELQKQLKDE	K585-ub	DTNIELQKLIQLKNE
14	K638-ub	AGDPRLAKMQFRSIL	K636-ub	CGDPRLSKIKFRTVL
1	S692-p*	KAAGAGLSQSLFERL	S684-p	KVARAGASQSLFERL
1	K786-ub	TEAANVEKITTKLLK	K778-ub	TEAANVEKIVTRFFK
1	K793-ub	KITTKLLKAGAKPDQ	K785-ub	KIVTRFFKAGVKPSQ
1	K797-ub	KLLKAGAKPDQIGII	K789-ub	RFFKAGVKPSQIGIV
2	K926-ub	ESLMQFSKPRKLVNT	K918-ub	PSMMQFPKPRRLQ-
81	Y958-p	AIIPGSVYDRSSQGR	Y943-p	DMVGVGYGASRGGR
0	S1089-p*	GLSQPELSQDSYLG	S1042-p	DRLHRRLSLSMSGGA
2	S1100-p	YLGDEFKSIDVALS	T1053-p°	MSGASDATSID_NYK
9	T1111-p	VALSQDSTYQGERAY	S1063-p°	_NYKGDGSYVGAHDD
20	Y1112-p	ALSQDSTYQGERAYQ	Y1064-p	NYKGDGSYVGAHDDD
1	T1124-p	AYQHGGVTGLSQY_	S1083-p°	FSAAISQSSFTTF_
65	S1127-p*	HGGVTGLSQY_	T1086-p°	AISQSSFTTF_

Figura 16. Modificaciones postraduccionales predichas para umUPF1. La posición relativa para los diferentes aminoácidos que podrían sufrir modificaciones en umUPF1 se muestra en la parte superior. En la parte inferior se incluye una tabla en donde se presenta la secuencia que alberga un sitio de fosforilación (p) o de ubiquitinación (ub) en *H. sapiens* y en *U. maydis*. (*) indica que la modificación se ha validado experimentalmente. (°) se utiliza para ilustrar las posiciones ocupadas por una treonina en *H. sapiens* mientras que la misma posición corresponde a una serina en *U. maydis*. El valor HTP corresponde al número de registros a los cuales se asignó este sitio empleando solamente estudios de proteómica acoplados a espectrometría de masas.

UPF1 es una fosfoproteína y se ha establecido que diversos ciclos de fosforilación y defosforilación regulan la vía de NMD en *C. elegans*, humano y plantas (Page *et al.*, 1999; Yamashita *et al.*, 2001; Kérenyi *et al.*, 2013). Por el contrario, pocos sitios de fosforilación han sido demostrados experimentalmente para UPF1 y mientras la fosforilación juega un papel relevante en animales, esta modificación no es determinante para NMD en *S. cerevisiae*. Los sitios de fosforilación conservados en umUPF1 corresponden principalmente a residuos de serina y ocupan las posiciones umS2, umS21, umS360, umS684, umY943, umS1042 y umY1064. Otros blancos de fosforilación identificados incluyen a umS473, umT1053, umS1063, umS1083 y umT1086. La mayoría de los sitios putativos identificados son nuevos y no han sido reportados anteriormente como blancos de fosforilación para UPF1. La participación de estas secuencias en la fosforilación de UPF1 podría ser evaluada en estudios futuros para *H. sapiens* y *U. maydis*.

Los sitios equivalentes a las posiciones umS1042 y umS1081 son funcionales en hUPF1 (Yamashita et al., 2001; Kashima et al., 2006). Ambos residuos se localizan en el extremo C-terminal dentro del dominio SQ, el cual está involucrado en la regulación de la actividad biológica de UPF1, incluyendo su capacidad para unirse a UPF2, su actividad ATPasa y el reclutamiento de otros factores de NMD (Yamashita, 2013; Kashima et al., 2006; Yamashita et al., 2005; Franks et al., 2010). En hUPF1 existen trece repetidos SQ, de los cuales solo los dos anteriores han mostrado actividad en estudios *in vitro* (Yamashita et al., 2001; Kashima et al., 2006). De manera relevante, estos dos sitios se encuentran conservados en umUPF1. Debido a su posición en la proteína fúngica y considerando los reportes previos, resulta posible proponer que umS1042 y umT1086 pudieran ser sitios de fosforilación activos en *U. maydis*, los cuales pudieran ser necesarios para la funcionalidad de NMD en el basidiomiceto.

Otro sitio activo de fosforilación reportado para hUPF1 es la posición T28, la cual es responsable del reclutamiento de SMG6 hacia una región conservada del N-terminal de hUPF1 cuando ésta se encuentra fosforilada (Okada-Katsuhata et al., 2012). Esta posición también se encuentra fosforilada en *A. thaliana* y es relevante para NMD en este organismo (Kerényi et al., 2013). Este blanco de fosforilación fue identificado por la herramienta PhosphositePlus, pero al alinear las secuencias de *H. sapiens* y *U. maydis* este sitio no siempre se conservaba. Este y otros residuos no han mostrado un efecto drástico sobre la fosforilación cuando han sido mutados *in vitro* (Kashima et al., 2006).

VI.I.II. RELACIÓN FILOGENÉTICA DE UPF1

El factor de NMD UPF1 se encuentra altamente conservado en eucariotes y se ha reportado que el porcentaje de identidad para los homólogos de UPF1 en *H. sapiens*, *A. thaliana*, *D. melanogaster*, *C. elegans* y *S. cerevisiae* oscila entre el 40 y el 62%, el cual es un porcentaje alto si se compara con el 59-67% reportado para proteínas ribosomales (Culbertson & Leeds, 2003), indicando la relevancia de este factor en la regulación postranscripcional de la expresión genética. Al comparar la secuencia aminoacídica de umUPF1 con el homólogo en *H. sapiens*, encontramos una identidad del 66% (Fig. 2), el cual es un valor más alto en relación con otros homólogos reportados para UPF1 (Culbertson & Leeds, 2003), lo que sugiere que umUPF1 podría realizar en *U. maydis* la importante función reguladora de NMD demostrada para esta proteína en *H. sapiens*.

Para analizar la relación evolutiva de umUPF1 con otras especies se construyó un árbol filogenético para 32 especies (Fig. 4, izquierda), empleando como organismo no relacionado a *D. purpureum*. Por otro lado se determinaron los dominios predichos por la base de datos InterPro para cada uno de los homólogos incluidos en el análisis (Fig 4, derecha). Como se esperaba, umUPF1 se relaciona filogenéticamente con otros hongos; sin embargo, muestra un tamaño y organización estructural similar a la que se encuentra en animales. En total encontramos 5 dominios predichos (IPR018999, IPR003593, IPR027417, IPR014001, IPR013083) que muestran diferentes arreglos en las 32 especies analizadas. La principal organización identificada incluye el dominio de interacción con UPF2

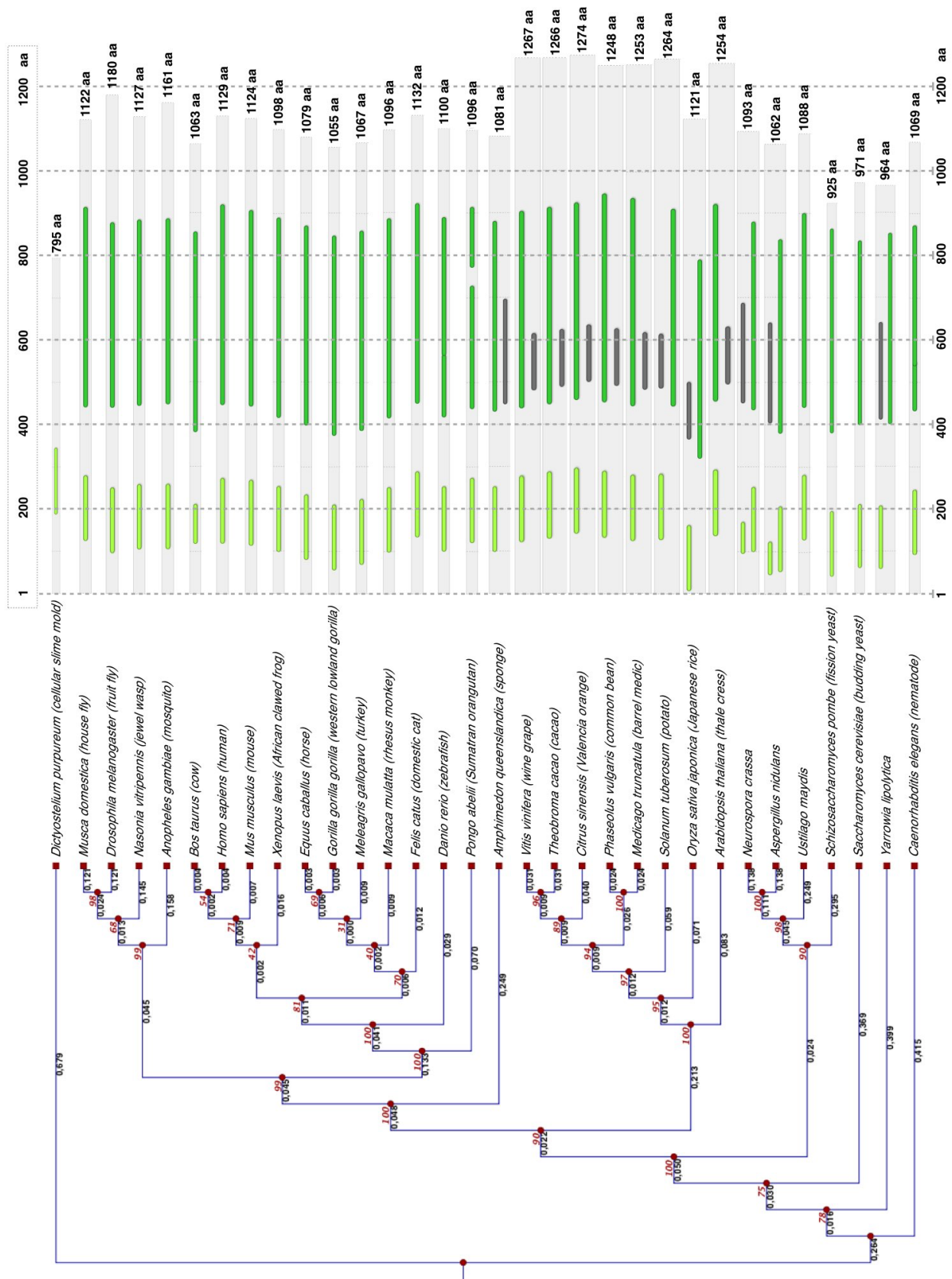


Figura 17. Relación evolutiva y organización de dominios para UPF1 en diferentes especies. El análisis involucró a 32 secuencias aminoacídicas de animales, plantas, hongos y protistas. A la izquierda se muestra la relación evolutiva de UPF1 inferida a partir del método UPGMA. El árbol se encuentra dibujado a escala y el análisis evolutivo se realizó empleando el CLC sequence viewer. Las arquitecturas para los dominios y el tamaño de la proteína de acuerdo con la base de datos InterPro se muestran a la derecha para cada especie. Se encontraron cinco diferentes dominios predichos, los cuales mostraron diferentes arquitecturas en las 32 especies. La principal organización para los diferentes dominios incluye al dominio de interacción con UPF2 en el extremo N-terminal (IPR018999, verde claro) mientras que el dominio de P-loop NTPasa (IPR027417) se identificó hacia el C-terminal para todas las especies a excepción de *D. purpureum*.

(IPR018999) localizado hacia el N-terminal y el dominio de P-loop NTP hidrolasa (IPR027417) hacia el C-terminal. Este dominio se caracteriza por dos secuencias conservadas que corresponden a los motivos “Walker A” y “Walker B”, los cuales son responsables de la unión con NTP y Mg²⁺, respectivamente y que se encuentran en varias proteínas, incluyendo las helicasas SF1/2. Este dominio fue identificado en todas las especies incluidas en el análisis comparativo, a excepción de *D. purpureum*. Sin embargo, el dominio IPR027417 fue identificado en otras especies de *Dictyostelium*. Por otro lado, existen algunas diferencias en cuanto a los arreglos encontrados empleando InterPro, en donde las secuencias aminoacídicas para *N. crassa* y *A. nidulans* presentan un dominio de dedo de zinc (tipo RING/FYVE/PHD; IPR013083) que se sobrelapa con el dominio IPR018999. El dominio de unión a ATP característico de las helicasas SF1/2 (IPR014001) se detectó en plantas, tres especies fúngicas y una animal. Finalmente, en casi todas las especies de hongos se encontró al dominio ATPasa AAA+ (IPR013593).

En relación con el tamaño de los homólogos comparados, se observaron algunas variaciones dependiendo del grupo de organismo en cuestión. Las proteínas de plantas son ligeramente mayores con un tamaño promedio de 1260 aa, a excepción de *Oryza sativa japonica* (1121 aa). Las proteínas animales oscilan entre 1032 y 1180aa, mientras que en hongos las proteínas son un poco más pequeñas (925 a 1093 aa).

VI.I.III. INTERACCIONES PROTEICAS DE *umUPF1*

UPF1 posee la habilidad de interactuar con una gran variedad de proteínas, algunas de las cuales contribuyen a su funcionalidad en NMD. Los componentes clave de la maquinaria de NMD y algunos factores relacionados han sido estudiados en humano y hemos identificado los ortólogos putativos para dichos factores en *U. maydis*, de los cuales 52% fueron considerados como posibles proteínas asociadas a UPF1 dentro de las 20 mejores interacciones empleando la herramienta STRING (Tabla 1). Estos factores putativos de NMD identificados en *U. maydis* incluyen el núcleo del NMD, SMG1, componentes del EJC, proteínas de unión al Cap y a la maquinaria de poliadenilación, factores de liberación y otras proteínas que participan en NMD o que interactúan de manera transitoria con alguno de los componentes de esta maquinaria.

UPF1 es la proteína central del proceso de NMD. Para analizar *in silico* las posibles interacciones proteicas que pudieran involucrar a *umUPF1* se empleó la herramienta STRING. Esta herramienta comprende una base de datos de interacciones proteicas tanto conocidas como predichas, las cuales incluyen asociaciones físicas y funcionales derivadas de cuatro fuentes principales: contexto genómico, validación experimental, co-expresión y reportes existentes. Con toda esta información se puede generar un puntaje particular para cada asociación proteica (Szklarczyk et al., 2015), mientras que la astringencia de los parámetros y el número de interacciones puede determinarse para cada análisis. En la figura 5 se muestran únicamente las diez interacciones para *hUPF1* y *umUPF1* que tuvieron el puntaje más alto, así como la nota obtenida para cada interacción. Todas las interacciones mostradas

Tabla 1. Factores de NMD putativos en *Ustilago maydis*.

	Factor	human protein size (aa)	<i>Ustilago maydis</i> ID	Over-lap (aa)	% Identity	STRING position
Núcleo del NMD	UPF1	1118	UM00237.1	878	66.3	input
	UPF2	1272	UM03549.1	1209	30.2	1
	UPF3 (UPFB)	470	UM00233.1	283	41	3
SMG's	SMG1	3661	UM03216.1	648	28.2	5
EJC (Exon-junction complex)	Y14	187	UM04564.1	174	40.6	9
	MAGOH	146	UM05829.1	144	66.7	19
	eIF4A3	411	UM06129.1	388	78.6	10
	RNPS1	305	UM03178.1	165	29.7	33
	REF/ALY	264	UM01000.1	263	32.3	117
Factores importantes de interacción con el EJC	Tap	626	UM01688.1	552	24.3	76
	p15	140	UM05497.1	100	33	-
	UAP56	428	UM04940.1	418	64.4	143
	PYM	203	UM04799.1	195	31.3	-
	SRm160	904	UM01915.1	119	64	132
CBC (Cap binding complex)	CBP20	103	UM04861.1	57	70.2	14
	CBP80	790	UM06211.1	790	23.6	7
Proteínas de poliadenilación	PABP1	660	UM03494.1	659	49.5	15
	PABP2	306	UM02401.1	186	50.5	38
Factores de liberación	eRF1	437	UM04192.1	440	71.6	4
	eRF3	628	UM05695.1	626	45.2	2
Otros factores importantes	Hbs1	642	UM04216.1	471	45.4	21
	Musashi	458	UM02420.1	425	30.4	29
	PP2A	309	UM03957.1	300	86.7	100

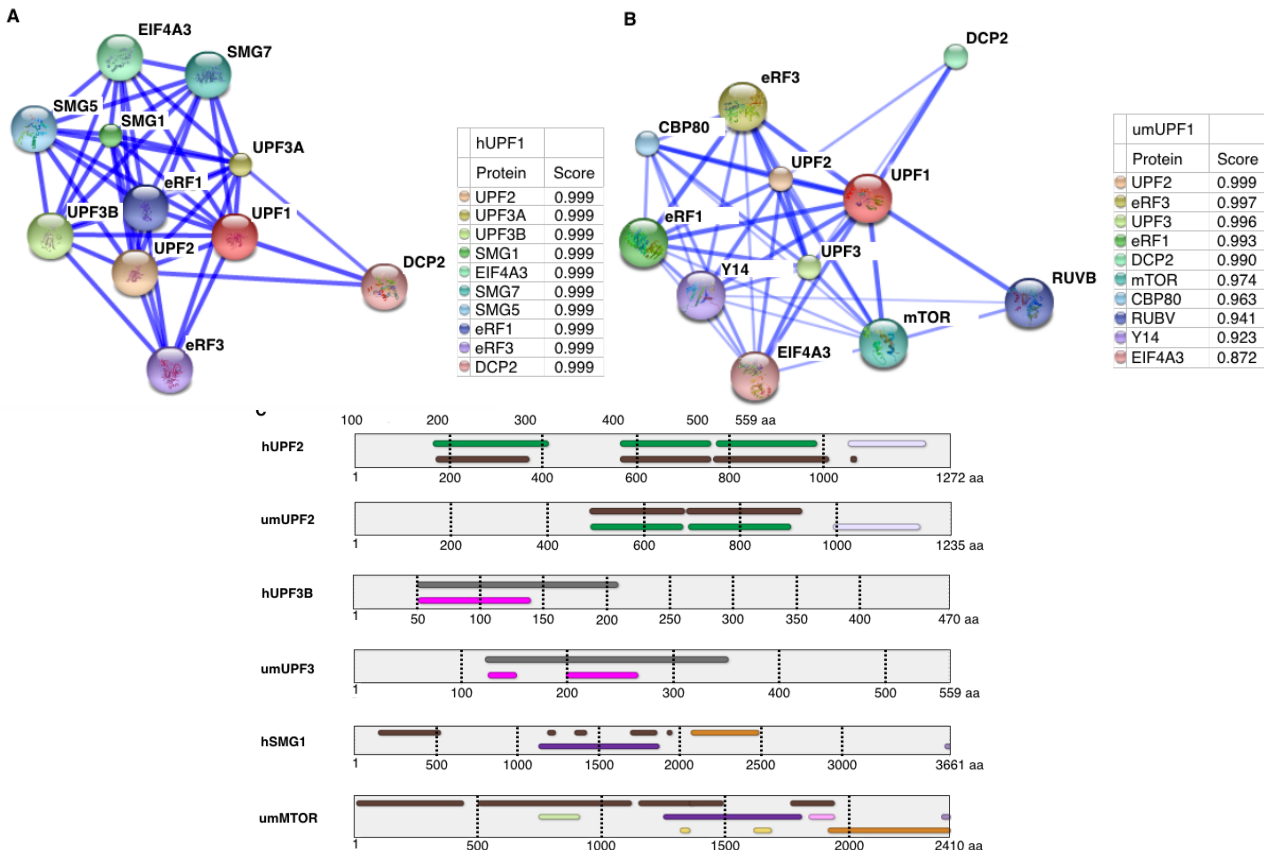


Figura 18. Interacción de UPF1 con diferentes proteínas. A) Interacciones proteicas conocidas para hUPF1 ilustradas empleando STRING. B) Interacciones predichas para umUPF1. En ambas representaciones, las líneas azules se emplean para unir a las dos proteínas involucradas en la interacción. Si la línea es sólida indica que la interacción es confiable. El puntaje determinado empleando STRING se muestra para cada interacción. C) Análisis de los dominios presentes en UPF2, UPF3 y SMG1 (mTOR) en *U. maydis* y *H. sapiens*. Se indica la arquitectura identificada por InterPro y el tamaño para cada proteína. En UPF2 se indican en verde y café los dominios tipo MIF4G y tipo 3 (IPRO16021), respectivamente; mientras que en lila se muestra el dominio supresor Upf2 (IPRO07193). Para UPF3 el pliegue alfa-beta de union a nucleótido (IPRO12677) se muestra en rosa mientras que el dominio Upf3 regulador de NMD está en gris. Para SMG1 y mTOR los dominios relevantes son el PI3/4 cinasa, el dominio catalítico en naranja, un dominio tipo Armadillo en café (IPRO16024), el dominio relacionado a PIK (IPRO14009) en púrpura y en lila el dominio FATC (IPRO03152).

para hUPF1 obtuvieron un puntaje de 0.999 e incluyen a las siguientes proteínas: UPF2, UPF3A, UPF3B, SMG1, EIF4A3, SMG7, SMG5, eRF1, eRF3 y DCP2 (Fig. 5A). Para umUPF1 las interacciones predichas fueron para los homólogos en *U. maydis* de las siguientes proteínas: UPF2, eRF3, UPF3, eRF1, DCP2, mTOR, CBP80, RUBV, Y14 and EIF4A3, presentando un puntaje entre 0.999 para UPF2 y 0.872 para EIF4A3. Todas las proteínas cuya interacción se predice con umUPF1 corresponden a proteínas involucradas con el proceso de NMD.

Tanto SMG1 como mTOR presentan actividad de cinasa. El análisis de ambas proteínas con InterPro mostró que los dos factores comparten un dominio cinasa conservado, el cual es característico de la familia PIKK. La posible interacción de umUPF1 con mTOR se encuentra sustentada en STRING por la interacción de UPF1 con SMG1 en *H. sapiens*. Estas observaciones podrían indicar que el factor putativo mTOR identificado en *U. maydis* podría ser el homólogo funcional de la proteína SMG1 de *H. sapiens*. Un análisis semejante mostró que umUPF2 y umUPF3 presentan los dominios funcionales conservados entre *H. sapiens* y *U. maydis* (Fig. 5C).

VI.I.IV. EL INTERRUPTOR PARA LA ACTIVIDAD HELICASA DE umUPF1

La estructura dinámica de UPF1 es importante para el proceso de NMD. En la conformación “cerrada” de la proteína, el ATP se une a la cavidad que se forma entre los dos dominios tipo RecA. En esta configuración, los dominios RecA se acercan, mientras que en presencia de ATP ocurre un ligero arreglo espacial en la dirección opuesta. También en la conformación “cerrada”, UPF1 contacta a la molécula de ARN de cadena sencilla (en orientación 5'->3') a través del canal que se forma con las proyecciones de los dominios RecA, el conector y los dominios 1B y 1C. En esta conformación pero en ausencia del dominio CH, el dominio 1B rota alejándose del extremo 3' del ARN, liberando a la molécula de ARN y favoreciendo las actividades ATPasa y helicasa de UPF1 (Chamieh *et al.*, 2008; Chakrabarti *et al.*, 2011). Por otro lado, la conformación “abierta” corresponde a un rearrreglo mayor de la proteína en donde el dominio CH (debido a su interacción con UPF2) cambia de posición adoptando un arreglo espacial más relajado. Esta actividad de helicasa es regulada por la unión con UPF2. Dicha interacción permite que UPF1 pase de una conformación cerrada en la que la actividad helicasa se encuentra inactiva a un estado activo de conformación abierta (Gleghorn & Maquat, 2011).

Empleando como templado el cristal 2XZL generamos un modelo para umUPF1 que nos permitiera analizar la estructura cerrada para umUPF1. En la figura 6 se observa que el modelo de umUPF1 exhibe la región hidrofóbica que estabiliza la conformación cerrada debido a las interacciones entre los residuos umV169, umF200 y umI241 del dominio CH (verde), los cuales interactúan con umI760 en el dominio RecA2 (amarillo). El residuo umF200 parece ser central en el mantenimiento de las interacciones hidrofóbicas, ya que se encuentra altamente conservado en diversas especies, además de que su función estructural ha sido evidenciada para los cristales de *H. sapiens* y *S. cerevisiae* (Chakrabarti *et al.*, 2011).

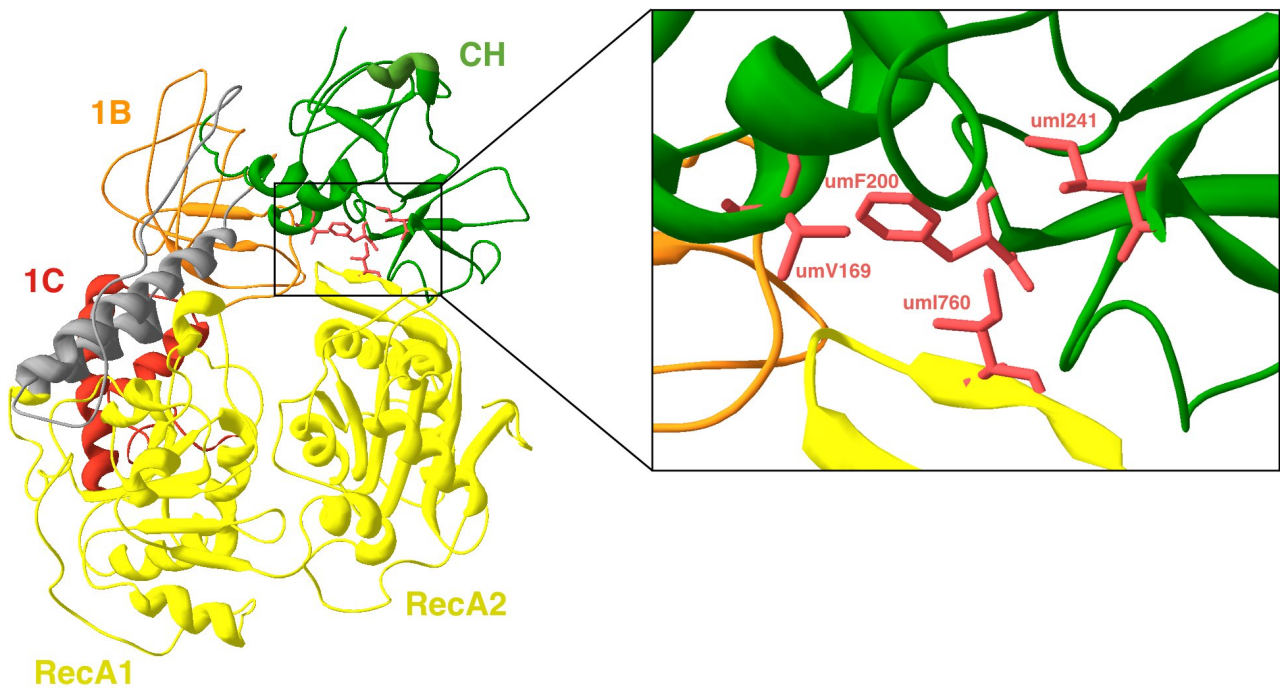


Figura 19. La conformación cerrada de umUPF1 se mantiene gracias a la interacción entre los dominios CH y RecA2. A la izquierda se muestra el modelo generado para umUPF1 empleando como templado el cristal 2XZL. El código de colores es el mismo que el de la figura 1. El detalle de la interacción se muestra en el acercamiento de la derecha. Los aminoácidos clave para esta cooperación se muestran en rosa.

