

**Universidad Pública de Navarra**

*Nafarroako Unibertsitate Publikoa*

**ESCUELA TECNICA SUPERIOR  
DE INGENIEROS AGRONOMOS**

*NEKAZARITZAKO INGENIARIEN  
GOI MAILAKO ESKOLA TEKNIKO*

**ANÁLISIS GENÓMICO Y TRANSCRIPTÓMICO DE GENES  
CAPTURADOS POR HELITRONES EN *Pleurotus ostreatus***

presentado por

**LETICIA LÓPEZ VARAS**

**GRADO EN INGENIERRÍA AGROALIMENTARIA Y DEL MEDIO RURAL**

**MENCIÓN EN INDUSTRIAS AGROALIMENTARIAS**

PAMPLONA, JUNIO 2014

## **AGRADECIMIENTOS**

Quiero agradecer a Lucía Ramírez y a Gerardo Pisabarro la oportunidad que me han dado de realizar este trabajo en el grupo de investigación de Genética y Microbiología de la Universidad Pública de Navarra.

También quiero dar las gracias a Raúl Castanera por el tiempo que ha dedicado a enseñarme a lo largo de todo el trabajo.

Y gracias también a todos los compañeros del grupo por hacer los días de trabajo más agradables.

# ÍNDICE

<b>1.- INTRODUCCIÓN.....</b>	<b>6</b>
1.1.- ANTECEDENTES .....	6
1.2.- OBJETIVOS DEL ESTUDIO .....	16
<b>2.- MATERIALES Y MÉTODOS.....</b>	<b>17</b>
2.1.- VARIEDAD DE Pleurotus ostreatus.....	17
2.2.- CONDICIONES DE CULTIVO .....	17
2.2.1.- Mantenimiento y expansión de las cepas.....	17
2.2.2.- Cultivos para la obtención de ADN y ARN .....	17
2.3.- EXTRACCIÓN DE ADN .....	18
2.4.- EXTRACCIÓN DE ARN Y RETROTRANSCRIPCIÓN.....	18
2.5.- DISEÑO DE CEBADORES .....	19
2.6.- RT-PCR EN TIEMPO REAL .....	21
2.7.- CAPTURA DE DATOS.....	22
2.8.- ESTRATEGIA DE CUANTIFICACIÓN DE DATOS DE qPCR.....	22
2.8.1.- Cuantificación de la dosis génica.....	22
2.8.2.- Cuantificación del perfil transcriptómico .....	23
<b>3.- RESULTADOS Y DISCUSIÓN.....</b>	<b>23</b>
3.1.- DOSIS GÉNICA .....	23
3.2.- PERFIL TRANSCRIPTÓMICO .....	29
<b>4.- CONCLUSIONES .....</b>	<b>33</b>
<b>5.- BIBLIOGRAFÍA.....</b>	<b>34</b>

## RESUMEN

El creciente número de organismos eucariotas secuenciados permite estudiar la aparición de elementos de transposición en gran cantidad de genomas. Dentro de los transposones de clase II se encuentran los transposones de círculo rodante (también conocidos como helitrones), que han sido descritos recientemente en plantas, animales y hongos utilizando herramientas bioinformáticas. Los helitrones se caracterizan fundamentalmente por codificar una proteína con un origen de replicación REP y un dominio helicasa PIF1 y por la capacidad de capturar genes enteros o fragmentos de genes, incrementando su número y dispersándolos por el genoma del huésped. En el genoma de *P. ostreatus* se han descrito dos familias de helitrones denominadas HELPO1 y HELPO2, que contienen elementos potencialmente autónomos que codifican helicasas RepHel, además de genes capturados de función desconocida. En este trabajo hemos analizado el efecto de la dosis génica (número de copias) y el perfil transcriptómico de estos helitrones portadores de genes utilizando PCR en tiempo real. Como resultados presentamos un protocolo basado en la PCR a tiempo real para cuantificar el número de copias de estos elementos en otras cepas monocarióticas de *P. ostreatus* de las que no se dispone de secuencia genómica y que existe expresión de algunos genes capturados cuya función aún no se conoce.

## PALABRAS CLAVE

Helitron, helicasa RepHel, genes capturados, *Pleurotus ostreatus*, Hongos, perfiles de expresión, dosis génica.

## ABSTRACT

The increasing number of sequenced eukaryotic organisms allows to study the occurrence of transposable elements (TE) in large amount of genomes. Within the class II transposons are rolling-circle transposons (also known as helitrons), which have recently been described in plants, animals and fungi using bioinformatic tools. Helitrons are mainly characterized by encoding a protein with a replication initiator Rep and a PIF1 helicase domain as well as by the ability to capture whole genes or fragments of genes, increasing the number and spreading by the host genome. Two families of helitrons have been described in the *P. ostreatus* genome, called HELPO1 and HELPO2, which contain putative autonomous elements encoding RepHel

helicases besides captured genes of unknown function. In this work we have analyzed the effect of the gene dosage (copy number) and the transcriptomic profile of these helitron captured genes using real-time PCR. The results showed that it is possible to use this experimental method of analysis by qPCR in other monokaryotic strains of *P. ostreatus*, whose genomic sequences are not available and that there is expression of some captured genes whose function is not yet known.

## **KEYWORDS**

Helitron, RepHel helicase, captured genes, *Pleurotus ostreatus*, Fungi, expression profiles, gene dosage.

# 1.- INTRODUCCIÓN

---

## 1.1.- ANTECEDENTES

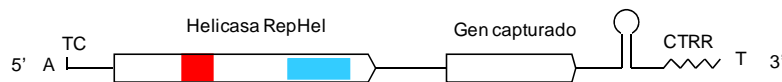
El creciente número de proyectos de secuenciación de organismos eucariotas ofrece la oportunidad de estudiar la presencia de elementos transponibles (TE) en un amplio rango de genomas de procariotas y eucariotas. Los elementos transponibles están involucrados en la organización del genoma, reordenamientos cromosómicos y cambios en la estructura y expresión de los genes. En base a su modo de transposición, los elementos transponibles se clasifican en dos clases (Wicker *et al.*, 2007). La Clase I incluye elementos que se transponen a través de intermediarios de ARN. La Clase II abarca elementos que se transponen directamente de ADN a ADN y la reacción es catalizada por una transposasa. Ambos grupos contienen subclases que se caracterizan en función de la estructura de sus secuencias. De esta manera, los elementos de la Clase I se diferencian entre ellos en base a la presencia o ausencia de Largas Repeticiones Terminales, (como elementos LTR y elementos no-LTR; LTR del inglés *Long Terminal Repeats*). Los elementos de la Clase II incluyen cuatro tipos de transposones de ADN: aquellos cuya secuencia se corta y pega en otra secuencia receptora (Ej: *MITE*), replicativos (Ej: *MuDR*), los transposones de círculo rodante (helitrones) y autosintetizantes (*Polintons o Mavericks*) (Kapitonov and Jurka, 2007). Los TE de Clase II se caracterizan porque tienen repeticiones invertidas terminales (TIRs: *terminal inverted repeats*) y duplicación de sitios de inserción (TSD: *target duplication sites*). Dentro de los elementos de Clase II se pueden encontrar elementos autónomos (que son aquellos que contienen todas las proteínas necesarias para su transposición) y elementos no autónomos, los cuales son copias defectuosas de los primeros debido a deleciones, inserciones o reordenamientos en la secuencia interna. Por lo tanto, los elementos no autónomos para transponerse dependen de las proteínas codificadas por elementos autónomos. Los transposones de círculo rodante, también conocidos como helitrones (uno de los cuatro tipos de transposones de ADN de Clase II), se detectaron y describieron primeramente en *Arabidopsis thaliana* y *Caenorhabditis elegans* mediante análisis bioinformáticos de regiones repetidas (Kapitonov and Jurka, 2001). Más tarde se encontraron en otras plantas, protozoos, hongos, cnidarios, insectos, gusanos, peces, ranas, reptiles y mamíferos (Kapitonov and Jurka, 2001; Poulter *et al.*, 2003; Zhou *et al.*, 2006; Kapitonov and Jurka, 2007; Thomas *et al.*, 2010). Los helitrones se insertan en el genoma entre los nucleótidos AT (Brunner *et al.*, 2005) y se caracterizan porque tienen extremos 5' TC y 3' CTRR conservados, así como por una secuencia de 16 a 20 nucleótidos,

que forma una horquilla palindrómica de al menos 6 pb localizada unos 12 nucleótidos aguas arriba del extremo 3' CTRR (Kapitonov and Jurka, 2001) y un bucle en la horquilla de 2 a 4 pb. Los posibles helitrones autónomos contienen genes que codifican una proteína RepHel con un dominio iniciador de la replicación (Rep) y un dominio helicasa (Hel). Ambos dominios se suponen esenciales para su transposición (Figura 1A). El dominio Rep está probablemente involucrado en las roturas endonucleolíticas del ADN, como la escisión y la religación del ADN transpuesto (Kapitonov and Jurka, 2007). El dominio Hel codifica una helicasa 5'-3' de ADN perteneciente a la familia PIF1/RRM3 (altamente conservada desde levaduras a humanos) que participa en el mantenimiento de la estabilidad del genoma (Boulé and Zakian, 2006). A pesar de que no se conoce el mecanismo exacto de transposición de helitrones, la hipótesis que actualmente se acepta es que se transponen mediante un mecanismo de círculo rodante (Kapitonov and Jurka, 2007) similar al de los elementos de transposición bacterianos IS91 (Toleman *et al.*, 2006).

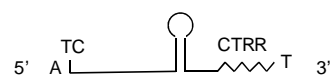
**A) ESTRUCTURA HELITRON AUTÓNOMO**



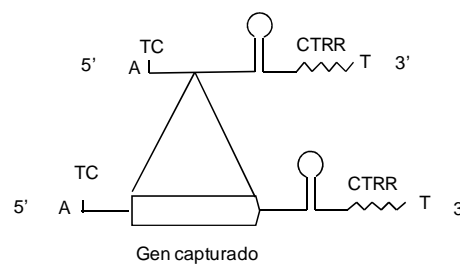
**B) ESTRUCTURA HELITRON AUTÓNOMO CON GEN CAPTURADO**



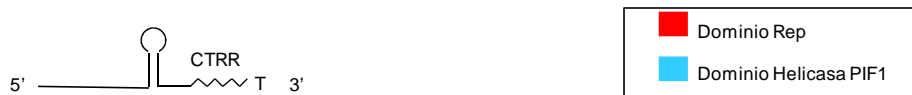
**C) ESTRUCTURA HELITRON NO AUTÓNOMO INTACTO**



**D) ESTRUCTURA HELITRON NO AUTÓNOMO INTACTO CON GEN CAPTURADO**



**E) ESTRUCTURA HELITRON TRUNCADO**



*Figura 1: Representación esquemática de las características estructurales de helitrones de eucariotas. (A) Helitron autónomo. (B) Helitron autónomo con gen capturado aguas abajo de la helicasa RepHel. (C) Helitron no autónomo intacto. (D) Helitron no autónomo intacto con gen capturado. (E) Helitron truncado en el extremo 5'. (Fuente: adaptada de (Castanera *et al.*, n.d.).*

