



UNIVERSITAT  
POLITÈCNICA  
DE VALÈNCIA



Instituto de Agroquímica  
y Tecnología de Alimentos



CONSEJO SUPERIOR DE INVESTIGACIONES CIENTÍFICAS

**UNIVERSITAT POLITÈCNICA DE VALÈNCIA**

**Escuela Técnica Superior de Ingeniería  
Agronómica y del Medio Natural**

“Estudio del mecanismo de acción de  
proteínas antifúngicas en *Penicillium digitatum*  
mediante análisis transcriptómico”

**GRADO EN BIOTECNOLOGÍA**

**2017-2018**

**AUTORA:**

Dña. Begoña Bolós Sierra

**TUTORA ACADÉMICA:**

Dra. Lynne P. Yenush

**DIRECTORES EXPERIMENTALES:**

Dr. Jose F. Marcos López

Dra. Paloma M. Manzanares Mir



Escola Tècnica Superior  
d'Enginyeria Agronòmica i del Medi Natural

Valencia, 9 de Julio de 2018



**TÍTULO:** Estudio del mecanismo de acción de proteínas antifúngicas en *Penicillium digitatum* mediante análisis transcriptómico

**RESUMEN:** Las infecciones causadas por hongos representan un importante riesgo para la salud humana, la producción de alimentos y la seguridad alimentaria. Sin embargo, se han introducido muy pocos compuestos antifúngicos en el mercado de fungicidas en los últimos años. Por ello, es necesario encontrar nuevas moléculas antifúngicas y a ser posible con modos de acción diferentes a los de los antifúngicos ya usados. Algunos hongos filamentosos producen y secretan grandes cantidades de proteínas AFPs (*AntiFungal Proteins*) con actividad antifúngica. Estas AFPs son pequeñas, catiónicas y ricas en cisteínas que forman puentes disulfuro, y tienen un gran interés biotecnológico por su elevada actividad antifúngica, especificidad y estabilidad. Sin embargo, su mecanismo de acción antifúngico y su función biológica en el hongo productor apenas se conocen.

Con el objetivo de caracterizar el modo de acción de la proteína antifúngica AfpB sobre el patógeno de frutos cítricos *Penicillium digitatum* del cual procede, se ha realizado un análisis transcriptómico basado en la secuenciación masiva o RNA-seq. Para ello, se han realizado distintos tratamientos de AfpB sobre la cepa de referencia *P. digitatum* CECT 20794 (ausencia de AfpB, tratamiento durante 1 ó 3 horas, y crecimiento en presencia de AfpB) y se ha incluido en el análisis la cepa modificada genéticamente PDSG420 que es sobreproductora de la proteína antifúngica. Se observaron los efectos de la proteína sobre el crecimiento del hongo en cada caso. Se determinaron los cambios en el transcriptoma, obteniéndose un elevado número de genes expresados de forma diferencial (DEG, *Differentially Expressed Gene*) en cada tratamiento (hasta 4.143 de los 9.133 anotados en el genoma) y con similitudes y diferencias en los tratamientos. Se realizó un análisis del enriquecimiento funcional en ontologías génicas (GO) de los DEGs, para lo que ha sido necesaria la mejora de la anotación funcional disponible del genoma de *P. digitatum*.

Como resultado de este trabajo, se han obtenido genes y procesos potencialmente implicados en la actividad de AfpB y que podrán ser validados experimentalmente en el futuro. Como ejemplo, el gen PDIG\_81760 que es el más inducido por la exposición continua de la proteína y que codifica una pequeña proteína secretable, rica en cisteínas, aniónica y con una estructura repetitiva muy peculiar, la cual podría antagonizar la acción de las AFPs catiónicas.

**PALABRAS CLAVE:** Proteína antifúngica, hongo filamentosos, expresión diferencial, *Penicillium digitatum*, RNA-seq

**AUTORA:** Dña. Begoña Bolós Sierra

**DIRECTORES:** Dr. Jose F. Marcos López, Dra. Paloma M. Manzanares Mir

**TUTORA:** Dra. Lynne P. Yenush

**LUGAR Y FECHA:** Valencia, Julio 2018

**TITLE:** Transcriptomic analysis to study the mode of action of antifungal proteins in *Penicillium digitatum*.

**SUMMARY:** Infections caused by fungi represent an important risk to human health, food production and food security. However, very few antifungal compounds have been introduced into the global fungicide market in recent years. There is thus an urgent need for the development of novel antifungal agents with properties and mechanisms of action different from existing ones. Some filamentous fungi produce and secrete large amounts of AFPs (AntiFungal Proteins) with antifungal activity. These AFPs are small, cationic and rich in cysteines that form disulfide bridges, and have a great biotechnological interest due to their high antifungal activity, specificity and stability. However, its mechanism of antifungal action and its biological function in the producer fungus are largely unknown.

In this project, we aim to study the mode of action of the antifungal protein AfpB against the citrus fruit pathogen *Penicillium digitatum*, from which the protein was identified, through a transcriptomic analysis based on RNA-seq. For this, we made different AfpB treatments on the reference strain *P. digitatum* CECT 20794 (absence of AfpB, treatment for 1 or 3 hours, and growth in the presence of AfpB) and the genetically modified strain PDSG420, which overproduces the antifungal protein, was also included in the analysis. The effects of the protein on the growth of the fungus were observed in each case. Changes in the transcriptome were determined, obtaining a high number of differentially expressed genes (DEG) in each treatment (up to 4,143 of the 9,133 annotated in the genome), with similarities and differences in the treatments. We made an analysis of the functional enrichment in gene ontologies (GO) of the DEGs, for which the improvement of the available functional annotation of the *P. digitatum* genome was necessary.

As result of this work, genes and processes potentially involved in AfpB activity have been identified and will be validated experimentally in the future. As an example, the gene PDIG\_81760 that is the most induced by AfpB and codes for a small, cysteine-rich secreted and anionic protein with a remarkable repetitive structure, that could antagonize the action of cationic AFPs.

**KEY WORDS:** Antifungal protein, filamentous fungi, differential expression, *Penicillium digitatum*, RNA-seq

**AUTHOR:** Dña. Begoña Bolós Sierra

**DIRECTORS:** Dr. Jose F. Marcos López, Dra. Paloma M. Manzanares Mir

**TUTOR:** Dra. Lynne P. Yenush

**PLACE AND DATE:** Valencia, July 2018

## AGRADECIMIENTOS

En primer lugar, quería agradecer este trabajo al Lab 016 del IATA, tanto a los primeros que me recibieron; Miguel y Marcos. Los que continúan desde que llegué, tan trabajadoras como siempre; Mónica y Sandra. Los que han venido recientemente; Helena, Rafael y la que se ha convertido mi compañera de piso, Ana. Por supuesto quiero agradecer este trabajo a los jefes de laboratorio; Jose y Paloma. Gracias a todos por crear un excelente ambiente de conocimiento y trabajo.

En concreto quiero agradecer este trabajo a Jose y Paloma, por acogerme en su laboratorio y por confiarme un trabajo de final de grado tan interesante como el que hemos hecho. Con este trabajo he aprendido muchísimo gracias a vosotros y ha sido la confirmación definitiva de que el ambiente de laboratorio quiero que sea mi ambiente de trabajo. A Paloma por todas las palabras de ánimo. Y a Jose por todas sus explicaciones y correcciones, por estar siempre atento, y ser un jefe cercano que busca la excelencia y la transmite.

En especial quiero agradecer todo el esfuerzo y tiempo que Mónica ha puesto en que aprendiera a hacerlo todo sola, no solo en este trabajo, sino también para poder continuarlo. Por ser nuestra profe 24 horas y ayudarme siempre. También a Sandra por contestarme a todas mis dudas, por la propia proteína que he usado y tu ayuda.

Le agradezco este trabajo a Lynne (IBMCP UPV-CSIC), porque gracias a ella pude hacer mi TFG en este laboratorio. Además, por ser mi tutora por excelencia en todo lo que hago y acogerme en su laboratorio hace un tiempo. También por los consejos y los siguientes pasos a dar para continuar con este trabajo. Un agradecimiento muy especial a Javier Forment (IBMCP UPV-CSIC), por sus directrices, consejos y facilidades para el análisis bioinformático de los datos y la utilización de los servidores de la UPV. A Nuria Andrés (IBMCP UPV-CSIC) por sus observaciones en un territorio desconocido como es para nosotros el editado de RNA. Y a Eladio Barrio y Laura Gutiérrez (Universidad de Valencia) por darnos pistas y abrir nuevas hipótesis sobre una posible duplicación del locus.

Quiero agradecer este trabajo al SCSIE de la Universidad de Valencia por el servicio en el RNA-seq, y en concreto a Amparo Martínez y Vicente Sentandreu por ayudarme con las dudas y análisis posteriores.

Quiero agradecer este trabajo al IATA por sus instalaciones y al Proyecto BIO2015-68790-C2-1-R "Nuevas proteínas antifúngicas de hongos: Producción en hongos filamentosos y caracterización de su mecanismo de acción", financiado por el MINECO y los Fondos FEDER.

Este trabajo se lo agradezco a mis padres, porque sin ellos está claro que no hubiera llegado hasta aquí. Gracias por creer y confiar en mí. Se lo dedico a Patrick, por estar siempre ayudándome en todo lo que puede y su apoyo continuo. A mis mejores amigos Richi y Salva, por hacerme ver el mundo de otra manera. Y a mi gato, que me ha soportado más que nadie.

Por supuesto este trabajo, se lo agradezco a mis biotecnólogas que siendo tan diferentes hemos estado juntas los cuatro años. Dicen que la universidad es un periodo que no se olvida y en mi caso será por vosotras. Gracias por todo.

<b>ÍNDICE GENERAL</b> .....	I
<b>ÍNDICE DE TRABAJO</b> .....	I
<b>ÍNDICE DE FIGURAS Y TABLAS</b> .....	II
<b>ABREVIATURAS</b> .....	IV

## ÍNDICE DEL TRABAJO

1. INTRODUCCIÓN .....	1
1.1. Los hongos patógenos y la necesidad de nuevas moléculas antifúngicas.....	1
1.2. <i>Penicillium digitatum</i> : el principal patógeno de frutos cítricos.....	2
1.3. Las proteínas antifúngicas de hongos. ....	3
1.4. Mecanismo de acción y función de proteínas AFP.....	4
2. OBJETIVOS .....	5
3. MATERIALES Y MÉTODOS.....	6
3.1. Cepas, medios y condiciones de cultivo. ....	6
3.2. Extracción y procesamiento de RNA total.....	6
3.3. RNA-seq	6
3.4. Mejora de la anotación funcional y enriquecimiento funcional.....	7
4. RESULTADOS.....	8
4.1. Respuesta de las cepas de <i>P. digitatum</i> a tratamientos con AfpB. ....	8
4.2. Obtención de RNA total de calidad apropiada .....	10
4.3. Determinación de genes que se expresan diferencialmente .....	11
4.4. Mejora de la anotación funcional .....	14
4.5. Enriquecimiento funcional.....	16
4.6. Selección de genes.....	17
5. DISCUSIÓN .....	20
6. CONCLUSIONES .....	26
7. BIBLIOGRAFÍA .....	27

## INDICE DE FIGURAS Y TABLAS

<b>Figura 1:</b> El hongo <i>P. digitatum</i> patógeno postcosecha de frutos cítricos.....	2
<b>Figura 1:</b> La proteína antifúngica AfpB de <i>P. digitatum</i> .....	4
<b>Tabla 1:</b> Especies de hongos usadas para la base de datos local en Blast2GO.....	8
<b>Figura 3:</b> Comportamiento de <i>P. digitatum</i> en presencia de la proteína antifúngica AfpB.....	9
<b>Figura 4:</b> Concentración, calidad e integridad del RNA obtenido del micelio de <i>P. digitatum</i> .....	10
<b>Tabla 2:</b> Calidad de las lecturas secuenciadas en RNA-seq.....	11
<b>Figura 5:</b> Variabilidad de las réplicas por tratamiento.....	12
<b>Figura 6:</b> Representación de los DEGs por tratamiento obtenidos de EdgeR (FDR<0,01).....	13
<b>Tabla 3:</b> Número de genes expresados diferencialmente por software.....	14
<b>Figura 7:</b> Esquema del procedimiento seguido para obtener la anotación propuesta en Blast2GO.....	15
<b>Figura 8:</b> Mejora de la anotación funcional de <i>P. digitatum</i> con Blast2GO.....	16
<b>Tabla 4:</b> Enriquecimiento funcional de por ontología génica (GO) de DEGs en tratamientos 05 y 06.....	17
<b>Tabla 5:</b> Enriquecimiento funcional de por ontología génica (GO) de DEGs en tratamientos 08 y 11.....	17
<b>Tabla 6:</b> Selección de genes expresados diferencialmente obtenidos con EdgeR.....	18
<b>Figura 9:</b> Gráficas de expresión diferencial de genes agrupados en los clusters seleccionados.....	19
<b>Figura 10:</b> Visualización con IGV 2.4 de las lecturas que mapean sobre el genoma de referencia en cada tratamiento.....	20
<b>Figura 11:</b> Ruta de biosíntesis de la acetoina.....	23
<b>Figura 12:</b> Gen PDIG_81760 en el visualizador IGV.....	24
<b>Figura 13:</b> Secuencia aminoacídica de la proteína codificada por PDIGG_81760.....	25





## **ABREVIATURAS**

AFP: Proteína antifúngica

AMP: Péptido antimicrobiano

BAM: Formato binario comprimido de alineamiento/mapeo

CC: Componente celular

cDNA: DNA copia

CPM o RPM: Cuentas o lecturas por millón

CRP: Proteína rica en cisteínas

DEGs: Genes expresados diferencialmente

DNA: Ácido desoxirribonucleico

EDTA: Ácido etilendiaminotetraacético

FDR: Tasa de descubrimientos falsos

FM: Función molecular

GO: Ontología génica

GTF: Formato de transferencia de genes.

IC50: Concentración inhibitoria 50%

LiCl: Cloruro de litio

MFS: Superfamilia facilitadora de membrana

MIC: Concentración mínima inhibitoria

ORF: Pauta abierta de lectura

PB: Proceso biológico

PC: Componente principal

PCA: Análisis de componentes principales

PDA: Patata dextrose agar

PDB: Caldo patata dextrosa

pI: Punto isoeléctrico

RIN: Número de integridad de RNA

RNA: Ácido ribonucleico

TAE: Tris acetate EDTA

## 1. INTRODUCCIÓN

### 1.1. Los hongos patógenos y la necesidad de nuevas moléculas antifúngicas.

Las infecciones causadas por hongos representan un importante riesgo para la salud humana, producción de alimentos y seguridad alimentaria (Fisher et al., 2013). El impacto de las infecciones fúngicas sobre el ser humano varía desde infecciones superficiales comunes que afectan a un 25% de la población mundial, hasta infecciones invasivas con índices de mortalidad de hasta el 50%. La importancia de las enfermedades fúngicas se ha incrementado debido al creciente número de pacientes inmunodeprimidos por trasplantes o terapias oncológicas severas y a la aparición de nuevas cepas patógenas y resistencias a los antifúngicos usados comúnmente (Brown et al., 2012; Denning et al., 2011). Por ejemplo, las aspergilosis invasivas provocan entre uno y un millón y medio de muertes anuales, superando a la mortalidad de la malaria.