VI.I.V POSIBLE INTERACCIÓN DE umUPF1 CON UPF2

Para estudiar *in silico* la posible interacción de umUPF1 con UPF2, se construyó el modelo tridimensional para umUPF1 empleando como molde el cristal 2WJV. Ambas estructuras fueron superpuestas pero se muestra solamente el modelo para umUPF1 y su interacción con hUPF2 (a partir de la estructura del cristal). En el modelo obtenido para *U. maydis* se identificaron las dos regiones críticas responsables de la interacción entre UPF1 y UPF2 separadas por una región espaciadora flexible, tal como se ha reportado para *H. sapiens* (Clerici *et al.*, 2009). Ambas regiones participan en interacciones hidrofóbicas con hUPF2 y los residuos involucrados se muestran en la Fig. 7A. La primera región corresponde a una cavidad hidrofóbica formada por los aminoácidos umL184, umY192, umF204, umV212, umV214, umW249 y umL250. Estos residuos forman un espacio abierto en donde UPF2 puede acoplarse. En particular, los residuos umV212 y umV214 interactúan con los residuos F1171 y con L1174 que forman parte de la lámina beta-A de UPF2 y se ha mostrado que si estos aminoácidos son mutados, se pierde la interacción (Kadlec *et al.*, 2006). La segunda región identificada en umUPF1 que pudiera estar involucrada en la interacción con UPF2 corresponde a una superficie hidrofóbica formada por los residuos umV169 y umF200, los cuales interactúan con los aminoácidos M1120 y M1121 de la alfa-hélice de UPF2 (Fig. 7A).

Para analizar con detalle el dominio de interacción con UPF2, se realizó el alineamiento del dominio CH para las mismas 32 especies empleadas en la comparación evolutiva de UPF1 encontrando que todos los aminoácidos involucrados en las interacciones hidrofóbicas están altamente conservados entre especies (Fig. 7B), lo que sugiere su importancia para la funcionalidad de UPF1. Las únicas

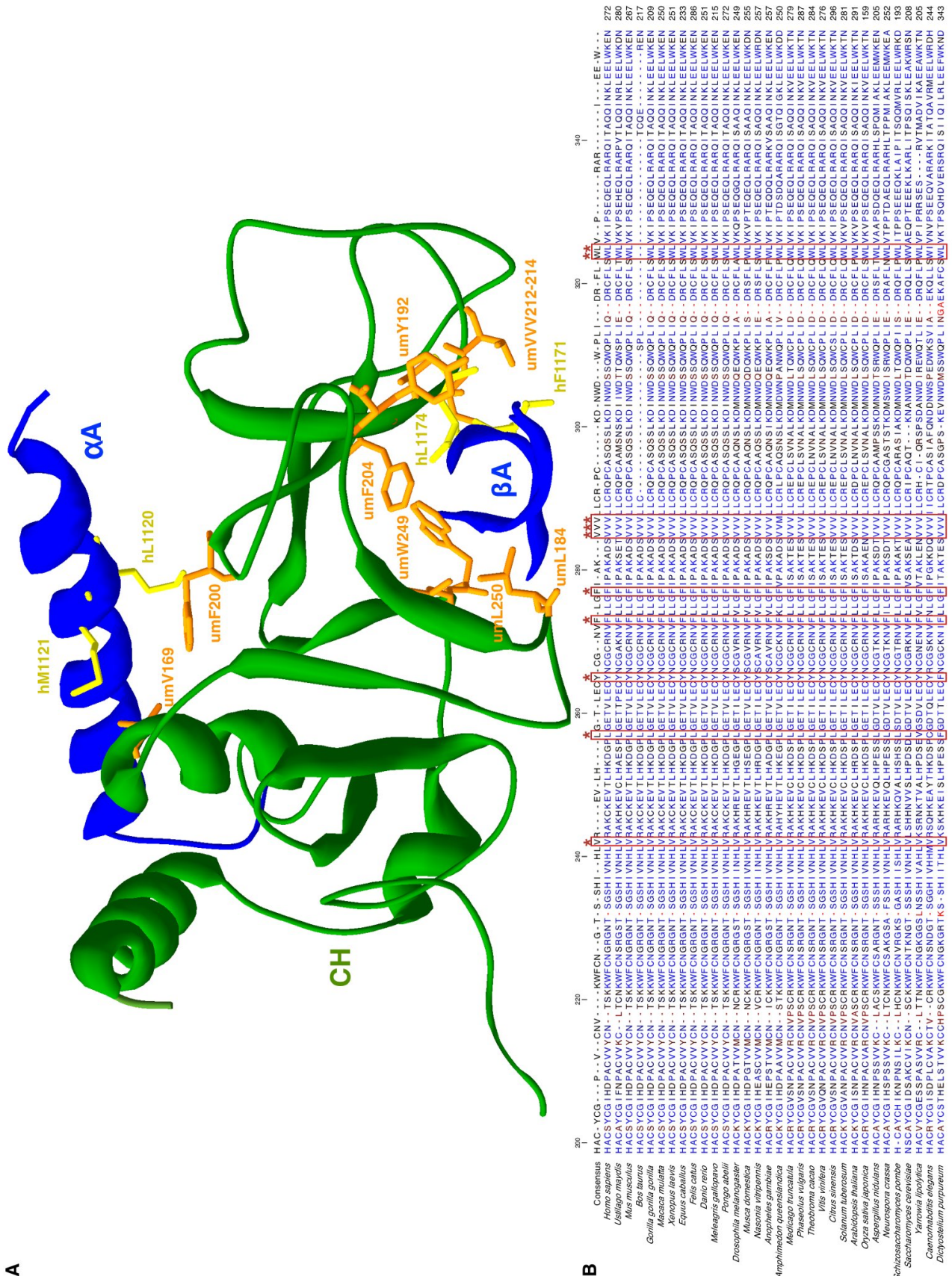


Figura 20. El dominio CH de umUPF1 media la interacción con UPF2. El cristal 2IYK se empleó como templado en donde el ligando corresponde a UPF2 de humano. El cristal y el modelo generado se superponen y se muestra solo el modelo para *U. maydis*. Se identificó una superficie hidrofóbica que participa en la interacción de umUPF1 con la hélice A de UPF2. Otros residuos forman una cavidad hidrofóbica involucrada en donde se ancla la lámina A de UPF2. Se muestran los residuos que participan en la interacción por parte de umUPF1 (naranja) y de UPF2 (amarillo). En la parte inferior se muestra el alineamiento para el dominio CH empleando la secuencia de las especies empleadas para el análisis filogenético de la figura 4. Todos los aminoácidos que forman las dos regiones hidrofóbicas se encuentran altamente conservados entre las especies (*).

excepciones corresponden a *S. cerevisiae*, *C. elegans*, *Y. lipolytica*, *D. purpureum* y *A. queenslandica* en los que una o dos de las posiciones identificadas dentro del dominio CH no se encuentran conservadas. Además de los residuos mencionados, encontramos que la posición umV213 se encuentra conservada en las 32 especies (Fig. 7B) y aunque no se ha reportado que esta posición sea relevante para la interacción entre UPF1 y UPF2, en este trabajo se propone que este aminoácido podría ayudar a estabilizar dicha interacción.

En un intento por confirmar la participación de los aminoácidos identificados en la interacción entre umUPF1 y hUPF2, se mutaron estos residuos *in silico*. En la figura 8 se puede observar una comparación entre la estructura original (Fig. 8A) con aquella que se genera por la mutación simultánea de las posiciones umV212D y umV214E (Fig. 8C), responsables de la interacción de umUPF1 con la lámina A de hUPF2. Al realizar esta doble mutación se pierden las interacciones hidrofóbicas (Fig. 8C). Estos resultados *in silico* correlacionan con los reportes en donde al mutar estos aminoácidos en hUPF1 se altera la interacción entre hUPF1 y UPF2 (Kadlec *et al.*, 2006).

En cuanto a la interacción de umUPF1 con la hélice A de hUPF2 (Fig. 8B), se realizó una doble mutación *in silico* cambiando los residuos conservados umV169 y umF200 por ácido glutámico. En la Fig. 8D se muestra cómo se pierde la superficie hidrofóbica con las mutaciones, alterando la estructura secundaria de ese dominio, debilitando posiblemente la interacción con UPF2. Este efecto puede deberse a alteraciones en la orientación espacial de los residuos y a un incremento en la distancia entre

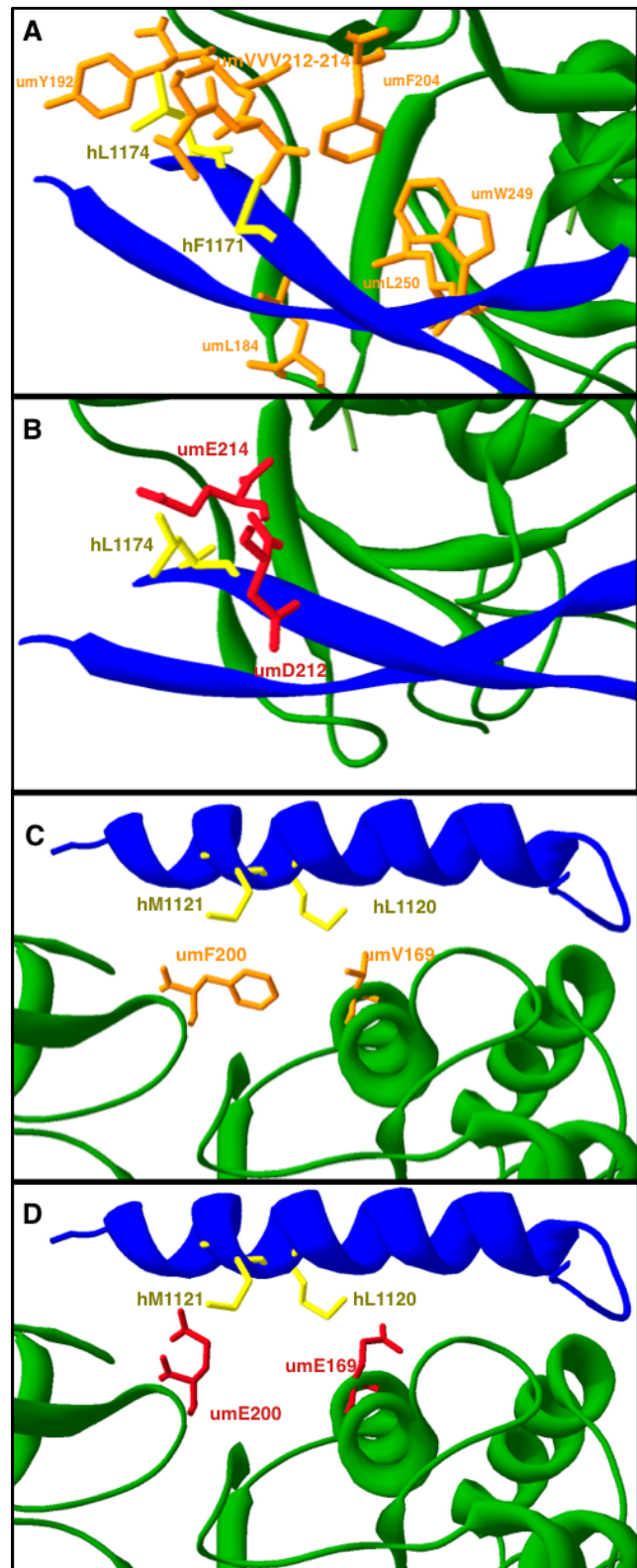
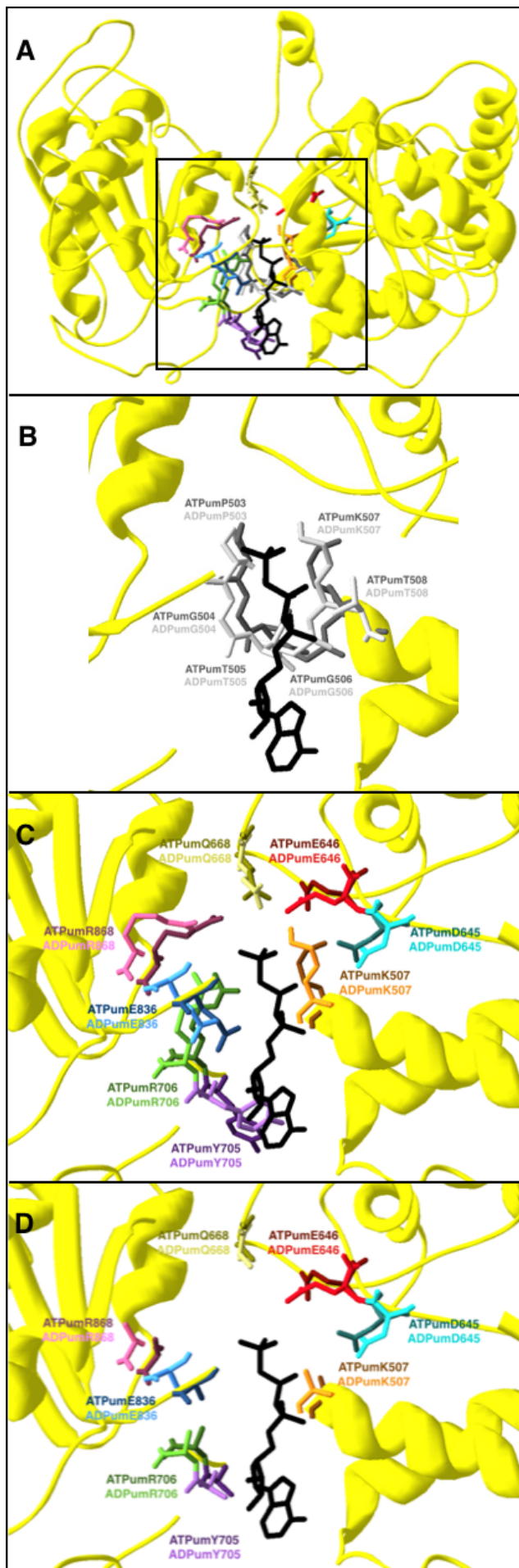


Figura 21. Alteración *in silico* de la asociación CH-UPF2. En la parte superior se muestra la conformación original que facilita las interacciones hidrofóbicas del dominio CH de umUPF1 ya sea con la lámina A (izquierda) o con la hélice A (derecha) de UPF2. En la parte inferior se muestran las mutaciones *in silico* para ambas regiones hidrofóbicas. Los aminoácidos hidrofóbicos identificados en el dominio CH de umUPF1 se indican en naranja y los residuos mutados en rojo.



los aminoácidos mutados y los residuos con los que interactúan en UPF2 (Tabla S4). El efecto de mutar la primera región hidrofóbica es más evidente que el que ocurre al mutar la segunda región de acuerdo con los valores de rms calculados antes y después de la mutación (Tablas S2 y S4). En conjunto, nuestras observaciones indican que las interacciones hidrofóbicas podrían ser responsables de la interacción entre umUPF1 y UPF2, la cual podría a su vez regular el proceso de NMD en *Ustilago maydis*.

VI.I.VI. CONFORMACIÓN ATPasa

La capacidad de umUPF1 para interactuar con ATP y ADP se estudió empleando los modelos generados empleando como templado los cristales 2GJK y 2GK6 (Cheng *et al.*, 2007). El cristal 2GJK se construyó empleando como ligando una molécula de ANP (éster adenilado del ácido fosfoaminofosfónico), el cual semeja el estado de transición correspondiente a la hidrólisis de ATP, mientras que el cristal 2GK6 corresponde a la unión con ADP. Una vez obtenidos los modelos para *U. maydis* se superpusieron con el cristal 2GJK (permitiendo que el ANP fuera visible) y se muestran solamente los principales aminoácidos de umUPF1 que participan en la interacción con el nucleótido (Fig. 9A). Los residuos que se unen

Figura 22. Arreglo conformacional de umUPF1 en presencia de ATP o de ADP. Los modelos para estudiar la unión de umUPF1 con ATP o ADP se construyeron a partir de los cristales 2GJK y 2GK6 respectivamente. A) En el centro del espacio que se forma con los dominios RecA (amarillo) se muestran todos los aminoácidos que podrían facilitar la interacción con el nucleótido. Los residuos del modelo que pudieran unir ATP se ilustran empleando colores sólidos, mientras que el color claro equivalente corresponde a la unión con ADP. B) Dominio de interacción con ATP común para las helicasas SF1/2. C) Representación ampliada de la región de interacción con ATP, mostrando la identidad de los residuos involucrados. D) Todos los residuos implicados en el arreglo estructural se mutaron por Ala.

con ATP se muestran en colores fuertes mientras que el color claro equivalente corresponde a la interacción con ADP; se muestra la estructura de umUPF1 modelada a partir del cristal 2GJK. Al analizar esta estructura encontramos que el sitio de unión al nucleótido se encuentra conformado por los residuos umP503, umG504, umT505, umG506, umK507 y umT508 (Fig. 9B), los cuales corresponden al motivo I conservado en las helicasas SF1/2 (Fig. 2). Otros residuos involucrados en la interacción con el ligando son umD645, umE647, umE646, umQ668, umR706, umE836 y umR686. Todos estos residuos se ubican en los motivos helicasa, además de que se encuentran conservados en la secuencia reportada para *H. sapiens* (Fig. 2); la posición y orientación espacial para estos aminoácidos correlaciona con la estructura reportada para humano (Cheng *et al.*, 2007). Para corroborar la participación de los residuos identificados en la unión de umUPF1 con ATP, éstos se mutaron por alanina y como resultado de esta mutación se observa que se pierde la cavidad en donde ocurre la interacción y en su lugar se puede observar un espacio sin interacciones aparentes en el que no se puede anclar el ANP. Todas estas observaciones sustentan el papel de los aminoácidos identificados en la actividad ATPasa de umUPF1.

VI.I.VII. INTERACCIÓN DE umUPF1 con ARN

Una característica clave de UPF1 es su capacidad para interactuar con ARN. Las propiedades de interacción de umUPF1 con el ARN fueron analizadas en el modelo generado a partir del cristal 2XZO, el cual carece del dominio CH y fue obtenido en presencia de seis ribonucleótidos (Chakrabarti *et al.*, 2011). En este modelo observamos que los dominios RecA son los responsables de todos los contactos de umUPF1 con el ARN (Fig. 10). En esta conformación se forma un canal central en el que el dominio 1C (rojo) sella un lado del canal haciendo que el ARN (líneas negras) se empaque contra el dominio 1B (naranja). Todas las interacciones con ARN involucran residuos (azul) que forman parte de los dominios RecA (amarillo). El recuadro de la figura 10 muestra en detalle a los 23 aminoácidos que conforman el canal de unión con ARN en el modelo construido para umUPF1. Este análisis detallado predice que el dominio 1B interactúa directamente con el extremo 3' del ARN por medio del residuo umK425, mientras que el extremo 5' interactúa con los aminoácidos umT765 y umS766.

VI.I.VIII. SIMULACIONES DE DINÁMICA MOLECULAR

Como se mencionó anteriormente, los residuos responsables de diferentes funciones de UPF1 han sido revelados principalmente mediante ensayos de mutagénesis, los cuales son importantes para entender la función de UPF1. El problema es que estos estudios nos permiten analizar solo un punto estático de una molécula altamente dinámica como lo es este factor central de la vía de NMD. En la actualidad, diversos estudios computacionales de moléculas peptídicas ayudan a predecir y analizar detalles funcionales y mecánicos de diversas biomoléculas (Khodadadi & Sokolov, 2015; Padariya *et al.*, 2015). Para continuar con el análisis de los resultados obtenidos del modelado estructural y las posibles implicaciones funcionales de ciertos dominios, se realizaron simulaciones de dinámica molecular tanto

para el dominio CH de humano correspondiente al cristal 2IYK, como para el modelo generado a partir de este templatado para *U. maydis*.

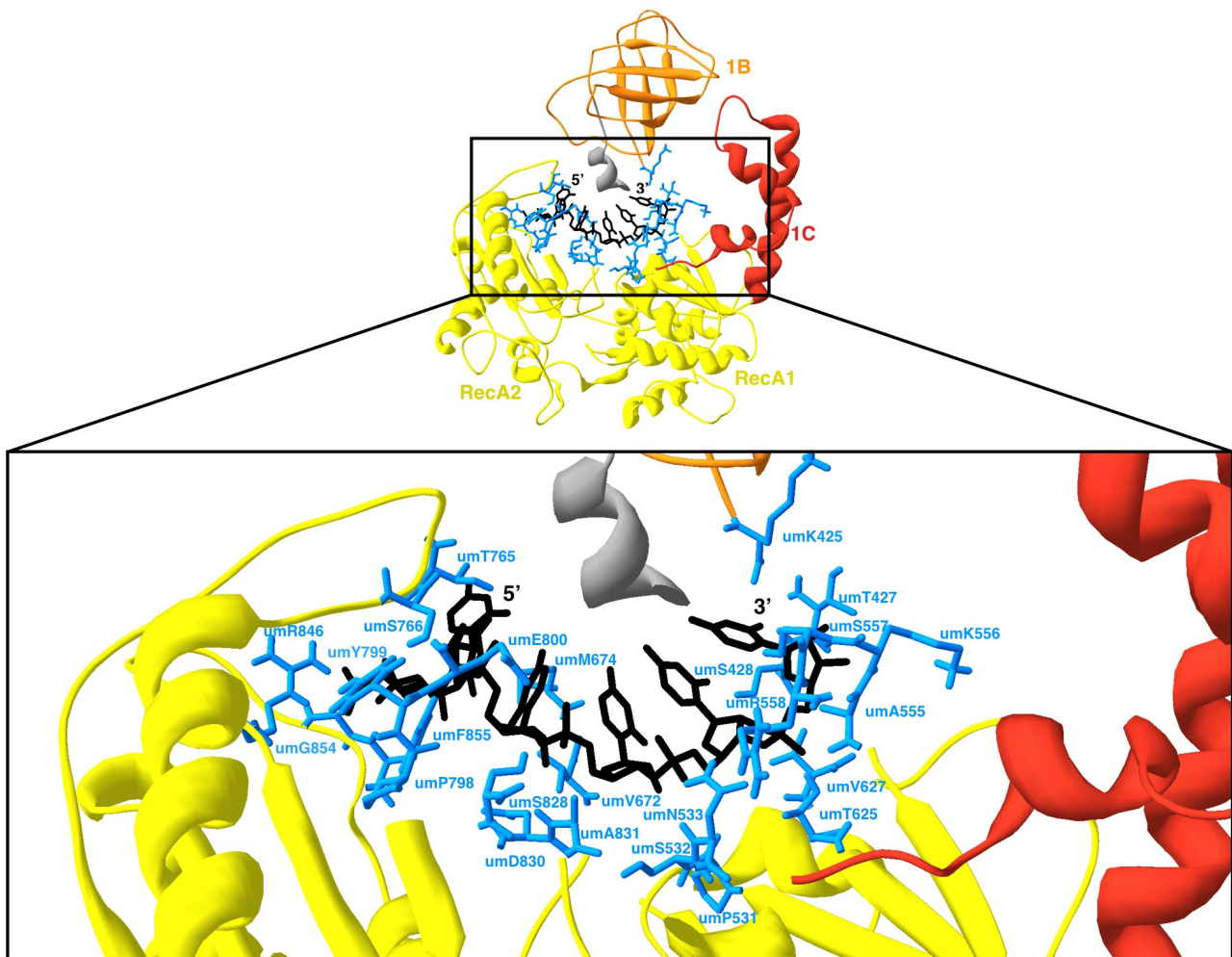


Figura 23. Interacción de umUPF1 con ARN. Se forma un canal en la molécula de umUPF1, el cual permite la interacción de esta proteína con el ARN. Los residuos que forma dicho canal se muestran en azul. La molécula de ARN se ejemplifica con líneas negras. En la parte inferior se presenta el detalle de esta interacción. Se incluye la identidad y posición de los principales aminoácidos de umUPF1 que participan en la unión al ARN. Algunos residuos ya han sido reportados como parte de la interacción, mientras que otros se evidenciaron en este estudio. El modelo presentado para umUPF1 se generó empleando como templatado el cristal 2XZO.

Inicialmente se calculó el valor de RMS para las estructuras de hUPF1 y umUPF1 y la representación gráfica se muestra en la figura 11A en la que se observa que la mayor parte de la molécula se muestra en azul, indicando que el par de aminoácidos comparados para cada posición se encuentran conservados y presentan la misma orientación espacial. A continuación se calculó el valor para el radio de giro (R_g), el cual permite medir el grado de compactación de una estructura (Lobanov *et al.*, 2008). Como se observa en la figura 11B, el valor del R_g para el dominio CH de 151aa permanece estable durante los 10ns que duró la simulación en *H. sapiens* y *U. maydis*, lo que sugiere que ambos péptidos tienen las mismas características de compactación y de plegamiento. A partir de la gráfica, podemos observar también que el valor R_g promedio para los dos péptidos es de 16.7 aproximadamente. Estas observaciones correlacionan con reportes previos en los cuales las proteínas que pertenecen a esta clase estructural (hélices más láminas) son más compactas que aquellas que se encuentran conformadas exclusivamente por alfa-hélices o láminas beta (Lobanov *et al.*, 2008).

Las proteínas como UPF1 que requieren de cambios conformacionales para alcanzar su forma activa poseen una estructura dinámica. Esta estructura permite cambios conformacionales frecuentes y un buen grado de flexibilidad, pero al mismo tiempo estas proteínas muestran un buen grado de compactación y una estabilidad termodinámica que les permitan llevar a cabo sus funciones catalíticas y fisiológicas.

Para verificar la estabilidad de ambos sistemas y para determinar de manera cuantitativa el grado de movimiento, se midió el valor de RMSD (root-mean-square deviation) para ambos péptidos durante trayectorias de 10 ns (Fig. 11C). A partir del análisis de los valores independientes de RMSD para las hélices y las asas, podemos observar que las distorsiones para las hélices son mucho más pequeñas que aquellas mostradas para las asas y que al final de los 10ns de simulación, los valores de RMSD son muy semejantes para ambas proteínas. Se observa una pequeña diferencia entre 3.3 y 5ns, indicando que en este punto es probable que ambos péptidos puedan sufrir un cambio conformacional, lo cual es altamente probable debido a la gran flexibilidad estructural y a la naturaleza versátil de la proteína. Además, dado que el dominio CH de UPF1 es el responsable de las interacciones proteicas relacionadas con la activación de la proteína, se espera que la estructura sea entrópica o al menos que presente dominios entrópicos.

Finalmente, la flexibilidad de la proteína se analizó midiendo también el valor de RMSF (root-mean-square fluctuation) para cada residuo durante toda la simulación.