Durante su transposición, los helitrones pueden capturar, amplificar y dispersar genes completos (Jameson *et al.*, 2008) y fragmentos de genes mediante un mecanismo aún desconocido (Morgante *et al.*, 2005; Lai *et al.*, 2005; Brunner *et al.*, 2005; Pritham and Feschotte, 2007; Yang and Bennetzen, 2009b) (Figuras 1B y 1D). Por lo tanto, contribuyen a la ruptura de colinearidad genética, como se ha descrito en haplotipos de maíz. Dada su capacidad para capturar y movilizar fragmentos de genes, los helitrones juegan un papel importante en la creación de nuevas proteínas a través de reordenamiento de exones y la duplicación de genes. Esta capacidad parece ser un mecanismo importante para la evolución de nuevos helitrones (Feschotte and Pritham, 2009). Según (Yang and Bennetzen, 2009b), los genes capturados por helitrones en maíz parecen conferir ventajas a los haplotipos portadores. La longitud de los helitrones (pueden ir desde 202 pb hasta 35,9 Kb en maíz), así como su abundancia en genomas eucariotas es muy variable. Por ejemplo, en la mosca de la fruta *Drosophila melanogaster*, los helitrones representan entre el 1 y el 5% de su genoma total (Kapitonov and Jurka, 2007) y en mamíferos un 3% (Pritham and Feschotte, 2007). En plantas su contribución al tamaño total del genoma es variable: en *A.thaliana* representan más del 2% (Kapitonov and Jurka, 2001), en *Oryza sativa* representan entre el 0,03 y el 1% (Zuccolo *et al.*, 2007), y en maíz representan menos del 0,01% (Messing *et al.*, 2004), aunque Yang y Bennetzen han demostrado que los helitrones representan alrededor del 2,2% del genoma en la línea de maíz B73 (Yang and Bennetzen, 2009b). En hongos, se han identificado secuencias de helitrones mediante análisis bioinformáticos en genomas de especies de basidiomicetos (*Phanerochaete chrysosporium*, *Coprinopsis cinereus*, *Ceriporiopsis subvermispora*, *Laccaria bicolor* y *Puccinia graminis*) ascomicetos (*Aspergillus nidulans*, *Chaetomium globosum*, *Fusarium oxisporum* y *Tolyocladium inflatum*) así como en el zigomiceto *Rhizopus oryzae* (Cultrone *et al.*, 2007; Kapitonov and Jurka, 2007; Kempken, 2008; Feschotte *et al.*, 2009; Martin *et al.*, 2010; Labbé *et al.*, 2012; Fernández-Fueyo *et al.*, 2012). Sin embargo, hasta el momento, sus características no se han analizado. *Pleurotus ostreatus* es un basidiomiceto responsable de la podredumbre blanca de la madera. Ampliamente utilizado como organismo modelo para el estudio de genomas de hongos degradadores de lignina. En los últimos años, el genoma de la cepa dicariótica N001 de *P. ostreatus* (que mide aproximadamente 34 Mb y está organizado en 11 cromosomas) ha sido analizado ampliamente, debido a que las secuencias de los dos genomas de las cepas monocarióticas PC9 y PC15 que la componen están actualmente disponibles en el JGI (N001 contiene los núcleos de PC9 y PC15 compartiendo un citoplasma común). El análisis de las secuencias de estas dos cepas reveló la presencia de helitrones en sitios específicos de sus



genomas (como se describió para los diferentes haplotipos de maíz) lo que produce una interrupción de la sintenia entre ambas cepas y la aparición de importantes reordenamientos. Con el objetivo de analizar el papel de los helitrones en el genoma de *P. ostreatus* se realizó un estudio completo sobre la abundancia de helitrones y las características de los mismos en las cepas monocarióticas PC9 y PC15 (Castanera *et al.*, n.d.). El presente trabajo constituye la continuación de los estudios realizados hasta el momento en el grupo de investigación de Genética y Microbiología y que se resumen a continuación:

Se identificaron los helitrones presentes en las cepas PC9 y PC15 mediante análisis bioinformáticos utilizando en la búsqueda parámetros relacionados con su estructura y en homología con proteínas RepHel de otros organismos. En este sentido se han identificado y definido helitrones intactos como aquellos elementos que contienen las secuencias 5' y 3' (Figuras 1A, 1B, 1C y 1D), y elementos truncados aquellos que contienen un extremo 3' intacto pero carecen de un extremo 5' conservado (Figura 1E). La clasificación de los helitrones se realizó de la siguiente manera: aquellos elementos que presentaban una similitud igual o mayor al 80% en las últimas 30 pb del extremo 3' se consideraron pertenecientes a la misma familia. Los elementos que pertenecían a la misma familia pero que mostraban una similitud inferior al 80% en las últimas 30 pb del extremo 5' se clasificaron como pertenecientes a una determinada subfamilia, de acuerdo con (Yang and Bennetzen, 2009a). Para nombrar los elementos transponibles se tiene en cuenta la especie a la que pertenecen y de qué tipo de elemento transponible se trata. Así HELPO significa que se habla de helitrones de *P. ostreatus*. Los números que acompañan la denominación indican la familia y subfamilia a la que pertenecen. Los elementos autónomos se nombran con letras mayúsculas y los elementos no autónomos con minúsculas. A los elementos truncados se les añade un asterisco. Utilizando letras normales cuando se refieren a familias y subfamilias y letras cursivas para copias específicas (por ejemplo la subfamilia HELPO1.1 vs el elemento *HELPO1.1*).

Basándose en la similitud de los extremos 5' y 3' los helitrones se dividen en dos familias: HELPO1 y HELPO2. Dentro de la familia HELPO1 se encuentran tres subfamilias: HELPO1.1, HELPO1.2 y HELPO1.3, que presentan elementos que varían entre 1,5 y 13,7 Kb de longitud (Figuras 2A y 2B). La familia HELPO2 tiene elementos de 3,9 a 10 Kb de longitud. Esta familia contiene un extremo 3' CTTG en lugar de CTAG y no contiene ninguna subfamilia (Figuras 2C y 2D). Ambas familias contienen elementos potencialmente autónomos. Estos elementos llevan una ORF que codifica una RepHel helicasa de aproximadamente 1.400 aminoácidos (Figuras 2B y 2D). Esta proteína contiene los tres motivos que definen el dominio Rep (Kapitonov and Jurka, 2007) así como los seis motivos conservados presentes en miembros de la superfamilia helicasa

SF1 descritos en otros helitrones (Pritham and Feschotte, 2007; Han *et al.*, 2013) y necesarios para la replicación y desenrollado de la molécula de ADN. La familia HELPO1 es la única que contiene copias no autónomas intactas, pertenecientes a la subfamilia HELPO1.3 (Figura 2B).

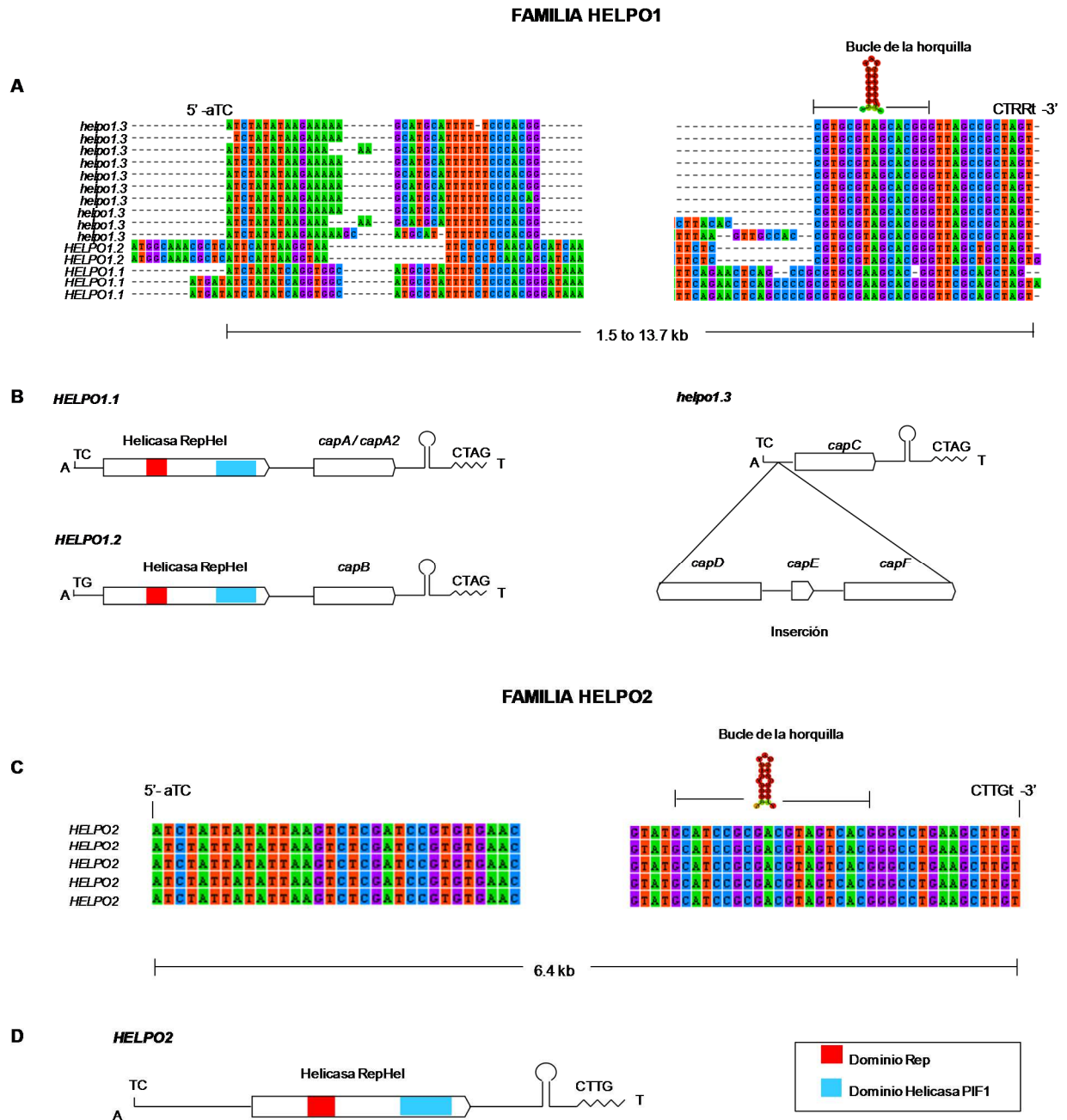


Figura 2: Representación esquemática de las características estructurales y codificantes de los distintos elementos pertenecientes a las familias de helitrones HELPO1 y HELPO2 de *P. ostreatus*. (Fuente: adaptada de (Castanera *et al.*, n.d.) ).

Los genes completos presentes dentro de helitrones intactos se analizaron usando el *browser* del JGI y los modelos de genes predichos (excepto helicasas RepHel) se consideraron como genes capturados (Castanera *et al.*, n.d.).

Se ha visto que los helitrones en *P. ostreatus* pueden llevar fragmentos de genes, de otros elementos transponibles y genes cuya secuencia se desconoce. Así, se ha observado que elementos autónomos de la familia HELPO1 contenían un gen capturado (*cap*) aguas abajo de la helicasa RepHel. Los helitrones de la familia HELPO1 mostraron una mayor tendencia que los pertenecientes a otras familias a la adquisición/creación de genes. Así, cada copia intacta llevaba de uno a cuatro genes (Figura 2B). Se convino en llamar a los genes capturados de la subfamilia HELPO1.1 *capA*, los de la subfamilia HELPO1.2 *capB*, y los de la subfamilia HELPO1.3 *capC*, *capD*, *capE* y *capF* (Figura 2B). En PC15, el gen capturado de la copia *HELPO1.1* (que se localiza en el cromosoma V) se llamó *capA2* para diferenciarlo de *capA* debido a la baja similitud que presentaba con otros genes *capA* (45%). El elemento no autónomo *helpo1.3* (localizado en el cromosoma I) es 10.7 Kb más larga que la media de las otras copias de *helpo1.3* presentes en el genoma de *P.ostreatus* (12.2 Kb vs 1.5 Kb) (Figura 2B) y es la única copia que lleva los cuatro genes capturados de HELPO1.3. El resto de copias *helpo1.3* mantienen sólo el gen *capC*. Finalmente la familia HELPO2 carece de genes capturados (Fig 2D).

La información antes descrita, que representa la localización de helitrones en las cepas PC9 y PC15 de *P. ostreatus* (Figura 3) puede ser utilizada para averiguar la dosis génica (el número de copias de cada helitron presentes en cada cepa), y la actividad transcripcional de las proteínas RepHel características de los helitrones y de los genes capturados por los mismos.