Las enfermedades fúngicas también suponen un gran reto en la agricultura, ya que los hongos son los patógenos principales de las especies vegetales cultivadas para la alimentación. También la producción de micotoxinas por parte de los hongos supone un problema de seguridad alimentaria, ya que estas sustancias pueden estar presentes en los alimentos ingeridos y ser peligrosos para la salud humana. Cabe resaltar el papel de los hongos en las pérdidas postcosecha. Su adaptación a un estilo de vida saprófito facilita la infección a través de heridas en el fruto ya cosechado producidas durante el transporte y el almacenamiento, traduciéndose en grandes pérdidas que pueden variar de un 15 a un 50% de la producción total (FAO, 1989).

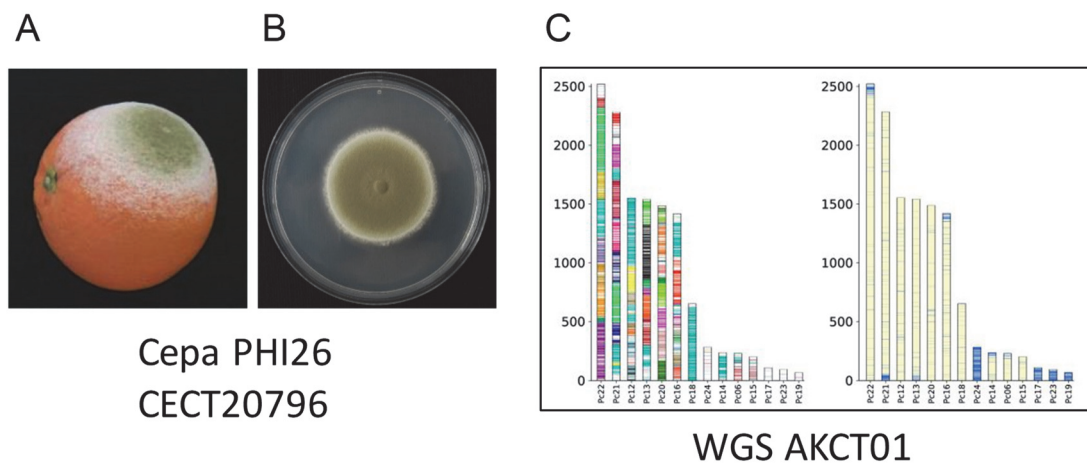
Por todo ello, el mercado global de fungicidas y medicamentos antifúngicos se encuentra en continuo crecimiento y se estima que antes del 2020 llegue al menos a 25.000 millones de euros anuales. Pero a pesar del enorme impacto sobre la salud humana y el suministro de alimentos, se han introducido muy pocos compuestos antifúngicos en el mercado global de fungicidas en los últimos treinta años. Además, las moléculas antifúngicas utilizadas en la actualidad pertenecen únicamente a tres clases diferentes, actuando fundamentalmente sobre la biosíntesis de ergosterol y pared celular, así como sobre la integridad de membrana plasmática (Campoy, S. and Adrio, 2017). Se están usando de forma masiva un número limitado de fungicidas químicos que presentan un limitado espectro de acción, favorecen la selección de cepas resistentes y además presentan toxicidad inespecífica poniendo en peligro la salud humana y el medioambiente. Además, se están empezando a describir resistencias cruzadas en clínica causadas por fungicidas utilizados en agricultura, lo que supone un enorme problema a medio plazo para el control de enfermedades en la agricultura. Por estos motivos existe una necesidad urgente de encontrar nuevas moléculas antifúngicas, y a ser posible con modos de acción diferentes a los ya conocidos (Perfect, 2016), y una crisis de antifúngicos completamente paralela a la más conocida de los antibióticos antibacterianos.

Los péptidos y pequeñas proteínas antimicrobianas (*AntiMicrobial Peptides*, AMP) tienen actividad antimicrobiana directa y son producidos por organismos de muy diversa escala filogenética, desde procariotas inferiores a eucariotas incluyendo insectos, plantas, hongos y animales superiores como el ser humano. Los AMPs forman parte de sistemas

de competición por nutrientes en microorganismos o mecanismos de defensa contra el ataque de virus, bacterias y hongos en organismos superiores. Los AMPs son de gran interés para el desarrollo de nuevos antibióticos incluyendo nuevos antifúngicos, y pueden ser producidos mediante estrategias biotecnológicas (Marcos et al., 2008). Todos los AMPs independientemente de su origen presentan características comunes, aunque su estructura es variable: son péptidos de pequeño tamaño (entre 10 a 50 aminoácidos), catiónicos y con alta proporción de residuos hidrofóbicos.

## 1.2. *Penicillium digitatum*: el principal patógeno de frutos cítricos.

La especie *Penicillium digitatum* y en menor medida *Penicillium italicum*, son los hongos patógenos postcosecha mayoritarios de los frutos cítricos, el principal cultivo frutal del mundo y de España. Ambos llegan a causar entre el 5 y 10% de pérdidas de la producción de cítricos durante su almacenamiento y comercialización, a pesar de la utilización masiva de tratamientos postcosecha con fungicidas. *P. digitatum* es un hongo necrótrofo causante de la enfermedad conocida como la podredumbre verde, específico de cítricos y que infecta a través de heridas preexistentes en la piel del fruto (Marcet-Houben et al., 2012; Palou, 2014) (Figura 2). Por su importancia, ha aumentado el interés por conocer mejor su patogenicidad (Gandía et al., 2014), su interacción con el fruto (González-Candelas et al., 2010) y se ha secuenciado el genoma de distintas cepas (Marcet-Houben et al., 2012; Sun et al., 2013), una de ellas la cepa de referencia utilizada por el laboratorio del IATA (Cepa PHI26, CECT 20796, Fig. 1). El genoma de esta cepa (WGS AKCT01) tiene un tamaño de 26 Mb y un total de 9.132 ORFs anotadas.



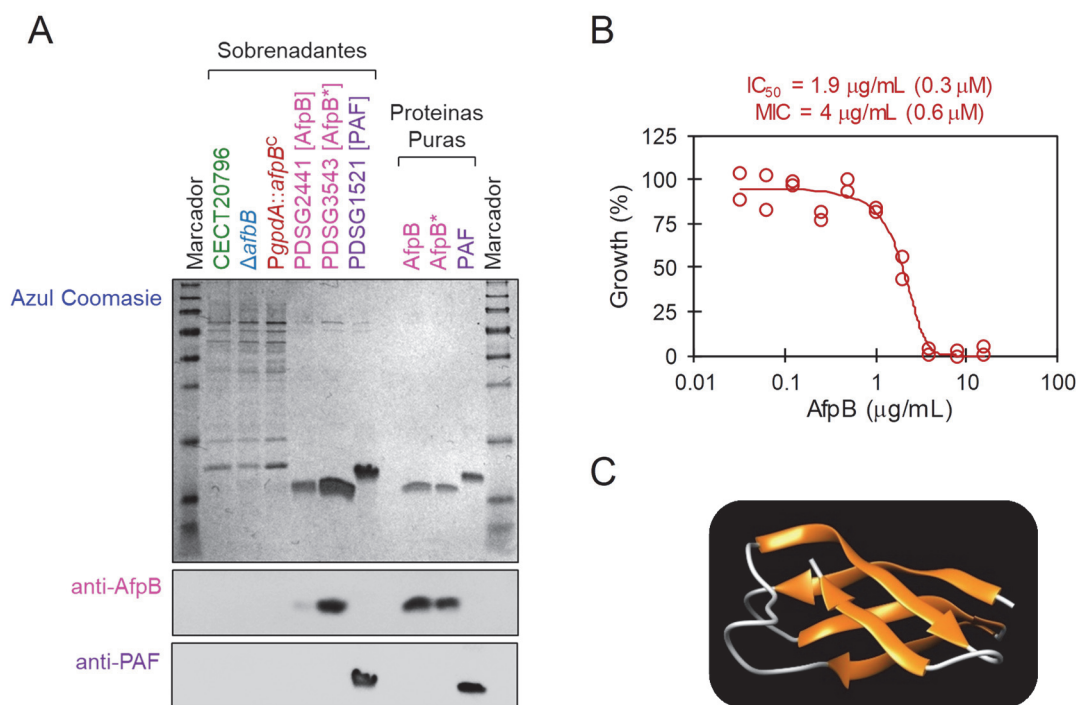
**Figura 2. El hongo *P. digitatum* patógeno postcosecha de frutos cítricos.** Fruto de naranja infectado por *P. digitatum* CECT20796 o PHI26 (A) y crecimiento del hongo en placa de cultivo de PDA (B). (C) Regiones genómicas compartidas entre *P. chrysogenum* y *P. digitatum* divididas en supercontigs obtenidos por WGS (*Whole Genome Shotgun*). Se muestran coloreados los genes ortólogos en ambos hongos (izquierda), y con el mismo color los que aparecen en un mismo *scaffold* en *P. digitatum*. Los genes coloreados en azul (derecha) representan a los genes parálogos de *P. chrysogenum*.

### 1.3. Las proteínas antifúngicas de hongos.

Un grupo de AMPs son las llamadas proteínas antifúngicas (*AntiFungal Proteins*, AFPs) producidas por hongos filamentosos ascomicetos, principalmente del género *Penicillium* y *Aspergillus* (Meyer, 2008). Las AFPs están filogenética y estructuralmente relacionadas con las proteínas antifúngicas de tipo defensinas, producidas por plantas o mamíferos. Son proteínas de alrededor de 50 aminoácidos, ricas en cisteínas (*Cysteine-Rich Proteins*, CRPs), catiónicas ( $pI > 9$ ) y con una estructura en láminas  $\beta$  estabilizada por puentes disulfuro (Hegedüs & Marx, 2013). Presentan un elevado interés biotecnológico por su alta actividad antifúngica, especificidad y estabilidad. Tienen actividad antifúngica a concentraciones micromolares y nula toxicidad frente a células de plantas o mamíferos (Beatriz & Álvaro, 2006; Garrigues, Gandía, Popa, et al., 2017; Rusznák et al., 2005). Siguiendo diferentes criterios, el laboratorio del IATA ha rastreado los genomas de hongos filamentosos encontrando una gran diversidad de proteínas AFP y proponiendo su clasificación en las clases A, B y C (Garrigues et al., 2016). En la clase A se encuentran la mayor parte de las AFPs mejor caracterizadas, como la PAF producida por *Penicillium chrysogenum* (Batta et al., 2009), la AFP de *Aspergillus giganteus* (Lacadena et al., 1995), y la NPAF de *Neosartorya fisheri* (Kovács et al., 2011; Virágh et al., 2015). Estas AFPs son secretadas en grandes cantidades por sus organismos productores que presentan inmunidad ante sus propias proteínas AFP. Distintos hongos tienen en su genoma un número variable de genes *afp*, entre uno y tres, y normalmente en una determinada cepa sólo se produce la proteína correspondiente a uno de ellos.

*P. digitatum* solo tiene en su genoma un gen *afpB* de clase B pero, sin embargo, la AfpB correspondiente no ha podido ser detectada en cultivos de cepas silvestres del hongo (cepa CECT20796, Figura 3A). El laboratorio del IATA ha logrado la producción recombinante de AfpB en el propio *P. digitatum* mediante una estrategia en la que la secuencia codificante se ha puesto bajo el control de secuencias reguladoras (promotor y terminador) del gen *paf* de *P. chrysogenum* (Garrigues et al., 2017; Sonderegger et al., 2016). Bajo estas condiciones, la proteína se produce en fases tardías/estacionarias de cultivo, cuando los nutrientes casi se han agotado, y se han conseguido altos rendimientos de 20 mg de proteína pura por L de cultivo en las cepas recombinantes de *P. digitatum* PDSG2441 (Fig. 2A) ó PDSG420 (Hernanz-Koers et al., 2018).

La disponibilidad de proteína pura está permitiendo su caracterización en profundidad. La AfpB tiene la particularidad de que es una proteína con alta actividad antifúngica frente a su propio hongo (concentración mínima inhibitoria, MIC = 4  $\mu\text{g}/\text{mL}$ , 0,6  $\mu\text{M}$ ) (Fig. 2B). Por otra parte, la AfpB de *P. digitatum* es muy activa frente otras especies de *Penicillium* e inactiva frente a bacterias y levaduras (Garrigues et al., 2017). Se ha modelizado su estructura terciaria, y es similar a las AFPs de otros hongos como PAF de *P. chrysogenum* (Fig. 2C). La AfpB presenta cinco láminas  $\beta$  antiparalelas que se conectan a través de cuatro lazos, donde uno de ellos es catiónico y está expuesto en la superficie de la proteína. Estas láminas generan dos hojas  $\beta$  con seis residuos de cisteína, donde se forman los puentes disulfuro que aportan estabilidad a la proteína (Garrigues et al., 2017). Cabe resaltar la presencia de una estructura tridimensional llamada motivo- $\gamma$ , cuya secuencia consenso es GXCX<sub>3-9</sub>C y que se encuentra generalmente en proteínas con actividad antimicrobiana (Yount & Yeaman, 2004).



**Figura 3. La proteína antifúngica AfpB de *P. digitatum*.** (A) Análisis mediante SDS-PAGE y Western blot del sobrenadante de cepas de *P. digitatum* productoras de AFPs (AfpB y PAF) (4-8 µg de proteína por muestra) y de las proteínas purificadas (1 µg por línea). Visualizadas por azul de Coomassie y tras un inmunoblot con anticuerpos anti-AfpB y anti-PAF. (B) Curva dosis-respuesta del crecimiento de *P. digitatum* en la presencia de AfpB. (C) Modelo tridimensional de la proteína AfpB. MIC: Concentración mínima inhibitoria (*Minimum Inhibitory Concentration*), IC<sub>50</sub>: Concentración a la que se inhibe crecimiento del hongo a la mitad (*50% Inhibitory Concentration*).

#### 1.4. Mecanismo de acción y función de proteínas AFP.

Las AFPs tienen un gran interés biotecnológico como nuevas moléculas antifúngicas específicas, muy activas, altamente estables y que se pueden producir en biofactorías fúngicas mediante biotecnología. Para la futura aplicación biotecnológica de estas AFPs, es necesario: (i) conocer su mecanismo de acción antifúngico, y (ii) la función biológica que desempeñan para el hongo productor. Además, el conocimiento de su modo de acción puede identificar nuevas dianas potenciales para el desarrollo de moléculas antifúngicas. Los escasos estudios sobre modo de acción y función biológica se han centrado en las proteínas de clase A, sobre todo en las más estudiadas: La AFP de *A. giganteus* y la PAF de *P. chrysogenum* (Meyer & Jung, 2018). La AFP se une a la pared celular o a la membrana plasmática de hongos sensibles, produciendo una pérdida de la integridad de la pared celular y causando la permeabilización de la membrana. Esta pérdida de la integridad de la pared celular puede estar relacionada con la inhibición de la actividad quitina sintasa por la presencia de AFP, que se ha descrito como una actividad importante para el correcto desarrollo, morfogénesis, patogenicidad y virulencia de los hongos filamentosos (Gandía et al., 2012, 2014). Otro posible efecto de la presencia de AFP puede ser su unión a ácidos nucleicos a través del motivo estructural de unión a oligonucleótidos

u oligosacáridos (OB), causando la neutralización de la carga de los ácidos nucleicos y la condensación del DNA (Marti et al., 2002). En cambio, la PAF de *P. chrysogenum* hiperpolariza la membrana plasmática y perturba la homeostasis de calcio, causando la apoptosis por el aumento de especies reactivas de oxígeno en hongos sensibles del género *Aspergillus*. De forma similar, la NAFP de *Neosartorya fischeri* induce la apoptosis en *A. nidulans* (Hegedüs & Marx, 2013; Meyer & Jung, 2018). Hasta la fecha, no se ha hecho un estudio transcriptómico para aportar datos sobre el modo de acción de proteínas AFP. El único trabajo relacionado es un meta análisis de todos los datos transcriptómicos disponibles del hongo *Aspergillus niger* y en el que se buscan correlaciones con la expresión del gen que codifica para AFP (Paege et al., 2016).