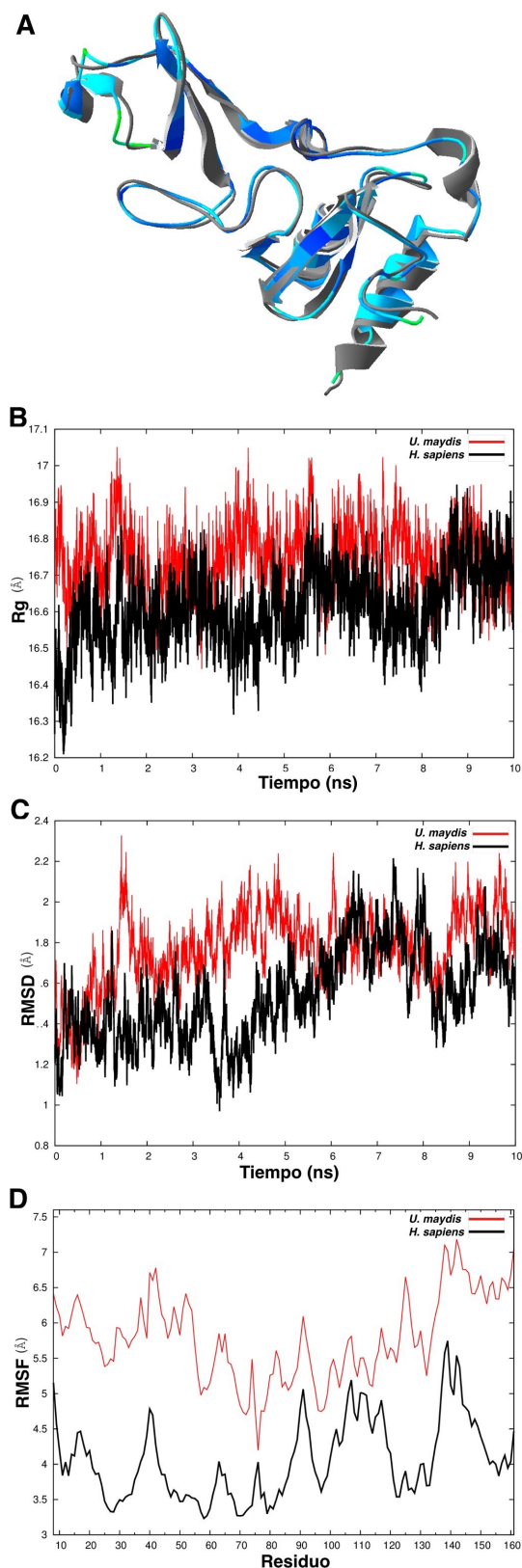


Figura 24. Dinámica molecular de las proteínas UPF1 de *H. sapiens* y *U. maydis*. A) Representación gráfica de los valores de RMS calculados para la superposición de las estructuras hUPF1 y umUPF1. Se generó el alineamiento estructural para el cristal 2IYK de humano y el modelo generado para umUPF1 empleando como templado dicho cristal. Cada aminoácido en umUPF1 presenta un color de acuerdo al valor de RMS obtenido en relación con el aminoácido correspondiente en hUPF1. La escala va de azul fuerte para una buena superposición hasta el rojo cuando la superposición es pobre (Swissmodel). B) Radio de giro (Rg). La línea negra corresponde al péptido de *U. maydis* y la línea roja para el de *H. sapiens* durante la simulación de 10ns a 300K. C) El valor de RMSD para los átomos es presentada en función del tiempo para la estructura reportada para hUPF1 (rojo, 2IYK) y para el modelo generado para umUPF1 (negro) durante la simulación de 10ns a 300K. D) El valor de RMSF para el carbón alfa de los residuos se calculó en función del tiempo durante la simulación de 10ns a 300K para hUPF1 (rojo) y para umUPF1 (negro).

Para este caso, se observaron más fluctuaciones en el péptido de *U. maydis*, las cuales correlacionan con los residuos presentes en el loop 6 y en el loop 10. A pesar de que el péptido fúngico muestra fluctuaciones más grandes, el comportamiento de los dos péptidos es semejante en las mismas posiciones, lo que sugiere que estas regiones son menos estables y más reactivas. Estas observaciones resultan interesantes al correlacionarlos con los estudios estructurales de UPF1, en donde se ha propuesto que el mecanismo mediante el cual se regula la actividad de esta proteína se relaciona con la transición de UPF1 entre las conformaciones “abierta” y “cerrada” durante el proceso de NMD como un resultado de la interacción o no con UPF2 (Gleghorn & Maquat, 2011).

VI.II. EL NÚCLEO DE NMD EN *USTILAGO MAYDIS*

VI.II.I. *umUPF1*, *umUPF2*, *umUPF3* DE GEN A PROTEÍNA

Para proceder con el estudio del núcleo del NMD en *Ustilago maydis*, una vez identificados en la anotación del genoma los locus para los posibles homólogos de las principales proteínas de la maquinaria de NMD (Tabla 7), se diseñaron los oligonucleótidos específicos para amplificar el Marco de Lectura Abierto (MAL) completo de todas estas proteínas. Las secuencias empleadas como cebadores se resumen en la tabla (Tabla 5). En este apartado nos centraremos en los factores del núcleo de NMD, sin embargo se investigaron todos los factores descritos en la tabla 7 (datos no mostrados).

Para la identificación en el genoma, la verificación de la expresión del transcrito y la expresión de las proteínas se empleó la cepa FB1 de *Ustilago maydis*, que corresponde a la forma haploide del basidiomiceto y fue la que se utilizó por el Broad Institute para ensamblar el genoma de *Ustilago maydis* 521.

Ustilago maydis FB1 se cultivó en el laboratorio empleando el medio YEPS y se permitió su crecimiento durante 18-20 horas a 30°C. Para corroborar la presencia en el genoma de las secuencias de interés se procedió a la extracción de ADN; para la comprobación de la expresión del transcrito se extrajo ARN total y para comprobar la presencia de la proteína correspondiente se obtuvieron extractos proteicos totales. En todos los casos se verificó la integridad y concentración aproximada del ADN, del ARN y de la proteína total obtenidos.

Empleando la técnica de PCR, a partir de ADN y empleando los oligonucleótidos enlistados en la tabla 5, se amplificó la secuencia correspondiente al marco de lectura abierta completo para cada uno de los tres factores del núcleo de NMD identificados (Fig. 25). Una vez verificada la integridad de los productos de PCR mediante electroforesis en gel de agarosa, los productos se purificaron para posteriormente ser utilizados en diferentes clonaciones.

Para verificar la expresión de los tres factores del núcleo de NMD se realizaron ensayos de RT-PCR en dos pasos: por una parte la retrotranscripción para la obtención de ADNc, seguida de una PCR de punto

final utilizando como templado dicho molde. El resultado de los ensayos de RT-PCR se muestra en la figura 26, donde podemos observar la expresión del transcrito para los tres factores.

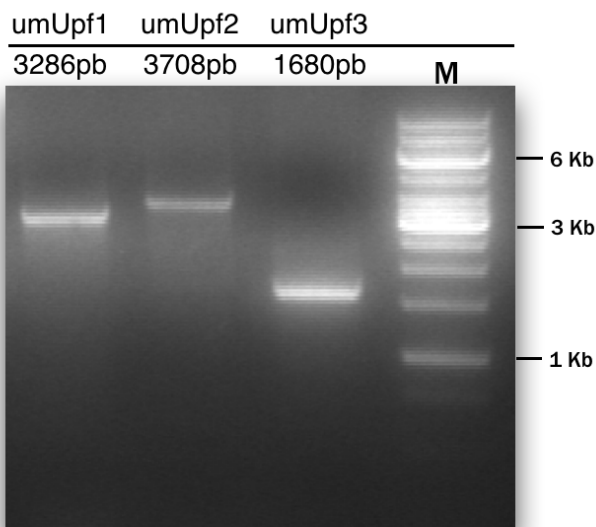


Figura 25 Presencia de los genes UM00237.1, UM03549.1, UM00233.1 en el genoma de Ustilago maydis FB1. En la imagen se muestra el gel de agarosa al 1% donde se observa como resultado de la electroforesis la visualización bajo UV con bromuro de etidio, los productos de PCR con los tamaños esperados para el MAL de los genes que codifican para las proteínas putativas UPF1, UPF2 y UPF3 respectivamente. A la derecha la descripción del marcados de peso molecular (100pb de Fermentas).

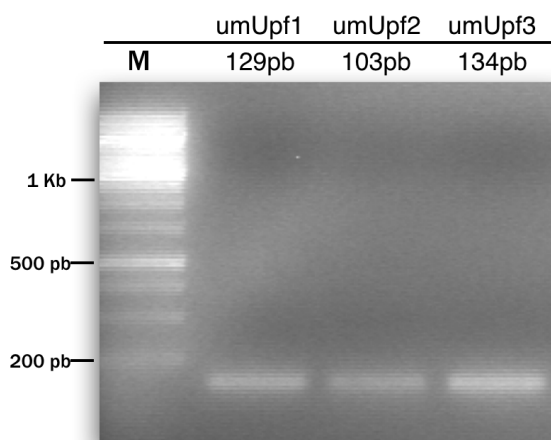


Figura 26. Expresión de los genes UM00237.1, UM03549.1, UM00233.1 en Ustilago maydis FB1. La imagen la electroforesis de un gel de agarosa al 1.5% donde se observan los amplificados de PCR a partir de cDNA de umUPF1, umUPF2 y umUPF3 con los tamaños esperados descritos en la parte superior. A la izquierda la descripción del marcados de peso molecular (100pb de Fermentas).

Para asegurar que los amplificados fueran los requeridos para nuestro estudio, los productos de PCR se clonaron en el vector pJet1.2 y se enviaron a secuenciar. Una vez recibidas, las secuencias fueron editadas y analizadas, con lo que se confirmó la identidad de los productos amplificados.

En cuanto a la expresión de las proteínas putativas, umUPF1, umUPF2 y umUPF3 a partir de extractos totales se identificó por medio de Western Blot la presencia de dichos factores utilizando anticuerpos policlonales comerciales (Fig. 27). Para ésto, se verificó que las secuencias de los epítomos se encontraran conservadas en las proteínas de *Ustilago maydis* ya que estos anticuerpos fueron construidos en base a los homólogos para humano (datos no mostrados).

En la tabla 8 se muestra el tamaño para los amplicones que corresponden al MAL para cada uno de los tres factores, también se indica el número de aminoácidos que componen a dichas proteínas en función a la secuencia amplificada, así como el peso en kDa calculado para estos factores (calculados con ProtParam tool de Expsy). En la tabla se muestra también el peso para la proteína recombinante generada, como se mostrará posteriormente.

Tabla 8. Tamaños de los factores del núcleo de NMD en *U. maydis* y *H. sapiens*

	pb	aa	kDa	+poliHis(kDa)
UPF1/CH	873	290	32	39
umUPF1	3285	1094	121	-
umUPF2	3708	1235	137	144
umUPF3	1680	559	59	66
hUPF1	3390	1129	124	-
hUPF2	3819	1272	148	-
hUPF3b	1413	470	56	-

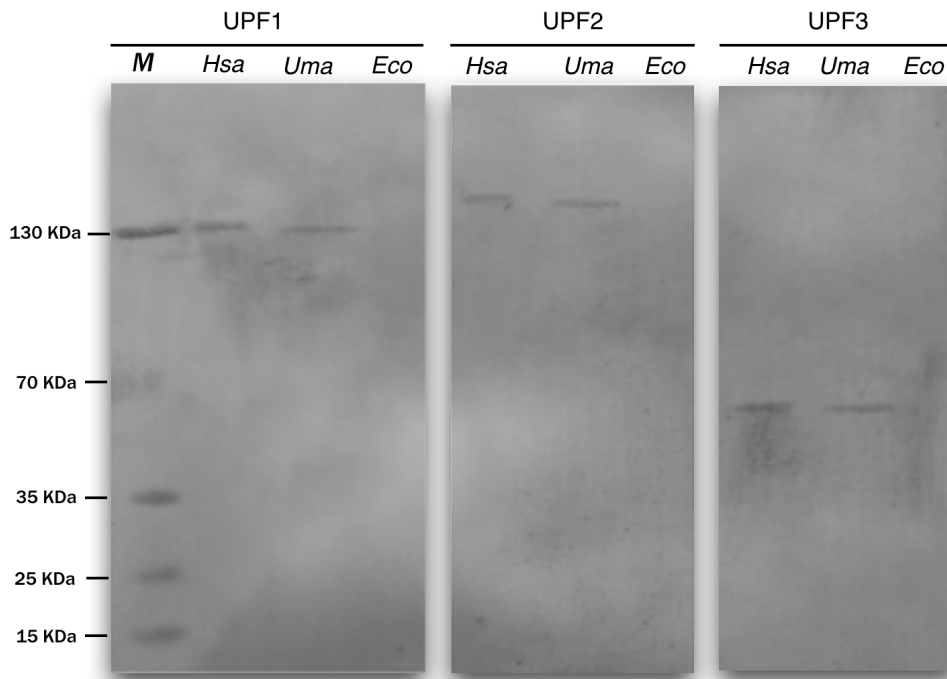


Fig. 27. Identificación de las proteínas UPF1, UPF2 y UPF3 en extractos totales de *U. maydis*. Se emplearon anticuerpos policlonales comerciales (Santa Cruz) dirigidos para cada una de las proteínas. Los anticuerpos secundarios se encontraban acoplados a fosfatasa alcalina. Como control positivo (Hsa) se empleó extracto total de la línea celular de cáncer de próstata PC-3 y como control negativo (Eco) se utilizó extracto total de un cultivo bacteriano (*E. coli* DH5a). El peso aproximado para las proteínas se muestra en la tabla 9. Los detalles del protocolo se describen en el apartado de material y métodos.

VI.II.II. Identificación de mensajeros blanco de la vía de NMD en *Ustilago maydis* FB1

Debido a la falta de información con respecto al mecanismo de NMD en nuestro modelo de estudio, una de las primeras tareas debía ser identificar algunos mensajeros que estuvieran entrando a esta vía de degradación y así, de manera indirecta empezar a tener indicios de que el mecanismo fuera funcional en el basidiomiceto. Además, esto sería de utilidad para poder valorar la alteración del mecanismo de NMD después de la interrupción de *umUpf1*.

Al ser éstos los primeros estudios de NMD en *U. maydis*, no se han reportado los ARNm que podrían ser susceptibles de NMD, por tanto no contábamos con marcadores que pudiéramos evaluar para verificar la funcionalidad de la vía en el hongo. Los primeros reportes en *S. cerevisiae* de mensajeros que pudieran ser monitoreados para ver los efectos de algunas mutaciones en la funcionalidad del sistema de NMD fueron *Hrp1* (He *et al.*, 1993) y *CYH2* (Cheng *et al.*, 1997).

Existen reportes de la inhibición de la vía de NMD mediante un fármaco que permite que los transcritos que contienen mutaciones “sin sentido” se “estabilicen” (los transcritos no se degrada por NMD si no que se se acumulan como se observa en la figura 28), lo que ayuda a su identificación ya que se acumulan aumentando su concentración y siendo perceptibles por diferentes métodos de medición (Mercola & Welsh, 2004). Este fármaco es la emetina, que es un producto natural alcaloide y tiene muchas aplicaciones, es usada principalmente como droga antiparasitaria pero también se ha visto que tiene buenos efectos para bloquear la ruta de NMD. Esta droga inhibe la síntesis de proteínas ribosomales y mitocondriales e interfiere con las actividades del ADN y el ARN. Noensie y Dietz en 2001

publicaron una estrategia para investigar algunas enfermedades genéticas asociadas a NMD con este fármaco. Su estudio se fundamentaba en la identificación de genes mutados (que presentaban codones de paro prematuro) mediante la inhibición de NMD con emetina; esta estrategia se denomina GINI (Gene Identification by NMD Inhibition) y por medio de la identificación de transcritos que en presencia del fármaco se “estabilizan” se han estudiado anomalías asociadas a cáncer de colon (Ionov et al., 2004), cáncer de próstata (Huusko et al. 2004) y de melanoma (Bloethner et al., 2008).

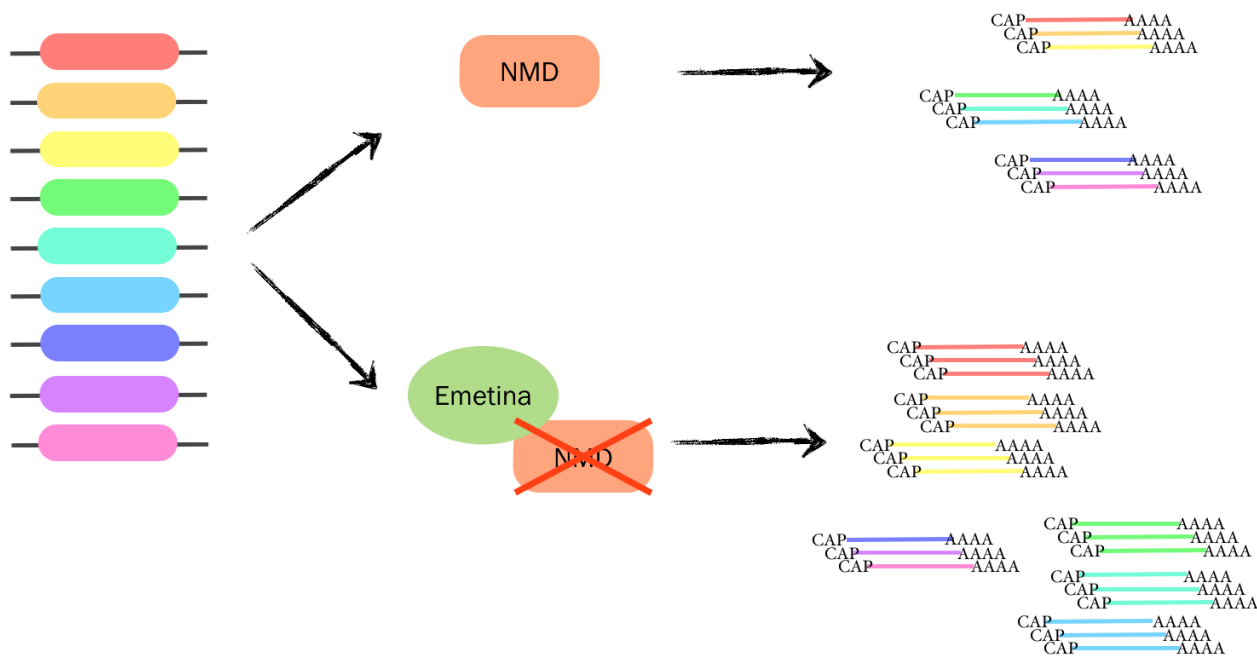


Figura 28. Esquema del Ensayo de Emetina. En la imagen se puede observar que algunos transcritos son susceptibles a NMD. Los ARNm mantienen una expresión regulada cuando el mecanismo de NMD es funcional, mientras que cuando éste es bloqueado con emetina se pierde la regulación, lo que hace que algunos mensajeros que tienen codones de paro prematuros (aunque no necesariamente) aumentan su concentración ya que no son llevados a degradación como se sugiere en presencia de NMD.

Inicialmente se implementaron las condiciones de cultivo de *U. maydis* para poder monitorear posteriormente el efecto del fármaco emetina sobre la vía de NMD al evaluar la presencia de ciertos mensajeros. Para ello, se estandarizó la curva de crecimiento en múltiples experimentos en los cuales además de cuantificar la DO (Densidad Óptica) en función del tiempo, se midió la sobrevivencia del hongo utilizando la técnica de DROP PLATE (datos no mostrados). La combinación de las dos técnicas nos permite asegurarnos de que la DO se refiere a células viables.

Una vez que se obtuvo la curva de crecimiento del hongo, se probaron diferentes condiciones para el tratamiento con emetina. En células de mamífero se ha reportado el empleo de diferentes concentraciones del fármaco para inhibir el NMD, las cuales varían desde 10 ug/ml hasta 1000 ug/ml (concentración final de emetina en el cultivo). Por otro lado, se probaron también diferentes tiempos de exposición al fármaco para asegurar tanto la viabilidad del microorganismo como el adecuado rendimiento para la extracción de ARN y la efectividad sobre el sistema de NMD. Evaluando todas estas variables, decidimos emplear las condiciones con mejores resultados para nuestro sistema. La

concentración de emetina fué 100 ug/ml, la cual se ha empleado también para células de mamífero en cultivo (Noensie & Dietz, 2001). En cuanto al tiempo de incubación con el fármaco, decidimos adicionarlo cuando el cultivo alcanzaba una DO de 0.4 (T=9h) y retirarlo después de 4 horas a una DO de 1 (T=13h) (Fig. 29). En nuestras observaciones tiempos más prolongados de incubación ocasionaron que las células perdieran viabilidad, además de que en condiciones de mayor cantidad de células no era posible recuperar de manera eficiente el ARN del cultivo. El empleo de concentraciones menores de emetina no mostró efecto alguno.

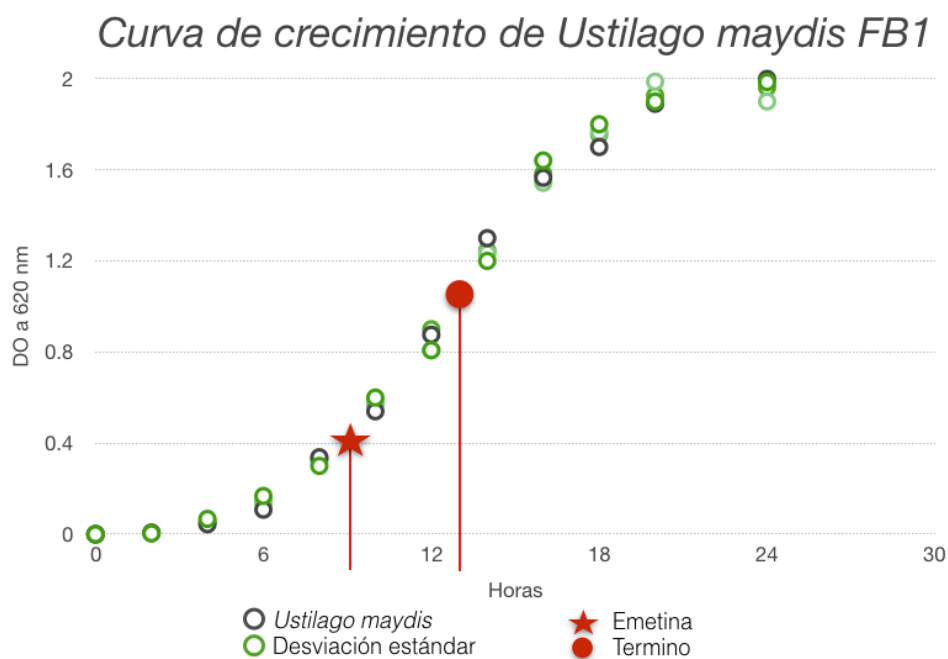


Figura 29. Curva de crecimiento y estandarización del Ensayo de Emetina en *Ustilago maydis* FB1. En la gráfica se observa el comportamiento para el crecimiento del hongo donde los puntos grises son los valores totales y los verdes las desviaciones estándar. El ensayo de emetina se realizó en medio de cultivo YEPS. La inoculación con el fármaco se realizó a las 9 horas (DO620 0.4 aproximada, estrella roja) a una concentración final de emetina de 100 ul/ml en el medio de cultivo; el hongo se incubó con la emetina durante 4 horas (DO620 1 aproximada, círculo rojo) y el cultivo fue recuperado a las 13 horas para proseguir con la extracción de ARN. El ensayos se realizó por triplicado.

Con la finalidad de medir cambios en la expresión de diferentes mensajeros al bloquearse la vía de NMD. A partir de los cultivos finales tanto en presencia como en ausencia de emetina, se extrajo ARN total para cuantificar mediante RT-PCR de punto final dichos cambios en la expresión de 17 mensajeros. Estos transcritos seleccionados se tomaron de la bibliografía como en el caso de *Cyh2* que se sabe entra a la ruta de NMD (Cheng *et al.*, 1997) y *ppi*, gen constitutivo que se reportado como control para medir la expresión de algunos transcritos en *Ustilago maydis* (Scherer *et al.*, 2006). También probamos con algunos como los descritos en Saville *et al.*, que en 2007 validó la expresión de algunos transcritos como: *Gsk*, *Par*, *Crmp*, *Frataxina*, *Zfp*. Además de todos los reportados tomamos algunos de nuestro interés como: *Magoh*, *Rnps1*, *Hbs1*, *Upf1*, *Upf2*, *Upf3*, *U1-70K*, *Smg1* y *Smgs*; todos relacionados con la vía de NMD.

Para este ensayo se verificó la concentración e integridad del ARN en un intento de asegurar que en todos los casos se adicionara al misma cantidad de templado al ensayo de RT y que éste no fuera un

Tabla 9. Resultados ensayo emetina

	RT-PCR (Emetina vs. Emetna)
Crmp	↓
Cyh2	↑
Frataxin	↑
Gsk	==
Hbs1	↓
Magoh	↑
Musashi	↓
Par	↑
ppi	==
Rnps1	↓
Smg1	==
Smg's	↑
U1-70K	↓
Upf1	==
Upf2	==
Upf3	↓
Zpf	↑

factor que influenciara la cantidad de producto obtenido. Para todos los mensajeros evaluados, el tamaño del amplicón iba de 100 a 400 pb. Los resultados de este ensayo los podemos observar de manera cualitativa en la tabla 9 donde tal y como lo esperábamos el mensajero para Cyh2 se acumuló mientras que el de ppi se mantuvo sin cambio. También aumentaron los ARNm para Frataxina, Magoh, Par, Smgs y Zfp mientras que otros mensajeros disminuyeron, como, Hbs1, Musashi, Rnps1, Upf3 y U1-70K, interesantemente todos estos factores relacionados con *splicing*; algunos que no tuvieron efecto fueron, Upf1, Upf2 y Smg1.

Para comprobar los resultados obtenidos en el ensayo anterior, se realizó otro experimento bajo las mismas condiciones de cultivo, tratamiento con emetina, extracción de ARN y RT-PCR de punto final, pero para este ensayo únicamente se monitorearon

cuatro transcritos, ppi como control; Cyh2, Magoh y Frataxin como posibles blancos de NMD. Lo que obtuvimos como resultado se encuentra en la figura 30 donde podemos observar que los datos obtenidos confirman las observaciones del ensayo anterior. Este experimento resultó interesante pues fue posible implementar una forma de estudiar la vía de NMD, además de que se cuenta a partir de ahora con una técnica para posteriores estudios con los diferentes fondos genéticos de *Ustilago maydis*.