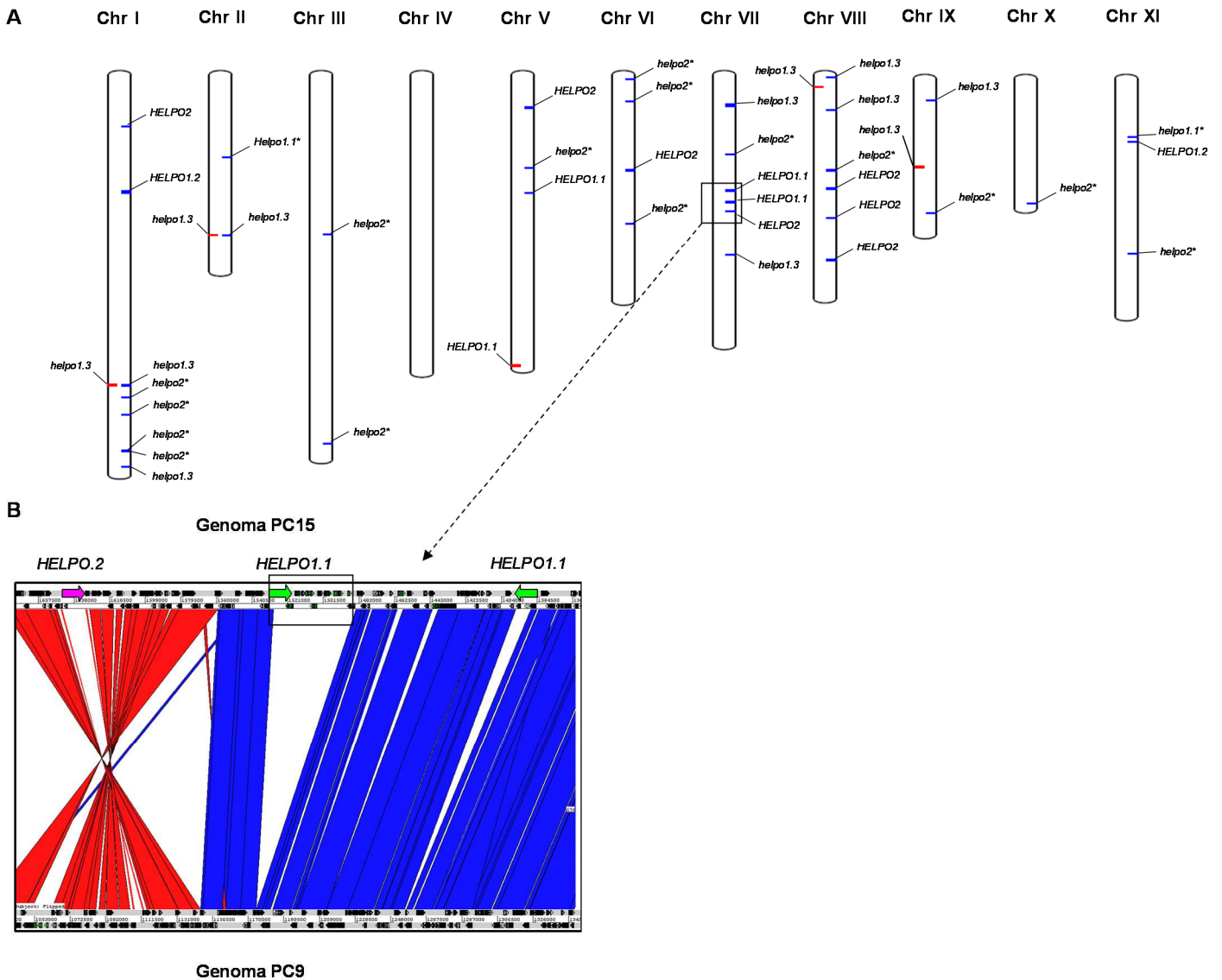


Figura 3: (A) Distribución de helitrones en los genomas de PC9 y PC15. Los elementos autónomos completos se muestran en letras mayúsculas y los elementos no autónomos en letras minúsculas. Los elementos truncados se representan con \*. Los helitrones de PC15 se representan con líneas horizontales azules y los de PC9 con líneas horizontales rojas. (B) Comparación ACT entre la región del cromosoma VII de PC9 y PC15, marcada con un recuadro. (Fuente: adaptada de (Castanera et al., n.d.)).

La aplicación de la técnica de PCR en tiempo real permite investigar los dos aspectos antes mencionados. Para ello se debe entender que la expresión génica es el proceso mediante el cual la información almacenada en el ADN es usada para dirigir la síntesis de un producto génico específico. Estrictamente el término abarca desde la activación del gen hasta que la

proteína madura se localiza en el lugar adecuado y realiza su función (Hernández *et al.*, 1994). En todos los organismos vivos las células regulan sus actividades activando o silenciando la expresión de sus genes y esta expresión es generalmente proporcional al número de copias de ARN mensajero de un gen determinado. Este hecho es crucial si tratamos de identificar la presencia de productos celulares específicos, ya que el ARNm se traduce en los ribosomas del citoplasma para formar proteínas (McPherson *et al.*, 2008).

Los métodos usados comúnmente para la cuantificación del ARN han sido: el Northern blot, el ensayo de protección de RNAsas, la supresión por hibridación sustractiva SSH y la RT-PCR. Recientemente se han descrito otras técnicas para el estudio de la expresión génica como la RT-qPCR (o RT-PCR a tiempo real) la PCR-digital (Vogelstein and Kinzler, 1999) y técnicas de alta productividad como los chips de ADN, la secuenciación de ESTs, la tecnología SAGE y el RNAseq. La RT-qPCR se puso a punto en 1993 (Higuchi *et al.*, 1993), y actualmente se considera como el estándar más fiable para validar los resultados de otros métodos como los basados en secuenciación (RNA-seq) o hibridación (microarrays). Su capacidad para detectar y medir cantidades pequeñas de ácidos nucleicos en un amplio rango de muestras de diferentes orígenes la convierte en una herramienta excelente para el diagnóstico molecular, las ciencias de la vida, agricultura y medicina (Bustin, 2000). La RT-qPCR posee una gran sensibilidad (detecta bajos niveles de ADNc molde), alta especificidad, buena reproducibilidad, un amplio rango de cuantificación (de hasta siete órdenes de magnitud) (Pfaffl, 2004), alta precisión en la cuantificación, procesamiento de un elevado número de muestras en paralelo y rapidez en la obtención de resultados. La alta sensibilidad en la detección de transcritos poco abundantes es crucial para una detección precisa en diversas situaciones como en tejidos con concentraciones de ARN bajas, cuando se dispone de poca muestra o cuando se pretende discriminar cambios pequeños en la expresión génica (Schmittgen *et al.*, 2000). Por el contrario, algunos de los inconvenientes de esta técnica son que no aporta información sobre los tamaños de ARNm, requiere información previa sobre la secuencia del transcrito, el resultado se ve afectado por el error cometido por la Taq ADN polimerasa y la cuantificación absoluta depende exclusivamente de la calidad del estándar utilizado (Smith and Osborn, 2009). Si comparamos entre las técnicas transcriptómicas disponibles, los microarrays y las basadas en la secuenciación masiva son muy útiles para analizar la expresión de muchos genes en pocas condiciones, mientras que la RT-qPCR y la PCR digital son muy adecuadas para detectar la expresión de pocos genes en muchas condiciones.

El fundamento del método de la RT-qPCR consiste en una extracción de ARN total seguida por la síntesis de ADNc a partir del ARNm utilizando una transcriptasa reversa (RT) y por una PCR que se monitoriza en tiempo real (qPCR) gracias a la adición de un fluorocromo. Existen dos clases principales de marcadores fluorescentes para esta técnica: i) marcadores específicos que emplean sondas marcadas con moléculas fluorescentes que se unen de manera específica al producto de PCR. Estos tienen la ventaja de ser muy precisos y evitan posibles amplificaciones no específicas o artefactos en el producto de la PCR, y ii) los marcadores inespecíficos, que son moléculas fluorescentes que se unen a todas las secuencias de ADN de doble cadena en una reacción de PCR. Una vez que el fluorocromo se une al ácido nucleico formado en la reacción, este emite una señal fluorescente que es detectada por el sistema óptico del termociclador y procesada en tiempo real (Walker, 2002). En cada ciclo de amplificación se produce un aumento de producto específico (la cantidad se duplica si la eficiencia es del 100%) que produce un aumento en la fluorescencia detectada (Figura 4).

Esto permite que las concentraciones de ADN o ADNc puedan ser cuantificadas en la fase exponencial de la reacción, donde la cantidad de ADN inicial y la cantidad de producto guardan una estrecha relación (Ririe *et al.*, 1997). Los marcadores inespecíficos más utilizados son los colorantes SYBR® Green y SYBR® Gold ya que son más baratos que los específicos y están disponibles como cócteles listos para usar que no requieren un diseño experimental adicional (Ririe *et al.*, 1997; Giglio, 2003).

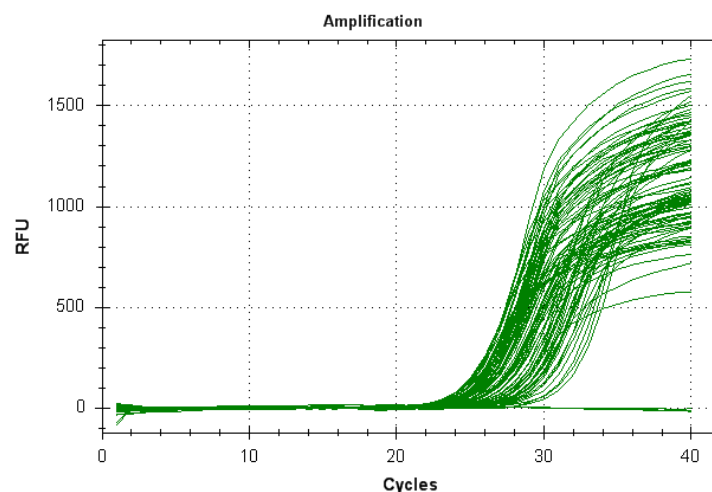


Figura 4: Curva de fluorescencia detectada en el termociclador durante la amplificación del ADNc. El eje de abscisas representa el número de ciclos de PCR y el de ordenadas la fluorescencia relativa (RFU).

La principal limitación de estos marcadores es que al unirse a todas las secuencias de doble cadena de ADN en la reacción de PCR, emiten una señal fluorescente tanto para productos específicos como para aquellos que no lo son (como dímeros de primer y amplificaciones inespecíficas). La co-amplificación de productos inespecíficos se detecta al final del experimento mediante la aplicación de un programa de curva de disociación del producto de PCR (Figuras 5A y 5B). De esta forma se obtiene una representación gráfica de la fluorescencia detectada en función de la temperatura, detectándose el punto en el que el 50% del producto se encuentra disociado (temperatura de fusión). La forma y posición de estas curvas depende del contenido de GC/AT y del tamaño de los amplicones. Para identificar los productos específicos de la PCR hay que tener en cuenta que estos se disocian a una temperatura más alta que los artefactos como dímeros de primer (Ririe *et al.*, 1997).

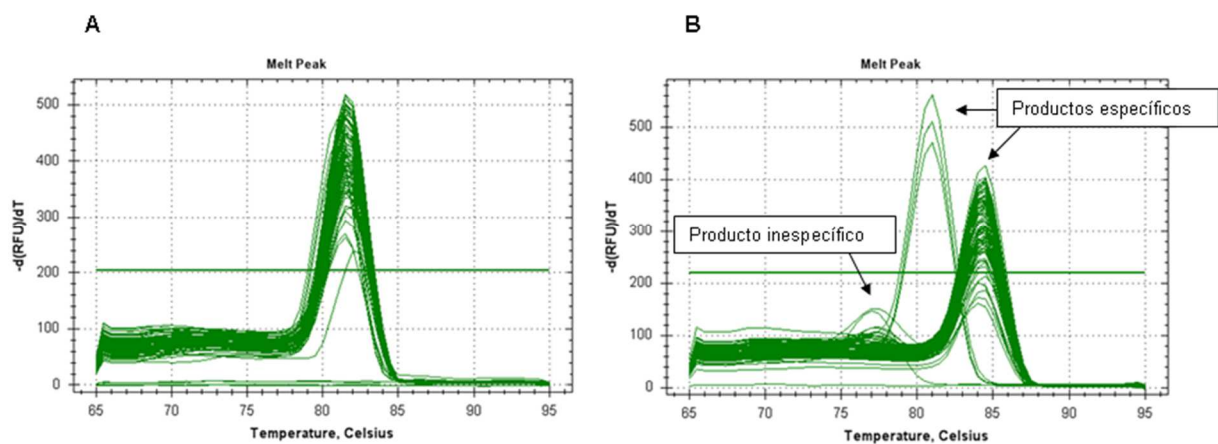


Figura 5: Curva de fusión de un sólo producto específico (A), y de dos productos específicos junto con artefactos de amplificación (B), detectadas en el termociclador CFX96 (BIO-RAD).