## 2. OBJETIVOS

El objetivo global de este trabajo es contribuir a la caracterización del modo de acción de la proteína antifúngica AfpB sobre *P. digitatum*, el hongo fitopatógeno en el que fue identificada la proteína. Para ello, se llevará a cabo un análisis del transcriptoma a través de secuenciación masiva por RNA-seq y posterior análisis bioinformático de los resultados obtenidos. Para alcanzar este objetivo global se plantean los siguientes objetivos individuales:

1. Obtención de muestras de micelio y de su RNA total de la cepa silvestre *P. digitatum* (CECT 20796) y de una cepa sobreproductora de AfpB (PDSG420), bajo distintos tratamientos de exposición a AfpB.
2. Determinación del transcriptoma mediante RNA-seq y de los genes expresados diferencialmente (DEGs, *Differentially Expressed Genes*) en los distintos tratamientos. Este objetivo se ha abordado con la ayuda del servicio de Genómica del SCSIE de la Universitat de València.
3. Análisis del enriquecimiento funcional de los DEGs. Para este objetivo ha sido necesario mejorar la anotación funcional por ontología génica (GO) del genoma de *P. digitatum* CECT 20796 que estaba disponible.
4. Identificación de genes potencialmente implicados en el modo de acción de AfpB.

### 3. MATERIALES Y MÉTODOS

#### 3.1. Cepas, medios y condiciones de cultivo.

Las cepas utilizadas fueron *P. digitatum* CECT20796 (aislado PHI26) (Marcet-Houben et al., 2012) y una cepa derivada de la misma (PDSG420) super productora de la proteína AfpB (Hernanz-Koers et al., 2018). Todas las manipulaciones se realizaron en condiciones de esterilidad. Para la obtención de conidios asexuales (esporas) ambas cepas se crecieron en medio sólido PDA (*Potato Dextrose Agar*) durante 7-10 días a 25°C. Las esporas se recogieron de la placa de PDA por raspado, se dispersaron en agua estéril, y se filtraron para eliminar restos de micelio. La concentración de conidios se determinó en una cámara Neubauer.

Los cultivos para los tratamientos se crecieron en 200 mL de medio líquido de PDB (*Potato Dextrose Broth*) al 25% (matraces de 1L) a una concentración de  $1 \times 10^6$  conidios/mL, durante 42-48 horas con una agitación de 180 rpm y a una temperatura de 25°C. Se realizaron tres réplicas biológicas para cada uno de los tratamientos detallados en la Figura 3A, los cuales varían en función de la ausencia o del tiempo que la proteína AfpB está presente en el medio y su concentración (ver sección 4.1 de Resultados).

#### 3.2. Extracción y procesamiento de RNA total

El micelio obtenido en los tratamientos se obtuvo por filtración y escurrido de los cultivos. Inmediatamente se congeló en nitrógeno líquido y almacenó a -80 °C. La extracción de RNA según el procedimiento de extracción fenólica en LiCl se realizó a partir del micelio congelado de las cepas de *P. digitatum* conforme al protocolo descrito (Gandía et al., 2012). Posteriormente, 10 µg de RNA total se trataron con DNasa libre de RNasa para eliminar los restos de DNA, mediante el kit AM1906 DNA-free de acuerdo con las instrucciones del fabricante (AMBION). La calidad y concentración del RNA total se determinaron inicialmente en el IATA por espectrofotometría mediante lectura en Nanodrop y gel de electroforesis TAE (Tris Acetato EDTA) y 1% agarosa teñido con bromuro de etidio (Sambrook, J. and Russell, 2001). Las muestras se enviaron al Servicio de Genómica del SCSIE de la Universitat de València. para su validación para utilización en RNA-seq mediante análisis por electroforesis capilar (Bioanalyzer 2100, Agilent).

#### 3.3. RNA-seq

La secuenciación de RNA mensajero de alto rendimiento fue realizado por el Servicio de Genómica del SCSIE de la Universitat de València. Para la construcción de la librería de cDNA se usó el protocolo *TruSeq Stranded mRNA Sample Prep Kit* (Illumina) y se analizó la calidad con el Bioanalyzer 2100 (Agilent). Se secuenciaron con lectura única (*single-reads*) las librerías con la plataforma NextSeq 500 de Illumina, obteniéndose lecturas de 75 nucleótidos en formato Fastq. El protocolo de procesamiento de las lecturas siguió las pautas recomendadas (Conesa et al., 2016) y fue el siguiente:

1. Se usó el programa FastQC para determinar la calidad de las lecturas.
2. Las secuencias de baja calidad (menos de 20 nucleótidos y las que presentan un elevado número de N) se eliminaron mediante el software Cutadapt que también excluye las secuencias de los oligonucleótidos adaptadores.
3. Se mapearon las lecturas sobre el genoma referencia de *P. digitatum* CECT 20796 (NCBI, WGS AKCT01) descargado de Ensembl con el software Tophat2 y se obtuvieron los ficheros BAM (*Binarily compressed sequence Alignment/Map*).
4. El software Samstats se usó para controlar la calidad de los ficheros BAM.
5. A partir de estos ficheros BAM y el fichero GTF (*Gene Transfer Format*) el software Qualimap se usó para controlar la calidad, detectando los sesgos en la secuenciación y en el mapeo de las lecturas, y se utilizó para la cuantificación del número de lecturas por gen.
6. Se detectaron los genes de expresión diferencial (DEGs) utilizando dos métodos y paquetes alternativos de R Bioconductor: EdgeR y DESeq2, basados en la distribución binomial negativa. Los datos se normalizaron para evitar sesgos relacionados con la profundidad de secuenciación entre las muestras. Para este último sesgo se usaron métodos diferentes en función de cada software, por TMM (*Trimmed Mean of M-values*) en el caso de EdgeR y por la relación entre media y varianza cuando se estima la dispersión de los datos (*sizeFactors*) en el caso de DESeq2.

### 3.4. Mejora de la anotación funcional y enriquecimiento funcional.

El software Blast2GO v5.0 (Conesa & Stefan, 2008) se utilizó para la mejora de la anotación funcional y el análisis de enriquecimiento funcional de los DEG. Se mejoró la anotación funcional del genoma de referencia de *P. digitatum* descargado de Ensembl ([www.ensembl.org](http://www.ensembl.org)) en formato FASTA. Se creó una base de datos local de las proteínas en formato FASTA de diez genomas de hongos cercanos filogenéticamente a *P. digitatum* (Tabla 1). La anotación se llevó a cabo con el siguiente procedimiento:

1. Se realizó una búsqueda BLAST en Blast2GO con el algoritmo BLASTP 2.4.0 y sus parámetros por defecto, con todas las proteínas anotadas en *P. digitatum* contra la base de datos local.
2. El resultado se utilizó para un mapeo online contra la base de datos PIR (*Non-redundant Reference Protein Database*), que incluye PSD, Uniprot, Swiss-Prot, TrEMBL, RefSeq, GenPept y PDB. Se obtuvieron términos GO (*Gene Ontology*) asociados a las proteínas de *P. digitatum*.
3. Los términos GO con mayores niveles de fiabilidad se seleccionaron a través del paso de anotación utilizando el criterio de filtrado de Blast2GO y los parámetros de defecto.
4. Se enriqueció la anotación en términos GO con la base de datos EMBL-EBI InterPro y con el ANNEX *annotation augmentation*.



5. Posteriormente se mapearon los códigos de enzimas que se asocian a las proteínas de la base de datos enriquecida de *P. digitatum*.

**Tabla 1. Especies de hongos usadas para la base de datos local en Blast2GO.**

<b>Especie</b>	<b>Ensamblado</b>	<b>Base de datos</b>
<i>Penicillium chrysogenum</i>	ASM71027v1	NCBI
<i>Penicillium expansum</i>	PEX1	Ensembl
<i>Penicillium italicum</i>	ASM76976v1	Ensembl
<i>Aspergillus fumigatus</i>	ASM265v1	Ensembl
<i>Aspergillus nidulans</i>	ASM1142v1	NCBI
<i>Aspergillus niger</i>	ASM285v2	NCBI
<i>Botrytis cinerea</i>	ASM83294v1	Ensembl
<i>Fusarium oxysporum</i>	FO2	Ensembl
<i>Magnaporthe oryzae</i>	MG8	Ensembl
<i>Neurospora crassa</i>	NC12	Ensembl

A partir de los ficheros que contienen los DEGs (*Differentially Expressed Genes*), junto con la anotación realizada con Blast2GO, se determinó el enriquecimiento funcional de los DEGs en Blast2GO, a través del paquete FatiGO que realiza un test estadístico basado en la prueba exacta de Fisher. Esta herramienta busca si la asociación entre los términos GO y los genes son significativas, obteniendo así los términos GO que representan las funciones que los DEGs presentan.

## 4. RESULTADOS

### 4.1. Respuesta de las cepas de *P. digitatum* a tratamientos con AfpB.

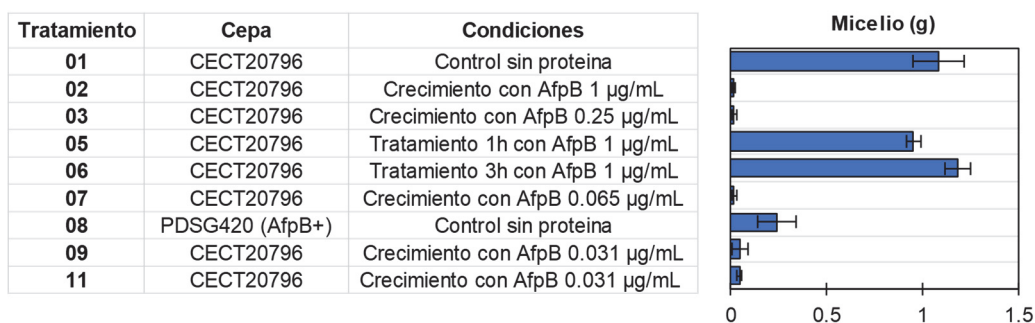
En este trabajo se han planteado tres tipos de tratamientos o muestras para obtener datos de RNA-seq relacionados con el modo de acción de AfpB sobre *P. digitatum*, los cuales se han comparado con la condición control sin tratamiento (resumidos en Fig. 3A). Los datos previos de inhibición de crecimiento (Fig. 2C, MIC de 4 µg/mL) se habían obtenido, como es habitual en el trabajo con AMPs, en cultivos miniaturizados con placas de microtítulo de 96 pocillos, con un volumen total de 100 µL en cultivo estático y sin agitación. Para este estudio se hacía necesario determinar la actividad inhibitoria de AfpB en condiciones de cultivo en matraz, con un volumen total de 200 mL que permitiera recoger suficiente micelio de *P. digitatum* para la extracción de RNA. Bajo estas condiciones experimentales, se obtuvieron para la cepa parental CECT20796 de *P. digitatum* alrededor de 1 g de micelio en la condición control (tratamiento 01, Figs. 3A y 3B), cantidad adecuada para la extracción de RNA.

El primer tipo de muestras a comparar con el control fueron tratamientos sobre el micelio ya crecido a una concentración sub-inhibitoria de AfpB (1 µg/mL) durante 1 ó 3 h (tratamientos 05 y 06, Fig. 3A). En estos casos, no se observó efecto de la proteína sobre el crecimiento tal y como estaba previsto, obteniéndose una biomasa similar a la condición control en ambos casos (Fig. 3A y Fig. 3B).

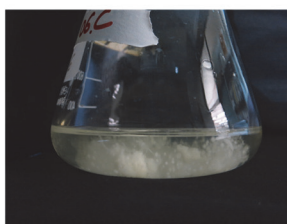
En el segundo tipo de tratamiento, el hongo se creció en presencia de la proteína desde el principio, como un componente más del medio de cultivo. Sorprendentemente, en estas condiciones de cultivo la proteína AfpB es mucho más activa que en las placas de microtítulo, ya que a muy bajas concentraciones (desde 1 µg/mL hasta 65 ng/mL) inhibe el crecimiento de *P. digitatum* casi en su totalidad (Fig. 3A, tratamientos 02, 03 y 07), obteniéndose una cantidad de micelio insuficiente para la extracción de RNA. Este resultado también queda reflejado cualitativamente en las imágenes de los matraces de los tratamientos 03 y 07 (Figs. 3C-D). En cambio, una disminución en la concentración hasta 31 ng/mL permitió un cierto crecimiento del hongo (tratamientos 09 y 11, Fig.3A) que, aunque muy limitado (<100 mg), hizo posible la extracción del RNA (ver apartado siguiente, 4.2).

En el tercer tipo de muestra se comparó la cepa control con la sobreproductora de AfpB. Las dos cepas de *P. digitatum* no se comportan igual en las mismas condiciones de cultivo, ya que la cepa parental CECT20796 (tratamiento 01) tiene un mayor crecimiento que la cepa PDSG420 (tratamiento 08) (Fig. 3A).