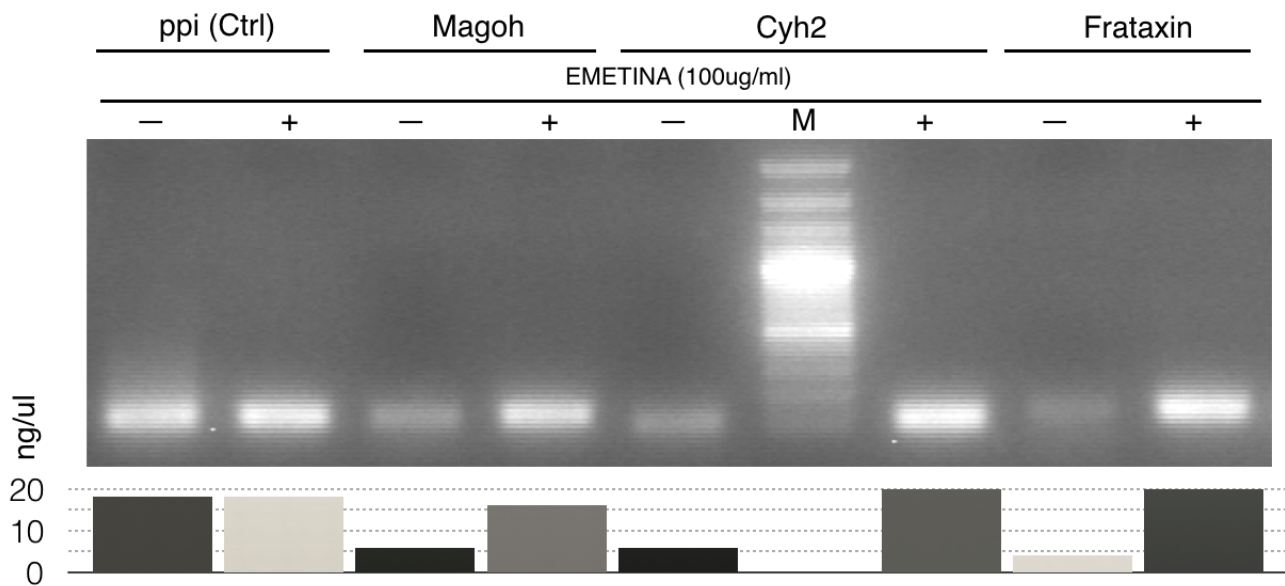


Figura 30. *Transcritos regulados por NDM en Ustilago maydis FB1*. Electroforesis de un gel de agarosa al 1.5% donde se observan los amplicados de PCR a partir de cDNA para ppi, Magoh, Cyh2 y Frataxin. El ARN proviene de cultivos de *U. maydis* en presencia (+) y en ausencia (-) de emetina. En la gráfica se observa la cuantificación de la concentración para los productos de PCR.

VI.II.III INTERRUPCIÓN DE *umUpf1* EN *U. maydis* FB1.

Una vez que se confirmó la presencia y la expresión de los factores UPF1, UPF2 y UPF3 en la cepa FB1 de *Ustilago maydis*, se realizó la interrupción del gen *umUpf1* para estudiar la funcionalidad del mecanismo de NMD en dicho modelo.

Estudios anteriores realizados en varios homólogos para UPF1 hacían interesante el estudio de las posibles implicaciones de esta proteína en la viabilidad del hongo, ya que se ha demostrado que UPF1 es esencial para la viabilidad de embriones de humano, pez zebra y ratón (Medghalchi *et al.*, 2001; Wittkopp *et al.*, 2009; Hwang *et al.*, 2011) así como para el desarrollo en *D. melanogaster* (Avery *et al.*, 2011), mientras que la mutante nula para UPF1 en *A. thaliana* resultó letal (Arciga-Reyes *et al.*, 2006). Por otro lado, en *S. cerevisiae* la pérdida de la función de UPF1 permite que el crecimiento sea casi normal (Altamura *et al.*, 1992), mientras que la delección de UPF1 en *Y. lipolytica* no mostró ningún efecto (Mekouar *et al.*, 2009). Teniendo en cuenta estos antecedentes, era importante realizar el estudio funcional de UPF1 en *Ustilago maydis* ya que se trata de un factor fundamental para que la vía de NMD se lleve a cabo. Las implicaciones de la interrupción del gen nos podrían ayudar a entender si este sistema es o no fundamental en el hongo, como se ha mostrado para otros organismos. De esta manera pretendíamos analizar el efecto de la ausencia del gen activo en la viabilidad del hongo. Otras observaciones con UPF1 en organismos fúngicos incluye su papel en la regulación del ritmo circadiano en *Neurospora crassa* y el reconocimiento de PTCs afectando la degradación del ARNm en *Aspergillus nidulans* (Morozov *et al.*, 2012; Compton *et al.*, 2013). Aunque el sistema de NMD es esencial en plantas, la participación de UPF1 en la funcionalidad de estos sistemas no se encuentra muy estudiada (Nyikó *et al.*, 2013; Kerényi *et al.*, 2013). De esta manera, un estudio posterior a la interrupción de UPF1 podría enfocarse al estudio de algunos aspectos fisiológicos del basidiomiceto.

Debido a que *U. maydis* tiene la capacidad de realizar recombinación homóloga de manera natural (Holliday, 1964), la generación de herramientas como interrupciones en el genoma es posible por intercambio alélico, donde se reemplaza la copia original del gen por la versión interrumpida generada en el laboratorio. Este mecanismo se han utilizado con éxito para introducir o eliminar marcadores en ADN cromosomal en *Ustilago maydis*. Para que dicho mecanismo se lleve a cabo es necesaria la presencia de secuencias suficientemente largas (1.5 Kb aprox) flanqueantes a la secuencia de interrupción, las cuales en este caso permitieron la sustitución por doble recombinación del gen nativo en el cromosoma por una copia interrumpida con un cassette Higromicina B. Así, la mutante conserva la copia del gen interrumpido en el cromosoma. De manera adicional, la inserción del cassette de resistencia (en este caso a higromicina) permite la recuperación de la mutante.

Para este ensayo se siguieron básicamente tres pasos: 1) la clonación del MAL de *umUpf1* en el vector pJet1.2, 2) la inserción del cassette de Higromicina B aproximadamente a la mitad del MAL clonado y 3) la inactivación del gen en el cromosoma de *Ustilago maydis* FB1.

En el primer paso, el producto de PCR obtenido inicialmente, que corresponde al MAL para umUpf1 (Fig. 25) se clonó en el vector pJet1.2 (clona pJet/umUpf1). Posteriormente, se seleccionó una enzima que presentará un sitio de corte único dentro de dicho MAL. El gen umUpf1 mostró una única secuencia de corte para la enzima HpaI (Figura 31A) aproximadamente a la mitad del gen. Por otro lado y de manera simultánea, se amplificó el cassette de higromicina a partir del vector pcDNA 3.1 (Invitrogen), el cual se clonó en el sitio de corte para HpaI. En la figura 31 podemos observar los mapas para las construcciones de pJet1.2/umUpf1 y pJet1.2/umUpf1::Hyg, así como los productos esperados para las digestiones con enzimas de restricción empleadas para verificar las clonas (Fig. 31C).

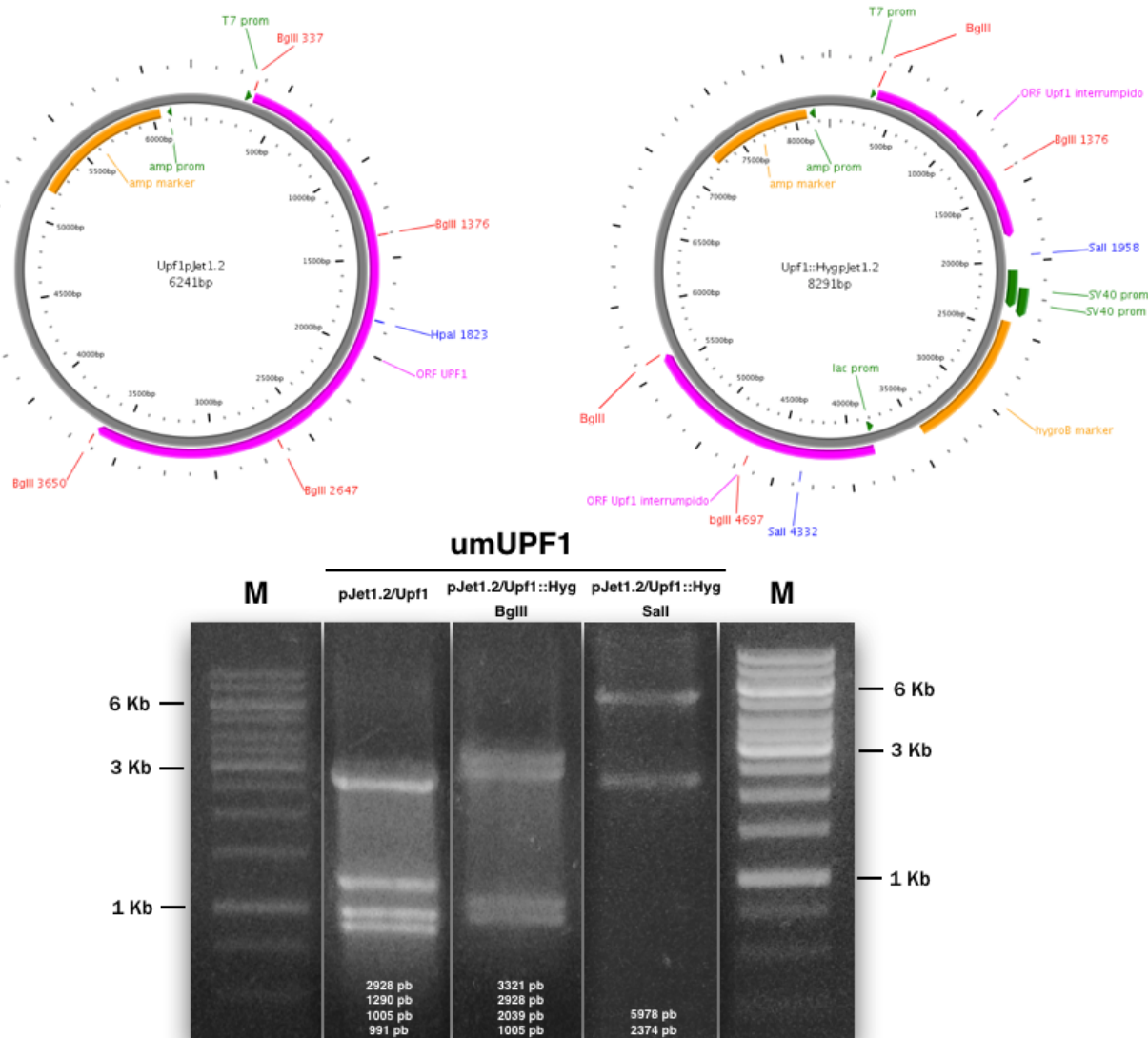


Fig. 31. Construcción de la interrupción de umUpf1. En la parte superior se muestran los mapas para cada una de las construcciones, indicando su peso y las principales características del plásmido, a la izquierda pJet1.2/umUpf1 y a la derecha el mismo plásmido con la inserción del cassette de higromicina en el gen (pJet1.2/umUpf1::Hyg). Abajo los ensayos de restricción que sirvieron para confirmar la clonación del inserto y su posterior interrupción con el cassette. Los tamaños para las bandas obtenidas se indican en cada carril.

Una vez obtenida la clona pJet1.2Upf1::Hyg, ésta se linealizó empleando la enzima ClaI, ya que es necesario introducir el ADN lineal para que pueda llevarse a cabo la recombinación homóloga. El plásmido lineal se purificó y se insertó mediante electroporación en células electrocompetentes de la

cepa FB1 de *U. maydis* (ver material y métodos). Mediante esta transformación, se esperaba cambiar el fondo génico de la cepa silvestre por la interrupción del gen *Upf1* debido a la recombinación homóloga realizada en la célula. Las células transformadas se sembraron en medio YEPS sólido adicionado con higromicina (200 ug/ml). En los primeros intentos, no fue posible recuperar transformantes con las condiciones indicadas. Debido a lo anterior, se decidió disminuir la concentración de antibiótico para permitir la recuperación de transformantes. Al emplear higromicina a una concentración de 20 ug/ml, pudimos recuperar células viables en las que se procedió a verificar la interrupción mediante PCR (Fig. 32). En el carril 4 podemos observar que al amplificar *umUpf1*, se obtienen dos productos. De acuerdo a su peso aproximado, el producto más grande podría corresponder a *Upf1* interrumpido con el cassette de higromicina (4 Kb), mientras que la banda pequeña tiene el peso esperado para el gen *Upf1* silvestre (2 Kb). Este resultado nos sugirió que a bajas concentraciones del marcador de selección podíamos recuperar una población mezclada que podía presentar al gen silvestre y la interrupción con higromicina. El resto de los carriles corresponden a la amplificación de *Upf1* y de higromicina, realizadas como control en diferentes condiciones.

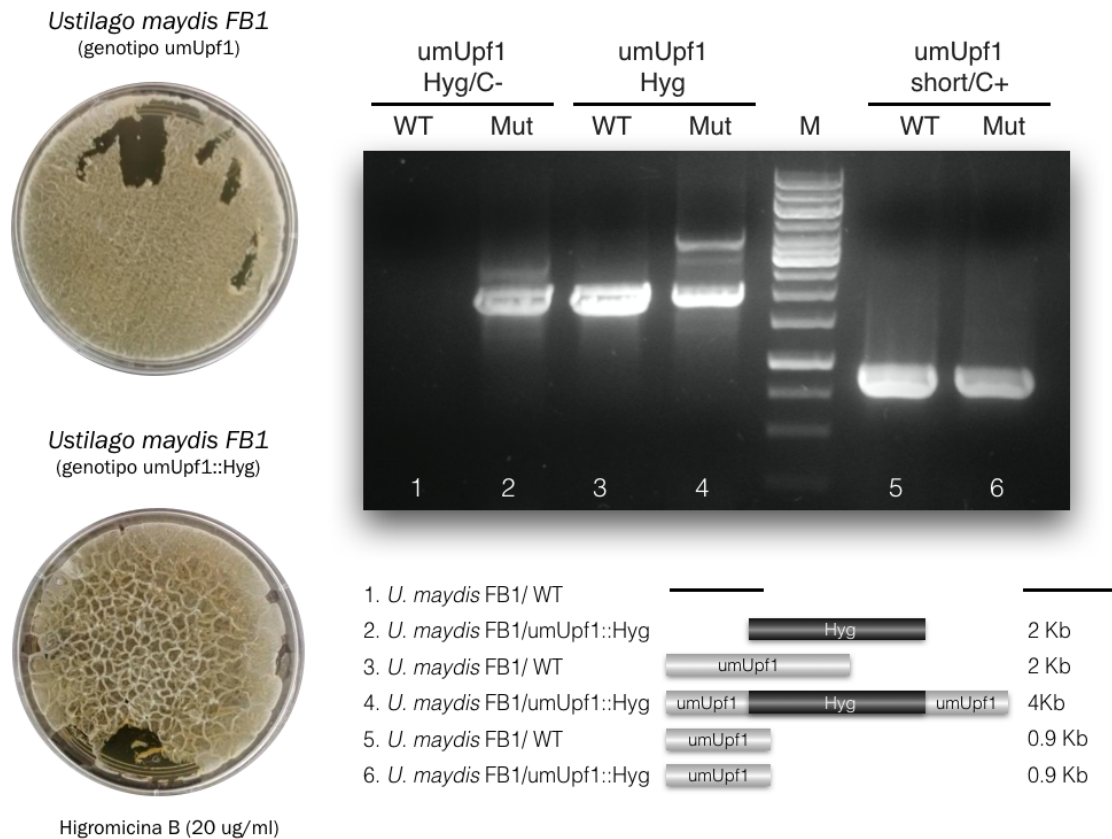


Fig. 32. Interrupción de *Upf1* en la cepa FB1 de *Ustilago maydis*. A la izquierda se muestran las cepas en medio sólido, tanto la cepa silvestre de *Ustilago maydis* FB1 (WT, arriba), como la cepa obtenida después de la electroporación y posterior crecimiento en medio de selección (20 ug/ml de Higromicina, Mut, abajo). A la derecha se presenta la imagen de un gel de agarosa al 1% donde podemos observar los productos de PCR amplificados a partir de ambas cepas WT y Mut. En el carril 1 y 2 el amplicon corresponde al cassette de Hyg (únicamente encontrado en la Mutante), en el carril 3 y 4 el amplicón se obtiene con los oligos FwdUpf1 y RevintronUpf1. En este caso en la WT únicamente se amplifica la región correspondiente al gen y para el caso de la mutante se amplifica además de dicha banda, una adicional correspondiente al generado a partir de la inclusión del gen de Higromicina. En los carriles 5 y 6 se muestra el amplicón para la región CH de *umUPF1* contenido en ambos fondos génicos. Abajo del gel se muestra la representación esquemática para describir el origen de las muestras, los tamaños esperados y las regiones que comprenden los amplificadas.

Las transformantes recuperadas a baja concentración de higromicina se enfrentaron ahora a la concentración inicial de 200 ug/ml y bajo esta condición se obtuvo un ligero crecimiento, el cual pudo emplearse para la extracción de ADN total. Este ADN se utilizó como templado para nuevos ensayos de PCR que permitieron confirmar la interrupción del gen *Upf1* con el cassette de higromicina (Fig. 33) y en este caso se obtuvo un producto único, sugiriendo que la población era homogénea y presentaba la mutación. Esta mutante no sobrevivió a un pase más en presencia de higromicina, después de este ensayo ya no era viable. Con estas observaciones podemos sugerir que *Upf1* es importante para la sobrevivencia del hongo y que la interrupción de dicho gen tiene un efecto subletal en el organismo. Debido a esta característica, ya no fue posible realizar ensayos posteriores con la mutante.

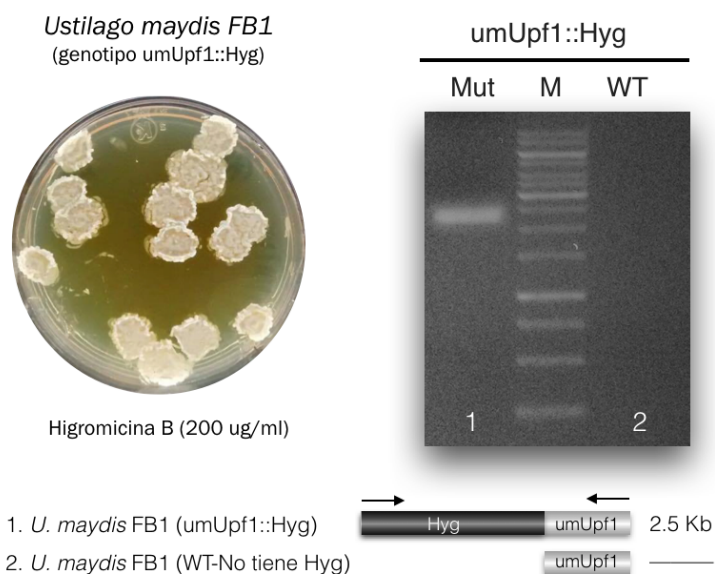


Fig. 33. Análisis de la mutante *U. maydis* umUpf1::Hyg. A la izquierda se muestra el crecimiento para la cepa mutante en medio YEPS adicionado con higromicina B a una concentración de 200 ug/ml y se puede observar que el crecimiento fue escaso. A la derecha se presenta la imagen de un gel de agarosa al 1% donde podemos observar los productos de PCR amplificados a partir de ambas cepas WT y Mutante. En el carril 1 y 2 el amplicón corresponde al a una región que abarca el cassette de Hyg y una región del gen umUpf1 (únicamente encontrado en la Mutante), en el carril 2 no se observa amplicon debido a que la reacción se llevó a cabo con el oligo fwd del cassette de hyg. Abajo se muestra la representación esquemática para describir el origen de las muestras, los tamaños esperados y las regiones que comprenden los amplificados.

VI.II.IV. ANÁLISIS DE LA INTERACCIÓN ENTRE LOS FACTORES DEL NÚCLEO DE NMD

Como hemos estudiado hasta ahora, los factores del UPF1, UPF2 y UPF3 son fundamentales para que se lleve a cabo el NMD y la interacción de dichos factores es de vital importancia en el sistema para el reconocimiento y degradación de transcritos anómalos. Bien sabemos que el dominio CH de UPF1 es el sitio de unión de la proteína UPF2 (Kervestin & Jacobson, 2012) y que este dominio se localiza en el extremo N-terminal de la proteína (Imamachi *et al.*, 2012). En diversos estudios se ha observado que este dominio es suficiente no solo para que ocurra la interacción con UPF2, sino que además es determinante para la función reguladora de UPF1 en la vía. Para el caso de la UPF3 las regiones reconocidas al centro de ambas proteínas hace posible la interacción (Kervestin & Jacobson 2012) como se observa en la figura 3.

Con estos antecedentes, decidimos amplificar la secuencia correspondiente al dominio CH en umUPF1 para generar una proteína recombinante que correspondiera sólo a este dominio, así, se esperaba que esta proteína recombinante fuera suficiente para realizar estudios de interacción con UPF2. Por otro lado, la proteína recombinante que incluye el MAL completo para UPF1 podría ser de utilidad en

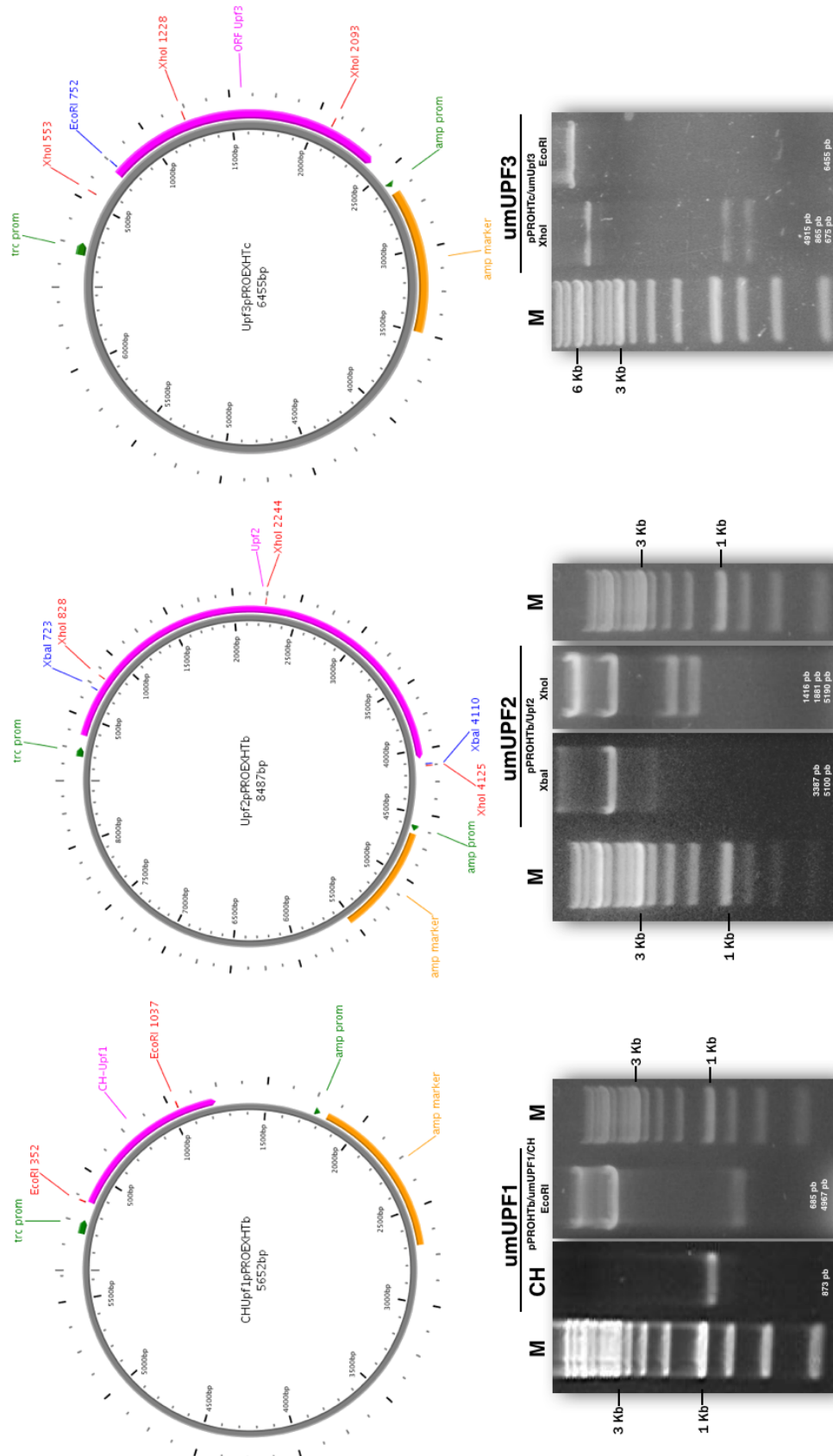


Fig. 34. Construcción de los vectores para las proteínas recombinantes *umUPF1-CH*, *umUPF2* y *umUPF3*. En la parte superior se muestran los mapas para cada una de las construcciones, indicando su peso y las principales características del plásmido. En la parte inferior se muestra el amplicón obtenido para CH y en cada caso el ensayo de restricción que sirvió para confirmar la clonación del inserto y la orientación de la clona. Los tamaños para las bandas obtenidas se indican en cada carril.

ensayos funcionales posteriores. En *Ustilago maydis* el dominio CH tiene un tamaño de 112 aminoácidos y abarca del residuo 129 al 241 aa; los oligonucleótidos utilizados para amplificar este dominio son específicos para *U. maydis* y amplificaron los primeros 873 pb del gen iniciando en el ATG, región que abarca el dominio CH.

Entonces para poder realizar ensayos de interacción, se construyeron las proteínas recombinantes para cada uno de los factores del núcleo del NMD y así, poder estudiar la interacción que éstas pudieran tener entre sí. La estrategia fue amplificar por PCR los marcos abiertos de lectura completos para umUpf2 y umUpf3 (previamente identificados en apartados anteriores) y para el caso de umUmpf1 la región CH. Estos amplicones se clonaron en el vector pPROEx-HT, el cual permite obtener una proteína recombinante que presenta una etiqueta de 6 histidinas en el extremo N-terminal. Una vez obtenidas las clonas se realizó la verificación por corte enzimático (figura 34) y se secuenciaron para verificar su identidad y orientación. Habiendo corroborado la inserción del MAL para cada una de las proteínas, asegurándonos de que la inserción fuera adecuada y en fase, se implementaron las condiciones para la inducción y purificación de las proteínas recombinantes. Dependiendo del tamaño de la proteína se corrieron en geles de poliacrilamida desnaturalizantes con una concentración del 8 al 12%. Los resultados representativos obtenidos en estos experimentos se presentan en la figura 35.