A lo largo de la reacción de PCR el termociclador detecta y cuantifica la fluorescencia emitida quedando registrada en forma de curva (Figura 4). En esta curva el investigador fija un punto por encima de la línea base que se denomina umbral. Los valores umbral del ciclo (valores Ct) se determinan identificando el ciclo en el cual la emisión de la intensidad del marcador fluorescente atraviesa el umbral por encima del ruido de fondo y en la fase exponencial de la reacción, y son los valores que se utilizarán para calcular la expresión génica. Los criterios más importantes para la optimización de esta técnica son la especificidad, eficiencia y reproducibilidad (Edwards, 2004). La especificidad depende de los cebadores usados y del protocolo de amplificación, mientras que la eficiencia de amplificación depende de factores

como el diseño de los cebadores, la concentración de inhibidores en las muestras y el protocolo de amplificación. El criterio de la reproducibilidad depende principalmente de las características del aparato y de la precisión en el pipeteo. La cuantificación de la expresión génica puede ser absoluta o relativa. En la cuantificación absoluta se relaciona la fluorescencia obtenida en la amplificación de la muestra con la obtenida en una recta de calibración realizada con ADN clonado, ADN genómico o producto de PCR (Pfaffl, 2004). De este modo, los niveles de expresión se obtienen habitualmente en número de copias (nº de copias por gramo de tejido, o por microgramo de ARN, etc). En cambio, en la cuantificación relativa se compara el nivel de expresión de cada gen a estudiar *versus* la expresión de dicho gen en una muestra de referencia. En este tipo de cuantificación las unidades en las que se mide la transcripción son arbitrarias ya que sólo expresan el aumento o disminución de transcrito entre muestras. Para compensar el error introducido por el pipeteo, la cuantificación del ARN, la eficiencia de la retrotranscripción y el nivel de degradación del ARN entre muestras, se debe incluir un control endógeno que pase por todas las etapas del proceso junto al gen que se desea amplificar. En este estudio se han utilizado los genes *pep* y *sar1* como controles endógenos debido a: i) la estabilidad de su expresión (Castanera *et al.*, 2013) y ii) que ambos presentan una única copia dentro de los genomas haploides de las cepas PC9 y PC15 que constituyen la cepa dicariótica N001.

## **1.2.- OBJETIVOS DEL ESTUDIO**

Los objetivos de este trabajo son: 1) poner a punto un método para determinar el número de copias de cada helitrón que están presentes en los monocariontes PC9 y PC15 (parentales de la cepa dicariótica N001) y su comparación con el número de helitrones presentes en ambas cepas determinado por métodos bioinformáticos, y 2) analizar el perfil transcriptómico de la proteína RepHel presente en helitrones de las familias HELPO1 y HELPO2 y de los diferentes genes capturados por los elementos transponibles de las dos familias antes citadas. Ambos objetivos serán de utilidad para la determinación del número de helitrones y de la actividad de los mismos y de sus genes capturados en cepas monocarióticas cuya secuencia genómica se desconoce.



## 2.- MATERIALES Y MÉTODOS

---

### 2.1.- VARIEDAD DE *Pleurotus ostreatus*

En este estudio se empleó la cepa dicariótica N001 de *Pleurotus ostreatus* var. *florida* (Colección Española de Cultivos Tipo, número de acceso CECT20600) y sus cepas parentales PC15 y PC9. Las secuencias genómicas de los protoclones haploides que constituyen la cepa N001 (PC9 y PC15) se encuentran disponibles en la página web del *Joint Genome Institute* (<http://genome.jgi-psf.org>).

### 2.2.- CONDICIONES DE CULTIVO

#### 2.2.1.- Mantenimiento y expansión de las cepas

El crecimiento de las cepas se realizó en placas Petri con medio sólido PDA (24 g de caldo de patata-dextrosa y 15 g de agar por litro) a 24°C y en oscuridad.

#### 2.2.2.- Cultivos para la obtención de ADN y ARN

Los cultivos de las cepas monocarióticas PC9, PC15 y de la cepa dicariótica N001 se prepararon por triplicado en Erlenmeyers que contenían 150 ml de medio líquido SMY (10 g de sacarosa, 10 g de extracto de malta y 4 g de extracto de levadura por litro), a los que se les agregaron cuatro fragmentos de micelio de 0,5cm x 0,5cm crecidos en medio sólido (PDA). Los cultivos se hicieron crecer en agitación orbital (130 rpm) a 24°C y en oscuridad durante 8 días.



Figura 6: Medio de cultivo líquido SMY utilizado para el crecimiento de las cepas de *P. ostreatus* (izquierda). Crecimiento cepa N001 en medio líquido (derecha).

### **2.3.- EXTRACCIÓN DE ADN**

El micelio recolectado se congeló, molió en un mortero estéril en presencia de nitrógeno líquido y se conservó a -80 °C. El ADN total se extrajo de unos 200 mg de muestra usando el kit comercial de extracción de ADN <sup>TM</sup>Fungal DNA Mini Kit (Omega Bio-Tek, Norcross, GA), de acuerdo con las instrucciones del fabricante. Para evitar la presencia de ARN en las muestras se realizó un tratamiento con RNase A (Omega Bio-Tek, Norcross, GA). La concentración de ADN fue medida usando un fluorómetro Qubit® 2.0 Fluorometer (Life Technologies) y la pureza de las muestras se estimó por la relación de absorbancia 260/280 nm (valores entre 2-2,15 indican una alta pureza del ADN) medida mediante un espectrofotómetro NanoDrop<sup>TM</sup> 2000 (Thermo-Scientific, Wilmington, DE). Posteriormente el ADN se conservó a -20° C hasta su utilización.

### **2.4.- EXTRACCIÓN DE ARN Y RETROTRANSCRIPCIÓN**

Al igual que en la extracción de ADN, el micelio recolectado se congeló, molió en un mortero estéril en presencia de nitrógeno líquido y se conservó a -80 °C, posteriormente el ARN total se extrajo de unos 200 mg de muestra usando el kit comercial de extracción de RNA E.Z.N.A. <sup>TM</sup>Fungal RNA (Omega Bio-Tek, Norcross, GA), de acuerdo con las instrucciones del fabricante. Para evitar la presencia de ADN genómico en las muestras se realizó un tratamiento con DNasa usando 1 U de RQ1 DNase (Promega, Madison, WI) por cada µg de ARN. La concentración de ARN se midió usando un fluorómetro Qubit® 2.0 Fluorometer (Life Technologies) y la pureza de las muestras se estimó de acuerdo con la relación de absorbancia 260/280 nm (valores entre 2-2,15 indican una alta pureza del ARN) medida mediante un espectrofotómetro NanoDrop<sup>TM</sup> 2000 (Thermo-Scientific, Wilmington, DE).

La integridad del ARN total obtenido se analizó mediante electroforesis en gel de agarosa desnaturante al 1% (peso/volumen). Se consideró que el ARN estaba degradado si no se distinguían nítidamente las bandas correspondientes al ARN ribosomal 28S y 18S (descartando en ese caso la muestra y repitiendo el proceso de extracción) (Figura 7). Para la reacción de retrotranscripción se utilizaron 225 ng de ARN total en un volumen de 20 µl usando el kit comercial iScript<sup>TM</sup> cDNA Synthesis (BIO-RAD Laboratories S.A., Alcobendas, España).

La reacción de retrotranscripción (RT) se llevó a cabo en un termociclador PTC-200 Peltier Thermal Cycler (MJ Research, Inc). El protocolo de retrotranscripción consistió en 5 minutos a 25°C, 30 minutos a 42°C y 5 minutos a 85°C. Posteriormente el ADNc se diluyó a una concentración 1:20 para su uso en la RT-qPCR y se conservó a -20° C.

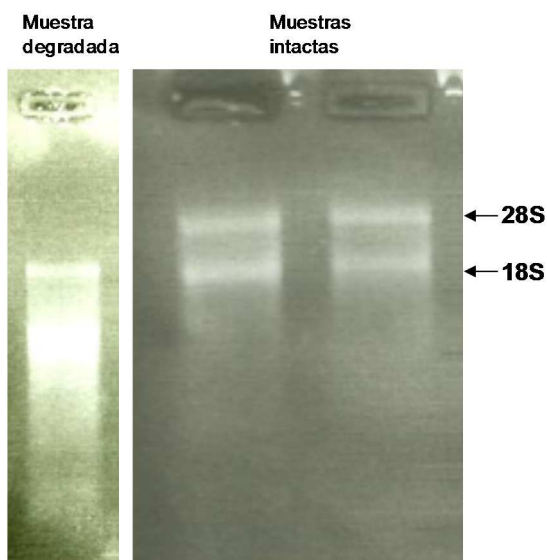


Figura 7: Ejemplo de gel de agarosa desnaturalizante: en la línea de la izquierda se muestra ARN degradado y en las líneas central y derecha se muestra ARN intacto con las bandas de ARN ribosomal 28S y 18S (señaladas por flechas).

## 2.5.- DISEÑO DE CEBADORES

Se diseñaron 4 pares de cebadores específicos para el análisis de la dosis génica (HELPO1.1\_dg, HELPO1.2\_dg, helpo1.3\_dg y HELPO2\_dg), otros 8 pares de cebadores específicos para el análisis de la expresión génica de los genes RepHel y los genes capturados (HELPO1.1rep, HELPO1.2rep, *capA*, *capA2*, *capB*, *capC*, *capD*, y *capF*) (Tabla1. Figura 8), usando el software *Primer3* (Koressaar and Remm, 2007; Untergasser *et al.*, 2012). Además se utilizaron otros siete pares de cebadores para la amplificación de siete genes de copia única: *lacc3*, *lacc11*, *mnp3*, *actin2*, *sar1*, *pep* y *cyph* (Castanera *et al.*, 2013). El objetivo era analizar la variabilidad en la cuantificación de ADN obtenida tras su amplificación (todos deberían presentar el mismo Ct). Las características de los cebadores y amplicones empleados se describen en la Tabla 1. La ausencia de productos de qPCR inespecíficos se confirmó mediante el análisis de la curva de fusión (Ririe *et al.*, 1997).

Tabla 1. Cebadores utilizados en experimentos de qPCR y RT-qPCR. La tabla muestra: la longitud del amplicón, la eficiencia de amplificación y la especificidad de los cebadores en números de copias. Con “dg” se representa los cebadores utilizados en el análisis de la dosis génica y “rep” representa los cebadores utilizados en el análisis del perfil de expresión.

Gen	ID Tránsito		Secuencia	Longitud amplicón (bp)	Eficiencia de amplificación (%) <sup>(*)</sup>	Especificidad (copias)	
	PC15 v2.0 <sup>a</sup>	PC9 v1.0 <sup>b</sup>				PC15	PC9
HELPO1.1_dg			Fw GGAAGGACGCCGAGTTCACATTCAC	79		2	1
			Rv CTTGTTGATGGTCATCGAGAACGCAA				
HELPO1.2_dg			Fw TCGTGCTACTGGTGATCTC	113		2	0
			Rv ATAGACGTCTCCTCGAGA				
helpo1.3_dg			Fw GCGCGCACATTCTTAATTTAC	83		6	5
			Rv CTGCTTACAGTACAGTGGTTATCCA				
HELPO2_dg			Fw AAAGTGGGACTCTGAAGA	127	1,87	6	1
			Rv CAGCTGTGGTGCTCCAGTA				
HELPO1.1rep			Fw CCAGATGCCGAGATCAAGCTTCG	99	1,87	4	
			Rv GGCATAATCATGGCAACCTC				
HELPO1.2rep			Fw CCACCAACTCCACAGAAT	84	1,86	2	
			Rv TCTAGCACCCCGATTATG				
capA			Fw AAATGGACCCCTCCGTTTAC	103	1,98	3 (HELPO1.1)	
			Rv TTTCTGCAAGGGACCCATAG				
capA2			Fw CCTGTTGTTGCATGATCCAG	91	1,89	1 (HELPO1.1)	
			Rv GATGTGCGCCTCAGTAGACA				
capB			Fw GGGCTTGCTGATTGGAAAA	89	1,91	2 (HELPO1.2)	
			Rv TGGGGAGCGAGATAGAATTG				
capC			Fw CACGAGCAATTTTGCAATG	85	1,87	6 (helpo1.3)	
			Rv GTAAGGGTCTGAGCAGCAG				
capD			Fw GCAGAGCAGCAGAGTTTCT	102	/	1 (helpo1.3)	
			Rv AAAATCCCGGTACGTGTTCA				
capF			Fw CATTGGACTGGGAATCTGCT	93	/	1 (helpo1.3)	
			Rv CCCTGCTTTTGACTTCAGC				
pep	1092697	115017	Fw TGATTCCAGAGGACAAGGACGCAA	148	1,89	1	1
			Rv AAATCTCCGCGATACGGTCACT				
sar1	1052294	69194	Fw GGATAGTCTTCTCGTCGATAG	133		1	1
			Rv GGGTGCGTCAATCTTGTTAC				
cyph	1058252	72928	Fw GACATTGCTATCGACTCCCAG	84		1	1
			Rv GAAATTCCTTGCACTCTGGG				
lacc3	1102751	123288	Fw TCGTTTCCGTCTCGTTTCTC	134		1	1
			Rv CTGCGAAGATTTGGATGCTG				
lacc11	1043488	90573	Fw CCTGAATGGTCTGATCTCTGC	93		1	1
			Rv CCTATGACTGGGCTCTTCG				
mnp3	1089546	51690	Fw ATGGACAAGTTGGTCACTCGGT	139		1	1
			Rv CACACGCTTGCTCGATGTTGTTCA				
actin2	1114037	90915	Fw GCCGTGATCTTACCGACTA	134		1	1
			Rv CTCCTGCTCAAAGTCCAA				

<sup>a,b</sup> Identificadores del modelo de tránsito correspondiente a la versión 2.0 de PC15 y la versión 1.0 de PC9, protoclones haploides que constituyen la cepa N001 ([www.jgi.doe.gov](http://www.jgi.doe.gov)).