A



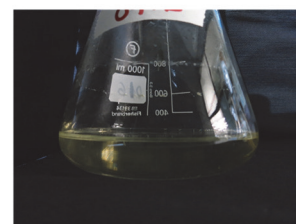
B



C



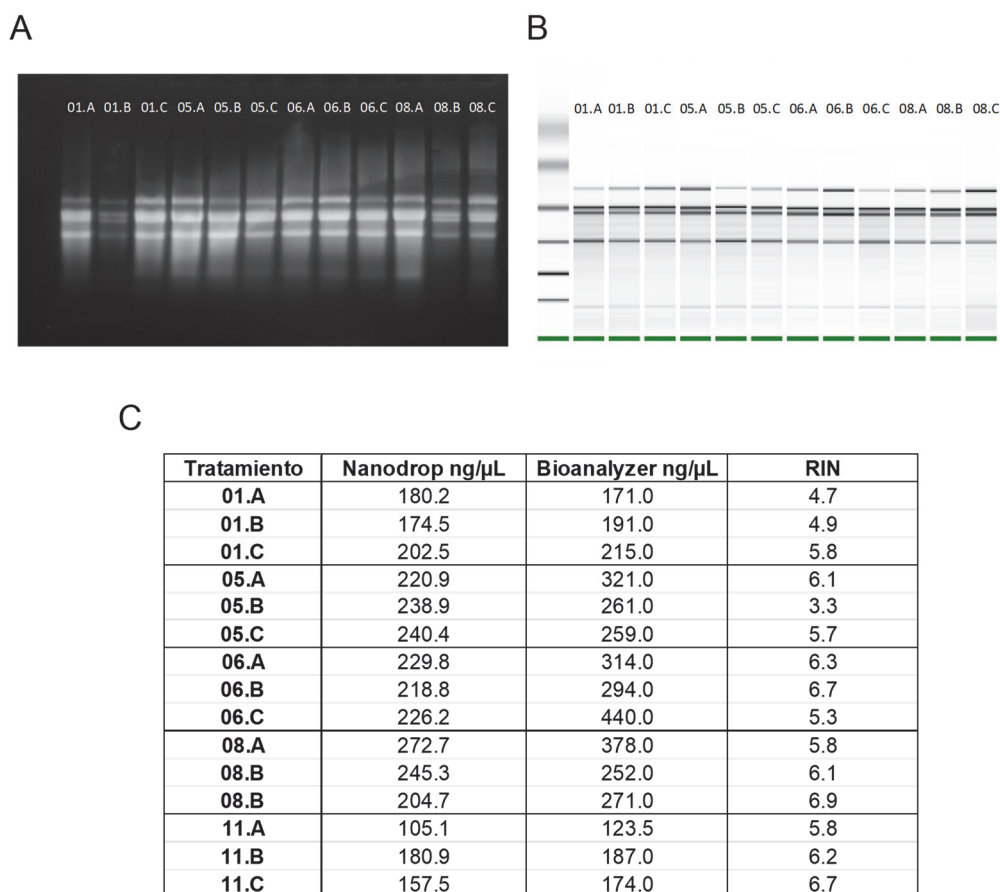
D



**Figura 3. Comportamiento de *P. digitatum* en presencia de la proteína antifúngica AfpB.** (A) Características de los tratamientos a los que ha sido sometido *P. digitatum* y su respuesta de crecimiento, calculado como peso medio de micelio recogido en tres réplicas biológicas realizadas por tratamiento. (B) *P. digitatum* (PHI26) crecido en ausencia de AfpB durante 42-48 horas, a 28°C y a 180 rpm en PDB. Antes de la recogida del micelio se añadió durante 1 hora la proteína antifúngica (1 µg/mL). (C) *P. digitatum* (PHI26) crecido en presencia de AfpB a una concentración de 0.065 µg/mL durante 42-48 horas, a 28°C y a 180 rpm en PDB. (D) *P. digitatum* (PHI26) crecido en presencia de AfpB a una concentración de 0.25 µg/mL durante 42-48 horas, a 28°C y a 180 rpm en PDB.

## 4.2. Obtención de RNA total de calidad apropiada

Se extrajo RNA total de cada una de las 15 muestras de micelio (tratamientos 01, 05, 06, 08 y 11; tres réplicas biológicas A, B, y C de cada uno). El RNA de las muestras presentó la calidad, concentración e integridad necesaria para el experimento de RNA-seq (Fig. 4). La calidad del RNA se comprobó inicialmente en gel de electroforesis de agarosa (Fig. 4A), observándose las bandas características del RNA ribosomal (28S y 18S). Estos resultados se confirmaron en la imagen virtual generada mediante electroforesis capilar de RNA (Bioanalyzer) (Fig. 4B). Los valores de concentración obtenidos por espectrofotometría (Nanodrop) fueron menores que los obtenidos por el Bioanalyzer (Fig. 4C), aunque consistentes entre muestras, debido a la variabilidad asociada a cada una de las técnicas. El número de integridad del RNA o RIN (Fig. 4C) no debe situarse por debajo de un valor de 4,0 para que una muestra sea idónea para su análisis por RNA-seq, lo que se cumple generalmente en las muestras analizadas, presentando una media de RIN de 5,8. El valor mínimo se encuentra en 3,3 (tratamiento 06.B) y el valor máximo en 6,9 (tratamiento 08.B), cercano al valor óptimo para el experimento de RNA-seq que se sitúa en 7,0 (Fig. 4A).



**Figura 4. Concentración, calidad e integridad del RNA obtenido del micelio de *P. digitatum*.** (A) Gel de electroforesis TAE (Tris Acetato EDTA) 1% de agarosa, siguiendo el orden de izquierda a derecha son las réplicas A, B y C de los tratamientos 01, 05, 06 y 08. (B) Imagen virtual del Bioanalyzer 2100 (Agilent), siguiendo el orden de izquierda a derecha son las réplicas A, B y C de los tratamientos 01, 05, 06 y 08. (C) La concentración del RNA fue medida por espectrofotometría en el Nanodrop y por electroforesis capilar en el Bioanalyzer para cada una de las réplicas por tratamiento realizadas. Esta última técnica permite la medición de la integridad del RNA o RIN.

### 4.3. Determinación de genes que se expresan diferencialmente

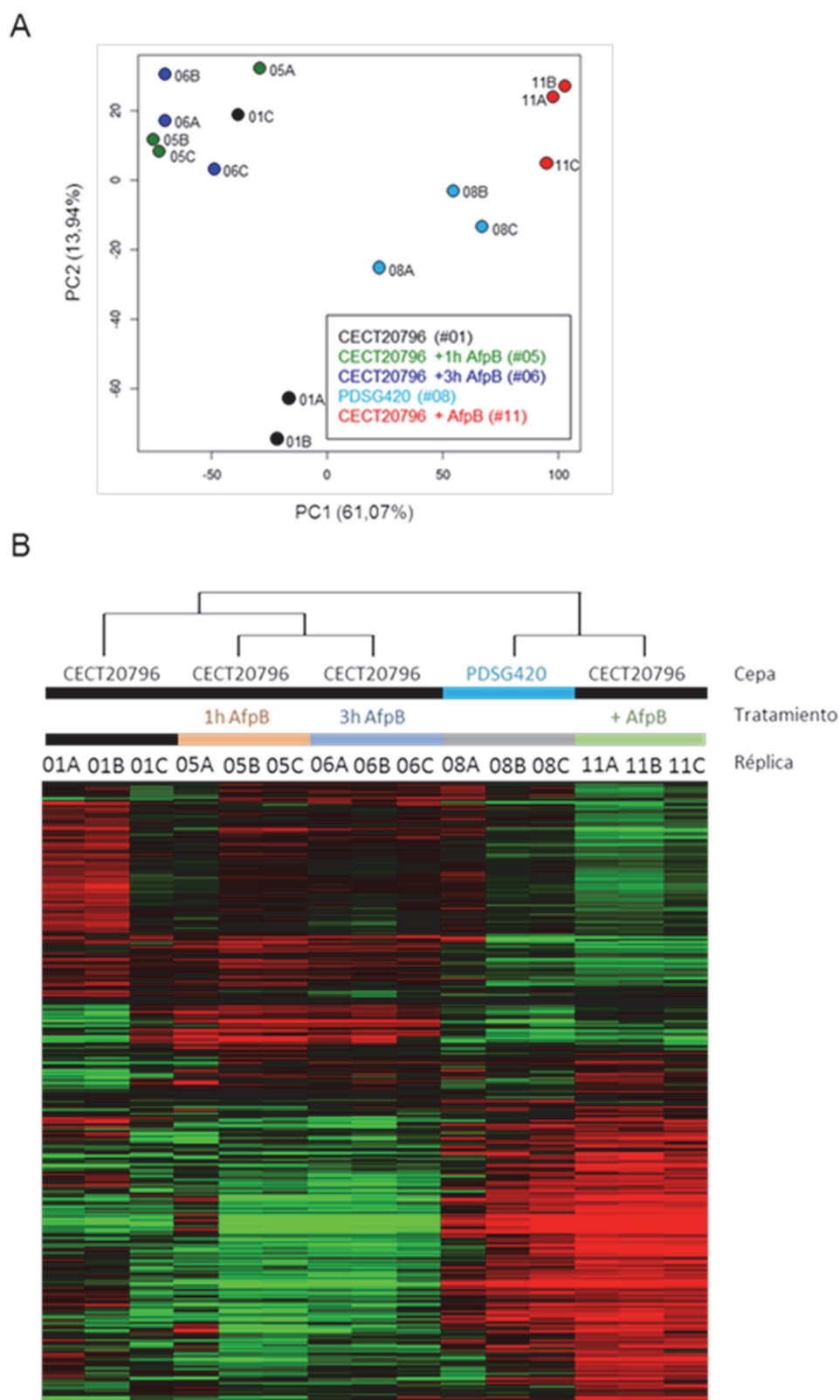
Tras ser evaluada la concentración, calidad e integridad del RNA en las muestras de los diferentes tratamientos, dichas muestras (15 en total) se enviaron al servicio de secuenciación masiva del SCSIE de la Universitat de València para su análisis por RNA-seq. En dicho servicio se obtuvieron las librerías de cDNA a partir de las muestras enriquecidas en RNA mensajero poli-A, que posteriormente se secuenciaron a través de NextSeq 500 de Illumina. El experimento se diseñó para una cantidad de lecturas crudas por réplica del orden de entre 20 y 30 millones de lecturas ( $2-3 \times 10^7$ ), con un tamaño de 75 nucleótidos por lectura, para dar una cobertura total del genoma del orden de 50-75 veces por muestra. Tras el control de calidad de las lecturas crudas que incluye la eliminación de los adaptadores de las librerías y las secuencias de baja calidad, el análisis de calidad mostró que no existían grandes diferencias entre las lecturas crudas y procesadas, ya que más del 99% de las secuencias crudas eran de calidad (Tabla 2).

**Tabla 2. Calidad de las lecturas secuenciadas en RNA-seq**

Tratamiento	Lecturas crudas	Lecturas procesadas	% Lecturas con calidad
01.A	3.326E+07	3.326E+07	99.99
01.B	2.899E+07	2.898E+07	99.98
01.C	2.275E+07	2.274E+07	99.97
05.A	2.456E+07	2.456E+07	99.99
05.B	4.062E+07	4.061E+07	99.99
05.C	3.629E+07	3.628E+07	99.99
06.A	3.314E+07	3.314E+07	99.99
06.B	2.895E+07	2.894E+07	99.97
06.C	3.330E+07	3.330E+07	99.99
08.A	3.195E+07	3.194E+07	99.98
08.B	2.180E+07	2.180E+07	99.98
08.B	2.576E+07	2.574E+07	99.93
11.A	3.496E+07	3.496E+07	99.99
11.B	2.706E+07	2.705E+07	99.97
11.C	2.514E+07	2.511E+07	99.88

Las lecturas procesadas de cada réplica se mapearon sobre el genoma de *P. digitatum* con el software Tophat2, obteniendo unos porcentajes de lecturas mapeadas variables entre las réplicas entre 91,8 y 94,0%, suponiendo una cobertura del genoma del 77X aproximadamente por réplica.

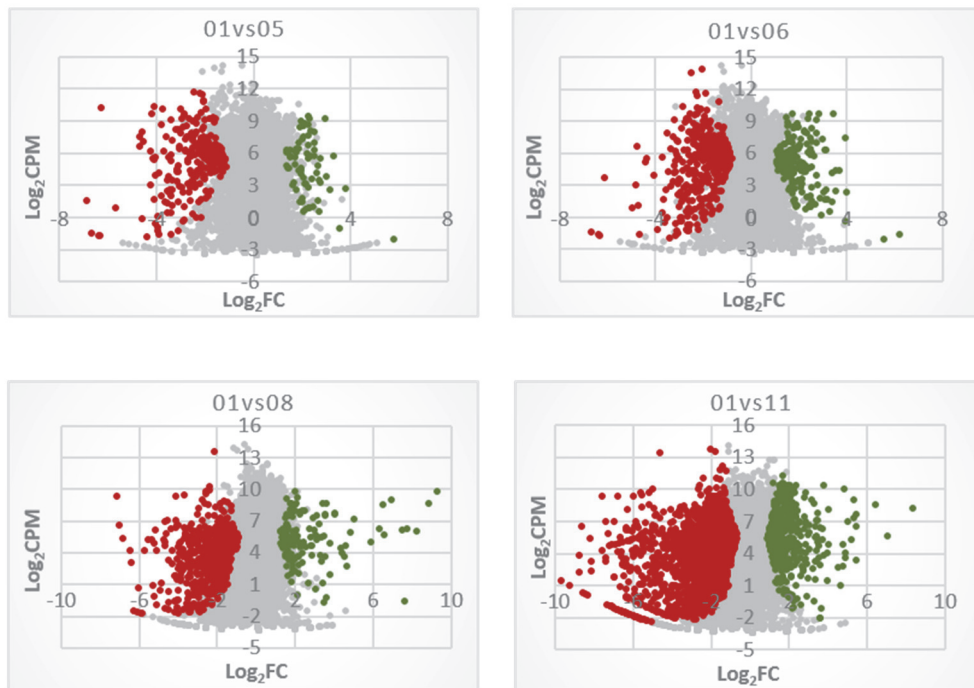
Con los datos normalizados del RNA-seq se pudo observar cómo las distintas réplicas se distribuyen en función de su varianza en un análisis de componentes principales o PCA (*Principal Component Analysis*). En el PCA obtenido, la primera componente principal explica un 61,07% de la varianza de las réplicas, y la segunda componente principal explica el 13,94% de la variabilidad de los datos. Los tratamientos a tiempos cortos con la proteína AfpB (tratamiento 05 y 06) junto con la condición control (tratamiento 01) se separan de los tratamientos largos (tratamiento 11) o la cepa productora de AfpB (tratamiento 08) por la primera componente principal. Las réplicas de un mismo tratamiento aparecen conjuntamente a excepción de una de las réplicas de la condición control (Fig. 5A). Dos de las tres réplicas del tratamiento control (01A y 01B) se separan del resto de muestras por la segunda componente principal (eje vertical). A partir de los datos de de las réplicas por tratamiento en un mapa de expresión (*heatmap*, Figura 5B) obtenido a partir de las lecturas por millón o CPM (*Counts Per Million*) que mapean sobre cada gen.



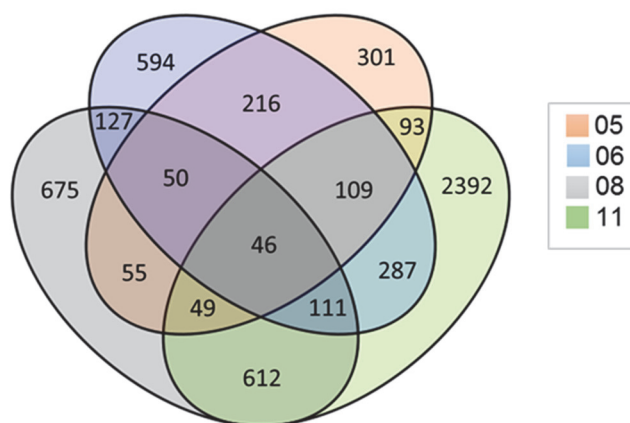
**Figura 5. Variabilidad de las réplicas por tratamiento.** (A) Análisis de componentes principales (PCA) de los datos normalizados (log<sub>2</sub>RPM) de cada réplica. (B) Agrupación jerárquica (*hierarchical clustering*) basada en la correlación de los datos de expresión normalizados (log<sub>2</sub>RPM) de los DEGs. Los colores representan el nivel de expresión por gen de más expresado (verde) a menos expresado (rojo), calculado como el valor medio de cada gen en el conjunto de las réplicas al que se resta el valor en cada una de las réplicas. PC1: Primera componente principal, PC2: Segunda componente principal

El dendrograma de los datos de expresión agrupa los distintos tratamientos de manera que los más parecidos entre sí son los tratamientos puntuales a 1 y 3 horas (tratamientos 05 y 06) por una parte, y los del tratamiento en continuo (11) junto con el crecimiento en presencia de la proteína (08) por otra. Los datos recogidos en la Figura 5 indican que, con la variabilidad experimental y limitaciones propias de la técnica, las muestras son apropiadas para el análisis transcriptómico de expresión diferencial.