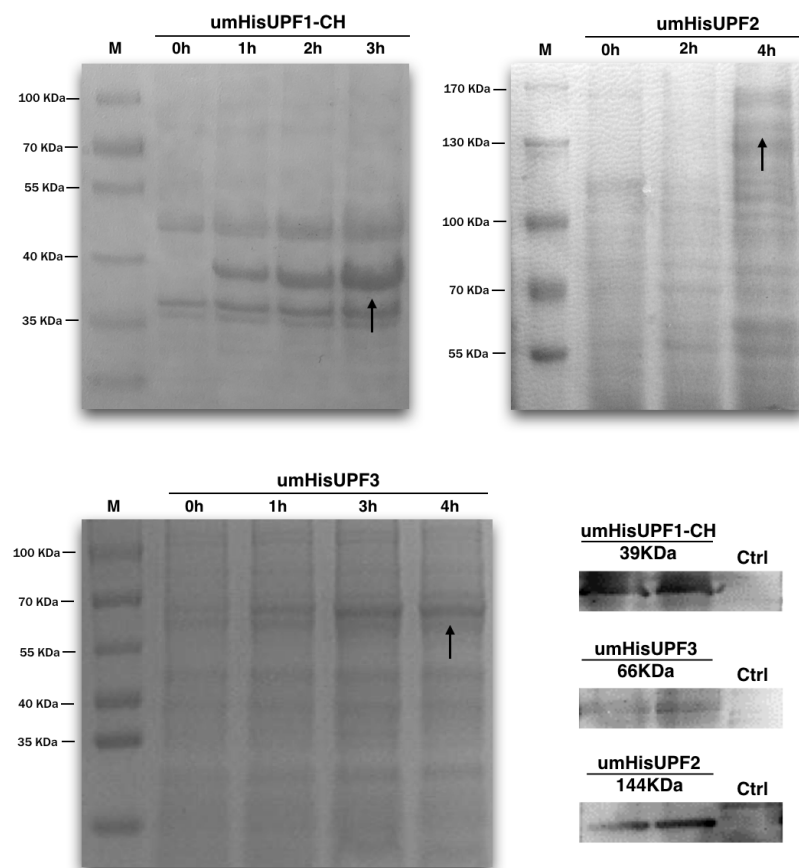


Fig. 35. Proteínas recombinantes homólogas a UPF1, UPF2 y UPF3 de *Ustilago maydis*. Se muestran las imágenes de los geles de poliacrilamida teñidos con azul de Coomassie en donde se verificó la inducción para cada una de las recombinantes. En todos los casos, el primer carril corresponde al marcador de peso molecular (PageRuler, Fermentas) seguido de las diferentes tiempos de la inducción. La banda correspondiente a la proteína recombinante se muestra con una flecha negra. En el recuadro inferior derecho se muestra el resultado del Western Blot donde podemos observar que las proteínas recombinantes son identificadas por el anticuerpo anti-His (SantaCruz). Como control negativo (Ctrl) se utilizaron

Para la obtención de las proteínas recombinantes se probaron diferentes condiciones de temperatura, tiempo, concentración de IPTG y sonicación (intensidad e intervalos); para el caso de umHisUPF1-CH en todas las condiciones pudo observarse sin problema alguno la inducción de una banda de 40 KDa que iba aumentando de concentración en cuanto al tiempo. La proteínas recombinantes homólogas a UPF2 y UPF3 presentan un tamaño aproximado de 141 y 66 KDa respectivamente y para el caso de umHisUPF2 y umHisUPF3 no fue tan evidente la inducción, sin embargo con la inducción se observa la aparición de una banda del tamaño esperado. Para comprobar que se tratara de las proteínas recombinantes, se realizó un *Western Blot* empleando un anticuerpo que reconoce la cola de poli-histidinas (Fig. 35). La identidad de las recombinantes generadas también se comprobó con el anticuerpo específico para cada proteína (datos no mostrados) y que en apartados anteriores habíamos comprobado que era capaz de reconocer nuestras proteínas.

Para la obtención de las proteínas recombinantes se probaron diferentes condiciones de temperatura, tiempo, concentración de IPTG y sonicación (intensidad e intervalos), para el caso de umHisUPF1-CH en todas las condiciones pudo observarse sin problema alguno la inducción de una banda de 40 KDa que iba aumentando de concentración en cuanto al tiempo. La proteínas recombinantes homólogas a UPF2 y UPF3 presentan un tamaño aproximado de 141 y 66 KDa respectivamente. Para el caso de umHisUPF2 y umHisUPF3 no es tan evidente la inducción, sin embargo, en el gel de poliacrilamida se logran ver bandas del tamaños esperados pero para comprobar que se tratara de nuestras proteínas se realizó un *Wester Blot* con anticuerpos contra la cola de poli-histidinas (Fig. 35) y con el anticuerpo específico para cada proteína (datos no mostrados) donde pudimos observar de manera más segura que la inducción de nuestras proteínas se había realizado de manera exitosa.

Una vez obtenidas las proteínas recombinantes (umHisUPF1-CH, umHisUPF2 y umHisUPF3) se procedió a investigar la interacción de umHisUPF1-CH y umHisUPF2. Este ensayo se basa en la interacción de proteínas por medio de recombinantes marcadas por el tracto de poli-histidinas, inicialmente una proteína (“anzuelo”) se anclan a una matriz de cobalto, lo que permiten su interacción con otras (“presas”) ya sea un conjunto o una única proteína, esta metodología nos permite unir de 10 a 25 mg de proteína marcada por ml de resina, lo que hace necesaria la identificación de dichas proteínas con tinción de plata.

Durante la realización del experimento se tomo como “anzuelo” a la umHisUPF1-CH y como “presa” a umHisUPF2 y se ajustaron las condiciones de tiempo para estandarizar la interacción. Los resultados de este ensayo se pueden observar en la figura 36 donde en el primer carril se observa a ambas proteínas y del carril 2 al 7 los controles para el experimento. Como se esperaba de acuerdo a los detalles del experimento y a los resultados obtenidos hasta el momento la interacción de ambas prometinas se llevo a cabo en el ensayo in vitro. De donde podemos suponer con este y otros resultados que la proteína umUPF1 es funcional y se une a umUPF2.

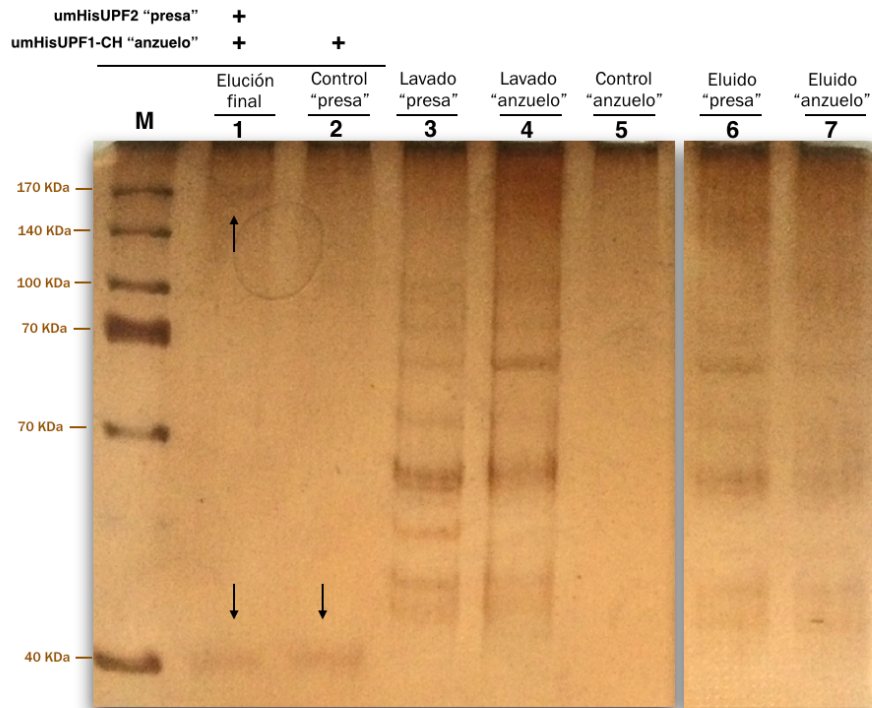


Fig. 36. Análisis de la interacción entre umHisUPF1-CH y umUPF2. Gel de poliacrilamida al 8% teñido con plata donde podemos observar en el carril 1 ambas proteínas el "anzuelo" y la "presa" comprobando la interacción de ambas umHisUPF1-CH y umUPF2. En los siguientes carriles los controles del experimento.

VI.III. ANÁLISIS COMPARATIVO DE LA MAQUINARIA DE NMD EN DIFERENTES ESPECIES

Los datos obtenidos a partir de los resultados bioinformáticos fueron muy valiosos ya que nos brindaron un panorama más amplio de las posibles estrategias para en el abordaje experimental y nos ayudaron a entender mejor el sistema y a las proteínas que actúan en el.

Se analizó el genoma de 299 eucariotes, entre los que se encuentran 99 metazoarios, 48 plantas, 109 hongos y 43 protistas. La distribución de los diferentes factores de NMD fue analizada y se presenta resumida en la tabla se encuentra en los anexos. A continuación se muestran los resultados en relación a los diferentes grupos de factores involucrados en el proceso de NMD.

VI.III.I. EL NÚCLEO CENTRAL DE LA VÍA DE NMD

En *S. cerevisiae*, el núcleo del NMD son las proteínas UPF1, UPF2 y UPF3 (Leeds *et al.*, 1991; Leeds *et al.*, 1992) y los mismos factores se han caracterizado también en *H. sapiens* (Perlick *et al.*, 1996; Applequist *et al.*, 1997; Mendell *et al.*, 2000; Serin *et al.*, 2000; Lykke-Andersen *et al.*, 2000; Pal *et al.*, 2001; Denning *et al.*, 2001; Yamashita *et al.*, 2001). En *C. elegans*, los homólogos para dichas proteínas son SMG2, SMG3 y SMG4 (Aronoff *et al.*, 2001; Page *et al.*, 1999). En la mayoría de las especies analizadas, se identificaron homólogos al menos para una proteína UPF en la mayoría de las especies. El grupo de organismos que carece de homólogos para estas proteínas corresponde a los protistas, en donde 12 especies diferentes no presentaron homólogo para ninguna proteína UPF.

Puesto que UPF1 tiene un papel clave en la ruta de NMD, resulta lógico que esta proteína esté presente en la mayoría de los organismos. Efectivamente, de los 299 organismos analizados, 265 presentan homólogo para UPF1. En metazoarios, UPF1 no fue identificada en 9 especies de mamíferos, incluyendo el conejo (*Oryctolagus cuniculus*) y el panda gigante (*Ailuropoda melanoleuca*), mientras que solo 2 especies de plantas no presentan homólogo para UPF1: un helecho y un alga roja. En hongos, 4 ascomicetos y 4 basidiomicetos aparentemente carecen de UPF1 y aproximadamente el 30% de los protistas no mostraron homólogo para dicha proteína.

En cuanto a UPF2, también se encontraron homólogos para metazoarios, a excepción de *Gorilla gorilla*, *Meleagris gallopavo* (pavo) y *Drosophila melanogaster*; la única planta que carece de UPF2 es *Oryza sativa japonica*, mientras que en hongos no se identificó a UPF2 para 2 ascomicetos (*Candida tenius* y *Tuber melanosporum*), así como para diferentes microsporidios, los cuales solo presentaron UPF1 y carecen del resto de la maquinaria de NMD, lo cual podría relacionarse con su naturaleza parasitaria.

En el caso de UPF3, se encontró un homólogo para todas las especies de mamíferos, aves, reptiles y peces secuenciados a la fecha y solo parece estar ausente en 6 especies animales, incluyendo *Xenopus laevis*. Con respecto a los hongos, no se identificó a UPF3 en 4 ascomicetos y 3 basidiomicetos, así como en los mismos microsporidios que carecen de UPF2.

VI.III.II. LAS PROTEÍNAS SMG

Homólogos para las proteínas SMG fueron identificadas prácticamente en todos los genomas analizados; algunas excepciones corresponden al chimpancé pigmeo (*Pan paniscus*) que carece de SMG1, el oso polar (*Ursus maritimus*) que solo tiene SMG8 y SMG9, así como *Ornithorhynchus anatinus* (ornitorrinco) que no tiene SMG5, confirmando que la regulación por fosforilación de la vía de NMD es empleada por la mayoría de los animales.

Todos los metazoarios mostraron un homólogo para SMG8, a excepción de *Xenopus laevis* que carece de SMG5, SMG6 y SMG8; mientras que ciertas especies no tiene SMG9, como *Caenorhabditis briggsae* y *Schistosoma mansoni*. En estos organismos, la actividad de SMG1 podría ser regulada por un mecanismo alternativo, mientras que la supresión de la actividad cinasa podría no ser necesaria. En general, los insectos carecen de SMG8 y SMG9, mientras que *Anopheles gambiae* y 12 diferentes especies de *Drosophila* carecen de SMG7, SMG8 y SMG9.

Las proteínas SMG5, SMG6 y SMG7 están involucradas en la defosforilación de UPF1 y al analizar los diferentes genomas encontramos que estas proteínas SMG son variables en eucariotes superiores. En este sentido, vertebrados y nemátodos conservan las 3 proteínas, que las moscas carecen de SMG7 y que las plantas presentan dos SMG7 aparentemente no-redundantes, al mismo tiempo que carecen de SMG5 y SMG6.

VI.III.III. EL COMPLEJO EJC

Las proteínas que conforman el complejo EJC se encuentran bien representadas en las diferentes especies analizadas, siendo EIF4A3 el factor más prevalente, ya que se encontró en 286 especies, lo cual resulta lógico considerando los diversos procesos en los que se involucra este factor además de NMD, incluyendo el inicio de la traducción, el ensamble del ribosoma y del *spliceosoma* (Zhang & Krainer, 2007). Por otro lado, Y14 y MAGOH fueron identificadas en 243 y 254 organismos, respectivamente y las especies que carecen de estos factores son principalmente hongos.

En general, los componentes del EJC prevalecen en la mayoría de las especies, pero en algunos casos algún factor no fue identificado. MAGOH parece ser un factor importante para la vía de NMD en metazoarios, ya que se identificó un homólogo para este factor en todas las especies animales. Y14 parece no ser indispensable, considerando que está ausente en 10 especies animales. De igual manera, MLN51 está ausente solo en 9 especies animales. Fuera del grupo de los animales, MLN51 fue encontrado solo en dos plantas y no lo contiene ningún hongo ni protista.

Las proteínas del EJC descritas hasta ahora se encuentran recubiertas por RNPS1 y REF/Aly, las cuales a su vez conforman junto con SAP18, PININ y ACIN1, la capa externa del complejo EJC. Estas proteínas mostraron una distribución homogénea en animales y plantas, a excepción de ACIN1, el cual no es prevalente en animales. Las proteínas de la capa externa del EJC no fueron encontradas en los Saccharomycetes, además de que no se encontraron homólogos para PININ y ACIN1 en hongos, mientras que los protistas prácticamente no poseen estos factores, sugiriendo que el proceso de NMD mediado por EJC no se lleva a cabo en estos organismos. La única excepción corresponde a *Phytophthora sojae*, en donde la única proteína del EJC que no identificamos fue MLN51.

Existen otras proteínas que interactúan con el EJC de manera transitoria, como UAP56, TAP:p15 y SRm160 (Tange *et al.*, 2005). Estos factores fueron identificados en la mayoría de las especies del grupo de los animales, a excepción de un pequeño grupo de animales y otros casos aislados. En plantas, TAP:p15 parece estar ausente. Encontramos el homólogo para p15 en hongos, excepto para 5 Saccharomycetes

Finalmente, PYM es otro transportador de ARN que interactúa con MAGOH y Y14, además de que favorece la traducción del ARNm que ha sufrido *splicing* (Diem *et al.*, 2007). En este trabajo se encontró que PYM está ausente únicamente en 2 especies de animales, en algunas plantas y en la mitad de las especies de hongos, exceptuando a Saccharomycetes y microsporidios.

VI.III.IV. FACTORES DEL PROCESAMIENTO DEL ARNm INVOLUCRADOS EN NMD

El factor eRF1 fue identificado en la mayoría de los animales, a excepción de 4 especies dentro de las que se incluye el panda gigante (*Ailuropoda melanoleuca*), mientras que eRF3 no se encontró en el pez *Takifugu rubripes* y en el mosquito de la fiebre amarilla (*Aedes aegypti*).

Por otro lado, se ha descrito la presencia de CBPC como un rasgo necesario para identificar a los transcritos que son blancos para NMD, como se mencionó en la introducción, de donde resulta interesante identificar a CBP20 y CBP80 en las diferentes especies. Para metazoarios, CBP80 no fue identificado en cerdo (*Sus scrofa*), pez zebra (*Danio rerio*), ascidia (*Ciona intestinalis*) y esponja (*Amphimedon queenslandica*), mientras que CBP80 no pudo identificarse en *Gorilla gorilla*, *Ornithorhynchus anatinus* y *C. elegans*. Es probable que en estos organismos una CBP sea suficiente para el procesamiento del ARNm. En general, CBP20 es el más prevalente de los dos factores y se ha encontrado incluso en protistas. Por otro lado, CBP80 se encontró en todas las plantas, a excepción del alga verde *Micromonas pusilla*; los hongos presentan un homólogo para CBP80 a excepción de los microsporidios y casi todos los protistas carecen de dicho factor.

Otro factor accesorio de NMD es PABP. En este análisis se identificó PABP2 en 297 de las 299 especies estudiadas y solo 2 protistas carecen de este factor. Por el contrario, PABP1 no fue identificada en 32 especies, incluyendo 5 animales, 3 plantas, 10 hongos y 14 protistas.

VI.III.V. VÍA DE DEGRADACIÓN

Para completar la vía de NMD, se requiere de la participación de complejos y factores de degradación (Conti & Izaurralde, 2005). A continuación se presenta el análisis de estas rutas en las diferentes especies.

El exosoma es un complejo multiproteico conservado que participa tanto en la maduración de moléculas estables de ARN, así como en la degradación de ciertas partículas de ARN. Este complejo fue descubierto inicialmente en *S. cerevisiae* y está compuesto por nueve proteínas, seis de las cuales son homólogas para la RNasa PH bacteriana (Rrp41, Rrp42, Rrp45, Rrp46 y Mtr3); los otros 3 componentes contienen un dominio putativo S1 de unión para ARN (Rrp4, Rrp40 y Csl4) (Raijmakers, et al., 2004; Allmang et al., 1999; Mitchell & Tollervey, 2000). El exosoma ha sido identificado en humano, plantas, parásitos, moscas y Arqueas y algunos factores asociados al exosoma también han sido identificados (Lrp1, Ski7, Ski2, Ski3, Ski8, Mtr4, Gsp1 y Nip7).

En este estudio se identificó a los componentes del exosoma en los diferentes grupos de organismos con una distribución homogénea. Una marcada excepción corresponde al componente MIR3, el cual no se identificó en 36 animales, 6 plantas, 6 hongos y en casi la mitad de los hongos. El resto de los componentes del exosoma presenta homólogo en la mayoría de las especies, salvo excepciones

particulares. Casos especiales corresponden al hongo *Postia placenta*, que no presenta ninguno de estos factores y *Monilophthora perniciososa*, en donde solo encontramos homólogo para RRP44.

Otros factores de la vía de degradación en el sentido 3'→5' son el complejo TRAMP (Trf4/Trf5, Air1/Air2, Mtr4), el complejo CCR4/NOT, el complejo PAN (Pan2/Pan3), así como las proteínas DCP5 y PARN.

El complejo TRAMP está conformado por los factores Trf4/Trf5, Air1/2 y Mtr4. Estas proteínas presentan una distribución variable en los diferentes organismos. La más prevalente es Mtr4, ausente en solo 6 organismos. Las proteínas Air muestran una distribución discreta, ya que no se identificaron en 4 vertebrados y están ausentes en ciertos grupos como el de los insectos, el de los nemátodos, el de las plantas y el de los Eurotiomicetos, así como en la mayoría de los protistas. Finalmente, Trf4/5 no presenta homólogo en animales ni en plantas y sólo se encontró en 12 especies de Saccharomycetes. Estas diferencias podrían ayudar a explicar las particularidades del sistema de NMD en dichos grupos.

Para el complejo CCR4/NOT se identificaron ambas proteínas en la mayoría de los organismos analizados, de tal manera que CCR4 no se identificó en 9 especies de *Drosophila*, en el nemátodo *C. elegans*, el hongo *Postia placenta* ni en 17 protistas. Por su parte, NOT no parece estar presente en 6 especies, entre las que se encuentra *X. laevis*.

El complejo PAN también es altamente prevalente y no se pudo identificar en alrededor de 70 especies. De los dos componentes de este complejo, PAN3 se identificó en 96 especies de animales y en 100 hongos, mientras que se pudo ubicar solo en 7 plantas y 17 protistas. Por otro lado, PAN2 se encontró en 93 animales, en 100 hongos, en 24 protistas y las únicas plantas que la contienen son musgos y algas. Cabe señalar que algunos hongos que no presentan a esta proteína son los microsporidios, los cuales carecen de varios factores, como se ha descrito anteriormente.

La proteína DCP5 presentó una amplia distribución y parece ser relevante en el grupo de animales, ya que sólo parece estar ausente en 2 especies de este grupo; lo mismo se observa para los hongos, ya que no se encontró en 5 especies de hongos. No ocurre lo mismo para las plantas y los protistas, ya que el homólogo para este factor se identificó solamente en 7 y 10 especies para dichos grupos, respectivamente. Por el contrario, el factor PARN no fue tan prevalente, ya que se encontró solamente en 162 organismos. Dentro de las especies en las que no se identificó a PARN se encuentran las especies de *Drosophila*, todos los nematodos y 27 hongos, incluyendo a *S. pombe*.

VI.III.VI. LA REMOCIÓN DEL CAP

Esta desprotección es mediada por el factor DCP2, el cual está presente en casi todos los organismos investigados, a excepción de 3 especies animales, 5 plantas incluyendo *Oryza sativa japónica*, 4 especies de hongos y 10 protistas, lo cual correlaciona con la importancia de este proceso para organismos eucariotes.

Una vez que se remueve el Cap, las exonucleasas XRN1 y XRN2 se activan para la degradación del transcrito que fue reconocido por la maquinaria de NMD. De estos dos factores, el más prevalente es XRN2, el cual no se encontró en tan solo 5 animales, 2 plantas, 21 hongos y 4 protistas. Por el contrario, XRN1 solo fue identificado en 41 animales, 3 plantas, 31 hongos y 20 protistas. De manera interesante, un total de 27 especies no presenta homólogo para ninguna de las 2 endonucleasas, sugiriendo que en estos organismos la degradación del ARNm ocurre por otras vías.

Factores generales de la eliminación del Cap son Ddx6 y Pat1 los cuales se encontraron de manera general en casi todos los organismos.

Finalmente, un agregado indispensable para activar la eliminación del Cap es el complejo Lsm, involucrado en NMD y otras vías de degradación (Coller & Parker, 2004). En este análisis, se encontraron homólogos para los componentes de este complejo en la mayoría de los organismos. Varias especies pueden carecer de una a cuatro de las 8 proteínas Lsm, aunque ninguno de ellos pertenece al grupo de los animales. Fuera de dicho grupo, encontramos algunos casos en los cuales varias Lsm no fueron identificadas. Por ejemplo, los organismos que solo tienen 3 proteínas Lsm son *Ostreococcus tauri* (planta), los hongos *Clavospora lusitaniae* y *Verticillium alfalfae*; además de los protistas *Cryptosporidium hominis*, *Tetrahymena thermophila* y *Naegleria gruberi*. Cabe resaltar que algunos hongos y protistas carecen incluso de más proteínas Lsm. Tal es el caso del hongo *Monilophthora perniciososa* en donde solo encontramos homólogos para Lsm1 y Lsm8, así como de los protistas *Trypanosoma brucei* y *Leishmania braziliensis*, los cuales presentan solo Lsm4 y Lsm7. Casos particulares corresponden a *Nannochloropsis gaditana*, que sólo tiene Lsm7 y *Nosema ceranae*, en la cual no se pudo identificar ningún homólogo para proteínas Lsm, aunque estas dos especies carecen de muchos de los factores involucrados en decaimiento, por lo que es posible proponer que este mecanismo no sea necesario para dichas especies.

VII. DISCUSIÓN

VII.1. Disertación de la vía de vigilancia del ARNm

Desde el descubrimiento de ambos fenómenos, el de splicing por Sharp en 1977 y NMD por Losson y Lacroute en 1979, se han hecho estudios para conocer cómo se lleva a cabo la regulación de estos mecanismos y aunque aún se encuentran pobremente entendidos, se conocen aspectos esenciales de ambos modelos. Los mecanismos detallados mediante los cuales se regulan estos y otros procesos aún no han sido dilucidados en su totalidad, pero los hallazgos descritos en diferentes organismos sugieren que estos procesos se encuentran altamente conservados en la escala evolutiva, muy probablemente por ser esenciales en las eucariotas.

Actualmente existen dificultades técnicas que no permiten conocer en detalle la regulación de procesos co- y postranscripcionales relacionados con la maduración del ARN mensajero y la regulación de la expresión genética en humanos. Una solución para este tipo de problemas lo constituye generalmente el empleo de un modelo que permita extrapolar conclusiones a organismos superiores. *Ustilago maydis* constituye una alternativa atractiva para el estudio de la regulación del splicing y del decaimiento del ARNm, así como de la modificación y degradación de proteínas ya que existe toda una serie de eventos celulares y moleculares conservados entre éste y el humano como se describió anteriormente. Dicha similitud ofrece una ventaja importante que nos puede permitir conocer el funcionamiento de la maquinaria responsable de regular estos eventos en humano, así como los mecanismos encargados de dicha actividad.

Los estudios bioinformáticos de la ruta vigilancia del ARNm en *Ustilago maydis* nos muestran la cercanía de estos dos organismos en cuanto a dichos procesos. En el caso particular de NMD, el análisis mostró la presencia de las proteínas conservadas para los factores centrales de NMD con un alto grado de homología en sus dominios funcionales, cuya función demostrada en humano y otros organismos presentan un homólogo en *Ustilago maydis* por lo que creemos que podría ser una proteína funcional para el mecanismo en el hongo. Las proteínas identificadas incluyen homólogos del mecanismo que involucra al EJC como parte principal en el reconocimiento de los codones prematuros así como en la coadyuvación del exporte del transcrito del núcleo al citoplasma, mecanismo que presentan principalmente los mamíferos. Por otro lado también se encontraron las proteínas capaces de marcar y llevar a degradación al transcrito que contiene secuencias río abajo del codón de paro ricas en serina (DSE) importantes para el reconocimiento del transcrito por Hrp1 y ésta a su vez por la proteína UPF1 (Figura 12), mecanismo descrito para *S. cerevisiae*. Los avances recientes en nuestra comprensión de estos procesos biológicos en *Ustilago maydis* muestran que el hongo podría ser una herramienta eficaz para la investigación de estos mecanismos de manera sencilla y menos costosa y que además los resultados podrían extrapolarse a lo que ocurre en mamíferos, pues el organismo está más relacionado con las células animales de lo que se pensaba anteriormente.

Las evidencias mostradas anteriormente dieron paso a estudios específicos sobre los factores principales de NMD: UPF1, UPF2 y UPF3 y sus homólogos en *Ustilago maydis*. Estas tres proteínas clave en el mecanismo, inicialmente descritas en *Saccharomyces cerevisiae* y posteriormente en otros eucariotas superiores (Leeds y cols, 1992; Cui y cols, 1995) han demostrado tener un alto grado de homología en todos los organismos con genomas descritos, donde pocos de estos organismos no cuentan con alguno de los tres factores principales y únicamente en algunos protozoarios resulta ausente el mecanismo.

En este trabajo analizamos *in silico* diversas características estructurales y bioquímicas de umUPF1. Inicialmente identificamos todos los dominios característicos descritos para UPF1 en el homólogo putativo para *U. maydis*, los cuales se resumen en la figura 1. Cuando alineamos este homólogo con la secuencia reportada para hUPF1, encontramos rasgos característicos adicionales que se comparten entre hUPF1 y umUPF1 (Fig. 2), incluyendo el alto grado de identidad para la secuencia completa. De manera importante, la identidad se incrementa en los dominios funcionales, siendo el dominio CH el más conservado (75%). Esta comparación también resaltó otros motivos importantes como los dominios conservados en las helicasas SF1/2, el asa 349-355 y el dominio SQ. Todos estos motivos podrían ser responsables de funciones e interacciones que involucren a umUPF1.