(\*) La eficiencia de amplificación sólo se calculó para los genes utilizados en el estudio de expresión.

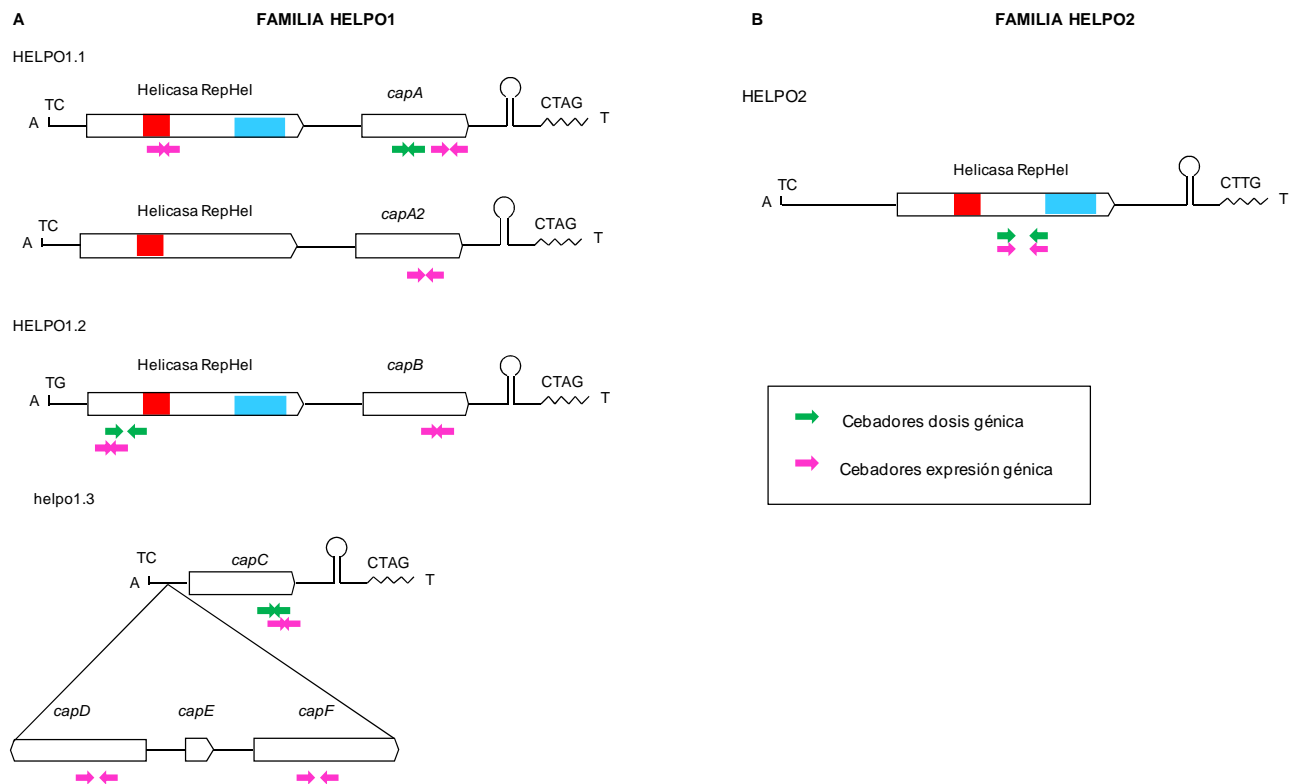


Figura 8: Localización de los cebadores utilizamos para el estudio de la dosis génica y expresión génica dentro de cada familia de helitrones. (A) Helitrones y genes capturados familia HELPO1. (B) Helitrones familia HELPO2. (Fuente: adaptada de (Castanera et al., n.d.))

## 2.6.- RT-PCR EN TIEMPO REAL

Las reacciones de RT-PCR en tiempo real (o RT-qPCR) se realizaron en un termociclador CFX96™ Real-Time System C1000™ (Bio-Rad Laboratories, S.A.). Para detectar los productos de amplificación se usó el agente intercalante SYBR Green. Cada reacción se ajustó a un volumen final de 20 µl que contenía: 10 µl iQ™ SYBR® Green Supermix (mezcla optimizada de iTaq™ DNA polimerasa, SYBR Green, dNTPs, buffer y estabilizadores) (Bio-Rad laboratories, S.A., Alcobendas, España), 2 µl de 3 µM de cada cebador (directo y reverso), 1 µl de ADNc de dilución 1:20 y 5 µl de agua estéril.

El programa de amplificación se realizó de acuerdo con el siguiente protocolo: 5 min a 95° C, seguido de 40 ciclos de 15 s a 95 °C y 30 s a 60 °C, seguidos de 1 min a 95° C y 1 min a 65 °C. Al final se añadió un programa de curva de fusión con incrementos de 0,5 °C cada 5 s en un gradiente lineal desde 65 °C hasta 95 °C. La lectura de fluorescencia se realizó a alta temperatura (72 °C) con el fin de eliminar el impacto de los artefactos de PCR en la cuantificación del ADNc, cuya ausencia se confirmó mediante el análisis de la curva de

fusión. Cada reacción se realizó por triplicado en placas de 96 pocillos y se incluyeron controles negativos sin ADNc (NTC, “*Non-template controls*”) y controles de no-retrotranscrito (NRT, *non-retrotranscribed controls*) para cada par de cebadores.

## 2.7- CAPTURA DE DATOS

Los datos de fluorescencia crudos (RFUs, *Relative Fluorescence Units*) se procesaron mediante el software Bio-Rad CFXManager del termociclador, quedando registrados en forma de curva de fluorescencia. Los valores Ct (valores umbral del ciclo, Cp o “Crossing Point”, (Bustin *et al.*, 2009) se calcularon realizando un ajuste matemático de cada curva de fluorescencia mediante regresión multivariable. Los resultados (RFUs y Cts) se exportaron a Microsoft Excel para su posterior análisis. La eficiencia de amplificación para cada gen se calculó usando el software LinRegPCR. Mediante esta aplicación se analiza la fase exponencial del histórico de fluorescencia (RFU) en escala logarítmica, realizando regresiones lineales sucesivas en “ventanas de linealidad” de cuatro ciclos (WOL, “window-of-linearity”, (Ramakers *et al.*, 2003). Sobre los 4 ciclos con mayor pendiente se calcula la eficiencia de amplificación en base a la fórmula 1. Los valores promedio para cada amplicón se muestran en la tabla 1.

Fórmula 1.  $E = 10^{[-1/\text{pendiente}]}$

## 2.8.- ESTRATEGIA DE CUANTIFICACIÓN DE DATOS DE qPCR

### 2.8.1.- Cuantificación de la dosis génica

Los valores Ct crudos (los valores Ct obtenidos directamente del termociclador) se procesaron utilizando el software Microsoft Excel. El número de copias de cada gen de interés (GOI) se calculó como cantidades relativas (Fórmula 2) usando los genes *pep* y *sar1* (Tabla 1) como controles internos.

Fórmula 2:  $RQ = 2^{-(Cp_{GOI} - \frac{1}{2}(Cp_{sar1} + Cp_{pep}))}$

En esta ecuación:

RQ: Cantidad Relativa (*Relative Quantity*).

Cp<sub>GOI</sub>: es el Ct crudo del gen de interés.

Cp<sub>sar1</sub> y Cp<sub>pep</sub>: son el Ct crudo del control interno (los genes *pep* y *sar1*).

### 2.8.2.- Cuantificación del perfil transcriptómico

Los valores Ct crudos se corrigieron con la eficiencia (Fórmula 3) y cualquier señal de fondo de ADN genómico se eliminó utilizando el software GenEx (GenEx User Guide, 2012). El nivel de transcripción de cada gen de interés (GOI) se calculó como cantidades relativas (Fórmula 4) usando el gen *pep* (Tabla 1) como control interno.

$$\text{Fórmula 3: } Cp_E = Cp_U \frac{\log(1 + E)}{\log 2}$$

Donde:

$Cp_E$ : representa el Ct corregido con la eficiencia de amplificación.

$Cp_U$ : es el Ct no corregido.

E: es la eficiencia de amplificación

$$\text{Fórmula 4: } RQ = 2^{-(Cp_E_{GOI} - Cp_E_{pep})}$$

Donde:

RQ: Cantidad Relativa (*Relative Quantity*).

$Cp_E_{GOI}$ : es el Ct corregido con la eficiencia del gen de interés.

$Cp_E_{pep}$ : es el Ct corregido con la eficiencia del control interno (el gen *pep*).

## 3.- RESULTADOS Y DISCUSIÓN

---

### 3.1.- DOSIS GÉNICA

Estudios bioinformáticos previos realizados por (Castanera *et al.*, n.d.) identificaron un total de 37 helitrones de las familias HELPO1 y HELPO2 en la cepa PC15 (Figura 3) que representan el 0,35% del tamaño total del genoma. De ellos, 19 eran elementos intactos (11 elementos autónomos completos y 8 elementos no autónomos intactos), y el resto (18 elementos) eran copias truncadas. En el genoma de PC9, se encontraron 10 helitrones (de los que sólo cinco se pudieron mapear usando como referencia los *scaffolds* de PC15) (Figura 3). Los helitrones de PC9 representan el 0,05% de su genoma.

Para comprobar si el número de helitrones de PC9 y PC15 detectados por métodos informáticos se mantenía o cambiaba se realizaron experimentos de PCR en tiempo real,

utilizándose cebadores específicos para amplificar los cuatro tipos de elementos intactos en muestras de ADN genómico procedentes de estas dos cepas cultivadas en medio líquido. Los cebadores utilizados fueron: HELPO1.1\_dg, HELPO1.2\_dg, helpo1.3\_dg y HELPO2\_dg (Tabla 1. Figura 8). Los resultados obtenidos se expresaron en cantidades relativas en relación con los genes *sar1* y *pep*, ya que estos son genes de única copia en cada monocariote. De esta manera, los resultados obtenidos en la amplificación de los helitrones podían relacionarse con los obtenidos en la amplificación de los genes *sar1* y *pep*. Pudiendo obtenerse de esta forma una aproximación del número de copias existentes (RQ) de estos cuatro helitrones en cada genoma. Al utilizarse muestras de ADN genómico y no de ADN complementario, la cantidad de fluorescencia cuantificada en la qPCR para cada amplicón depende únicamente del número de copias de cada helitron presente en cada genoma y no del número de copias de ARN mensajero presente en la muestra.