A



B



**Figura 6. Representación de los DEGs por tratamiento obtenidos de EdgeR (FDR<0,01).** (A) Diagrama de dispersión (*volcano plot*) donde cada punto representa a un gen en función del  $\text{log}_2\text{FC}$  y  $\text{log}_2\text{CPM}$  (valor normalizado de las lecturas por millón o RPM). Los DEGs se representan en rojo para los genes reprimidos y en verde para los genes inducidos en cada tratamiento respecto al control (tratamiento 01). (B) Diagrama de Venn del número de DEGs en cada tratamiento y de los compartidos entre tratamientos.



**Tabla 3. Número de genes expresados diferencialmente por software.**

Software	DESeq2			EdgeR			Comunes				
	Muestras	Inducidos	Reprimidos	Total	Inducidos	Reprimidos	Total	Inducidos	Reprimidos	Total	% Inducidos
01 vs 05	202	493	695	134	408	542	126	371	497	25,35	74,65
01 vs 06	477	907	1384	335	689	1024	293	633	926	31,64	68,36
01 vs 08	576	929	1505	297	792	1089	285	723	1008	28,27	71,73
01 vs 11	1877	2266	4143	1257	2167	3424	1231	2052	3283	37,5	62,50

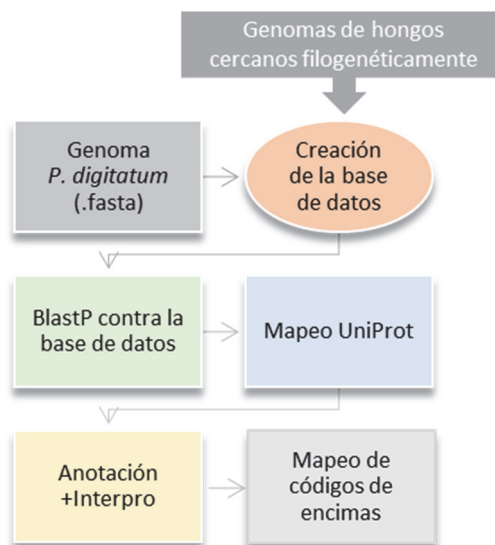
Para la obtención de los genes DEGs que de forma significativa se expresan diferencialmente, se utilizaron los softwares EdgeR y DESeq2. Cada uno de los tratamientos se comparó con el control o tratamiento 01 para determinar los DEGs y distinguir entre genes inducidos y reprimidos en cada caso. El número de DEGs varía con respecto al software utilizado, siendo EdgeR más restrictivo que DESeq2 con los mismos parámetros (FDR < 0.05) (Tabla 3). Esto se debe a que los softwares utilizan métodos diferentes para la normalización de los datos, aunque se basan ambos en la binomial negativa y suelen ser los métodos que aportan resultados más similares. Con respecto al control, el tratamiento 11 presenta un mayor número de DEGs que los demás tratamientos llegando a 4.143 genes (casi la mitad del genoma), seguido del tratamiento 08 y tratamiento 06, mientras que el tratamiento 05 es el que menor número de DEGs presenta no superando los 695 genes. Los genes reprimidos se encuentran en mayor número que los genes inducidos en todos los casos estudiados. El porcentaje de genes reprimidos comunes en ambos softwares no disminuye del 62,5%, y los genes inducidos no superan el 37,5% (Tabla 3). Este hecho también se observa en los Volcano plots (FDR < 0.01), donde los DEGs reprimidos e inducidos (puntos de color rojo y verde respectivamente) se representan sobre los genes no expresados diferencialmente en cada uno de los casos (Fig. 6A). Conforme la proteína AfpB permanece más tiempo con el hongo, se puede observar un aumento de DEGs (Fig. 6A y 6B), y también un mayor número de genes compartidos entre los tratamientos 05 y 06 o tratamientos cortos, y entre los tratamientos 08 y 11 o tratamientos en continuo. (Fig. 6B). Por otra parte, el cambio en el parámetro de significancia (FDR) de menor que 0.05 a menor que 0.01, se relaciona directamente con la reducción de los DEGs en cada tratamiento, así como un menor número de los DEGs que son compartidos entre los tratamientos.

#### 4.4. Mejora de la anotación funcional

La anotación funcional basada en ontologías génicas o términos GO (*Gene Ontology*) se clasifica en tres categorías: proceso biológico (PB), función molecular (FM) y componente celular (CC). También se divide en niveles, siendo los niveles más altos los que describen de forma más detallada la categoría de un gen en cuestión. Un procedimiento habitual en un proyecto de RNA-seq es determinar si las listas de DEGs tienen un enriquecimiento específico en alguna categoría GO, lo que indicaría que esa categoría funcional está implicada en el proceso biológico estudiado. En nuestro caso, los estudios preliminares que realizamos nos revelaron que la anotación funcional GO del genoma de *P. digitatum* CECT 20796 (WGS AKCT01) disponible en las bases públicas (JGI *Joint Genome Institute* y UniProt) y que data del año 2012 (Marcet-Houben et al., 2012; Sun et al., 2013) era muy pobre, y con ella no se obtuvo ninguna categoría funcional significativa en ninguno de los tratamientos estudiados. Por tanto, nos planteamos mejorar dicha anotación funcional. La mejora de la anotación funcional en términos GO se llevó a

cabo con el paquete Blast2GO v5.0, siguiendo el procedimiento recomendado (Fig. 7) (ver apartado 4.4 de Materiales y Métodos).

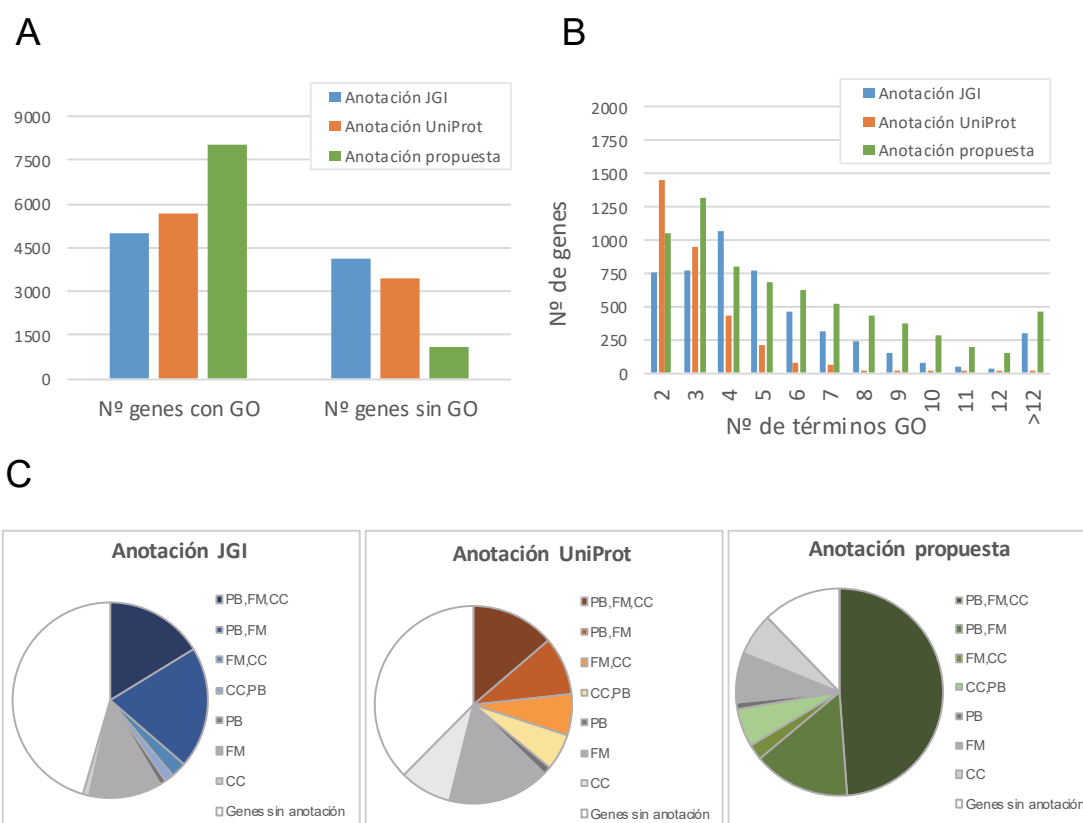
La nueva anotación propuesta en este trabajo recoge 8.014 genes anotados que representa un 87,7% del total (Fig. 8). En comparación con la anotación de JGI, se ha aumentado el número de genes anotados en un 33,2% y en el caso de UniProt en un 25,3%, quedando un 12,2% de genes sin anotación frente al 45,5% que presentaba el JGI y 37,5% de UniProt (Fig. 8A). Se ha representado el número de genes con respecto al número de términos GO, tendiendo a disminuir conforme aumenta el número de términos GO en ambas anotaciones, es decir que cada vez son menos los genes que tienen un alto número de términos GO acumulados (Fig. 8B). Los genes pueden estar anotados con términos GO de las tres, dos, una o ninguna de las categorías funcionales mencionadas anteriormente. En la nueva anotación propuesta en este trabajo observamos cómo se ha aumentado el porcentaje de genes con las tres categorías y disminuido el número de genes sin anotación con respecto a la anotación de JGI y UniProt (Fig. 8C). La nueva anotación funcional obtenida en este estudio se encuentra disponible en el laboratorio del IATA en formato Blast2GO (fichero “penicillium\_digitatum\_phi26\_blast2go\_v3\_6.b2g”, tamaño 69,4 Mb).



**Figura 7. Esquema del procedimiento seguido para obtener la anotación propuesta en Blast2GO.**

El genoma de referencia de *P. digitatum* fue descargado de Ensembl en formato FASTA. La base de datos local utilizada fue realizada con Blast2GO con los parámetros por defecto y a partir de los genomas de diez hongos cercanos filogenéticamente a *P. digitatum* (Tabla 1).





**Figura 8. Mejora de la anotación funcional de *P. digitatum* con Blast2GO.** Representación del número de genes anotados y no anotados en las bases de datos JGI, UniProt y la propuesta obtenida por Blast2GO (A), y en función del número de términos GO (B). En cada una de las anotaciones analizadas se muestra el número de genes que tienen una, dos y tres combinaciones de las categorías GO o ninguna de ellas (C). PB: 'Proceso Biológico', FM: 'Función Molecular', CC: 'Componente Celular'.

#### 4.5. Enriquecimiento funcional

Se realizó un enriquecimiento funcional de los DEGs a través de la herramienta Blast2GO mediante un test de Fisher y la nueva anotación funcional propuesta como referencia. El análisis proporciona información sobre aquellas funciones o términos GO que se encuentran sobre o infrarrepresentados en nuestro conjunto de DEGs en comparación con las funciones anotadas en la referencia. Se han acotado los resultados a los GO de niveles más altos, que en el caso de los tratamientos 05 y 06 son los niveles 7 y 8 (Tabla 4), mientras que para los tratamientos 08 y 11 se llega a niveles de 9 y 10 (Tabla 5). Como la mayoría de DEGs en cada tratamiento son genes reprimidos, los términos GO de estos genes son los que aparecen mayormente representados en el enriquecimiento. Del mismo modo, la categoría de proceso biológico predomina en la anotación funcional propuesta, por lo que mayoritariamente los términos GO sobre o infrarrepresentados son de proceso biológico.

**Tabla 4. Enriquecimiento funcional de por ontología génica (GO) de DEGs en tratamientos 05 y 06.**

Se muestran las funciones enriquecidas de los tratamientos 05 y 06 en los niveles GO 7 y 8, los cuales fueron los más altos que dieron valores significativos (FDR<0,01) en estos tratamientos. En rojo las categorías sobre-representadas y en verde las infra-representadas. PB: Proceso Biológico, FM: Función Molecular; CC: Componente Celular, FDR: *False Discovery Ratio*, DEG: *Differentially Expressed Gene*, G Ref.: Genes de la referencia. Los ficheros completos de anotación funcional se encuentran disponibles en el laboratorio del IATA.

Tratamiento	Expresión	Nivel	GO ID	Nombre GO	Tipo GO	FDR	Nº DEG Test	Nº G Ref.
05	Reprimidos	7	GO:0006572	Proceso catabólico de la tirosina	PB	2.3E-02	3	0
		7	GO:0043648	Proceso metabólico del ácido dicarboxílico	PB	3.2E-02	14	90
		7	GO:0006094	Gluconeogénesis	PB	3.1E-02	10	47
		7	GO:0009071	Proceso catabólico de aminoácidos de la familia serina	PB	4.5E-02	3	1
		7	GO:0046496	Proceso metabólico del nucleótido nicotinamida	PB	3.1E-02	14	88
		7	GO:0008106	Actividad alcohol deshidrogenasa (NADP +)	FM	2.5E-02	4	3
06	Inducidos	8	GO:0006090	Proceso metabólico de piruvato	PB	6.6E-03	14	70
		7	GO:0034470	Procesamiento de ncRNA	PB	2.5E-02	0	231
	Reprimidos	7	GO:0005506	Unión a iones hierro	FM	3.7E-05	18	98
		7	GO:0000460	Maduración de 5.8S rRNA	PB	3.6E-02	10	23
		8	GO:0016072	Proceso metabólico de rRNA	PB	1.0E-02	29	136
		8	GO:0000466	Maduración de 5.8S rRNA del rRNA tricistrónico	PB	1.6E-02	9	15

**Tabla 5. Enriquecimiento funcional de por ontología génica (GO) de DEGs en tratamientos 08 y 11.**

Se muestran las funciones enriquecidas de los tratamientos 08 y 11 en los niveles GO 7 y 8, los cuales fueron los más altos que dieron valores significativos (FDR<0,01) en estos tratamientos. En rojo las categorías sobre-representadas y en verde las infra-representadas PB: Proceso Biológico, FM: Función Molecular; CC: Componente celular, FDR: *False Discovery Ratio*, DEG: *Differentially Expressed Gene*, G Ref.: Genes de la referencia Los ficheros completos de anotación funcional se encuentran disponibles en el laboratorio del IATA.