Un rasgo importante que regula la conformación y funcionalidad de UPF1 se relaciona con su estado de fosforilación. En este trabajo identificamos 12 sitios de fosforilación conservados en *H. sapiens* y *U. maydis*. Estos sitios incluyen a los aminoácidos S/Y/T, en donde la mayoría fueron residuos de serina localizados principalmente hacia el extremo C-terminal, mientras que solo identificamos algunos repetidos SQ en umUPF1. Algunos de estos residuos, como umS1042 y umS1081 han sido reportados previamente como sitios de fosforilación funcionales que participan en la actividad de UPF1. Cuando los aminoácidos equivalentes a estas posiciones son mutadas en humano (hS1089 y hS1127), se elimina la fosforilación de UPF1, así como su capacidad para interactuar con UPF2, SMG1, SMG7 (Yamashita *et al.*, 2001; Kashima *et al.*, 2006). Aunque el mecanismo responsable de este efecto no se ha dilucidado completamente, se sabe que UPF1 es fosforilada por SMG1 y que se requiere que se complete un ciclo de fosforilación y desfosforilación completo para que el proceso de NMD se active (Yamashita *et al.*, 2001; Paillusson *et al.*, 2005). También se ha demostrado que la activación de SMG5, SMG6 o SMG7 ocasiona la acumulación de UPF1 fosforilada e incrementa la cantidad de ARNm que contienen PTCs en mamíferos (Kashima *et al.*, 2006; Paillusson *et al.*, 2005) y en *C. elegans* (Pulak & Anderson, 1993; Page *et al.*, 1999). Por otro lado, cuando se inhibe la desfosforilación de hUPF1 se suprime la vía de NMD (Okada-Katsuhata *et al.*, 2012; Ohnishi *et al.*, 2003). Con todas estas consideraciones, sería posible hipotetizar que el proceso de NMD también es regulado mediante eventos de fosforilación en *U. maydis*, tal como se ha descrito para el proceso en humano.

En nuestro estudio también descubrimos once blancos de ubiquitinación que se encuentran conservados entre *H. sapiens* y *U. maydis*. A pesar de que solo encontramos un reporte en el que se sugiere que la ubiquitinación de UPF1 puede estar involucrada en la regulación del proceso de NMD

(Takahashi *et al.*, 2008), sería interesante abordar la importancia de esta modificación postraduccional en estudios futuros.

UPF1 interactúa principalmente con UPF2 (Mendell *et al.*, 2000; Yamashita *et al.*, 2001; Kadlec *et al.*, 2006; Kashima *et al.*, 2006; Chamieh *et al.*, 2008; Clerici *et al.*, 2009; Chakrabarti *et al.*, 2011) y desde el inicio de nuestro estudio encontramos que el dominio responsable de la interacción entre UPF1 y UPF2 se encuentra altamente conservado en *U. maydis*. Por otro lado, los homólogos putativos identificados en nuestro estudio para UPF1 y UPF2 poseen la capacidad de interactuar de acuerdo con las predicciones determinadas por la herramienta STRING. Otros factores capaces de interactuar con UPF1 incluyen al factor central de NMD UPF3 (Chamieh *et al.*, 2008), al factor de poliadenilación PABP (Ivanov *et al.*, 2008) y a los factores de liberación eRF1 y eRF3 (Kashima *et al.*, 2006; Ivanov *et al.*, 2008). En este análisis se identificaron también los homólogos en *U. maydis* para estos factores que participan en la vía de NMD y también se predijo su interacción con umUPF1 de acuerdo con el análisis de STRING. Además, se encontró que SMG1 de humano y el homólogo para mTOR en *U. maydis* comparten un dominio cinasa conservado (InterPro), sugiriendo que ambas proteínas podrían llevar a cabo la misma función. SMG1 también tiene la capacidad de interactuar con UPF1 (Kashima *et al.*, 2006; Yamashita *et al.*, 2013) y nosotros encontramos que el homólogo de mTOR también podría estar interactuando con umUPF1. En conjunto, estas observaciones sugieren que el homólogo de mTOR identificado podría llevar a cabo en *U. maydis* la función descrita para SMG1. UPF1 también tiene la capacidad de interactuar con otras proteínas SMG (Okada-Katsuhata *et al.*, 2012; Ohnishi *et al.*, 2003), los cuales también fueron identificados con el programa STRING en este estudio. De manera interesante, se predice la posible interacción de umUPF1 con la mayoría de los homólogos identificados en *U. maydis* que se presentan en la tabla 1, a excepción de p15 y PYM. Otro hallazgo relevante fue la identificación en *U. maydis* de homólogos para componentes del EJC, los cuales no se han descrito en otras levaduras como *S. cerevisiae*.

De acuerdo con las propiedades catalíticas que presenta y el mecanismo que involucra, las helicasas del tipo UPF1 pertenecen al grupo SF1/2 (Singleton *et al.*, 2007). Estas helicasas se unen a la región de cadena sencilla del ARN y desenrollan a esta molécula en un mecanismo facilitado por la hidrólisis de ATP (Jankowsky, 2011; Jarmoskaite & Russell, 2014). En humano, la actividad de helicasa de UPF1 se regula por la estructura de la proteína, la cual cambia de acuerdo a un gran cambio conformacional inducido por la interacción con UPF2 responsable de activar el interruptor de activación de UPF1 (Gleghorn & Maquat, 2011) descrito en la introducción. En este trabajo identificamos en umUPF1 a los residuos umV169, umF200 y umI241 presentes en el dominio CH, así como al aminoácido umI760 en el dominio RecA, los cuales parecen ser responsables de mantener la conformación cerrada de la proteína.

En este trabajo encontramos además las dos regiones hidrofóbicas del dominio CH, las cuales parecen ser responsables de la interacción con UPF2. La primera región hidrofóbica identificada en umUPF1 corresponde a la cavidad a la que se ancla UPF2 (Fig. 6A), la cual se ha visto que es determinante para

la interacción entre UPF1 y UPF2 (Kadlec *et al.*, 2006; Clerici *et al.*, 2009). Dentro de esta cavidad, las posiciones relevantes corresponden a los residuos umY192, umV212 y umV214. Se ha observado que en hUPF1, al mutar la posición hY184 (umY192) se elimina la interacción con UPF2, pero únicamente si la mutación se realiza de manera simultánea con hE182 (Kadlec *et al.*, 2006). La mutación independiente de hV204 (umV212) o de hV206 (umV214) también elimina la interacción entre UPF1 y UPF2 (Kadlec *et al.*, 2006). La posición intermedia entre estos dos aminoácidos mutados también corresponde a un residuo de valina conservado; sin embargo, la participación de este residuo en la interacción no ha sido probado experimentalmente. Estos tres residuos consecutivos de valina se encuentran conservados en vertebrados pero están ausentes en protistas en donde no se ha descrito el sistema de NMD (Kadlec *et al.*, 2006). Como en vertebrados, estos tres residuos también están conservados en *U. maydis*, sugiriendo su posible implicación funcional (Fig. 6B).

La segunda región hidrofóbica corresponde a la superficie que presenta a los residuos conservados umV169 y umF200. Los residuos equivalentes han sido mutados en hUPF1 (hV161, hF192) y como resultado de este cambio se pierde la interacción entre UPF1 y UPF2. El efecto más dramático se observó para la mutante hF192E, la cual presentó una elevada actividad de helicasa acompañada de una disminución en su capacidad para unir ARN en comparación con hUPF1 silvestre (Kadlec *et al.*, 2006). En esta mutante se vió afectada también la actividad ATPasa en donde los niveles catalíticos observados fueron semejantes a los que se observan cuando se deletó el dominio CH completo (Chakrabarti *et al.*, 2011). En este trabajo realizamos las mismas mutaciones *in silico* y en todos los casos se observó que se afecta la interacción con UPF2, sustentando la importancia de esta posición para la funcionalidad de UPF1.

En este estudio se determinaron *in silico* los principales aminoácidos que pudieran estar involucrados en la actividad de ATPasa y en la unión a ATP para umUPF1. Algunos de los aminoácidos identificados aquí han demostrado ser relevantes para la funcionalidad de hUPF1, los cuales se presentan a continuación. Los residuos umD645 y umE646 (hD647, hE648) se han mutado para el factor de humano y como resultado se observó una pérdida de la actividad ATPasa y helicasa (Battacharya *et al.*, 2000; Cheng *et al.*, 2007). Tanto la actividad de ATPasa como la capacidad para unir ATP se ven perjudicadas cuando las posiciones equivalentes a los residuos umR706 (hR714) o umR868 (hR876) son mutadas en el factor de humano (Cheng *et al.*, 2007). Por otro lado, se ha observado que la mutación del residuo hQ676 (umQ668) daña la actividad de ATPasa sin alterar la capacidad de unión a ATP (Cheng *et al.*, 2007). Finalmente, se ha observado que cuando se muta el residuo hK509 (umK507) se inhibe la degradación del ARNm para las histonas, así como la actividad ATPasa y la unión con ATP (Altamura *et al.*, 1992; Phillips *et al.*, 2005; Cheng *et al.*, 2007).

Aquellos residuos que han mostrado un fuerte impacto tanto en la capacidad para unir ATP como en la actividad de ATPasa fueron mutados *in silico* en este estudio. El modelo generado para umUPF1 (Fig. 9C) ilustra la manera en la que este ligando estabiliza una serie de interacciones entre los dominios RecA1 y RecA2 de umUPF1, las cuales se pierden al realizar la mutación (Fig. 9D). Los aminoácidos

característicos del Motivo I conservado en las helicasas del tipo SF1 (Fig. 9B) también son importantes para la funcionalidad de UPF1, por lo que al mutar los sitios h506-508 (um504-506) se ha observado que se inhibe la desforforilación de la proteína, la cual es direccionada a los cuerpos de procesamiento (Cho *et al.*, 2009).

En cuanto a la unión con ARN, los resultados obtenidos muestran que se forma un canal central en umUPF1 el cual permite la interacción con ARN. En este modelo, el dominio 1C bloquea un extremo del canal y esta conformación es semejante a la reportada para hUPF1 (Chakrabarti *et al.*, 2011).

Los estudios estructurales y bioquímicos constituyen instrumentos poderosos aunque presenta la desventaja de que nos permiten analizar moléculas estáticas. Por el contrario, las simulaciones de dinámica molecular representan estudios clave que facilitan el estudio de una molécula en un ambiente dinámico. Como se observa en este trabajo, ambos análisis se complementan para realizar un análisis integral. En este sentido, las simulaciones de dinámica molecular nos permitieron concluir que umUPF1 es una proteína estable y compacta, características que sugieren que esta proteína podría ser funcional.

En conjunto, nuestro análisis *in silico* mostró que los aminoácidos funcionales descritos para UPF1 se encuentran conservados en *U. maydis*, sugiriendo que estas posiciones podrían ser activas también para el homólogo putativo de *U. maydis*. Las semejanzas estructurales observadas para UPF1 de *H. sapiens* y de *U. maydis* sugieren fuertemente que estos dos homólogos podrían presentar funciones bioquímicas y catalíticas semejantes, indicando la posible actividad de umUPF1 como helicasa dependiente de ARN, ATPasa y reguladora de la vía de NMD. Este análisis nos permitió profundizar en la estructura y funcionalidad del homólogo estudiado, lo que hizo posible entender las características dinámicas de este homólogo mediante las simulaciones realizadas, así como las implicaciones de aminoácidos particulares gracias a las mutaciones realizadas *in silico*. Se espera que esta información pueda ser aprovechada para orientar los próximos experimentos *in vitro*. Al existir un mayor número de genomas completamente secuenciados, resulta atractiva la posibilidad de realizar estudios *in silico* más completos que proporcionen información relevante para el abordaje experimental.

VII.II. El núcleo de la maquinaria de NMD en *U. maydis*.

La vía de NMD se encuentra altamente conservada en los organismos eucariotes. Inicialmente, en *S. cerevisiae* se identificaron los 3 genes que constituyen el núcleo del NMD y que corresponden a Upf1, Upf2 y Upf3 (Leeds *et al.*, 1991; Leeds *et al.*, 1992). Los ortólogos respectivos identificados posteriormente en *C. elegans* fueron Smg2, Smg3 y Smg4 (Pulak & Anderson, 1993; Aronoff *et al.*, 2001). Ortólogos para estos genes fueron encontrados también para otros eucariotes en donde se ha explorado la vía de NMD, incluyendo *H. sapiens* (Perlick *et al.*, 1996; Applequist *et al.*, 1997; Mendell *et al.*, 2000; Serin *et al.*, 2001), *D. melanogaster* (Avery *et al.*, 2011) y *A. thaliana* (Arciga-Reyes *et al.*,

2006). Estas observaciones nos permitían suponer que los homólogos para el núcleo central de la maquinaria de NMD estarían presentes en *U. maydis*.

El análisis comparativo realizado como parte del estudio bioinformático al inicio de este estudio nos ayudó a determinar la identidad entre humano y *U. maydis* para los factores proteicos UPF1, UPF2 y UPF3, los cuales mostraron una identidad del 65%, 32% y 33%, respectivamente. Estudios similares se llevaron a cabo en *D. melanogaster*, en donde los homólogos para el núcleo de NMD se identificaron inicialmente de manera bioinformática (Culbertson & Leeds, 2003) pero su validación funcional se realizó posteriormente (Avery et al., 2011). En este sentido, una vez realizado in silico el estudio de las proteínas mencionadas, era necesario probar la presencia y expresión de estos factores en el basidiomiceto; para lograrlo, se realizaron tres verificaciones. Inicialmente se confirmó la presencia en el genoma de los homólogos identificados mediante ensayos de PCR a partir de ADN genómico de la cepa FB1 de *U. maydis*. Por otro lado, se determinó la expresión del ARNm para los genes identificados y finalmente se evaluó la producción de la proteína para los tres factores de NMD mencionadas que constituyen el núcleo central de la regulación y catálisis de esta vía de decaimiento. Con los resultados obtenidos para este abordaje pudimos confirmar la presencia de los genes, la expresión del transcrito correspondiente y la generación de la proteína para UPF1, UPF2 y UPF3 en *U. maydis*. Estos resultados no solamente confirman algunas de las observaciones obtenidas con el estudio bioinformático, sino que además sugieren que la vía de NMD puede ser funcional en el basidiomiceto.

En general, la ruta de vigilancia y degradación del ARN se encuentra ampliamente distribuida en los organismos eucariotes, en los que se ha observado que el mecanismo conserva características generales en todas las especies, por lo que la identificación de los factores de la vía de NMD era de esperarse en *U. maydis*. Sin embargo, se han establecido también diferencias particulares para ciertos organismos. En este estudio fue posible identificar a los homólogos para los factores UPF en *U. maydis* empleando los anticuerpos generados contra las proteínas de humano. Estos resultados refuerzan las observaciones *in silico* y en conjunto sugieren que rasgos particulares de la estructura y función de los factores UPF1, UPF2 y UPF3 de humano están conservados en *U. maydis*.

A pesar de la distancia evolutiva que existe entre *Homo sapiens* y *Ustilago maydis* encontramos que los factores centrales de NMD se encuentran altamente conservados para ambos organismos. Esto concuerda además con un análisis previo en donde se muestra el alto grado de conservación entre proteínas de *H. sapiens* y *U. maydis*, la mayoría de las cuales se relacionaron incluso con procesos de regulación de la expresión genética y con la modificación y degradación de proteínas (Münsterkötter & Steinberg, 2007).

VII.III. UPF1 en la viabilidad de *U. maydis*

La influencia de los factores de NMD en el desarrollo del organismo es variable para las diferentes especies (Vicente-Crespo & Palacios, 2010). En *S. cerevisiae*, la pérdida de UPF1 no mostró ningún

efecto aparente en el crecimiento del hongo (Leeds *et al.*, 1991; Cui *et al.* 1995), mientras que en *C. elegans* la inactivación de este gen genera organismos viables pero con anomalías morfológicas en la cola y en los genitales (Hodgkin *et al.*, 1989). Por el contrario, la pérdida de este factor en el pez zebra y en ratón tiene fuertes implicaciones en el crecimiento, desarrollo y viabilidad de los organismos. En el pez zebra la inactivación de *upf1* se llevó a cabo mediante ensayos tipo “knock-down” empleando morfolinós, causando letalidad embrionaria (Wittkopp *et al.*, 2009). En ratón, el *knockout* para *upf1* también causa letalidad embrionaria, mientras que los blastocitos nulos para *upf1* no tienen la capacidad de generar líneas estables de células madre embrionarias (Medghalchi *et al.*, 2001; Weischenfeldt *et al.*, 2008). Estos resultados concuerdan con experimentos realizados en células humanas en cultivo en donde una reducción de *upf1* ocasionó arresto celular (Azzalin & Lingner, 2006).

Debido a que nuestro modelo de estudio es un hongo, inicialmente se esperaba que la mutante con la interrupción de *upf1* generara un efecto semejante al observado para *S. cerevisiae*; es decir, que no tuviera efecto aparente en la viabilidad del organismo. Sin embargo, obtuvimos el resultado opuesto ya que la mutante obtenida por interrupción fue sub-letal. De esta manera, el resultado obtenido corresponde al efecto observado en pez zebra y ratón. Este resultado es interesante ya que indica la relevancia de UPF1 para la viabilidad de *U. maydis*.

Otras observaciones relacionadas con el papel de UPF1 en organismos fúngicos incluye su papel en la regulación del ritmo circadiano en *Neurospora crassa* y el reconocimiento de PTCs afectando la degradación del ARNm en *Aspergillus nidulans* (Morozov *et al.*, 2012; Compton *et al.*, 2013). Aunque sería interesante realizar experimentos futuros para evaluar la funcionalidad de *upf1* en otros aspectos de la fisiología de *U. maydis*, sería necesario implementar las condiciones para poder realizar dichas valoraciones.

Por otro lado, aunque en pez zebra se ha evidenciado la relevancia fisiológica de *Upf1* y se obtuvieron resultados congruentes con las observaciones en ratón, los experimentos realizados en el pez pueden verse afectados por los niveles alcanzados para el *knockdown*, de donde sería interesante contar con otros organismos en los que se estudien los efectos del *knockout* de *upf1*. Nuestros resultados semejan el contexto para un *knockout* de *upf1* y tal como se ha observado para mamíferos, este gen resultó ser determinante para la viabilidad de este modelo.

VII.IV. Interacción umUPF1-umUPF2

UPF1 es el regulador central de NMD y su función reguladora es controlada a través de sus interacciones con otras moléculas, principalmente con UPF2 (Clerici *et al.*, 2009; Chakrabarti *et al.*, 2011; He & Jacobson, 2013). La unión a UPF2 ocurre a través del dominio CH de UPF1 (Leeds *et al.*, 1992). Mediante el análisis *in silico* realizado en este trabajo, determinamos que la interacción entre umUPF1 y hUPF2 es viable. A partir de estas observaciones, estudiamos la posible interacción entre los homólogos para UPF1 y UPF2 de *U. maydis* mediante ensayos tipo *pull-down*. Dichos experimentos se

realizaron empleando la proteína recombinante que contiene el MAL completo para umUPF2 y una recombinante que incluye solamente el dominio CH identificado en umUPF1.

Los detalles mecánicos de las interacciones entre UPF1 y UPF2 se han dilucidado mediante ensayos de cristalografía de rayos X, empleando cristales obtenidos principalmente para *H. sapiens* y *S. cerevisiae* (Kadlec *et al.*, 2006; Clerici *et al.*, 2009; Chakrabarti *et al.*, 2011). Al inicio de este estudio establecimos que las características bioquímicas necesarias para la interacción de los homólogos identificados en *U. maydis* se conservaban en nuestro modelo y sugerían que la unión entre los dos factores era posible. Los resultados del experimento tipo *pull-down* confirman las observaciones de los ensayos *in silico*, demostrando la colaboración de las dos proteínas, presumiblemente en la vía de NMD. Además, este resultado indica que tal y como se ha observado para las proteínas de *S. cerevisiae* y *H. sapiens*, el dominio CH es suficiente para que se lleve a cabo la unión de UPF1 con UPF2. Por otro lado, el resultado contribuye a mostrar la funcionalidad de las proteínas al confirmarse la principal actividad reportada para estos factores. En conjunto, estos resultados confirman la funcionalidad de umUPF1 y umUPF2, al tiempo que sugieren que el mecanismo de NMD puede ocurrir en *U. maydis* como se ha reportado en otros organismos.

VII.V. Funcionalidad de la vía de NMD en *U. maydis*

A pesar de que se había determinado *in silico* la presencia en *U. maydis* de la mayoría de los factores descritos para la vía de NMD, era necesario determinar la funcionalidad de este sistema en nuestro modelo. En este sentido, se han desarrollado e implementado diferentes estrategias para identificar transcritos que efectivamente puedan ser sustratos de NMD, para lo cual se han utilizado algunos fármacos que tienen la capacidad de bloquear esta vía. El fármaco más empleado para esta aplicación es la emetina, aunque también se han empleado otras moléculas como la cafeína, pero con menores resultados (Larson *et al.*, 2009).

La emetina es un producto natural de estructura alcaloide nativo de Brazil. Se han descrito diferentes efectos fisiológicos para esta molécula, las cuales incluyen la inhibición de la síntesis de ADN, de ARN y de proteínas; aunque también se le han atribuido actividades antivirales, antihelmínticas y anticancerígenas (Akinboye & Bakare, 2011). Además de estas funciones, se ha demostrado la actividad de la emetina como inhibidor del proceso de NMD (Mercola & Welsh, 2011). En estos ensayos, los transcritos que aumentan en presencia de emetina se consideran blanco de la vía de NMD. Empleando este abordaje experimental, observamos que algunos transcritos que habían sido reportados como blanco para NMD, también eran sustratos de la vía en *U. maydis*. Tal es el caso de *Chy2*, empleado comúnmente para estudiar NMD en *S. cerevisiae* (He *et al.*, 1997). En los ensayos realizados, observamos varios transcritos que aumentaron en presencia de emetina; los que mostraron una acumulación más evidente fueron *Frataxina* y *Magoh*. Estos resultados indican que estos transcritos muy probablemente sean sustratos de NMD en *U. maydis*. Por el contrario, se observó una disminución para los transcritos que están relacionados con el proceso de splicing. Esta observación

resulta interesante ya que podría sugerir la relación entre ambos procesos en *U. maydis*, como se ha reportado para otros organismos (McGincy & Smith, 2008). Sin embargo, será necesario realizar experimentos posteriores para explorar esta posible interconexión. Finalmente, algunos de los transcritos permanecieron sin cambio, como se esperaba para ppi que corresponde a un gen *housekeeping*. De manera más importante, estos resultados en conjunto nos indican que la vía de degradación mediada por ARNm sin sentido es activa y funcional en este modelo, con lo que se valida la implementación de este hongo para estudiar la vía de NMD.

Al combinar estos ensayos de inhibición con microarreglos, fue posible identificar genes que fueran blanco de la vía de NMD, los cuales estaban relacionados a enfermedades humanas (Noensie & Dietz, 2001). Esta estrategia se denominó GINI y se ha empleado con éxito para identificar mutaciones específicas ligadas a diferentes tipos de cáncer (Huusko et al., 2004; Ionov et al., 2004; Bloethner et al., 2008). Sería interesante para estudios futuros realizar ensayos tipo GINI para profundizar en el estudio de los transcritos que están siendo regulados por NMD en *U. maydis*.

En conjunto, los hallazgos del presente trabajo permiten hipotetizar que el mecanismo de NMD se encuentra altamente conservado y es probable que ocurra de manera semejante entre *H. sapiens* y *U. maydis* y que los factores involucrados en la catálisis y regulación de este proceso sean los mismos en ambos organismos.

VII.VI. El núcleo central del proceso de NMD

El núcleo central de la vía de NMD son las proteínas UPF1, UPF2 y UPF3 (Leeds et al., 1991; Pulak & Anderson, 1993; Culbertson & Leeds, 2003). En la mayoría de las especies analizadas, pudimos identificar homólogos al menos para una proteína UPF. El grupo de organismos que carece de homólogos para estas proteínas corresponde a los protistas, en donde 12 especies diferentes no presentaron homólogo para ninguna proteína UPF. Considerando el papel central de UPF1 en la ruta de NMD, era de esperar que esta proteína fuera identificada en la mayoría de los organismos. En este estudio se encontró que de los 299 organismos analizados, 265 presentan homólogo para UPF1. El principal grupo que carece de UPF1 correspondió a ciertos protistas como diferentes especies de *Leishmania*, en donde la maquinaria de NMD parece estar ausente. Cabe resaltar que hubo especies dentro de todos los grupos (animales, plantas, hongos y protistas) en donde no se determinó el homólogo para UPF1. En estos organismos particulares, será necesario realizar estudios puntuales para determinar la presencia y funcionalidad de la vía de NMD.

En cuanto al núcleo completo de la vía de NMD (UPF1, UPF2, UPF3), se encontró que los hongos pertenecientes a la familia *Encephalitozoonidae*, que solo presentaron UPF1 y carecen UPF2 y UPF3. En general, estos organismos carecen de la mayor parte de la maquinaria de NMD, lo cual podría relacionarse con su naturaleza parasitaria. En estos organismos, sería interesante determinar si el

homólogo putativo para UPF1 tiene una función diferente a la de llevar a cabo NMD, debido a la ausencia del resto de la maquinaria de NMD.

VII.VII. Las proteínas SMG

En este estudio, se identificaron homólogos para las proteínas SMG prácticamente en todos los genomas analizados, confirmando que la regulación por fosforilación de la vía de NMD es empleada por la mayoría de las especies (Yamashita *et al.*, 2013; Kerényi *et al.*, 2013).

Las proteínas SMG5, SMG6 y SMG7 prevalecen en vertebrados y nemátodos, en donde se ha demostrado para ciertos organismos que el ciclo de fosforilación y defosforilación realizado por estas proteínas es relevante para la vía de NMD (Ohnishi *et al.*, 2003; Anders *et al.*, 2003). En varias especies de *Drosophila*, no se encontró un homólogo para SMG7, lo que concuerda con los reportes de los estudios realizados en *Drosophila melanogaster* (Gatfield *et al.*, 2003), mientras que en plantas por lo general se presentan dos SMG7 aparentemente no-redundantes, al mismo tiempo que carecen de SMG5 y SMG6. En concordancia, se ha demostrado la funcionalidad para SMG7 en plantas, aunque no se ha explorado la funcionalidad de más de una proteína en estos sistemas (Benkovics *et al.*, 2011; Kerényi *et al.*, 2013).

VII.VIII. El complejo EJC

Las proteínas que conforman el complejo EJC se encuentran bien representadas en las diferentes especies analizadas. El factor más prevalente resultó EIF4A3, en congruencia con el papel regulador que realiza también en el inicio de la traducción, el ensamble del ribosoma y del *spliceosoma* (Zhang & Krainer, 2007). Por otro lado, Y14 y MAGOH no parecen estar presentes en hongos, lo que correlaciona con el modelo propuesto que establece que el EJC no es indispensable para que ocurra el proceso de NMD en *S. cerevisiae* (Conti & Izaurralde, 2005). La única excepción corresponde a *Yarrowia lypolitica*, el cual es un hemiascomiceto atípico (Mekouar *et al.*, 2010).