Los resultados obtenidos mostraron una buena correlación ( $R^2=0,94$ ) entre el número de helitrones pertenecientes a la familia HELPO1 detectados por métodos informáticos y aquellos derivados de PCR en tiempo real (Tabla 2. Figura 9). Resulta llamativo, sin embargo, que el número de copias obtenido para los helitrones de la familia HELPO2 mediante qPCR (13-14) es más del doble de las 6 copias obtenidas en BLASTN. Esta elevada diferencia en el número de copias cuantificadas con respecto a las que aparecen como resultado de experimentos de PCR a tiempo real sugeriría que estos elementos podrían haberse amplificado en el genoma de PC15. Teniendo en cuenta que el posible número de helitrones de la familia HELPO2 detectado por métodos informáticos representa el número posible de helitrones hallados en el genoma de PC15 en el momento de su secuenciación (año 2007), la diferencia que se observa con el número de helitrones detectados en PC15 por métodos experimentales podría deberse al aumento que se produce en los genomas como respuesta a la situación de estrés que ha sufrido el genoma de PC15 en los múltiples repiques a los que ha sido sometido desde el año 2007. Incremento en el número de copias de TE como respuesta a estrés ambiental se ha descrito en (Chadha and Sharma, 2014).

En la cepa PC9 el número de copias encontradas mediante búsquedas BLASTN difieren bastante de los resultados de los experimentos de qPCR (Tabla 2. Figura 9). Así los datos obtenidos para los elementos de las subfamilias HELPO1.1 y HELPO1.2 mediante qPCR (2,75 RQs y 2,79 RQs respectivamente) contrastan con los obtenidos con BLASTN (1 y 0 copias respectivamente). Una posible explicación a la discordancia observada en los datos es que la secuencia del genoma de PC9 no está correctamente ensamblada. Hay una multitud de *scaffolds* en los que se observan regiones del genoma de gran complejidad sin ensamblar, lo



que ha dificultado en gran medida la detección bioinformática de los helitrones. Esta situación puede dar lugar a que helitrones no identificados informáticamente sí sean amplificados en el experimento de qPCR. Por lo que el número de copias obtenido mediante qPCR sería mayor que las identificadas informáticamente. Por ejemplo, en el caso de los elementos de la subfamilia HELPO1.2 BLASTN identifica 0 copias, mientras que por qPCR sí se detecta amplificación (según los cálculos de qPCR podría haber entre 2 y 3 copias). Otro motivo por el cual se produciría este desequilibrio en el número de copias en estos genes es que el genoma de PC15 está ensamblado en 11 *scaffolds* que corresponden con los 11 cromosomas (Figura 3) y sobre él se han podido identificar los helitrones presentes en esta cepa. Por el contrario, el genoma secuenciado de PC9 está ensamblado en 572 *scaffolds*, la mayoría de ellos de tamaño muy pequeño lo que hace que la estimación de la cantidad de helitrones en un genoma por medios informáticos sea más precisa en un genoma bien ensamblado como PC15. Búsquedas informáticas realizadas sobre la secuencia genómica de PC9 muestran que los helitrones aparecían fragmentados entre dos *scaffold* diferentes haciendo muy difícil la identificación como helitrones. De esto se deriva que un genoma pobremente ensamblado afecta la búsqueda de helitrones conduciendo a una subestimación de su contenido (Castanera *et al.*, n.d.). Esta podría ser también la razón por la que sólo cinco de los diez helitrones encontrados en PC9 se han podido mapear con los correspondientes *scaffolds* de PC15, mapeo realizado en base a la homología de las regiones flanqueantes de los helitrones. En el caso de PC9 algunos helitrones ocupaban un *scaffold* completo, desconociéndose así las secuencias de las regiones flanqueantes, no pudiéndose localizarlos sobre el genoma de PC15. Con respecto a la familia HELPO2, el dato obtenido de la qPCR (0,27 RQs) es menor que el obtenido mediante BLASTN (1 copia). Pero como se mencionó anteriormente, al detectarse amplificación de dicha secuencia en la qPCR, significa que debería haber al menos 1 copia dentro del genoma de PC9. Esto coincide con el dato de 1 copia de HELPO2 identificada con BLASTN.

*Tabla 2: Número de copias de cada helitron amplificadas por los cebadores en las cepas PC15 y PC9 de P. ostreatus en función de los resultados obtenidos mediante BLASTN y experimentos de qPCR.*

Helitron	PC15		PC9	
	qPCR	BLASTN	qPCR	BLASTN
HELPO1.1	1-2	2	2-3	1
helpo1.3	5-6	6	3-4	5
HELPO1.2	2-3	2	2-3	0
HELPO2	13-14	6	1	1

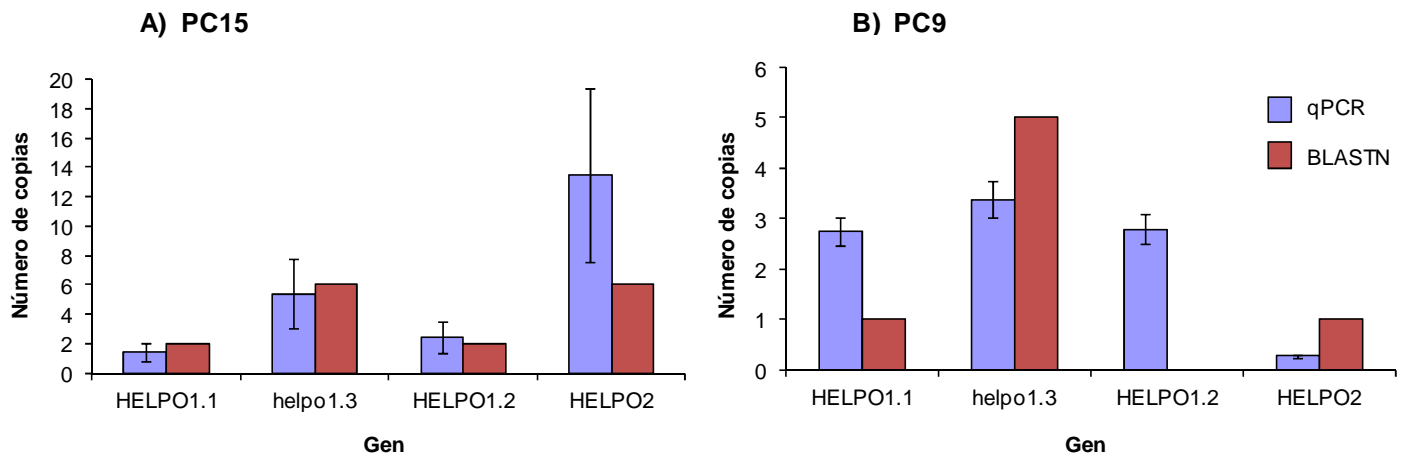


Figura 9: Representación gráfica del número de copias de helitrones amplificadas por los cebadores en las cepas PC15 (A) y PC9 (B). Las copias resultantes de experimentos de qPCR se muestran en azul y las copias encontradas mediante búsquedas BLASTN se representan en rojo. Las barras de error indican la desviación estándar basada en la variabilidad de los genes de copia única *sar1* y *pep*.

Como se puede observar en la figura 9, las barras de error son mayores en PC15, esto es debido a que la variabilidad en los Cts entre los genes de copia única *sar1* y *pep* fueron mayores en PC15 que en PC9.

El número total de helitrones encontrados en cada cepa y el porcentaje que representan de su genoma total es pequeño, si lo comparamos con otros organismos (Han *et al.*, 2013), esto es debido principalmente a dos motivos. El primero es el tamaño del genoma de *P. ostreatus*, de unas 34 Mb, un tamaño pequeño si se compara por ejemplo con el genoma de plantas (2,3 Gb el maíz) (Schnable *et al.*, 2009), insectos (432 Mb Gusano de seda) (Xia *et al.*, 2008) o mamíferos (3,2 Gb humanos) (Venter *et al.*, 2001). Y el segundo es que la abundancia de helitrones en genomas eucariotas es muy variable, incluso entre especies cercanas (Zuccolo *et al.*, 2007; Kapitonov and Jurka, 2007; Yang and Bennetzen, 2009a). Podría también achacarse a que la diferencia entre el número de helitrones detectados por cada uno de los métodos utilizados se debe a que los cebadores utilizados para su amplificación no eran específicos. Pero la secuencia de los amplicones derivados de dichos cebadores se correspondía totalmente con las secuencias de los helitrones que se utilizaron para su diseño (Figura 10).

## HELPO1.2

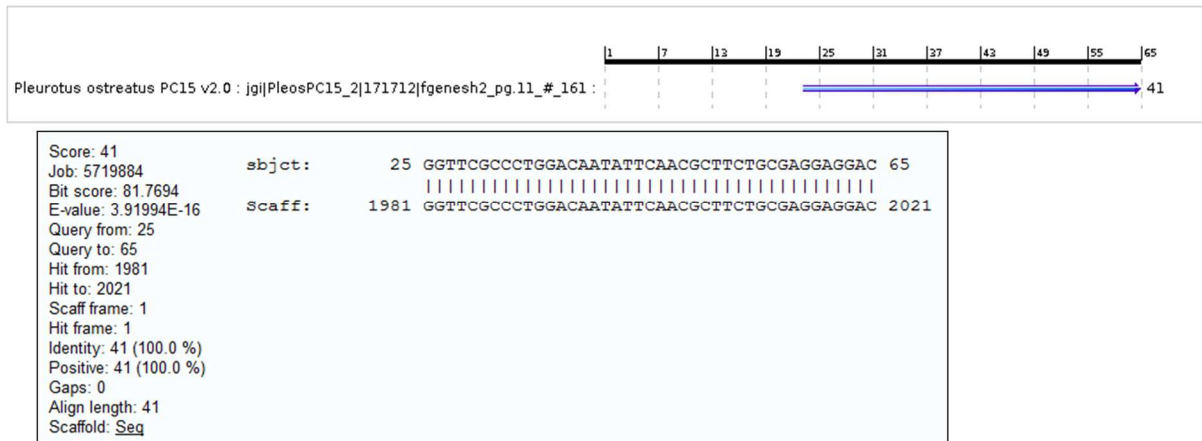


Figura 10: Alineamiento obtenido mediante el programa BLASTN para el elemento HELPO1.2.

Para comprobar si el análisis mediante qPCR podía ser un método válido para detectar el número de copias de helitrones en un genoma se realizaron las siguientes comprobaciones:

i) Experimentos de qPCR con diluciones seriadas de ADN genómico de la cepa PC9: Se utilizaron 6 diluciones con las siguientes concentraciones de ADN: 2 ng/μl, 1 ng/μl, 0,5 ng/μl, 0,25 ng/μl, 0,125 ng/μl y 0,06 ng/μl. En todas las concentraciones se amplificó el elemento de copia única en PC9: HELPO2. Los datos obtenidos se representaron gráficamente y se realizó el ajuste de la recta (Figura 11).

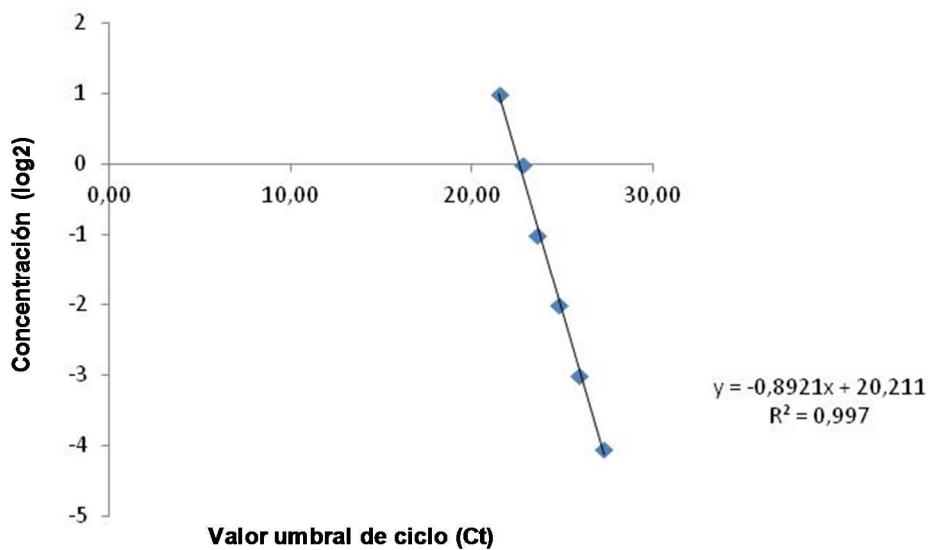


Figura 11: Representación gráfica y ajuste de la curva de los datos obtenidos en experimentos de qPCR (Cts) en seis diluciones secuenciadas de ADN genómico de la cepa PC9 amplificando el helitron HELPO2.