Tratamiento	Expresión	Nivel	GO ID	Nombre GO	Tipo GO	FDR	Nº DEG Test	Nº G Ref.
08	Reprimidos	9	GO:1904659	Transporte transmembrana de glucosa	PB	9.1E-08	20	28
		10	GO:0046323	Importe de glucosa	PB	2.5E-05	15	24
11	Inducidos	9	GO:0009206	Proceso biosintético de ribonucleósido triP de purina	PB	1.1E-03	24	49
		9	GO:0009168	Proceso biosintético de ribonucleósido monoP de purina	PB	4.4E-06	33	58
	Reprimidos	9	GO:0009168	Proceso biosintético de ribonucleósido monoP de purina	PB	8.2E-03	8	83
		9	GO:0016567	Ubiquitinación de proteínas	PB	3.6E-02	6	62
		9	GO:0000377	Procesamiento de RNA, reacción de transesterificación	PB	6.4E-05	4	85
		9	GO:1904659	Transporte transmembrana de glucosa	PB	2.7E-06	29	19
		10	GO:0046323	Importe de glucosa	PB	3.2E-04	22	17
		10	GO:0008094	Actividad ATPasa dependiente de DNA	FM	2.0E-02	1	35

#### 4.6. Selección de genes

Se han seleccionado algunos genes relevantes por su anotación funcional y patrones significativos de represión o inducción (Tabla 6). Los genes se han filtrado según el parámetro de cambio con respecto a la condición control ( $\log_2\text{FoldChange}$ ), según las lecturas por millón de cada gen ( $\log_2\text{CountsPerMillion} > 3$ ) y el parámetro de significancia (FDR o *False Discovery Ratio*). En los genes seleccionados aparecen proteínas hipotéticas (*hypothetical protein*) que suponen un 50% de los genes seleccionados. También se han obtenido genes relacionados con la conidiogénesis (PDIG\_59420), y genes involucrados en la síntesis de la pared celular (PDIG\_37510). También se encuentra entre los genes

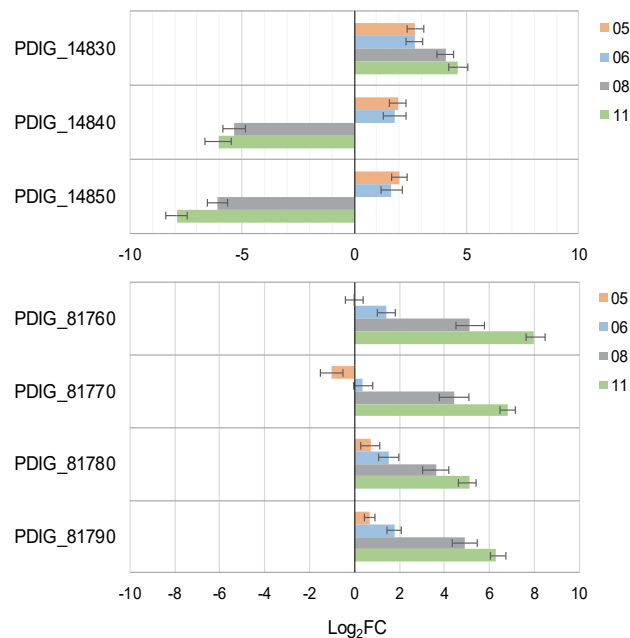
seleccionados la proteína antifúngica AfpB, sobreexpresada por la cepa superproductora PDSG420 (tratamiento 08).

**Tabla 6. Selección de genes expresados diferencialmente obtenidos con EdgeR.** Se muestra una selección de DEG obtenidos con EdgeR que incluye a los cinco genes más inducidos y los cinco más reprimidos en cada uno de los cuatro tratamientos. Los valores de abundancia de expresión (escala logarítmica en base 2 de las lecturas por millón, log2CPM) se filtraron para los genes superiores a 3. Los valores de cambio de expresión (log2FC) varían su color desde inducido (verde) a reprimido (rojo). El parámetro de significancia (FDR: *False Discovery Ratio*) se marca en color cuando es significativo (FDR < 0,01) en cada tratamiento.

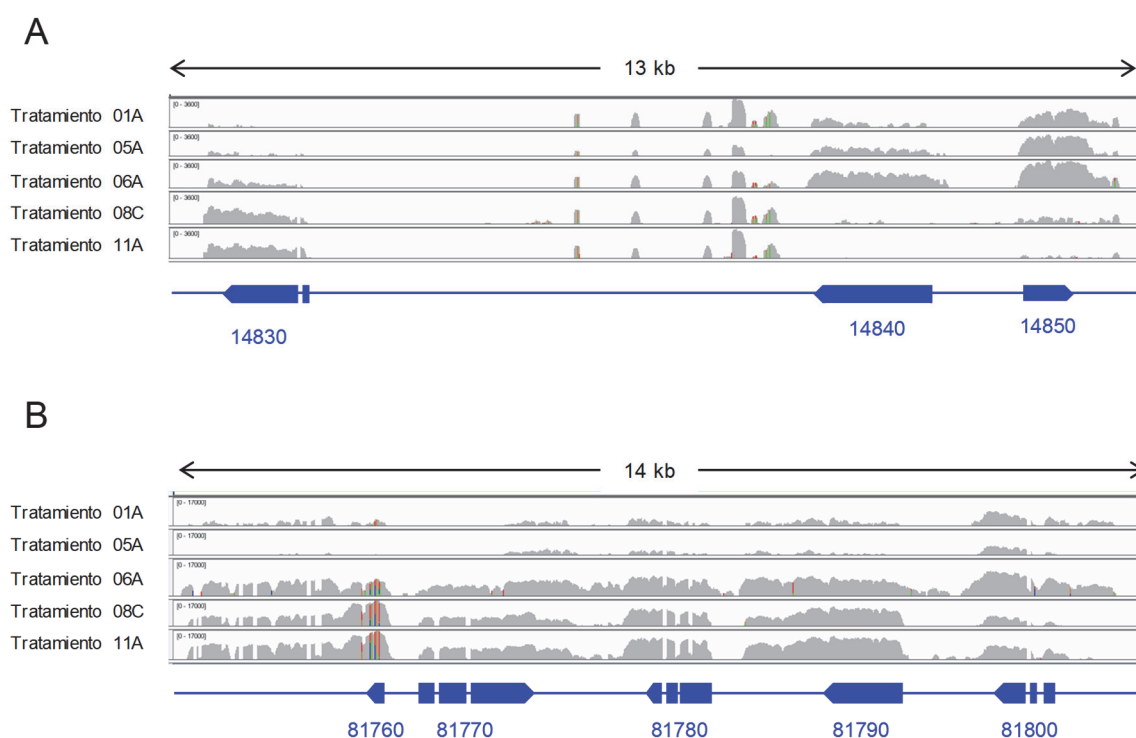
Gene ID	Description Ensembl	1vs5			1vs6			1vs8			1vs11		
		logFC	logCPM	FDR	logFC	logCPM	FDR	logFC	logCPM	FDR	logFC	logCPM	FDR
PDIG_01560	Hypothetical protein	2.58	3.76	7.7E-05	3.31	4.42	2.8E-03	-6.19	0.89	1.5E-09	-6.49	0.87	1.2E-12
PDIG_09380	Alternative oxidase	-4.79	7.74	3.9E-07	-1.91	8.04	8.7E-02	0.61	8.87	6.6E-01	0.35	8.73	7.1E-01
PDIG_10270	2,4-dienoyl-CoA reductase	-4.62	7.26	3.4E-06	-3.15	7.37	1.0E-03	0.51	8.33	7.7E-01	1.47	8.98	8.8E-02
PDIG_14170	MFS monocarboxylate transporter, putative	1.19	5.36	6.4E-01	1.26	5.42	8.8E-02	-4.64	3.55	2.8E-08	-8.27	3.51	1.7E-20
PDIG_14830	Hypothetical protein	3.07	2.81	9.9E-07	3.03	2.81	1.2E-08	4.42	3.90	3.5E-16	4.78	4.28	6.5E-22
PDIG_14840	Acetolactate synthase	2.23	5.96	1.2E-05	2.39	6.12	4.7E-04	-6.54	3.33	1.6E-19	-6.56	3.36	1.7E-22
PDIG_14850	Alpha-acetolactate decarboxylase, putative	2.26	8.30	2.0E-04	2.12	8.21	2.2E-03	-6.98	5.63	1.0E-34	-8.31	5.67	1.4E-57
PDIG_25880	Hypothetical protein	-4.78	6.91	5.8E-07	-4.85	6.92	1.7E-07	-1.58	7.12	8.5E-02	-2.76	6.92	2.8E-04
PDIG_33070	Benzoate 4-monooxygenase cyt. P450	1.92	7.50	1.0E-03	2.62	8.10	7.1E-03	-5.36	5.12	4.9E-29	-7.01	5.13	3.3E-51
PDIG_34470	GABA permease, putative	2.08	3.43	9.9E-03	3.45	4.64	8.1E-15	-1.14	1.48	2.8E-01	-1.52	1.39	2.3E-02
PDIG_35740	Steroid monooxygenase (CpmA), putative	2.42	7.23	1.2E-03	2.21	7.08	1.8E-03	-6.61	4.44	7.2E-19	-7.73	4.47	4.1E-25
PDIG_36860	NRPS-like enzyme, putative	2.15	9.71	6.2E-03	2.09	9.69	1.0E-02	-4.82	7.17	2.8E-13	-6.02	7.19	8.1E-23
PDIG_36880	Hypothetical protein	2.39	8.19	7.5E-04	2.23	8.08	3.0E-03	-5.52	5.43	1.0E-24	-5.69	5.47	1.0E-28
PDIG_37510	Endo-1,3(4)-beta-glucanase, putative	-0.62	6.61	7.5E-01	-0.42	6.72	7.4E-01	-1.80	6.11	1.1E-02	-0.31	6.65	6.9E-01
PDIG_37530	Hypothetical protein	2.85	9.49	2.7E-05	3.34	9.95	1.7E-07	-5.32	6.34	3.6E-29	-6.20	6.36	2.1E-61
PDIG_41710	Hypothetical protein	0.44	5.48	9.1E-01	0.83	5.73	9.1E-02	-4.06	4.18	2.8E-10	-8.90	4.12	1.8E-37
PDIG_42500	Hypothetical protein	3.18	5.98	1.1E-07	-0.76	3.34	3.4E-01	-0.37	3.34	7.8E-01	-2.68	2.76	3.0E-07
PDIG_45940	Hypothetical protein	2.90	4.04	3.7E-06	3.56	4.66	7.0E-06	-1.62	1.27	8.1E-03	-2.16	1.16	6.4E-06
PDIG_46550	Formate dehydrogenase	-1.71	4.49	2.3E-03	-1.32	4.61	1.0E-02	3.92	7.96	8.5E-10	5.19	9.21	1.0E-25
PDIG_46910	ERG3_Sterol delta 5,6-desaturase	-2.63	9.21	8.1E-03	-2.86	9.19	5.3E-04	-0.18	9.75	9.4E-01	-0.10	9.81	9.5E-01
PDIG_47930	SNF2 family helicase, putative	-4.33	3.25	1.5E-05	-4.91	3.23	7.2E-13	-4.06	3.12	5.1E-08	-4.81	3.10	4.1E-14
PDIG_48140	S-adenosyl-Met-sterol-C-methyltransferase	-2.13	1.98	2.8E-02	-2.03	2.01	1.6E-02	0.67	2.90	6.0E-01	1.49	3.48	5.1E-02
PDIG_49600	MFS peptide transporter, putative	1.94	8.75	3.2E-03	1.83	8.69	1.9E-04	1.10	7.98	4.8E-02	0.62	7.71	1.5E-01
PDIG_50010	Hypothetical protein	-4.67	8.19	9.9E-07	-2.66	8.37	1.1E-02	-0.23	8.87	9.2E-01	-0.65	8.72	4.5E-01
PDIG_50670	Endopolygalacturonase 1	-0.78	10.37	7.5E-01	-0.72	10.41	6.4E-01	-7.27	9.56	8.1E-13	-7.74	9.58	2.8E-14
PDIG_50820	Hypothetical protein	2.19	9.41	1.8E-04	2.75	9.91	1.1E-04	-7.16	6.80	4.5E-38	-8.76	6.83	9.8E-67
PDIG_52590	Hypothetical protein	2.64	5.72	9.9E-07	2.08	5.28	2.1E-04	1.00	4.31	1.3E-01	-2.15	3.06	9.4E-05
PDIG_59420	Conidiation-specific protein (Con-10), putative	-4.23	5.72	6.2E-15	-4.46	5.73	1.3E-12	-2.06	5.80	5.6E-05	-2.42	5.77	4.8E-04
PDIG_65770	Hypothetical protein	-3.62	4.04	2.4E-03	-6.22	3.95	1.2E-09	-1.92	4.11	9.2E-02	-1.82	4.15	1.3E-02
PDIG_68840	Antifungal protein AfpB	0.91	2.56	3.0E-01	0.80	2.52	2.0E-01	9.19	10.05	6.1E-39	-2.95	1.11	8.7E-09
PDIG_70130	Hypothetical protein	2.16	6.16	1.4E-02	3.84	7.67	3.6E-12	1.55	5.54	6.5E-03	-0.09	4.56	9.9E-01
PDIG_71910	Sodium/phosphate symporter, putative	2.63	3.44	2.8E-04	1.51	2.58	1.1E-02	0.58	1.79	5.0E-01	0.24	1.64	6.9E-01
PDIG_76310	Hypothetical protein	-3.49	5.50	1.1E-07	-4.52	5.45	5.5E-13	-1.07	5.78	1.7E-01	-1.20	5.76	2.1E-02
PDIG_81760	Hypothetical protein	-0.01	1.35	1.0E+00	1.63	2.40	6.2E-03	8.70	8.86	2.5E-12	8.22	8.41	3.4E-48
PDIG_81770	Putative isoaspartyl peptidase/L-asparaginase	-1.33	-0.29	2.8E-01	0.47	0.48	6.6E-01	7.58	6.52	4.0E-10	6.92	5.90	6.4E-69
PDIG_81780	Hypothetical protein	0.88	4.11	4.3E-01	1.87	4.86	1.0E-03	4.90	7.40	6.6E-08	5.34	7.88	1.3E-19
PDIG_81790	Hypothetical protein	0.72	4.03	3.7E-01	1.95	4.94	1.2E-06	6.35	8.84	4.0E-13	6.32	8.85	3.8E-42
PDIG_89380	Hypothetical protein	-6.39	10.49	2.5E-05	-3.26	10.63	3.1E-02	-0.17	11.23	9.7E-01	0.00	11.34	1.0E+00
PDIG_90910	Hypothetical protein	-2.34	-0.46	6.0E-02	-0.95	-0.11	4.0E-01	7.39	6.38	9.5E-07	4.77	3.85	4.4E-17
PDIG_90930	Hypothetical protein	-1.07	-0.83	6.8E-01	0.03	-0.39	1.0E+00	8.10	6.33	5.3E-08	5.31	3.61	3.4E-21

Hay dos grupos de genes DEG que destacan porque incluyen genes que son continuos en el genoma y co-expresan en los distintos tratamientos según su patrón de expresión y el valor de cambio con respecto de la condición control (log2FC) (Tabla 6). El grupo de genes o cluster 1 está formado por PDIG\_14830, PDIG\_14840, PDIG\_14850. Como se puede observar en los valores de log2FC, en los tratamientos 05 y 06 los tres genes están altamente inducidos sobreexpresándose con respecto a la condición control (tratamiento 01) entre más de cuatro a seis veces. En cambio, en los tratamientos 08 y 11 el patrón cambia permaneciendo inducido el gen PDIG\_14830 y siendo altamente reprimidos PDIG\_14840 y PDIG\_14850. Alcanzando unos niveles de represión que

superan en doce veces a la condición control (PDIG\_14840), e incluso en dieciséis veces (PDIG\_14850) en el tratamiento 11 (Fig. 9A). Destaca otro grupo de DEGs por su elevada sobreexpresión en tres de los cuatro tratamientos estudiados y su cercanía en el genoma. Este grupo o cluster 2 está formado por PDIG\_81760, PDIG\_81770, PDIG\_81780, PDIG\_81790. En los tratamientos 08 y 11, los genes tienden a aumentar significativamente su expresión hasta llegar a superar entre ocho a dieciséis veces la expresión de la condición control (Fig 9B). Estos patrones de expresión son visibles cuando se representan las lecturas de cada réplica sobre el genoma de referencia. El visualizador IGV permite observar cómo los genes con más lecturas forman un pico en la región genómica sobre la que mapean, como se puede observar al realizar el análisis con los clusters de genes comentados anteriormente (Fig 10A y 10B).