En eucariotes superiores, la participación del EJC es relevante para la identificación de los mensajeros anómalos y la adecuada regulación de la vía de NMD (). La estructura tridimensional del EJC establece que el núcleo central de este complejo se compone por las proteínas eIF4AIII, Y14, Magoh y MLN51 (Stroupe *et al.*, 2006). Y14 y Magoh forman un heterodímero, el cual estabiliza la interacción entre la molécula de eIF4AIII (acoplada a ATP) y el ARN de cadena sencilla. La unión de MLN51 a eIF4AIII completa la formación de este complejo. En este estudio observamos que los componentes del EJC prevalecen en la mayoría de las especies, aunque en algunos casos ciertos factores no pudieron ser identificados. Un caso notable corresponde a la proteína MAGOH, el cual fue identificado en todas las especies animales, lo que sugiere que este factor es importante para la vía de NMD en metazoarios. Por el contrario, es probable que Y14 no sea indispensable, considerando que está ausente en 10 especies animales. Puesto que estas proteínas se encuentran formando un dímero, es posible que en

ciertas especies la presencia de uno de los dos factores sea suficiente para la formación del EJC. Fuera del grupo de los animales, MLN51 fue encontrado solo en dos plantas y no lo contiene ningún hongo ni protista, por lo que pudiera ser un factor exclusivo de la vía de NMD en metazoarios y correlaciona con su posición externa en la formación del núcleo del EJC.

Se ha descrito que en *S. cerevisiae*, la vía de NMD requiere de un elemento denominado DSE para promover el decaimiento, en lugar de la formación del complejo EJC (González *et al.*, 2000). En concordancia, todos los *Saccharomyces* analizados en este estudio carecen del homólogo para los componentes del EJC Y14 y Magoh, sugiriendo que en dicho grupo de hongos la vía de NMD ocurrirá muy probablemente en ausencia de EJC. Las proteínas que constituyen la capa externa del EJC (RNPS1, REF/Aly, SAP18, PININ y ACIN1; Tange *et al.*, 2005) tampoco fueron encontradas en los *Saccharomyces*, además de que no se encontraron homólogos para PININ y ACIN1 en hongos, mientras que los protistas prácticamente no poseen estos factores, sugiriendo que el proceso de NMD mediado por EJC no se lleva a cabo en estos organismos.

Un modelo interesante corresponde al protista *Phytophthora sojae*, uno de los pocos protistas que presenta buena parte de la maquinaria de NMD y un EJC casi completo en donde el único factor faltante es MLN51.

Algunas proteínas interactúan con el EJC durante el proceso de NMD de manera transitoria, como UAP56, TAP:p15 y SRm160. Alguno o algunos de estos factores parecen estar ausentes en pocas especies de animales y otros casos aislados, en donde factores que presenten la misma actividad bioquímica podrían suplir la funcionalidad requerida para NMD. En plantas no se encontró el dímero TAP:p15, por lo que es posible sugerir que la exportación nuclear ocurra mediante un mecanismo diferente en este grupo, lo que concuerda con otros factores de transporte descritos para estos organismos. Encontramos el homólogo para p15 en hongos, excepto para 5 *Saccharomyces*, reforzando la observación de que un mecanismo particular de NMD se lleva a cabo para este grupo de organismos.

Las observaciones relacionadas con la distribución de los factores que forman el EJC o que interactúan con éste son interesantes, ya que ésta es uno de las principales rasgos asociados a las principales vías alternativas descritas para NMD (Conti & Izaurralde, 2005). De manera interesante, nuestras observaciones permiten sugerir que dejando de lado los *Saccharomyces*, en el resto de los hongos pueden ocurrir ambos mecanismos de NMD, como se empieza a demostrar en plantas (Kéryni *et al.*, 2013; Nyikó *et al.*, 2013). El análisis funcional de estas observaciones deberá ser demostrada en estudios posteriores.

VII.IX. Factores del procesamiento del ARNm involucrados en NMD

Considerando el papel global que tienen los factores de liberación en la terminación de la traducción, es posible asumir que la mayoría de los organismos presentarán un homólogo para estas proteínas, los cuales encontramos en todas las plantas y en la mayoría de los hongos. Los factores eRF1 y eRF3 fueron identificados en la mayoría de las especies animales, confirmando el papel global de los factores de liberación en procesos moleculares eucariotes, aunque su función particular en NMD para las diferentes especies todavía debe ser demostrado.

Se ha descrito la presencia de CBPC como un rasgo necesario para identificar a los transcritos que son blancos para NMD (Lejeune *et al.*, 2002; Kashima *et al.*, 2006; Sato *et al.*, 2008), de donde resulta interesante identificar a CBP20 y CBP80 en las diferentes especies. Para metazoarios, se observó que algunas especies presentan un homólogo solamente para uno de los dos factores, de donde resulta probable que para estos organismos una CBP sea suficiente para el procesamiento del ARNm, la cual podría corresponder a CBP20 ya que es el más prevalente de los dos factores en todos los grupos y se identificó incluso en protistas. Por otro lado, CBP80 se encontró en plantas y hongos (a excepción de los microsporidios) pero está ausente en protistas, en donde podría ser dispensable.

Otro factor de procesamiento involucrado en la vía de NMD es PABP; en este trabajo se identificó a PABP2 en 297 especies y las dos excepciones fueron protistas. Por el contrario, PABP1 no fue identificada en 32 especies en donde se representan a todos los grupos. Dado que se ha establecido la relevancia de PABP1 en el proceso de NMD (Amrani *et al.*, 2004; Behm-Ansmant *et al.*, 2007; Ivanov *et al.*, 2008), sería interesante estudiar en el futuro las particularidades del mecanismo que regula la vía de NMD en ausencia de este factor, empleando como modelo alguna de las especies identificadas en este trabajo.

VII.1X. Vía de degradación

Para completar la vía de NMD, se requiere de la participación de complejos y factores de degradación (Conti & Izaurralde, 2005), los cuales incluyen al exosoma, el complejo TRAMP, el complejo CCR4/NOT, el complejo PAN y los factores PARN y DCP5.

El exosoma es un complejo que fue descubierto inicialmente en *S. cerevisiae* (Raijmakers *et al.*, 2004; Allmang *et al.*, 1999; Mitchell & Tollervey, 2000), que ha sido identificado en humano, plantas, parásitos, moscas y arqueas (Raijmakers *et al.*, 2004), el cual presenta un núcleo central y otros factores asociados. En este estudio identificamos a los componentes del exosoma en los diferentes grupos de organismos con una distribución homogénea, lo cual podía esperarse si se considera que varias de las proteínas centrales del exosoma son homólogas para la RNAsa PH bacteriana, sugiriendo un origen antiguo en la evolución. El factor menos prevalente corresponde al homólogo para MIR3. Casos particulares que podemos resaltar corresponden al hongo *Postia placenta*, que no presenta

ninguno de los factores del exosoma y *Monilophthora perniciososa*, en donde solo encontramos homólogo para RRP44.

El complejo TRAMP no es tan prevalente en las diferentes especies. De manera interesante, la presencia o ausencia de ciertos factores parece estar asociada a grupos particulares de organismos. Estas diferencias podrían ayudar a explicar las particularidades del sistema de NMD en dichos grupos. Por el contrario, el complejo CCR4/NOT parece estar conservado en la mayoría de los organismos analizados, en donde se encontró el homólogo para las dos proteínas, mientras que el complejo PAN también es altamente prevalente y se encontró ampliamente distribuido en todos los grupos de organismos analizados.

Por su parte, la proteína DCP5 presentó una amplia distribución y parece ser relevante en los grupos de animales y hongos; cosa que no ocurre en las plantas y los protistas, en donde casi no se encontró el homólogo para este factor. La proteína PARN no fue tan prevalente aunque muestra una marcada distribución asociada al grupo de organismo, estando ausente en todas las especies de *Drosophila* incluidas en este estudio, en todos los nemátodos y en varios hongos, en donde su función podría ser dispensable para NMD.

VII.XI. La remoción del Cap

De acuerdo con su importante función en la desprotección del ARNm, DCP2 está presente en la mayoría de los organismos investigados, en donde la degradación del transcrito al parecer puede llevarse a cabo principalmente por la exonucleasa XRN2, la cual fue la más prevalente, mientras que el homólogo para XRN1 se encontró en menos de un tercio de las especies estudiadas. De manera interesante, un total de 27 especies no presenta homólogo para ninguna de las 2 endonucleasas, sugiriendo que en estos organismos la degradación del ARNm podría ser catalizado por otros factores. Otros factores generales de la eliminación del Cap son Ddx6 y Pat1, los cuales se encontraron igual que DCP2 en casi todos los organismos.

El complejo Lsm también se encuentra involucrado en la vía de NMD y en otras vías de degradación (Coller & Parker, 2004). En este análisis, encontramos homólogos para los componentes de este complejo aunque en varias especies encontramos de cuatro a siete de las 8 proteínas Lsm, sugiriendo que quizá en dichos organismos algunas proteínas Lsm pudieran presentar redundancia de función. Cabe resaltar que algunos hongos y protistas carecen pueden presentar solo dos proteínas Lsm7. En estos organismos, sería interesante estudiar no solo el funcionamiento de la vía de NMD en ausencia de estos factores, sino además las implicaciones de la ausencia de estas proteínas en otros mecanismos en los que participan las proteínas Lsm. Como casos extremos puede considerarse a *Nannochloropsis gaditana*, que sólo presenta homólogo para Lsm7 y *Nosema ceranae*, que aparentemente carece de proteínas Lsm, aunque estas dos especies carecen de muchos de los

factores que participan en la vía de NMD, por lo que es posible proponer que este mecanismo no sea necesario para dichas especies.

VII.XII. El caso de *U. maydis*

Cabe mencionar de manera independiente los resultados encontrados en este análisis en relación con *Ustilago maydis*, ya que constituye el principal modelo de estudio para este trabajo. De manera general, este basidiomiceto presentó homólogo para 47 de los 60 factores buscados en este análisis, indicando que la maquinaria que participa en la vía de NMD se encuentra en el hongo, aunque la demostración funcional de todos estos factores y de la manera en la que se regulan requerirá muchos estudios posteriores.

El núcleo central del NMD ya se ha presentado y discutido ampliamente en este trabajo. En cuanto a las proteínas SMG, no se encontraron homólogos aparentes en el análisis general. Sin embargo, un estudio específico nos permitió determinar que el homólogo para mTOR podría llevar a cabo la función de SMG1 en el hongo. Para el caso de SMG5, SMG6 y SMG7, el homólogo para los tres factores parece corresponder al locus UMO4177.1. Este resultado correlaciona con las observaciones relacionadas con las proteínas Est1 en diferentes organismos. Por ejemplo, la proteína Est1 de *S. cerevisiae*, es muy similar a los 3 factores Hsa EST1A, Hsa EST1B y Hsa EST1C, los cuales participan en unión a telómeros. De manera interesante, estas proteínas humanas son idénticas a los factores SMG5, SMG6 y SMG7 de humano, las cuales participan en NMD. Por otro lado, Est1 también presenta homología con Ebs1p, ambas de *S. cerevisiae*. Esta última es similar a SMG5-7 de humano, presentando la mayor homología con SMG7 y mostrando actividad en NMD (Luke *et al.*, 2007). Esta evidencia aunada a los resultados del análisis estructural realizado para estas proteínas (datos no mostrados), es posible que la proteína codificada en el locus UMO4177.1 pueda llevar a cabo la función descrita para las proteínas SMG5-7 en *U. maydis*, sin embargo, la demostración funcional de esta observación requiere de la realización de estudios posteriores.

En cuanto a las proteínas del EJC, *U. maydis* presenta homólogos tanto para proteínas del núcleo de NMD, como Magoh y para componentes de la capa externa, incluyendo a RNPS1, REF/Aly, TAP, UAP56, PYM y SRm160, sugiriendo la posibilidad de que en este hongo también se lleve a cabo la vía de NMD mediada por EJC.

Finalmente, *U. maydis* muestra casi todos los factores involucrados en el decaimiento del ARNm tanto en el sentido 3'→5' como en la dirección opuesta. Todas estas observaciones sustentan no solo que la vía descrita para NMD es funcional en nuestro modelo de estudio, sino que además los factores reportados para que este mecanismo ocurra en humano están presentes en el basidiomiceto, reforzando la posibilidad de emplear a este hongo para estudios posteriores cuyas observaciones puedan correlacionar con lo que ocurre en humano.

VIII. CONCLUSIONES

1. Identificamos los homólogos del núcleo del NMD además de otras proteínas accesorias importantes en la ruta del decaimiento y vigilancia del ARN en *Ustilago maydis*.
2. En general, la ruta de vigilancia y degradación de ARN es diversa en las diferentes especies de eucariotes sin embargo y a pesar de la distancia evolutiva entre el *Homo sapiens* y *Ustilago maydis* las proteínas de estos mecanismos son altamente conservados.
3. Las características principales responsables de la funcionalidad de UPF1 en humano se encuentran conservadas en su homólogo para *Ustilago maydis*.
4. Se comprobó la presencia en el genoma, la expresión del ARNm y la producción de las proteínas para los homólogos de UPF1, UPF2 y UPF3 en *Ustilago maydis*.
5. La vía de NMD es un proceso funcional en *Ustilago maydis*.
6. La mutación del homólogo para UPF1 es subletal en la cepa FB1 de *Ustilago maydis*.
7. El dominio CH de umUPF1 tienen la capacidad de interactuar con umUPF2.
8. Los anticuerpos policlonales comerciales dirigidos a las proteínas UPF1, UPF2 y UPF3 de humano son capaces de reconocer al homólogo respectivo en la cepa FB1 de *Ustilago maydis*.

IX. BIBLIOGRAFIA

1. Akinboye ES, Bakare O (2011) Biological Activities of Emetine. *The Open Natural Products Journal* 4: 8-15
2. Anders KR, Grimson A, Anderson P (2003) SMG-5, required for *C.elegans* nonsense-mediated mRNA decay, associates with SMG-2 and protein phosphatase 2A. *EMBO J* 22: 641-50
3. Addington AM, Gauthier J, Piton A, Hamdan, *et al.* (2011) A novel frameshift mutation in UPF3B identified in brothers affected with childhood onset schizophrenia and autism spectrum disorders. *Mol Psychiatry* 16: 238-9
4. Albers CA, Paul DS, Schulze H, Freson K, *et al.* (2012) Compound inheritance of a low-frequency regulatory snp and a rare null mutation in exon-junction complex subunit RBM8A causes tar syndrome. *Nat Genet* 44: 435-9
5. Allmang, C., Petfalski, E., Podtelejnikov, A., Mann, M., Tollervey, D., and Mitchell, P. (1999) The yeast exosome and human PM-Scl are related complexes of 3' 5' exonucleases. *Genes Dev.* 13, 2148-58
6. Altamura N, Groudinsky O, Dujardin G, Slonimski PP (1992) NAM7 nuclear gene encodes a novel member of a family of helicases with a Zn-ligand motif and is involved in mitochondrial functions in *Saccharomyces cerevisiae*. *J Mol Biol* 224: 575-87
7. Amrani N, Ganesan R, Kervestin S, Mangus DA, *et al.* (2004) A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. *Nature* 432: 112-118
8. Amrani N, Dong S, He F, Ganesan R, *et al.* (2006) Aberrant termination triggers nonsense-mediated mRNA decay. *Biochem Soc Trans* 34: 39-42
9. Anders KR, Grimson A, Anderson P (2003) SMG-5, required for *C.elegans* nonsense-mediated mRNA decay, associates with SMG-2 and protein phosphatase 2A. *EMBO J.* 22: 641-50
10. Applequist SE, Selg M, Raman C, Jack HM: Cloning and characterization of HUPF1, a human homolog of the *Saccharomyces cerevisiae* nonsense mRNA-reducing UPF1 protein. *Nucleic Acids Res* 1997, 25:814-824.
11. Arciga-Reyes L, Wootton L, Kieffer M, Davies B (2006) UPF1 is required for nonsense-mediated mRNA decay (NMD) and RNAi in *Arabidopsis*. *Plant J* 47: 480-9
12. Arias-Palomo E, Yamashita A, Fernandez IS, Nunez-Ramirez R, *et al.* (2011) The nonsense-mediated mRNA decay SMG-1 kinase is regulated by large-scale conformational changes controlled by SMG-8. *Genes Dev* 25: 153-164
13. Aronoff R, Baran R, Hodgkin J: Molecular identification of Smg-4, required for mRNA surveillance in *C. elegans*. *Gene* 2001, 268: 153-164
14. Avery P, Vicente-Crespo M, Francis D, Nashchekina O, Alonso C, Palacios IM (2011) *Drosophila* Upf1 and Upf2 loss of function inhibits cell growth and causes animal death in a Upf3-independent manner. *RNA* 17: 624-38.
15. Azzalin CM, Lingner J (2006) The human RNA surveillance factor UPF1 is required for S phase progression and genome stability. *Curr Biol* 16: 433-9
16. Ballut L, Marchadier B, Baguet A, Tomasetto C, *et al.* (2005) The exon junction core complex is locked onto RNA by inhibition of EIF4AIII ATPase activity. *Nat Struct Mol Biol* 12: 861-9
17. Barberan-Soler S, Lambert NJ, Zahler AM (2009) Global analysis of alternative splicing uncovers developmental regulation of nonsense-mediated decay in *C. elegans*. *RNA* 15: 1652-60
18. Bhattacharya A, Czaplinski, K, Trifillis P, He F, Jacobson A, Peltz SW (2000) Characterization of the biochemical properties of the human Upf1 gene product that is involved in nonsense-mediated mRNA decay. *RNA* 6: 1226-1235
19. Behm-Ansmant I, Gatfield D, Rehwinkel J, Hilgers V, Izaurralde E (2007) A conserved role for cytoplasmic poly(A)-binding protein 1 (PABPC1) in nonsense-mediated mRNA decay. *EMBO J* 26: 1591-601
20. Benkovics AH, Nyiko T, Merai Z, Silhavy D, Bisztray GD (2011) Functional analysis of the grapevine paralogs of the SMG7 NMD factor using a heterolog VIGSbased gene depletion-complementation system. *Plant Mol. Biol.* 75, 277-290
21. Bloethner S, Mould A, Stark M, Hayward NK (2008) Identification of ARHGEF17, DENND2D, FGFR3, and RBI mutations in melanoma by inhibition of nonsense-mediated mRNA decay. *Gene Chromosome Cancer* 47: 1076-85
22. Brunetti-Pierri N, Berg JS, Scaglia F, Belmont J, *et al.* (2008) Recurrent reciprocal 1q21.1 deletions and duplications associated with microcephaly or macrocephaly and developmental and behavioral abnormalities. *Nat Genet* 40: 1466-71
23. Cao D, Parker R (2003) Computational modeling and experimental analysis of nonsense-mediated decay in yeast. *Cell* 113: 533-45
24. Chakrabarti S, Jayachandran U, Bonneau F, Fiorini F, Basquin C, Domcke S *et al.* (2011) Molecular mechanisms for the RNA-dependent ATPase activity of Upf1 and its regulation by Upf2. *Mol Cell* 41: 693-703
25. Chamieh H, Ballut L, Bonneau F, Le Hir H (2008) NMD factors UPF2 and UPF3 bridge UPF1 to the exon junction complex and stimulate its RNA helicase activity. *Nat Struct Mol Biol* 15: 85-93
26. Cheng Z, Muhlrad D, Lim MK, Parker R, Song H (2007) Structural and functional insights into the human Upf1 helicase core. *The EMBO Journal* 26: 253-64
27. Chiu SY, Serin G, Ohara O, Maquat LE (2003) Characterization of human Smg5/7a, a protein with similarities to *Caenorhabditis elegans* SMG5 and SMG7 that functions in the dephosphorylation of Upf1. *RNA* 9:77-87
28. Cho H, Kim KM, Kim YK (2009) Human proline-rich nuclear receptor coregulatory protein 2 mediates an interaction between mRNA surveillance machinery and decapping complex. *Mol Cell* 33: 75-86
29. Chuang TW, Peng PJ, Tarn WY (2011) The exon junction complex component Y14 modulates the activity of the methylosome in biogenesis of spliceosomal small nuclear ribonucleoproteins. *J Biol Chem* 286: 8722-8.
30. Chuang TW, Chang WL, Lee KM, Tarn WY (2013) The RNA-binding protein Y14 inhibits mRNA decapping and modulates processing body formation. *Mol Biol Cell* 24: 1-13.
31. Chuang TW, Lee KM, Tarn WY (2015) Function and Pathological Implications of Exon Junction Complex Factor Y14. *Biomolecules* 5: 343-55.
32. Clerici M, Mourão A, Gutsche I, Gehring NH, Hentze MW, Kulozik A, *et al.* (2009). Unusual bipartite mode of interaction between the nonsense-mediated decay factors, UPF1 and UPF2. *The EMBO J.* 28: 2293-2306.
33. Collier J, Parker R. Eukaryotic mRNA decapping. *Annu Rev Biochem* 2004; 73:861-890.
34. Conti E, Izaurralde E (2005) Nonsense-mediated mRNA decay: molecular insights and mechanistic variations across species. *Curr Opin Cell Biol* 17: 316-25.
35. Cuccurese M, Russo G, Russo A, Pietropaolo C (2005) Alternative splicing and nonsense-mediated mRNA decay regulate mammalian ribosomal gene expression. *Nucleic Acids Res.* 33, 5965-77.

36. Cui Y, Hagan KW, Zhang S, Peltz SW (1995) Identification and characterization of genes that are required for the accelerated degradation of mRNAs containing a premature translational termination codon. *Genes Dev* 9: 423-36.
37. Culbertson MR, Leeds PF (2003) Looking at mRNA decay pathways through the window of molecular evolution. *Curr Opin Genet Dev* 13:207–214.
38. Czaplinski K, Ruiz-Echevarria MJ, Paushkin SV, Han X, *et al.*, (1998) The surveillance complex interacts with the translation release factors to enhance termination and degrade aberrant mRNAs. *Genes Dev* 12: 1665–77.
39. Dagleish R. (2014) Osteogenesis Imperfecta Genotypes and Genotype–Phenotype Relationships. Elsevier. DOI: <http://dx.doi.org/10.1016/B978-0-12-397165-4.00010-1>.
40. Danckwardt S, Neu-Yilik G, Thermann R, Frede U, *et al.* (2002) Abnormally spliced beta-globin mRNAs: a single point mutation generates transcripts sensitive and insensitive to nonsense-mediated mRNA decay. *Blood* 99: 1811-6
41. Denning G, Jamieson L, Maquat LE, Thompson EA, *et al.*, (2001) Cloning of a novel phosphatidylinositol kinase-related kinase: Characterization of the human SMG-1 RNA surveillance protein. *J. Biol. Chem.* 276: 22709–14
42. Diem MD, Chan CC, Younis I, Dreyfuss G (2007) PYM binds the cytoplasmic exon-junction complex and ribosomes to enhance translation of spliced mRNAs. *Nat Struct Mol Biol* 14: 1173–9
43. Dostie J, Dreyfuss G (2002) Translation is required to remove Y14 from mRNAs in the cytoplasm. *Curr Biol* 12: 1060–7.
44. Durand S, Lykke-Andersen J (2013) Nonsense-mediated mRNA decay occurs during eIF4F-dependent translation in human cells. *Nat Struct Mol Biol* 20:702–709
45. Eberle AB, Stalder L, Mathys H, Orozco RZ, *et al* (2008) Posttranscriptional gene regulation by spatial rearrangement of the 3' untranslated region. *PLoS Biol* 6:e92
46. Eberle AB, Lykke-Andersen S, Muhlemann O, Jensen TH (2009) SMG6 promotes endonucleolytic cleavage of nonsense mRNA in human cells. *Nat Struct Mol Biol* 16, 49–55
47. Fairman-Williams ME, Guenther UP, Jankowsky E (2010) SF1 and SF2 helicases: family matters. *Curr Opin Struct Biol* 20: 313–24
48. Förch P, Valcárcel J (2003) Splicing regulation in *Drosophila* sex determination. *Prog Mol Subcell Biol* 31: 127-51
49. Franks TM, Singh G, Lykke-Andersen J (2010) Upf1 ATPase-dependent mRNP disassembly is required for completion of nonsense-mediated mRNA decay. *Cell* 143: 938–950
50. Frischmeyer PA, Dietz HC (1999) Nonsense-mediated mRNA decay in health and disease. *Hum Mol Genet* 8: 1893–900.
51. Fukuhara N, Ebert J, Unterholzner L, Lindner D, Izaurralde E, Conti E (2005) SMG7 is a 14-3-3-like adaptor in the nonsense-mediated mRNA decay pathway. *Mol Cell* 17: 537–47
52. Gatfield D, Unterholzner L, Ciccarelli FD, Bork P, Izaurralde E. (2003) Nonsense-mediated mRNA decay in *Drosophila*: at the intersection of the yeast and mammalian pathways. *EMBO J.* 22, 3960–70.
53. Gehring NH, Neu-Yilik G, Schell T, Hentze MW, *et al.* (2003) Y14 and hUpf3b form an NMD-activating complex. *Mol Cell* 11: 939–49.
54. Gleghom ML, Maquat LE (2011) UPF1 learns to relax and unwind. *Mol Cell* 41: 621-3.
55. Gonzalez CI, Ruiz-Echevarria MJ, Vasudevan S, Henry MF, *et al.* (2000) The yeast hnRNP-like protein Hrp1/Nab4 marks a transcript for nonsense-mediated mRNA decay. *Mol Cell* 5: 489-9
56. Green RE, Lewis BP, Hillman RT, Blanchette M, *et al.* (2003) Widespread predicted nonsense-mediated mRNA decay of alternatively-spliced transcripts of human normal and disease genes. *Bioinformatics* 19(Suppl. 1), i118–i121
57. Gulsuner S, Walsh T, Watts AC, Lee MK, *et al.* 2013. Spatial and temporal mapping of de novo mutations in Schizophrenia to a fetal prefrontal cortical network. *Cell* 154: 518–29
58. Hwang J, Sato H, Tang Y, Matsuda D, Maquat LE (2010) UPF1 Association With the Cap-binding Protein, CBP80, Promotes Nonsense-mediated mRNA Decay at Two Distinct Steps. *Mol Cell* 39: 396-409
59. He F, Jacobson A (1995) Identification of a novel component of the nonsense-mediated mRNA decay pathway by use of an interacting protein screen. *Genes Dev* 9: 437–454
60. He F, Brown AH, Jacobson A (1997) Upf1p, Nmd2p, and Upf3p are interacting components of the yeast nonsense-mediated mRNA decay pathway. *Mol Cell Biol* 17: 1580-94
61. Hodgkin J, Papp A, Pulak R, Ambros V, Anderson P (1989) A new kind of informational suppression in the nematode *Caenorhabditis elegans*. *Genetics* 123: 301–13
62. Hodgkin J, Papp A, Pulak R, Ambros V, Anderson P (1989) A new kind of informational suppression in the nematode *Caenorhabditis elegans*. *Genetics* 123: 301–13
63. Holliday R (1974) *Ustilago maydis*. In R. C. King (ed.), *Handbook of genetics*, vol. 1. Plenum Press, New York, N.Y. pp 575–95.
64. Hosoda N, Kim YK, Lejeune F, Maquat LE (2005) CBP80 promotes interaction of Upf1 with Upf2 during nonsense-mediated mRNA decay in mammalian cells. *Nat Struct Mol Biol* 12: 893–901.
65. Hori K, Watanabe Y (2005) UPF3 suppresses aberrant spliced mRNA in *Arabidopsis*. *Plant J* 43: 530–40
66. Hornbeck PV, Zhang B, Murray B, Kornhauser JM, Latham V, Skrzypek E (2014) PhosphoSitePlus, 2014: mutations, PTMs and recalibrations. *Nucleic Acids Res* 43: D512-520
67. Huntzinger E, Kashima I, Fauser M, Sauliere J, Izaurralde E (2008) SMG6 is the catalytic endonuclease that cleaves mRNAs containing nonsense codons in metazoan. *RNA* 14: 2609–17
68. Huusko P, Ponciano-Jackson D, Wolf M, Kiefer JA, Azorsa DO, Tuzmen S, *et al.* (2004) Nonsense-mediated decay microarray analysis identifies mutations of EPHB2 in human prostate cancer. *Nat. Genet.* 36: 979-83
69. Hwang J, Maquat LE. (2011) Nonsense-mediated mRNA decay (NMD) in animal embryogenesis: To die or not to die, that is the question. *Curr Opin Genet Dev* 21: 422- 430
70. Ideue T, Sasaki YT, Hagiwara M, Hirose T (2007) Introns play an essential role in splicing-dependent formation of the exon junction complex. *Genes Dev* 21: 1993–1998
71. Ionov Y, Nowak N, Perucho M, Markowitz S, Cowell JK (2004) Manipulation of nonsense mediated decay identifies gene mutations in colon cancer cells with microsatellite instability. *Oncogen* 23: 639-45
72. Ishigaki Y, Li X, Serin G, Maquat LE (2001) Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell* 106: 607–17.
73. Ivanov I, Lo KC, Hawthorn L, Cowell JK, Ionov Y (2007) Identifying candidate colon cancer tumor suppressor genes using inhibition of nonsense-mediated mRNA decay in colon cancer cells. *Oncogene* 26, 2873–84
74. Jankowsky E (2011) RNA helicases at work: Binding and rearranging. *Trends Biochem Sci* 36: 19–29