Los resultados obtenidos demostraron que cuando la concentración entre dos muestras se duplica, la diferencia que debe existir entre sus Cts resultantes es de 1 ciclo. Así, si la dilución de concentración 2 ng/μl presenta un Ct de 21 ciclos la dilución de 1 ng/μl deberá presentar un Ct de 22, un ciclo más porque se parte de la mitad de ADN y en cada ciclo de amplificación se duplica el producto específico (si la eficiencia es del 100%) y así sucesivamente. Dando lugar a una línea recta. En este caso la línea resultante presenta una  $R^2=0,997$ , lo que significa que el ajuste es bueno. Esto se realizó para comprobar que la cuantificación de copias es robusta en un amplio rango de concentraciones. Finalmente se calculó la desviación estándar de las diferencias entre el Ct obtenido con una concentración y el Ct que se obtendría para una eficiencia del 100% (un ciclo de diferencia en cada concentración). Obteniendo una desviación de 0,25 ciclos. Que es un error sistemático que se introduce en la cuantificación inherente a la metodología.

ii) Experimentos de qPCR en las que se utilizó la misma concentración de ADN genómico de PC9 en todas las reacciones amplificándose siete genes de copia única en el genoma de PC9: *lacc3*, *lacc11*, *mnp3*, *actin2*, *sar1*, *pep* y *cyph* (Castanera *et al.*, 2013) (Tabla 1). Partiendo de la base de que todos los genes deberían dar el mismo resultado (ya que los siete son de copia única y la concentración en ADN genómico es la misma), el objetivo es analizar la variabilidad en los resultados. Posteriormente se registraron los Cts de cada gen y se calculó la desviación estándar entre ellos, obteniendo como resultado una desviación estándar de 0,44 ciclos.

Sumando ambas desviaciones obtenemos 0,69 ciclos de error sistemático en el peor escenario posible (máxima variabilidad). Usando este error sistemático se puede calcular una tabla con los márgenes de error en el número de copias (máximo y mínimo esperado) (Tabla 3).

Tabla 3: Número de copias esperado y sus márgenes de error en Cts.

Nº copias esperado	Márgenes de error (Cts)	
1	0,6	1,6
2	1,2	3,2
3	1,9	4,9
4	2,5	6,5
5	3,1	8,1
6	3,8	9,8
8	5,0	12,9

Estos resultados validan la idea de poder utilizar este método experimental de análisis mediante RT-qPCR para realizar un *screening* en otras cepas monocariontes de *P. ostreatus*, de las que no se dispone de secuencia genómica. Para encontrar cepas con alto y bajo contenido en helitrones y poder estudiar el efecto que estos tienen sobre el genoma en las cepas.

### 3.2.- PERFIL TRANSCRIPTÓMICO

Se utilizaron datos procedentes de experimentos de RT-qPCR realizados con ADNc procedentes de las cepas PC9, PC15 y N001 cultivadas en medio líquido SMY para analizar la actividad transcripcional de los genes que controlan la expresión de la proteína RepHel así como de los genes capturados por lo helitrones. Se estudiaron los perfiles de transcripción de 3 helicasas RepHel: HELPO1.1, HELPO1.2, y HELPO2 y de 6 genes capturados: *capA*, *capA2*, *capB*, *capC*, *capD*, y *capF*, utilizando para ello los cebadores específicos diseñados para el estudio de su perfil de expresión: HELPO1.1\_rep, HELPO1.2\_rep, HELPO2\_dg, *capA*, *capA2*, *capB*, *capC*, *capD*, y *capF* (Tabla 1. Figura 8).

Los datos obtenidos se utilizaron para complementar el análisis transcripcional en cultivos sólidos llevado a cabo mediante la técnica de secuenciación ARNseq (datos de la cepa N001 cultivada en medio sólido SMY procedentes de experimentos previos llevados a cabo en nuestro grupo (Castanera *et al.*, n.d.).

Los resultados obtenidos del análisis de los datos procedentes de experimentos de RT-qPCR permiten conocer la expresión de las helicasas RepHel y de los genes capturados de manera independiente. Conviene señalar que mientras los perfiles relativos de expresión son comparables entre cepas, las cantidades obtenidas no lo son, porque se desconoce el número exacto de copias del gen *pep* en cada muestra. Así para las helicasas RepHel, se observaron perfiles de expresión relativa similares en las tres cepas aunque los rangos de los niveles de transcripción eran diferentes (Fig. 12) . La helicasa RepHel de HELPO1.2 fue generalmente la más expresada en las tres cepas, aunque en cantidades diferentes (0,31; 20,9 y 6,9 RQs en PC9, PC15 y N001 respectivamente). La RepHel de HELPO1.1 mostró mucha menor expresión que HELPO1.2 (0,25; 1,8 y 0,2 RQs en PC9, PC15 y N001 respectivamente) y HELPO2 no mostró expresión en PC9 y N001 (0; 2,6 y 0 RQs en PC9, PC15 y N001 respectivamente) (Figura 12).

Respecto a los genes capturados que en su secuencia llevaban dominios de virus y retrotransposones (*capD* y *capF*), no se observó transcripción en ninguna cepa (aunque sí estaban presentes en el genoma de PC15 y N001). Los genes de función desconocida *capA*, *capA2*, *capB* y *capC* mostraron un perfil de expresión específico de la cepa. En la cepa PC9 el gen *capA* era el más expresado (52,3 RQs); *capA2*, *capD* y *capF* no estaban presentes en PC9. En PC15, los valores de expresión más altos correspondieron a *capC* y *capA2* (30,4 y 22,1 RQs). En N001, el gen más expresado fue *capA* seguido de *capC* y *capA2* (8,51; 2,4 y 2,2 RQs) (Figura 12).

El análisis conjunto y comparativo de las dos cepas parentales monocarióticas PC9 y PC15 con respecto a la dicariótica N001 derivada de ellas, en lo que se refiere a expresión génica (de RepHel y genes capturados) muestra que la expresión de los genes en la cepa dicariótica podría explicarse de acuerdo con un modelo aditivo de control de la expresión resultante de la sumatoria de la acción individual de cada uno de los genes de los parentales.

En relación con el nivel de expresión se observó que las helicasas RepHel de la familia HELPO1 tenían mayor expresión que las de la familia HELPO2, aun presentando HELPO2 mayor número de copias.

Cabe señalar que los cebadores se diseñaron para amplificar más de una copia de los genes estudiados a excepción de *capA2*, *capD* y *capF* (Tabla 1) de los que existía sólo una copia dentro del genoma de PC15. Así, los niveles de transcripción obtenidos para los genes (HELPO1.1, HELPO1.2, HELPO2, *capA*, *capB* y *capC*) fueron el resultado de la suma de la expresión de todas las copias de cada helicasa RepHel y de cada gen capturado presentes en cada subfamilia de helitrones. Por lo tanto no fue posible saber si una copia se expresaba más que otra, o si solo se expresan ciertas copias y otras no.

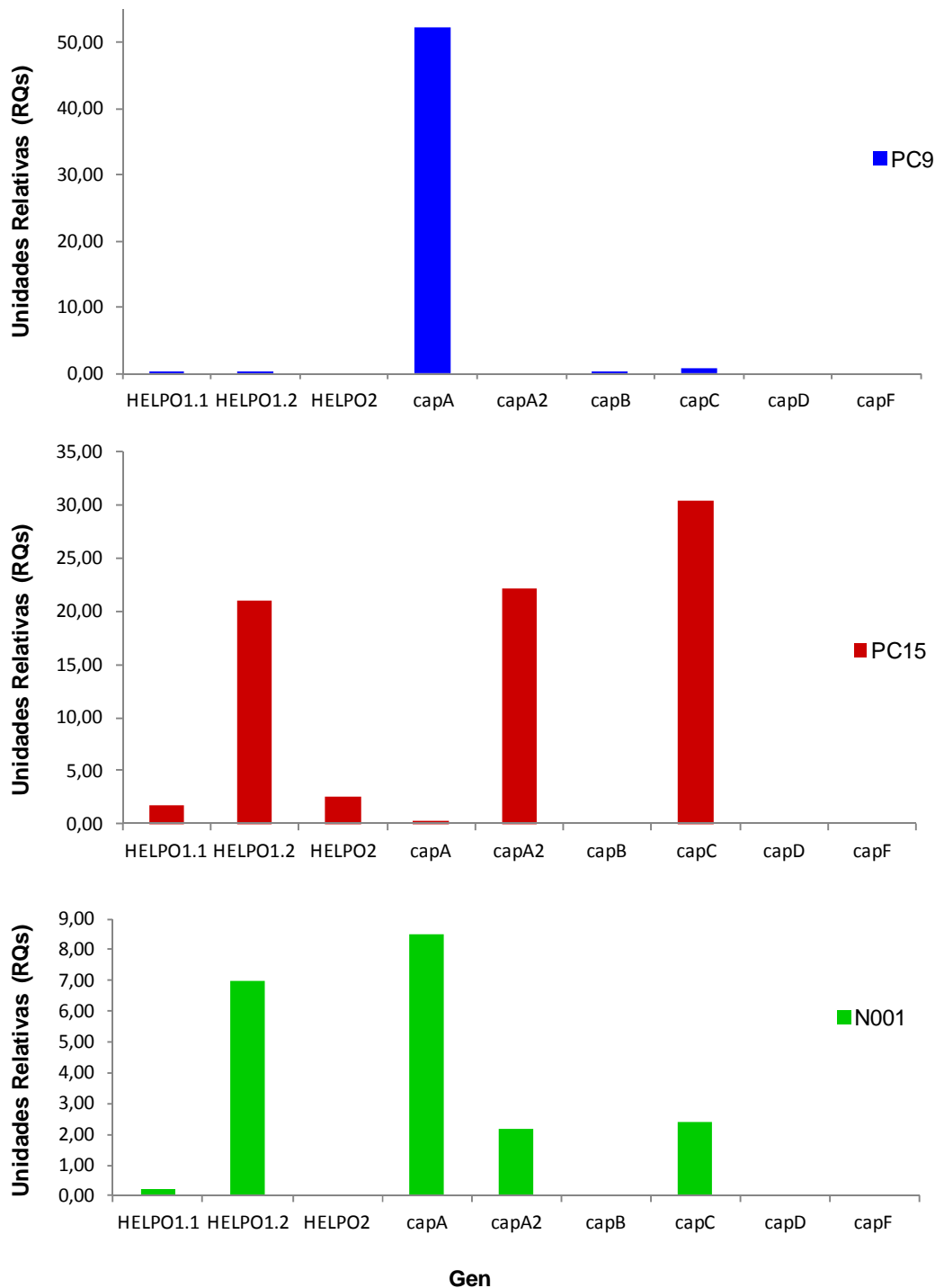


Figura 12: Perfiles de expresión de las helicasas RepHel y de los genes capturados en las cepas PC9, PC15 y N001 en función de los datos obtenidos de experimentos de RT-qPCR. El eje vertical indica los valores arbitrarios (RQs) referidos a la expresión del gen de referencia pep.

Estos datos obtenidos con PCR a tiempo real confirman los obtenidos con ARNseq (Castanera *et al.*, n.d.) que mostraron que los helitrones de la familia HELPO1 presentaban los niveles más altos de transcripción (basado en los valores RPKM del helitron completo,

incluyendo la helicasa RepHel y los genes capturados) en comparación con los elementos de la familia HELPO2. La subfamilia HELPO1.1 mostró niveles muy altos de expresión (hay que resaltar que el gen *capA* representaba parte sustancial de la expresión total), hasta 38,64 RPKM, en comparación con HELPO1.3 (máximo de 7,69 RPKM) y los miembros de la subfamilia HELPO1.2 (máximo de 2,41 RPKM).

En este estudio, se analizó también la similitud entre las proteínas RepHel de HELPO1.1 y HELPO1.2 que fue del 68,5%. Este valor de similitud entre ambas proteínas indica el grado de conservación que presentan, necesario para llevar a cabo la función específica del helitrón.