**Figura 9. Gráficas de expresión diferencial de genes agrupados en los clusters seleccionados.** La expresión diferencial de los genes ( $\log_2FC$ ) se obtuvo por DESeq2



**Figura 10. Visualización con IGV 2.4 de las lecturas que mapean sobre el genoma de referencia en cada tratamiento.** A partir de los ficheros del mapeo de los transcritos (BAM; *Binarily compressed sequence Alignment/Map*), genoma de referencia (FASTA) y ficheros de anotación (GTF: *Gene Transfer Format*), se puede observar la expresión en los diferentes tratamientos de los genes de cada cluster en escala logarítmica; log3600 para el primer cluster (A) y log1700 para el segundo cluster (B).

## 5. DISCUSIÓN

Las proteínas AFP tienen un gran interés biotecnológico en el ámbito de desarrollo de nuevos compuestos antifúngicos, tanto para ser utilizadas por sí mismas como porque la caracterización de su mecanismo de acción podría identificar nuevas dianas celulares para el desarrollo de antifúngicos. Por ello, el modo de acción de las proteínas AFPs y las funciones biológicas que tienen para el hongo productor están siendo objeto de estudio actualmente. Aunque se ha avanzado en la identificación de nuevas AFPs en distintos hongos, sus efectos en hongos sensibles y sus mecanismos de acción todavía no se conocen lo suficiente ya que los estudios son escasos y restringidos a unas pocas AFPs (Meyer & Jung, 2018). Estos estudios apuntan a que algunas de estas proteínas están involucradas en la autofagia, movilización de nutrientes y una respuesta temprana a la falta de los mismos (Paege et al., 2016), mientras que otras podrían afectar al desarrollo del hongo durante la diferenciación asexual o conidiación (Hegedüs et al., 2011).

Los análisis transcriptómicos basados en la secuenciación masiva de RNA generan datos a gran escala con una alta reproducibilidad, lo que permite estudiar el dinamismo del transcriptoma. Esta metodología tiene una cobertura del genoma mayor que otras anteriores como la secuenciación por Sanger o métodos basados en microarrays, mejorando la cuantificación de los genes que se expresan de forma diferencial y otras aplicaciones como la identificación de procesamiento alternativo o de nuevos transcritos.

Se ha seleccionado esta aproximación transcriptómica para este estudio, cuyo objetivo es determinar el modo de acción de la proteína antifúngica AfpB sobre los hongos filamentosos, así como la función biológica para el organismo productor.

Un primer resultado sorprendente de nuestro estudio fue que en las condiciones de cultivo necesarias para obtener cantidad suficiente de biomasa (cultivo en matraz y agitación), la proteína AfpB fue mucho más activa (64 veces) que lo que se había descrito (cultivo estático en placas de microtítulo), determinándose un MIC de 0,0625 µg/mL (aproximadamente 10 nM) frente a los 4 µg/mL conocidos anteriormente. En el futuro se intentará estudiar las causas de este comportamiento. En cualquier caso, esto hace que los resultados y conclusiones extraídos de nuestro estudio no sean necesariamente extrapolables a todas las condiciones de cultivo, y por ello deberán ser confirmados en todas ellas.

A partir del análisis transcriptómico se han obtenido un elevado número de DEGs que se asocian a la presencia de AfpB, cuyo número varía entre tratamientos, predominando los genes reprimidos. Son las condiciones de cultivo con presencia constante de AfpB (tratamiento 11 y 08) las que presentan un mayor número de DEGs en comparación con los tratamientos a tiempos cortos (tratamiento 05 y 06). Además, el análisis de PCA, mapa de calor y diagrama de Venn (Fig. 5 y Fig. 6) nos indican que se agrupan por un lado los tratamientos a tiempos cortos 05 y 06, y por otro lado los tratamientos a tiempos largos 08 y 11. Esto indica que el efecto de tratamientos cortos es cuantitativamente menor (como era de esperar) y que el crecimiento continuo en presencia de proteína (tratamiento 11) se asemeja al efecto observado en la cepa sobreproductora de proteína.

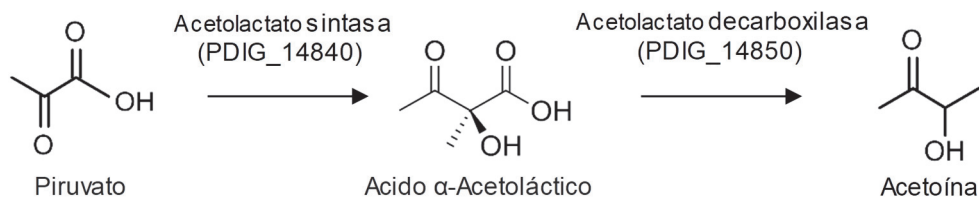
Como siguiente paso del análisis de los DEGs, llevamos a cabo el enriquecimiento funcional. Pero nuestros resultados no eran informativos, ya que las bases de datos eran pobres en cuanto a la anotación funcional con un 45.5% y 37.5% de genes sin anotar en el caso de JGI y UniProt respectivamente. Por lo tanto, se mejoró la anotación funcional hasta obtener un 87,7% de los genes anotados (12,2% sin anotar), y aumentando también el porcentaje de genes con las tres categorías GO (Fig. 8). Con ello se ha conseguido el enriquecimiento funcional de los DEGs en cada tratamiento, predominando en los genes reprimidos funciones relacionadas con el transporte de glucosa y el metabolismo primario (piruvato, nucleótidos o ácidos dicarboxílicos). Esto podría asociarse a que la AfpB limita la toma de nutrientes y el metabolismo celular, afectando negativamente al crecimiento del hongo, aunque deberemos determinar si esto es una causa o una consecuencia del efecto nocivo de la proteína.

La Tabla 6 es una selección arbitraria de los DEG con mayor inducción o represión, y con muy alta significación estadística, en cada uno de los tratamientos. Este listado es en cualquier caso, un buen ejemplo de la alta potencialidad de los resultados obtenidos en este estudio. Un 50% (20 de 40) de los DEGs enumerados en la Tabla 6 tienen función desconocida, entre ellos y de forma muy especial PDIG\_81760 (ver más abajo). También destacan proteínas relacionadas con la conidiogénesis (PDIG\_59420) y que aparecen reprimidas, indicando que la proteína afecta a este proceso esencial del hongo. Otros genes relacionados con conidiogénesis también se han identificado como reprimidos en nuestro estudio, aunque con menor cambio de expresión (datos no mostrados). Esto es interesante porque el mutante nulo *paf* de *P. chrysogenum* tiene afectada la esporulación y por ello se indicó que las proteínas AFP pueden estar relacionadas con la conidiogénesis

(Hegedüs et al., 2011). Sin embargo, el mutante nulo *afpB* de *P. digitatum* esporula con normalidad y su número de conidios no difiere de la cepa parental, por tanto se descartó una implicación de AfpB en la conidiogénesis (Garrigues et al., 2016). Los resultados de nuestro estudio abren de nuevo esta cuestión, indicando que AfpB sí puede de alguna forma estar relacionada con la conidiogénesis, aunque su mutante no esté afectado. Está por comprobar hasta qué punto el efecto que observamos es primario y directamente relacionado con la función de la proteína o más bien secundario y motivado por su efecto deletéreo sobre el hongo.

Las proteínas AFP se ha descrito que interaccionan primero con la pared celular del hongo al que inhiben y luego se internalizan al interior de la célula (Meyer & Jung, 2018). Nuestros datos indican que la AfpB afecta a la síntesis de la pared celular, reprimiendo genes relacionados con la síntesis de glucanos (PDIG\_37510). Como componente de pared celular y primera barrera de defensa, su represión hace más susceptible al hongo frente a la proteína. El mecanismo de acción de AfpB podría estar relacionado con la inhibición de la síntesis de glucanos, que formando parte de la pared son una diana común para antifúngicos y otras proteínas, como la AFP de *A. giganteus* cuyo modo de acción es la inhibición de la quitina, componente importante de la pared celular de hongos. Otros genes destacados son los que codifican proteínas transportadoras de membrana de la familia MFS (Major Facilitator Superfamily), que aparecen en nuestro estudio con distintos patrones de expresión estando dos de ellos entre los que más cambian su expresión (Tabla 6). El MFS transportador de péptidos (PDIG\_49600) podría estar implicado en la internalización de AfpB.

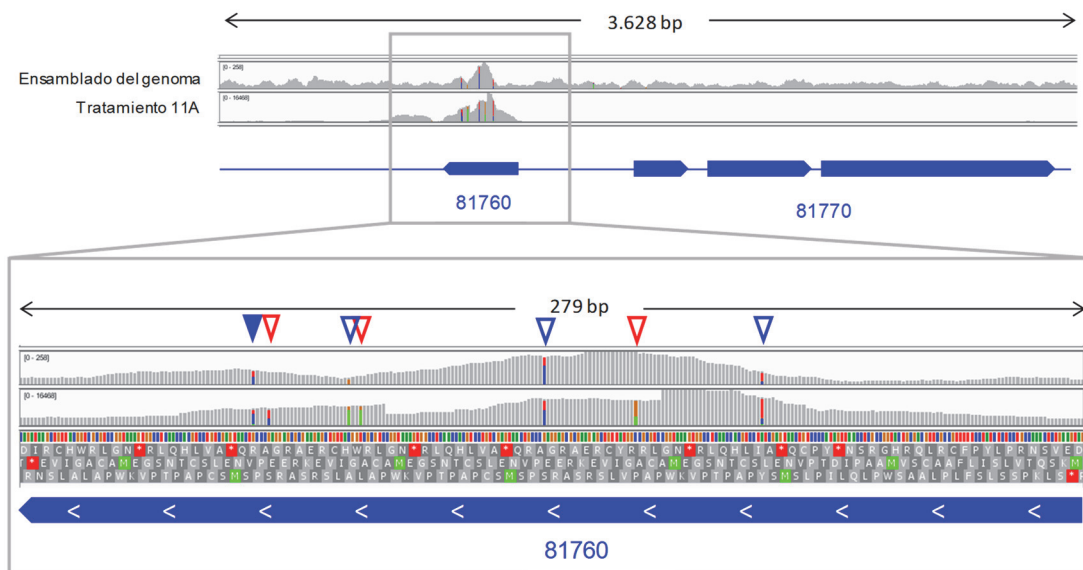
Son relevantes los genes PDI\_14840 y PDI\_14850 que forman parte de la ruta biosintética de la acetoina (Fig. 11), aunque se desconoce la función del gen contiguo PDI\_14830. A tiempos de exposición cortos de AfpB (tratamientos 05 y 06), los genes se encuentran sobreexpresados, pero este patrón cambia, y se reprimen en la cepa PDSG420 (tratamiento 08) y también cuando la exposición a AfpB se extiende a tiempos largos (tratamiento 11). Una posible explicación a estos patrones de expresión podría ser que la proteína AfpB, como proteína producida por hongos, induce la sobreexpresión de esta ruta por causas que nuestro estudio no aclara. Bacterias del género *Bacillus* son capaces de producir compuestos antifúngicos, y en concreto los compuestos volátiles (VOCs) han sido objeto de estudio como potenciales moléculas antifúngicas (Leyva Salas et al., 2017) e incluso como posible herramienta de biocontrol de enfermedades fúngicas postcosecha. La acetoina es uno de estos compuestos volátiles que presentan esta actividad. Se ha observado frente al patógeno postcosecha de la cereza (Gotor-Vila et al., 2017), y aunque en otros estudios la actividad antifúngica se ha relacionado con lipopéptidos, la acetoina también es el principal VOC sintetizado por este tipo de bacterias que inhiben el crecimiento del hongo (Lim et al., 2017). Además el efecto de la acetoina ya se ha descrito en los hongos patógenos de frutos cítricos del género *Penicillium* incluido *P. digitatum* (Arrebola et al., 2010). Su actividad antifúngica podría explicar el cambio de patrón de expresión que se observa a tiempos largos de exposición a AfpB. Su acumulación durante tiempos más cortos de exposición, afectan al crecimiento del propio hongo que responde reprimiendo los genes de síntesis del compuesto volátil. Por ello, como siguiente paso para el estudio del efecto de la acetoina en el crecimiento del hongo, proponemos la eliminación del gen PDI\_14850 que codifica para la última enzima de la ruta biosintética de compuesto.



**Figura 11.** Ruta de biosíntesis de acetoina.

El gen PDIG\_81760 es un gen que nos ha llamado la atención por varios motivos, es el gen más sobreexpresado de todo el genoma en los tratamientos 08 y 11. La función de la proteína PDIG\_81760 es desconocida. Cuando se visualizó la expresión de los genes del segundo cluster en los diferentes tratamientos (Fig. 10B), resultó muy llamativa la concentración de variabilidad nucleotídica en el RNA-seq del gen PDIG\_81760. Centrándonos más en esta región, se observaron hasta siete puntos variables en las lecturas correspondientes a la inducción de expresión en los tratamientos (ejemplificado con flechas azules y rojas en el RNA-seq del tratamiento 11A, Fig. 12). La presencia de esta variabilidad nos hizo plantearnos que era posible que el gen se encuentre duplicado en el genoma, no detectándose correctamente como copias de un mismo gen cuando se llevó a cabo el ensamblado inicial (Marcet-Houben et al., 2012). Si hubiera un error de ensamblado en esta región, se observaría un aumento de cobertura en el ensamblado inicial del genoma, debido a que se habrían mapeado dos o tres genes en la misma región. Para corroborarlo hemos visualizado el ensamblado del genoma en la región correspondiente a PDIG\_81760 (Fig. 12). Sorprendentemente se observó que existía un aumento de cobertura llamativo con respecto al resto del genoma, que permanece con una cobertura homogénea. Las proporciones de las variantes nucleotídicas (1:2) podrían apoyar la hipótesis de una doble duplicación que daría lugar a tres copias en el genoma, donde dos copias presentarían la variante mayoritaria y la otra copia la variante minoritaria. La presencia de tres copias en el genoma del gen PDIG\_81760 podría contribuir a su elevada expresión, ya que cada copia individualmente estaría contribuyendo con efecto aditivo a la expresión final del gen que ha sido detectada. Además, en el ensamblado del genoma había menor variabilidad nucleotídica (solo cuatro residuos, flechas azules en Fig. 12) que la observada en el transcriptoma (siete residuos). Esta variabilidad no se puede explicar por duplicación, debido a que no se encuentra en el ensamblado del genoma, y sí podría ser debida a un fenómeno de editado de RNA descrito en eucariotas superiores y también en hongos (Eisenberg & Levanon, 2018). Evaluando el efecto de los cambios nucleotídicos a nivel aminoacídico, solo uno de ellos da lugar a un cambio no sinónimo (flecha azul rellena, Fig. 12) de valina a isoleucina, por lo que desconocemos el motivo por el cual ocurre este posible editado en el resto de puntos variables. Tanto la hipótesis de la duplicación como la del editado de RNA son posibles y no excluyentes.





**Figura 12. Gen PDIG\_81760 en el visualizador IGV.** Se encuentra representado el ensamblado del genoma frente al transcriptoma obtenido del tratamiento 11A. Las rayas verticales son los nucleótidos del gen, y los coloreados representan un cambio con respecto a la referencia. También se encuentran las tres pautas de lectura que dan lugar a la secuencia aminoacídica, donde la segunda es la que da lugar a la proteína codificada por el gen (en verde el codón de inicio y en rojo el codón de stop).