75. Jarmoskaite I, Russell R (2014) RNA helicase proteins as chaperones and remodelers. *Annu Rev Biochem* 83: 697–725
76. Johnson JM, Castle J, Garrett-Engele P, Kan Z, *et al.* (2003) Genome-wide survey of human alternative pre-mRNA splicing with exon junction microarrays. *Science* 302: 2141–44
77. Jouanguy E, Lamhamedi-Cherradi S, Lammass D, Dorman SE, *et al.* (1999) A human IFNGR1 small deletion hotspot associated with dominant susceptibility to mycobacterial infection. *Nat Genet* 21: 370–8
78. Juneau K, Palm C, Miranda M, Davis RW. (2007) High-density yeast-tiling array reveals previously undiscovered introns and extensive regulation of meiotic splicing. *Proc Natl Acad Sci USA* 104: 1522–7
79. Kadlec J, Guilligay D, Ravelli RB, Cusack S (2006) Crystal structure of the UPF2-interacting domain of nonsense-mediated mRNA decay factor UPF1. *RNA* 12: 1817–1824
80. Kashima I, Yamashita A, Izumi N, Kataoka N, *et al.* (2006) Binding of a novel SMG-1-Upf1-eRF1-eRF3 complex (SURF) to the exon junction complex triggers Upf1 phosphorylation and nonsense-mediated mRNA decay. *Genes Dev* 20: 355–67
81. Keeling KM, Bedwell DM (2011) Suppression of nonsense mutations as a therapeutic approach to treat genetic diseases. *Wiley Interdiscip. Rev. RNA* 2: 837–52.
82. Keeling KM, Wang D, Conard SE, Bedwell DM (2012) Suppression of premature termination codons as a therapeutic approach. *Crit Rev Biochem Mol Biol* 47: 444–63
83. Kerényi Z, Merai Z, Hiripi L, Benkovics A, *et al.* (2008) Interkingdom conservation of mechanism of nonsense-mediated mRNA decay. *EMBO J* 27: 1585–95
84. Kerényi F, Wawer I, Sikorski PW, Kufel J, Silhavy D (2013) Phosphorylation of the N- and C-terminal UPF1 domains plays a critical role in plant nonsense-mediated mRNA decay. *Plant J* 76: 836–848.
85. Kertesz S, Kerényi Z, Merai Z, Bartos I, *et al.* (2006) Both introns and long 3'-UTRs operate as cis-acting elements to trigger nonsense-mediated decay in plants. *Nucleic Acids Res* 34: 6147–6157
86. Khodadadi S, Sokolov AP (2015) Protein dynamics: from rattling in a cage to structural relaxation. *Soft Matter* 11: 4984–98
87. Kisselev L, Ehrenberg M, Frolova L (2003) Termination of translation: interplay of mRNA, rRNAs and release factors? *EMBO J*. 22: 175–82
88. Klopocki E, Schulze H, Strauss G, Ott CE, *et al.* (2007) Complex inheritance pattern resembling autosomal recessive inheritance involving a microdeletion in thrombocytopenia-absent radius syndrome. *Am J Hum Genet* 80: 232–40
89. Krawczak M, Ball EV, Fenton I, Stenson PD, *et al.* (2000) Human gene mutation database—a biomedical information and research resource. *Hum Mutat* 15: 45–51
90. Kurihara Y, Matsui A, Hanada K, Kawashima M, *et al.* (2009) Genome-wide suppression of aberrant mRNA-like noncoding RNAs by NMD in Arabidopsis. *Proc Natl Acad Sci USA*, 106, 2453–2458
91. Lareau LF, Inada M, Green RE, Wengrod JC, Brenner SE (2007) Unproductive splicing of SR genes associated with highly conserved and ultraconserved DNA elements. *Nature* 446, 926–9
92. Larsson DE, Hassan S, Larsson R, Oberg K, Granberg D (2009) Combination analysis of anti-cancer drugs on human neuroendocrine tumor cell lines. *Cancer Chemother. Pharmacol* 65: 5–12
93. Laumonier F, Shoubridge C, Antar C, Nguyen LS, *et al.* (2010) Mutations of the UPF3B gene, which encodes a protein widely expressed in neurons, are associated with nonspecific mental retardation with or without autism. *Mol Psychiatry* 15: 767–76
94. Le Hir H, Izaurralde E, Maquat LE, Moore MJ (2000) The spliceosome deposits multiple proteins 20–24 nucleotides upstream of mRNA exon-exon junctions. *EMBO J* 19: 6860–69
95. Le Hir H, Gatfield D, Izaurralde E, Moore MJ (2001) The exon-exon junction complex provides a binding platform for factors involved in mRNA export and nonsense-mediated mRNA decay. *EMBO J* 20: 4987–97
96. Le Jeune F, Maquat LE (2005) Mechanistic links between nonsense-mediated mRNA decay and pre-mRNA splicing in mammalian cells. *Curr Opin Cell Biol.* 17: 309–15.
97. Le Jeune F, Ishigaki Y, Li X, Maquat LE (2002) The exon junction complex is detected on CBP80-bound but not eIF4E-bound mRNA in mammalian cells: dynamics of mRNP remodeling. *EMBO J* 21: 3536–45.
98. Lee HC, Choe J, Chi SG, Kim YK (2009) Exon junction complex enhances translation of spliced mRNAs at multiple steps. *Biochem Biophys Res Commun.* 384: 334–40.
99. Lee HL, Dougherty JP (2012) Pharmaceutical therapies to recode nonsense mutations in inherited diseases. *Pharmacol Ther* 136: 227–66.
100. Leeds PF, Wood J, Lee B, Culbertson MR (1992) Factors that promote mRNA turnover in *Saccharomyces cerevisiae*. *Mol Cell Biol* 12: 2165–77.
101. Leeds P, Peltz SW, Jacobson A, Culbertson MR (1991) The product of the yeast UPF1 gene is required for rapid turnover of mRNAs containing a premature translational termination codon. *Genes Dev* 5: 2303–14.
102. Longman D, Plasterk RH, Johnstone IL, Caceres JF (2007) Mechanistic insights and identification of two novel factors in the *C. elegans* NMD pathway. *Genes Dev* 21: 1075–85.
103. Losson R, Lacroute F (1979) Interference of nonsense mutations with eukaryotic messenger RNA stability. *Proc Natl Acad Sci USA* 76, 5134–5137.
104. Lobanov MY, Bogatyreva NS, Galzitskaya OV (2008) Radius of gyration as an indicator of protein structure compactness. *Molecular Biology* 42: 623–628
105. Lykke-Andersen J, Shu MD, Steitz JA (2000) Human Upf proteins target an mRNA for nonsense mediated decay when bound downstream of a termination codon. *Cell* 103: 1121–1131
106. Lynch SA, Nguyen LS, Ng LY, Waldron M, *et al.* (2012) Broadening the phenotype associated with mutations in UPF3B: two further cases with renal dysplasia and variable developmental delay. *Eur J Med Genet* 55: 476–9.
107. Maderazo AB, Belk JP, He F, Jacobson A (2003) Nonsense-containing mRNAs that accumulate in the absence of a functional nonsense-mediated mRNA decay pathway are destabilized rapidly upon its restitution. *Mol Cell Biol* 23: 842–51.
108. Mango SE (2001) Stop making nonSense: the *C. elegans* smg genes. *Trends Genet* 17: 646–53.
109. Maquat LE (1995) When cells stop making sense: effects of nonsense codons on RNA metabolism in vertebrate cells. *RNA* 1: 453–65.
110. Martínez-Contreras R, Galindo JM, Aguilar-Rojas A, Valdés J (2003) Two exonic elements in the flanking constitutive exons control the alternative splicing of the alpha exon of the ZO-1 pre-mRNA. *Biochim Biophys Acta* 1630: 71–83.
111. Martínez-Contreras R, Fiset JF, Nasim FU, Madden R, *et al.* (2006) Intronic binding sites for hnRNP A/B and hnRNP F/H proteins stimulate pre-mRNA splicing. *PLoS Biol.* 4: e21.

112. Martínez-Contreras R, Cloutier P, Shkreta L, Fiset JF, et al. (2007) hnRNP proteins and splicing control. *Adv Exp Med Biol* 623: 123-47.
113. Martínez-Contreras R, Martínez-Montiel N. (2015) Alternative Splicing Modification as a Treatment for Genetic Disorders. *Gene Technol* 4: 3.
114. Martínez-Contreras R, Martínez-Montiel N (2016) The role of splicing factors in cancer prognosis and treatment. En: "Alternative splicing and disease". E. Massey, ed. Nova science publishers. New York, USA. In press.
115. Martínez-Montiel N, Rosas-Murrieta N, Martínez-Contreras R (2015) Regulación del splicing alternativo: Implicaciones en el diagnóstico y tratamiento del cáncer. *Med Clin* 144: 317-23.
116. Martínez-Montiel N, Rosas-Murrieta N, Martínez-Contreras R (2015) Peptidic tools applied to redirect alternative splicing events. *Peptides* 67: 1-11.
117. Martínez-Montiel N, Martínez-Contreras R (2016) Nonsense-mediated decay and human disease. En: "Alternative splicing and disease". E. Massey, ed. Nova science publishers. New York, USA.
118. McGlincy NJ, Smith CW. (2008) Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? *Trends Biochem Sci* 33: 385-93
119. Meaux S, van Hoof A, Baker KE (2008) Nonsense-mediated mRNA decay in yeast does not require PAB1 or a poly(A) tail. *Mol Cell* 29: 134-40
120. Medghalchi SM, Frischmeyer PA, Mendell JT, Kelly AG, Lawler AM, Dietz HC (2001) Rent1, a trans-effector of nonsense-mediated mRNA decay, is essential for mammalian embryonic viability. *Hum Mol Genet* 10: 99-105
121. Meers P, Neville M, Malinin V, Scotto AW, et al. (2008) Biofilm penetration, triggered release and in vivo activity of inhaled liposomal amikacin in chronic *Pseudomonas aeruginosa* lung infections. *J Antimicrob Chemother* 61: 859-68.
122. Mekouar M, Blanc-Lenfle I, Ozanne C, Da Silva C, Cruaud C, Wincker P, et al. (2009) Detection and analysis of alternative splicing in *Yarrowia lipolytica* reveal structural constraints facilitating nonsense-mediated decay of intron-retaining transcripts. *Genome Biology* 11: R65
123. Mendell JT, Sharifi NA, Meyers JL, Martinez-Murillo F, Dietz HC (2004) Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. *Nat Genet* 36, 1073-1078.
124. Mercola D, Welsh J. (2004) From mRNA to tumor suppressor. *Nat Genet* 36: 937-8
125. Michelle L, Cloutier A, Toutant J, Shkreta L, et al. (2012) Proteins associated with the exon junction complex also control the alternative splicing of apoptotic regulators. *Mol Cell Biol* 32: 954-67
126. Millar DS, Kemball-Cook G, McVey JH, Tuddenham EG, et al. (2000) Molecular analysis of the genotype-phenotype relationship in factor VII deficiency. *Hum. Genet.* 107: 327-42
127. Miller JN, Pearce DA (2014) Nonsense-mediated decay in genetic disease: Friend or foe? *Mutation Research* 762: 52-64
128. Mitchell P, Tollervey D (2000) Musing on the structural organization of the exosome complex. *Nat Struct Biol* 7, 843-6
129. Mitrovich QM, Anderson P (2000) Unproductively spliced ribosomal protein mRNAs are natural targets of mRNA surveillance in *C. elegans*. *Genes Dev* 14: 2173-84
130. Morita T, Yamashita A, Kashima I, Ogata K, Ishiura S, Ohno S (2007) Distant N- and C-terminal domains are required for intrinsic kinase activity of SMG-1, a critical component of nonsense-mediated mRNA decay. *J Biol Chem* 282, 7799-808
131. Morozov IY, Jones MG, Gould PD, Crome V, Wilson JB, Hall AJ, et al (2012) mRNA 3' tagging is induced by nonsense-mediated decay and promotes ribosome dissociation. *Mol Cell Biol* 32: 2585-95
132. Mühlemann O, Lykke-Andersen J (2010) How and where are nonsense mRNAs degraded in mammalian cells? *RNA Biol.* 7:28-32
133. Mühlemann O (2008) Recognition of nonsense mRNA: towards a unified model. *Biochem Soc Trans* 36: 497-501.
134. Münsterkötter M, Steinberg G (2007) The fungus *Ustilago maydis* and humans share disease-related proteins that are not found in *Saccharomyces cerevisiae*. *BMC Genomics* 8: 473.
135. Nguyen LS, Kim HG, Rosenfeld JA, Shen Y, et al. (2013) Contribution of copy number variants involving nonsense-mediated mRNA decay pathway genes to neuro-developmental disorders. *Hum Mol Genet* 22: 1816-25
136. Nguyen LS, Wilkinson MF, Gecz J (2014) Nonsense-mediated mRNA decay: Inter-individual variability and human disease. *Neurosci Biobehav Rev* 46: 175-86
137. Noensie EN, Dietz HC (2001) A strategy for disease gene identification through nonsense-mediated mRNA decay inhibition. *Nat Biotechnol* 19: 434-9
138. Nyikó T, Kerényi F, Szabadkai L, Benkovics AH, et al. (2013) Plant nonsense-mediated mRNA decay is controlled by different autoregulatory circuits and can be induced by an EJC-like complex. *Nucleic Acids Res* 41: 6715-28.
139. Ohnishi T, Yamashita A, Kashima I, Schell T, et al. (2003) Phosphorylation of hUPF1 induces formation of mRNA surveillance complexes containing hSMG-5 and hSMG-7. *Mol. Cell* 12: 1187-200.
140. Okada-Katsuhata Y, Yamashita A, Kutsuzawa K, Izumi N, et al (2012) N- and C-terminal Upf1 phosphorylations create binding platforms for SMG-6 and SMG-5: SMG-7 during NMD. *Nucleic Acids Res* 40: 1251-66.
141. Padariya M, Kalathiya U, Baginski M (2015) Structural and dynamic changes adopted by EmrE, multidrug transporter protein—Studies by molecular dynamics simulation. *Biochim Biophys Acta* doi: 10.1016/j.bbame.2015.05.014.
142. Page MF, Carr B, Anders KR, Grimson A, Anderson P (1999) SMG-2 is a phosphorylated protein required for mRNA surveillance in *Caenorhabditis elegans* and related to Upf1p of yeast. *Mol Cell Biol* 19: 5943-51
143. Paillusson A, Hirschi N, Vallan C, Azzalin CM, Mühlemann O (2005) A GFP-based reporter system to monitor nonsense-mediated mRNA decay. *Nucleic Acids Res* 33: e54
144. Pal M, Ishigaki Y, Nagy E, Maquat LE (2001) Evidence that phosphorylation of human Upf1 protein varies with intracellular location and is mediated by a wortmannin-sensitive and rapamycin-sensitive PI 3-kinase-related kinase signaling pathway. *RNA* 7: 5-15.
145. Pan Q, Saltzman AL, Kim YK, Misquitta C, et al (2006) Quantitative microarray profiling provides evidence against widespread coupling of alternative splicing with nonsense-mediated mRNA decay to control gene expression. *Genes Dev* 20: 153-8.
146. Peltz SW, Brown AH, Jacobson A. (1993) mRNA destabilization triggered by premature translational termination depends on at least three cis-acting sequence elements and one trans-acting factor. *Genes Dev* 7: 1737-54.
147. Peltz SW, Morsy M, Welch EM, Jacobson A (2013) Ataluren as an agent for therapeutic nonsense suppression. *Annu Rev Med* 64: 407-25.

148. Perlick HA, Medghalchi SM, Spencer FA, Kendzior RJJ, Dietz HC: Mammalian orthologues of a yeast regulator of nonsense transcript stability. *Proc Natl Acad Sci USA* 1996, 93:10928-32.
149. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al (2005) Scalable molecular dynamics with NAMD. *J Comput Chem* 26: 1781–1802
150. Pulak R, Anderson P (1993) mRNA surveillance by the *Caenorhabditis elegans* smg genes. *Genes Dev* 7: 1885–97.
151. Raijmakers, R., Schilders, G., and Pruijn, G. J. (2004) The exosome, a molecular machine for controlled RNA degradation in both nucleus and cytoplasm. *Eur. J. Cell Biol.* 83, 175–183
152. Ramani AK, Nelson AC, Kapranov P, Bell I, Gingeras TR, et al. (2009) High resolution transcriptome maps for wild-type and nonsense-mediated decay-defective *Caenorhabditis elegans*. *Genome Biol* 10: R101.
153. Rayson S, Arciga-Reyes L, Wootton L, De Torres-Zabala M., et al. (2012) A role for nonsense-mediated mRNA decay in plants: pathogen responses are induced in *Arabidopsis thaliana* NMD mutants. *PLoS One*, 7, e31917.
154. Rebbapragada I, Lykke-Andersen J (2009) Execution of nonsense-mediated mRNA decay: what defines a substrate? *Curr Opin Cell Biol* 21: 394-402
155. Rehwinkel J, Letunic I, Raes J, Bork P, et al. (2005) Nonsense-mediated mRNA decay factors act in concert to regulate common mRNA targets. *RNA* 11: 1530–44.
156. Reiter RJ, Tan DX, Korkmaz A, Fuentes-Broto L (2011) Drug-mediated ototoxicity and tinnitus: Alleviation with melatonin. *J Physiol Pharmacol* 62: 151–7.
157. Rivolta C, Berson EL, Dryja TP (2001) Dominant Leber congenital amaurosis, cone-rod degeneration, and retinitis pigmentosa caused by mutant versions of the transcription factor CRX. *Hum Mutat* 18: 488-98.
158. Rufener SC, Mühlemann O (2013) eIF4E-bound mRNPs are substrates for nonsense-mediated mRNA decay in mammalian cells. *Nat Struct Mol Biol* 20: 710–7.
159. Sato H, Hosoda N, Maquat LE. 2001. Efficiency of the pioneer round of translation affects the cellular site of nonsense-mediated mRNA decay. *Mol Cell.* 29: 255-62.
160. Schneppenheim R, Budde U, Obser T, Brassard J, et al. (2001) Expression and characterization of von Willebrand factor dimerization defects in different types of von Willebrand disease. *Blood* 97: 2059-66.
161. Schwabe GC, Tinschert S, Buschow C, Meinecke P, et al. (2000) Distinct mutations in the receptor tyrosine kinase gene ROR2 cause brachydactyly type B. *Am J Hum Genet* 67: 822-31
162. Schweingruber C, Rufener SC, Zünd D, Yamashita A, Mühlemann O (2013) Nonsense-mediated mRNA decay – Mechanisms of substrate mRNA recognition and degradation in mammalian cells. *Biochim Biophys Acta* 1829: 612-23
163. Serin G, Gersappe A, Black JD, Aronoff R, Maquat LE (2001) Identification and characterization of human orthologues to *Saccharomyces cerevisiae* Upf2 protein and Upf3 protein (*Caenorhabditis elegans* SMG-4). *Mol Cell Biol* 21: 209-23.
164. Silver DL, Watkins-Chow DE, Schreck KC, Pierfelice TJ, et al. (2010) The exon junction complex component Magoh controls brain size by regulating neural stem cell division. *Nat Neurosci* 13: 551–58.
165. Singh G, Rebbapragada I, Lykke-Andersen J. (2008) A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. *PLoS Biol.* 6:e111.
166. Skorka A, Bielicka-Cymermann J, Gieruszczak-Bialek D, Korniszewski L. (2005) Thrombocytopenia-absent radius (TAR) syndrome: A case with agenesis of corpus callosum, hypoplasia of cerebellar vermis and horseshoe kidney. *Genet Couns* 16: 377–82.
167. Smyth AR, Bhatt J (2012) Once-daily versus multiple-daily dosing with intravenous aminoglycosides for cystic fibrosis. *Cochrane Database Syst Rev* 2: CD002009.
168. Sun X, Moriarty PM, Maquat LE (2000) Nonsense-mediated decay of glutathione peroxidase 1 mRNA in the cytoplasm depends on intron position. *EMBO J* 19: 4734–44.
169. Szklarczyk D, Franceschini A, Wyder S, Forslund K, Heller D, Huerta-Cepas J, et al (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43(Database issue):D447-52.
170. Takahashi S, Araki Y, Ohya Y, Sakuno T, Hoshino SI, Kontani K, et al (2008) Upf1 potentially serves as a RING-related E3 ubiquitin ligase via its association with Upf3 in yeast. *RNA* 14: 1950-1958.
171. Tange TO, Shibuya T, Jurica MS, Moore MJ (2005) Biochemical analysis of the EJC reveals two new factors and a stable tetrameric protein core. *RNA* 11: 1869–75.
172. Tani H, Imamachi N, Salam KA, Mizutani R, et al. (2012) Identification of hundreds of novel UPF1 target transcripts by direct determination of whole transcriptome stability. *RNA Biol* 9: 1370–9.
173. Tarpey PS, Nguyen LS, Raymond FL, Rodriguez J, et al. (2007) Mutations in UPF3B, a member of the nonsense-mediated mRNA decay complex, cause syndromic and nonsyndromic mental retardation. *Nat Genet* 39: 1127–33.
174. Tazi J, Bakkour N, Stamm S (2009) Alternative splicing and disease. *Biochim Biophys Acta* 1792: 14-26.
175. Thibault N, Grenier L, Simard M, Bergeron MG, et al. 1995. Protection against gentamicin nephrotoxicity by daptomycin in nephrectomized rats. *Life Sci* 56: 1877–87.
176. Thoren LA, Norgaard GA, Weischenfeldt J, Waage J, et al. (2010) UPF2 is a critical regulator of liver development, function and regeneration. *PLoS ONE* 5, e11650.
177. Unterholzner L, Izaurralde E (2004) SMG7 acts as a molecular link between mRNA surveillance and mRNA decay. *Mol Cell* 16, 587–96.
178. Usuki F, Yamashita A, Kashima I, Higuchi I, et al. (2006) Specific inhibition of nonsense-mediated mRNA decay components, SMG-1 or Upf1, rescues the phenotype of Ullrich disease fibroblasts *Mol Ther* 14: 351–60
179. Usuki F, Yamashita A, Fujimura M. (2011) Post-transcriptional defects of antioxidant selenoenzymes cause oxidative stress under methylmercury exposure. *J Biol Chem* 286, 6641–49
180. Vicente-Crespo M, Palacios IM (2010) Nonsense-mediated mRNA decay and development: shoot the messenger to survive? *Biochem Soc Trans* 38: 1500–5
181. Viegas MH, Gehring NH, Breit S, Hentze MW, Kulozik AE (2007) The abundance of RNPS1, a protein component of the exon junction complex, can determine the variability in efficiency of the Nonsense Mediated Decay pathway. *Nucleic Acids Res.* 35, 4542–51
182. Weischenfeldt J, Damgaard I, Bryder D, Theilgaard-Monch K, Thoren LA, Nielsen FC, et al. (2008) NMD is essential for hematopoietic stem and progenitor cells and for eliminating by-products of programmed DNA rearrangements. *Genes Dev* 22: 1381–96

183. Weischenfeldt J, Waage J, Tian G, Zhao J, et al. (2012) Mammalian tissues defective in nonsense-mediated mRNA decay display highly aberrant splicing patterns. *Genome Biol.* 13: R35
184. Weng Y, Czaplinski K, Peltz SW (1996) Genetic and biochemical characterization of mutations in the ATPase and helicase regions of the Upf1 protein. *Mol Cell Biol* 16: 5477-90
185. Wilkinson MF (2003) The cycle of nonsense. *Mol Cell* 12:1059-61
186. Wittkopp N, Huntzinger E, Weiler C, Saulière J, Schmidt S, et al. (2009) Nonsense-mediated mRNA decay effectors are essential for zebrafish embryonic development and survival. *Mol Cell Biol* 29: 3517-28.
187. Yamashita A, Ohnishi T, Kashima I, Taya Y, Ohno S (2001) Human SMG-1, a novel phosphatidylinositol 3-kinase-related protein kinase, associates with components of the mRNA surveillance complex and is involved in the regulation of nonsense-mediated mRNA decay. *Genes Dev* 15: 2215-28.
188. Yamashita A, Chang TC, Yamashita Y, Zhu W, Zhong Z, Chen CY, et al. (2005) Concerted action of poly(A) nucleases and decapping enzyme in mammalian mRNA turnover. *Nat Struct Mol Biol* 12: 1054-1063.
189. Yamashita A, Izumi N, Kashima I, Ohnishi T, et al. (2009) SMG-8 and SMG-9, two novel subunits of the SMG-1 complex, regulate remodeling of the mRNA surveillance complex during nonsense-mediated mRNA decay. *Genes Dev* 23: 1091-105.
190. Yamashita A (2013) Role of SMG-1-mediated Upf1 phosphorylation in mammalian nonsense-mediated mRNA decay. *Genes to Cells* 18: 161-175.
191. Yepiskoposyan H, Aeschmann F, Nilsson D, Okoniewski M, Mühlemann O (2011) Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. *RNA* 17: 2108-18.
192. Yukihiro M, Ito K, Tanoue O, Goto K, et al. 2011. Effective drug delivery system for duchenne muscular dystrophy using hybrid liposomes including gentamicin along with reduced toxicity. *Biol Pharm Bull* 34:712-6.
193. Zhang Z, Krainer AR (2004) Involvement of SR proteins in mRNA surveillance. *Mol Cell* 16: 597-607.
194. Zhang Z, Krainer AR. 2007. Splicing remodels messenger ribonucleoprotein architecture via eIF4A3-dependent and -independent recruitment of exon junction complex components. *Proc Natl Acad Sci U S A.* 104: 11574-9.