En cuanto a los genes capturados que aparecen en *Pleurotus*, pocas coincidencias se han observado en las bases de datos utilizando la secuencia de *capB*, y la ausencia total de coincidencias usando los otros genes *cap*. Esto podría indicar que estas secuencias son nuevas estructuras creadas mediante la recombinación con ADN de origen diverso, o el resultado de la captura de un gen completo en un huésped que no sea un hongo, cuya secuencia genómica aún no está disponible. Los genes de función desconocida *capA*, *capA2*, *capB* y *capC* se observó que carecían de dominios conservados. Una búsqueda realizada sobre la base de datos MycoCosm (Grigoriev *et al.*, 2012) reveló que los genes *capA*, *capA2*, y *capC* constituyen genes nuevos de hongos específicos de *P. ostreatus*. Este no es el caso de *capB* que se encontró que estaba presente en muy pocos miembros de la división Basidiomycetos. Las similitudes entre los genes *capA*, *capA2* y *capB* (alrededor del 45%) sugieren probablemente que se podría derivar de un gen ancestral que se duplicó y en los que ocurrió una divergencia funcional, llevando al mantenimiento o supresión de su actividad en función de cualquier posible ventaja que pudiera proporcionar al genoma huésped (Castanera *et al.*, n.d.). Los genes *capD* y *capF* no se expresan. Por tanto podrían representar pseudogenes o bien estar silenciados por algún mecanismo, dado que los elementos de transposición pueden ser activados o inhibidos bajo ciertas condiciones de estrés como infección por patógenos, daños físicos o estrés abiótico.

Para terminar, es importante indicar que no se han podido realizar comparaciones de expresión de helitrones y de sus genes capturados dado que hasta el momento no se han realizado estudios similares al descrito en este trabajo en otras especies de hongos.



## 4.- CONCLUSIONES

---

- Se ha puesto a punto un método basado en qPCR para determinar el número de helitrones presentes en otras cepas monocarióticas de *P. ostreatus* de las que no se dispone de secuencia genómica.
- Se ha comprobado que las helicasas RepHel de las dos familias de helitrones presentan actividad transcripcional en la cepa PC15, lo que sugiere que estos elementos podrían estar activos.
- Se ha comprobado que algunos de los genes capturados se expresan en las tres cepas, aunque con perfiles de transcripción diferentes.
- La expresión de los genes capturados sugiere que podrían conferir alguna ventaja al genoma en el que se encuentran.
- Los genes capturados con dominios virales no se expresan en ninguna cepa, lo que indica que podrían ser fragmentos de virus endogenizados y no activos.
- Los genes capturados y con alta expresión *capA* y *capA2* no tienen funciones conocidas por el momento.
- Se desconoce si todas las copias de los genes presentes en los helitrones se expresan por igual.

## 5.- BIBLIOGRAFÍA

---

Boulé, J.-B., and Zakian, V.A. (2006) Roles of Pif1-like helicases in the maintenance of genomic stability. *Nucleic Acids Res* **34**: 4147–53.

Brunner, S., Pea, G., and Rafalski, A. (2005) Origins, genetic organization and transcription of a family of non-autonomous helitron elements in maize. *Plant J* **43**: 799–810.

Bustin, S. (2000) Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *J Mol Endocrinol* **25**: 169–193.

Bustin, S. a, Benes, V., Garson, J. a, Hellemans, J., Huggett, J., Kubista, M., *et al.* (2009) The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clin Chem* **55**: 611–22.

Castanera, R., Omarini, A., Santoyo, F., Pérez, G., Pisabarro, A.G., and Ramírez, L. (2013) Non-Additive Transcriptional Profiles Underlie Dikaryotic Superiority in *Pleurotus ostreatus* Laccase Activity. *PLoS One* **8**: e73282.

Castanera, R., Pérez, G., López, L., Sancho, R., Santoyo, F., Alfaro, M., *et al.* Highly expressed captured genes and cross-kingdom domains present in helitrons create novel diversity in *Pleurotus ostreatus* and other fungi. *Plos Genet* (En Revisión).

Chadha, S., and Sharma, M. (2014) Transposable elements as stress adaptive capacitors induce genomic instability in fungal pathogen *Magnaporthe oryzae*. *PLoS One* **9**: e94415.

Cultrone, A., Domínguez, Y.R., Drevet, C., Scazzocchio, C., and Fernández-Martín, R. (2007) The tightly regulated promoter of the *xanA* gene of *Aspergillus nidulans* is included in a helitron. *Mol Microbiol* **63**: 1577–1587.

Edwards, K.J. (2004) Performing Real-Time PCR. In *Real-Time PCR: An Essential Guide*. K.J.Edwards et al. (ed.). Horizonbioscience. Wymondham, pp. 71–84.

Fernández-Fueyo, E., Ruiz-Dueñas, F.J., Miki, Y., Martínez, M.J., Hammel, K.E., and Martínez, A.T. (2012) Lignin-degrading peroxidases from genome of selective ligninolytic fungus *Ceriporiopsis subvermispora*. *J Biol Chem* **287**: 16903–16.

Feschotte, C., Keswani, U., Ranganathan, N., Guibotsy, M.L., and Levine, D. (2009) Exploring repetitive DNA landscapes using REPCCLASS, a tool that automates the classification of transposable elements in eukaryotic genomes. *Genome Biol Evol* **1**: 205–20.

Feschotte, C., and Pritham, E.J. (2009) A cornucopia of Helitrons shapes the maize genome. *Proc Natl Acad Sci U S A* **106**: 19747–8.

GenEx User Guide (2012) 1–20 <http://www.multid.se>.

Giglio, S. (2003) Demonstration of preferential binding of SYBR Green I to specific DNA fragments in real-time multiplex PCR. *Nucleic Acids Res* **31**: 136e–136.

Grigoriev, I. V, Nordberg, H., Shabalov, I., Aerts, A., Cantor, M., Goodstein, D., *et al.* (2012) The genome portal of the Department of Energy Joint Genome Institute. *Nucleic Acids Res* **40**: D26–32.

- Han, M.-J., Shen, Y.-H., Xu, M.-S., Liang, H.-Y., Zhang, H.-H., and Zhang, Z. (2013) Identification and evolution of the silkworm helitrons and their contribution to transcripts. *DNA Res* **20**: 471–84.
- Hernández, A., Martín Vasallo, P., Torres Ramirez, A., and Salido, E. (1994) RNA analysis: The study of genic expression. *NEFROLOGÍA XIV*: 145–162.
- Higuchi, R., Fockler, C., Dollinger, G., and Watson, R. (1993) Kinetic PCR analysis: real-time monitoring of DNA amplification reactions. *Biotechnology (N Y)* **11**: 1026–30.
- Jameson, N., Georgelis, N., Fouladbash, E., Martens, S., Hannah, L.C., and Lal, S. (2008) Helitron mediated amplification of cytochrome P450 monooxygenase gene in maize. *Plant Mol Biol* **67**: 295–304.
- Kapitonov, V. V., and Jurka, J. (2001) Rolling-circle transposons in eukaryotes. *Proc Natl Acad Sci U S A* **98**: 8714–9.
- Kapitonov, V. V., and Jurka, J. (2007) Helitrons on a roll: eukaryotic rolling-circle transposons. *Trends Genet* **23**: 521–9.
- Kempken, F. (2008) The *Tolypocladium inflatum* CPA element encodes a RecQ helicase-like gene. *J Basic Microbiol* **48**: 496–499.
- Koressaar, T., and Remm, M. (2007) Enhancements and modifications of primer design program Primer3. *Bioinformatics* **23**: 1289–91.
- Labbé, J., Murat, C., Morin, E., Tuskan, G.A., Tacon, F. Le, and Martin, F. (2012) Characterization of transposable elements in the ectomycorrhizal fungus *Laccaria bicolor*. *PLoS One* **7**: e40197.
- Lai, J., Li, Y., Messing, J., and Dooner, H.K. (2005) Gene movement by Helitron transposons contributes to the haplotype variability of maize. *Proc Natl Acad Sci U S A* **102**: 9068–73.
- Martin, F., Kohler, A., Murat, C., Balestrini, R., Coutinho, P.M., Jaillon, O., *et al.* (2010) Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature* **464**: 1033–8.
- McPherson, M.J., Hames, B.D., and Taylor, G.R. (2008) *PCR: a Practical Approach*. Oxford University Press, .
- Messing, J., Bharti, A.K., Karlowski, W.M., Gundlach, H., Kim, H.R., Yu, Y., *et al.* (2004) Sequence composition and genome organization of maize. *Proc Natl Acad Sci U S A* **101**: 14349–54.
- Morgante, M., Brunner, S., Pea, G., Fengler, K., Zuccolo, A., and Antoni, R. (2005) Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize. *Nat Genet* **37**: 997–1002.
- Pfaffl, M.W. (2004) Quantification strategies in real-time PCR. In *A-Z of quantitative PCR*. S.A. Bustin (ed.). International University Line. La Jolla, pp. 87–112.
- Poulter, R.T.M., Goodwin, T.J.D., and Butler, M.I. (2003) Vertebrate helitrons and other novel Helitrons. *Gene* **313**: 201–212.
- Pritham, E.J., and Feschotte, C. (2007) Massive amplification of rolling-circle transposons in the lineage of the bat *Myotis lucifugus*. *Proc Natl Acad Sci U S A* **104**: 1895–900.

- Ramakers, C., Ruijter, J.M., Deprez, R.H.L., and Moorman, A.F.. (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. *Neurosci Lett* **339**: 62–66.
- Ririe, K.M., Rasmussen, R.P., and Wittwer, C.T. (1997) Product Differentiation by Analysis of DNA Melting Curves during the Polymerase Chain Reaction. *Anal Biochem* **245**: 154–160.
- Schmittgen, T.D., Zakrajsek, B.A., Mills, A.G., Gorn, V., Singer, M.J., and Reed, M.W. (2000) Quantitative Reverse Transcription–Polymerase Chain Reaction to Study mRNA Decay: Comparison of Endpoint and Real-Time Methods. *Anal Biochem* **285**: 194–204.
- Schnable, P.S., Ware, D., Fulton, R.S., Stein, J.C., Wei, F., Pasternak, S., *et al.* (2009) The B73 maize genome: complexity, diversity, and dynamics. *Science* **326**: 1112–5.
- Smith, C.J., and Osborn, a M. (2009) Advantages and limitations of quantitative PCR (Q-PCR)-based approaches in microbial ecology. *FEMS Microbiol Ecol* **67**: 6–20.
- Thomas, J., Schaack, S., and Pritham, E.J. (2010) Pervasive horizontal transfer of rolling-circle transposons among animals. *Genome Biol Evol* **2**: 656–64.
- Toleman, M.A., Bennett, P.M., and Walsh, T.R. (2006) Common regions e.g. orf513 and antibiotic resistance: IS91-like elements evolving complex class 1 integrons. *J Antimicrob Chemother* **58**: 1–6.
- Untergasser, A., Cutcutache, I., Koressaar, T., Ye, J., Faircloth, B.C., Remm, M., and Rozen, S.G. (2012) Primer3--new capabilities and interfaces. *Nucleic Acids Res* **40**: e115.
- Venter, J.C., Adams, M.D., Myers, E.W., Li, P.W., Mural, R.J., Sutton, G.G., *et al.* (2001) The sequence of the human genome. *Science* **291**: 1304–51.
- Vogelstein, B., and Kinzler, K.W. (1999) Digital PCR. *Proc Natl Acad Sci U S A* **96**: 9236–41.
- Walker, N.J. (2002) Tech.Sight. A technique whose time has come. *Science* **296**: 557–9.
- Wicker, T., Sabot, F., Hua-Van, A., and Jeffrey L. Bennetzen, *et al* (2007) A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* **8**: 973–982.
- Xia, Q., Wang, J., Zhou, Z., Li, R., Fan, W., Cheng, D., *et al.* (2008) The genome of a lepidopteran model insect, the silkworm *Bombyx mori*. *Insect Biochem Mol Biol* **38**: 1036–1045.
- Yang, L., and Bennetzen, J.L. (2009a) Structure-based discovery and description of plant and animal Helitrons. *Proc Natl Acad Sci U S A* **106**: 12832–7.
- Yang, L., and Bennetzen, J.L. (2009b) Distribution, diversity, evolution, and survival of Helitrons in the maize genome. *Proc Natl Acad Sci U S A* **106**: 19922–7.
- Zhou, Q., Froschauer, A., Schultheis, C., Schmidt, C., Bienert, G.P., Wenning, M., *et al.* (2006) Helitron Transposons on the Sex Chromosomes of the Platyfish *Xiphophorus maculatus* and Their Evolution in Animal Genomes. *Zebrafish* **3**: 39–52.
- Zuccolo, A., Sebastian, A., Talag, J., Yu, Y., Kim, H., Collura, K., *et al.* (2007) Transposable element distribution, abundance and role in genome size variation in the genus *Oryza*. *BMC Evol Biol* **7**: 152.