Hay otro aspecto extraordinariamente llamativo en PDIG\_81760. La proteína codificada por el gen PDIG\_81760 presenta un péptido señal para su secreción, es de pequeño tamaño (91 aminoácidos) y aniónica, con un alto contenido en cisteínas. Observando la secuencia aminoacídica (Fig. 13), destacan repeticiones en tandem que conforman la proteína y destaca la presencia del motivo-Y en cada una de estas repeticiones de secuencia. Además, proteínas homólogas a la codificada por el gen PDIG\_81760, se encuentran presentes en los genomas de hongos del género *Penicillium* en los cuales se mantienen todas las características nombradas anteriormente y se desconoce su función.

```

>tr|K9F9Z8|K9F9Z8_PEND2 Uncharacterized protein
OS=Penicillium digitatum (strain PHI26 / CECT 20796)
GN=PDIG_81760

      10      20      30      40      50
MKSQTVLSIL FAACSVMAAP IDTPVNELSC TNSGEMACAG IVEKREEPVN
      60      70      80      90
ELSC TNSGEM ACAGIVEKRE EPVNELSCTN SGEMACAGIV E

SP (Secretion)
MKSQTVLSILFAACSVMAAPIDTPVNELSCTNSGEMACAGIVEKREE
                        PVNELSCTNSGEMACAGIVEKREE
                        PVNELSCTNSGEMACAGIVE

```

**Figura 13. Secuencia aminoacídica de la proteína codificada por PDIGG\_81760.** En verde se destaca el péptido de secreción. Encuadrado en rojo se encuentra el motivo gamma en las repeticiones en tandem que conforman la proteína.

Debido a que este gen se sobreexpresa a tiempos largos en presencia de AfpB y debido a las propiedades de la proteína codificada (pequeña proteína secretada y aniónica), planteamos la hipótesis de que se trata de una posible proteína de inmunidad que se secretaría para contrarrestar el efecto de la AfpB. Es posible que el hongo, ya que es sensible a la propia proteína antifúngica que produce, tenga un sistema para poder sobrevivir en presencia de la misma. Como primeros pasos para abordar el estudio del posible 'antídoto' frente a AfpB, se plantea la eliminación del gen que codifica para la proteína. Sin embargo, hay que tener en cuenta que si el gen se encuentra duplicado podría no observarse el efecto tras su eliminación.

## 6. CONCLUSIONES

1. Cuando se realizan los ensayos de crecimiento en matraz y agitación, la proteína antifúngica AfpB es mucho más activa frente a *P. digitatum* que lo que se había descrito, llegando hasta un MIC de 10 nM, 64 veces inferior al observado en placas de microtítulo.
2. A través del RNA-seq se ha observado un gran efecto de AfpB en el transcriptoma de *P. digitatum*, obteniéndose desde 542 hasta 4.143 genes expresados diferencialmente del total de 9135, dependiendo del tratamiento.
3. El efecto inhibitorio sobre el crecimiento y el número de genes expresados diferencialmente es más elevado en el hongo crecido continuamente en presencia de AfpB o en la cepa sobreproductora de AfpB, en comparación con los tratamientos puntuales.
4. Se ha mejorado la anotación funcional del genoma de *P. digitatum* CECT 20796 respecto a la disponible en las bases de datos públicas como JGI y Uniprot, pasando de un 54,5% o 62,4% respectivamente a un 87,7% de genes con anotación GO, y aumentando el número y profundidad de términos GO.
5. El efecto mayoritario de AfpB sobre el transcriptoma de *P. digitatum* es la represión génica, sobre todo de genes relacionados con el transporte de azúcares y el metabolismo primario, evidenciando el efecto negativo sobre el crecimiento del hongo.
6. Se han identificado una serie de genes que podrían estar más específicamente relacionados con el modo de acción y el efecto de la proteína AfpB sobre *P. digitatum*. Como ejemplo, los genes PDIG\_14840 y PDIG\_14850 que codifican los últimos pasos en la biosíntesis de acetoina; o PDIG\_81760 que codifica una pequeña proteína secretable, aniónica y rica en cisteínas, que podría antagonizar la acción de las AFPs catiónicas

## 7. BIBLIOGRAFÍA

- Arrebola, E., Sivakumar, D., & Korsten, L. (2010). Effect of volatile compounds produced by *Bacillus* strains on postharvest decay in citrus. *Biological Control*, 53(1), 122–128. <https://doi.org/10.1016/j.biocontrol.2009.11.010>
- Batta, G., Barna, T., Gáspári, Z., Sándor, S., Kövér, K. E., Binder, U., ... Marx, F. (2009). Functional aspects of the solution structure and dynamics of PAF - A highly-stable antifungal protein from *Penicillium chrysogenum*. *FEBS Journal*, 276(10), 2875–2890. <https://doi.org/10.1111/j.1742-4658.2009.07011.x>
- Beatriz, A., & Álvaro, M. (2006). Biotechnologically relevant enzymes and proteins Antifungal mechanism of the *Aspergillus giganteus* AFP against the rice blast fungus *Magnaporthe grisea*, 883–895. <https://doi.org/10.1007/s00253-006-0362-1>
- Brown, G. D., Denning, D. W., Gow, N. A. R., Levitz, S. M., Netea, M. G., & White, T. C. (2012). Hidden killers: Human fungal infections. *Science Translational Medicine*, 4(165). <https://doi.org/10.1126/scitranslmed.3004404>
- Campoy, S. and Adrio, J. L. (2017). Antifungals. *Biochem Pharmacol*, 133, 86-96.
- Conesa, A., Madrigal, P., Tarazona, S., Gomez-Cabrero, D., Cervera, A., McPherson, A., ... Mortazavi, A. (2016). A survey of best practices for RNA-seq data analysis. *Genome Biology*, 17(1), 1–19. <https://doi.org/10.1186/s13059-016-0881-8>
- Conesa, A., & Stefan, G. (2008). Blast2GO : A Comprehensive Suite for Functional Analysis in, 2008. <https://doi.org/10.1155/2008/619832>
- Denning, D. W., Park, S., Lass-Flörl, C., Fraczek, M. G., Kirwan, M., Gore, R., ... Perlin, D. S. (2011). High-frequency triazole resistance found in nonculturable *aspergillus fumigatus* from lungs of patients with chronic fungal disease. *Clinical Infectious Diseases*, 52(9), 1123–1129. <https://doi.org/10.1093/cid/cir179>
- Eisenberg, E., & Levanon, E. Y. (2018). A-to-I RNA editing — immune protector and transcriptome diversifier. *Nature Reviews Genetics*, 1–18. <https://doi.org/10.1038/s41576-018-0006-1>
- FAO: Prevention of post-harvest food losses: food, vegetables and root crops - a training manual, FAO Training series. vol. 17. Rome: FAO. (1989).
- Fisher, M. C., Henk, D. A., Briggs, C. J., Brownstein, J. S., Madoff, L. C., McCraw, S. L., & Gurr, S. J. (2013). NIH Public Access. *Nature*, 484(7393), 1–18. <https://doi.org/10.1038/nature10947>.Emerging
- Gandía, M., Harries, E., & Marcos, J. F. (2012). Identification and characterization of chitin synthase genes in the postharvest citrus fruit pathogen *Penicillium digitatum*. *Fungal Biology*, 116(6), 654–664. <https://doi.org/10.1016/j.funbio.2012.03.005>
- Gandía, M., Harries, E., & Marcos, J. F. (2014). The myosin motor domain-containing chitin synthase PdChsVII is required for development, cell wall integrity and virulence in the citrus postharvest pathogen *Penicillium digitatum*. *Fungal Genetics and Biology*, 67, 58–70. <https://doi.org/10.1016/j.fgb.2014.04.002>
- Garrigues, S., Gandía, M., Borics, A., Marx, F., Manzanares, P., & Marcos, J. F. (2017). Mapping and identification of antifungal peptides in the putative antifungal protein AfpB from the filamentous fungus *Penicillium digitatum*. *Frontiers in Microbiology*, 8(APR), 1–13. <https://doi.org/10.3389/fmicb.2017.00592>
- Garrigues, S., Gandía, M., & Marcos, J. F. (2016). Occurrence and function of fungal antifungal proteins: a case study of the citrus postharvest pathogen *Penicillium digitatum*. *Applied Microbiology and Biotechnology*, 100(5), 2243–2256. <https://doi.org/10.1007/s00253-015-7110-3>
- Garrigues, S., Gandía, M., Popa, C., Borics, A., Marx, F., Coca, M., ... Manzanares, P. (2017). Efficient production and characterization of the novel and highly active antifungal protein AfpB from *Penicillium digitatum*. *Scientific Reports*, 7(1), 1–13. <https://doi.org/10.1038/s41598-017-15277-w>

- González-Candelas, L., Alamar, S., Sánchez-Torres, P., Zacarías, L., & Marcos, J. F. (2010). A transcriptomic approach highlights induction of secondary metabolism in citrus fruit in response to *Penicillium digitatum* infection. *BMC Plant Biology*, *10*. <https://doi.org/10.1186/1471-2229-10-194>
- Gotor-Vila, A., Teixidó, N., Di Francesco, A., Usall, J., Ugolini, L., Torres, R., & Mari, M. (2017). Antifungal effect of volatile organic compounds produced by *Bacillus amyloliquefaciens* CPA-8 against fruit pathogen decays of cherry. *Food Microbiology*, *64*, 219–225. <https://doi.org/10.1016/j.fm.2017.01.006>
- Hegedüs, N., & Marx, F. (2013). Antifungal proteins: More than antimicrobials? *Fungal Biology Reviews*, *26*(4), 132–145. <https://doi.org/10.1016/j.fbr.2012.07.002>
- Hegedüs, N., Sigl, C., Zadra, I., Pócsi, I., & Marx, F. (2011). The paf gene product modulates asexual development in *Penicillium chrysogenum*. *Journal of Basic Microbiology*, *51*(3), 253–262. <https://doi.org/10.1002/jobm.201000321>
- Hernanz-Koers, M., Gandía, M., Garrigues, S., Manzanares, P., Yenush, L., Orzaez, D., & Marcos, J. F. (2018). FungalBraid: A GoldenBraid-based modular cloning platform for the assembly and exchange of DNA elements tailored to fungal synthetic biology. *Fungal Genetics and Biology*, *116*, 51–61. <https://doi.org/10.1016/j.fgb.2018.04.010>
- Kovács, L., Virágh, M., Takó, M., Papp, T., Vágvölgyi, C., & Galgóczy, L. (2011). Isolation and characterization of *Neosartorya fischeri* antifungal protein (NFAP). *Peptides*, *32*(8), 1724–1731. <https://doi.org/10.1016/j.peptides.2011.06.022>
- Lacadena, J., del Pozo, A. M., Gasset, M., Patiño, B., Campos-Olivas, R., Vázquez, C., ... Gavilanes, J. G. (1995). Characterization of the Antifungal Protein Secreted by the Mould *Aspergillus giganteus*. *Archives of Biochemistry and Biophysics*, *324*(2), 273–281. <https://doi.org/10.1006/abbi.1995.0040>
- Leyva Salas, M., Mounier, J., Valence, F., Coton, M., Thierry, A., & Coton, E. (2017). Antifungal Microbial Agents for Food Biopreservation—A Review. *Microorganisms*, *5*(3), 37. <https://doi.org/10.3390/microorganisms5030037>
- Lim, S. M., Yoon, M. Y., Choi, G. J., Choi, Y. H., Jang, K. S., Shin, T. S., ... Kim, J. C. (2017). Diffusible and volatile antifungal compounds produced by an antagonistic bacillus velezensis G341 against various phytopathogenic fungi. *Plant Pathology Journal*, *33*(5), 488–498. <https://doi.org/10.5423/PPJ.OA.04.2017.0073>
- Marcet-Houben, M., Ballester, A. R., de la Fuente, B., Harries, E., Marcos, J. F., González-Candelas, L., & Gabaldón, T. (2012). Genome sequence of the necrotrophic fungus *Penicillium digitatum*, the main postharvest pathogen of citrus. *BMC Genomics*, *13*(1). <https://doi.org/10.1186/1471-2164-13-646>
- Marcos, J. F., Muñoz, A., Pérez-Payá, E., Misra, S., & López-García, B. (2008). Identification and Rational Design of Novel Antimicrobial Peptides for Plant Protection. *Annual Review of Phytopathology*, *46*(1), 273–301. <https://doi.org/10.1146/annurev.phyto.121307.094843>
- Marti, Á., Lacadena, V., Mancheño, J. M., Olmo, N., Gavilanes, J. G., Lacadena, V., ... Gavilanes, G. (2002). PROTEIN STRUCTURE AND FOLDING : The Antifungal Protein AFP of *Aspergillus giganteus* Is an Oligonucleotide / Oligosaccharide Binding ( OB ) Fold-containing Protein That Produces Condensation of DNA The Antifungal Protein AFP of *Aspergillus giganteus* Is an . <https://doi.org/10.1074/jbc.M207472200>
- Meyer, V. (2008). A small protein that fights fungi: AFP as a new promising antifungal agent of biotechnological value. *Applied Microbiology and Biotechnology*, *78*(1), 17–28. <https://doi.org/10.1007/s00253-007-1291-3>
- Meyer, V., & Jung, S. (2018). Antifungal peptides of the AFP family revisited: Are these cannibal toxins? *Microorganisms*, *6*(2), 50. <https://doi.org/10.3390/microorganisms6020050>
- Paage, N., Jung, S., Schäpe, P., Müller-Hagen, D., Ouedraogo, J. P., Heiderich, C., ... Meyer, V. (2016). A transcriptome meta-Analysis proposes novel biological roles for the antifungal protein anafp in *Aspergillus Niger*. *PLoS ONE*, *11*(11), 1–35. <https://doi.org/10.1371/journal.pone.0165755>

- Palou, L. (2014). *Penicillium digitatum*, *Penicillium italicum* (Green Mold, Blue Mold). In: Postharvest Decay: Control Strategies., *Palou, L.*
- Perfect, J. R. (2016). Expert Opinion on Emerging Drugs “ Is there an emerging need for new antifungals ?,” *8214*(June), 2–5. <https://doi.org/10.1517/14728214.2016.1155554>
- Rusznák, Z., Sz, G., & Rajnavölgyi, É. (2005). The *Penicillium chrysogenum*-derived antifungal peptide shows no toxic effects on mammalian cells in the intended therapeutic concentration, 122–132. <https://doi.org/10.1007/s00210-004-1013-7>
- Sambrook, J. and Russell, D. W. (2001). Molecular cloning. A laboratory manual. *Cold Spring Harbor, New York: Cold Spring Harbor Laboratory Press.*
- Sonderegger, C., Galgoczy, L., Garrigues, S., Fizil, A., Borics, A., Manzanares, P., ... Marx, F. (2016). A *Penicillium chrysogenum*-based expression system for the production of small, cysteine-rich antifungal proteins for structural and functional analyses. *Microbial Cell Factories*, *15*(1), 1–14. <https://doi.org/10.1186/s12934-016-0586-4>
- Sun, X., Ruan, R., Lin, L., Zhu, C., Zhang, T., Wang, M., ... Yu, D. (2013). Genomewide investigation into DNA elements and ABC transporters involved in imazalil resistance in *penicillium digitatum*. *FEMS Microbiology Letters*, *348*(1), 11–18. <https://doi.org/10.1111/1574-6968.12235>
- Virágh, M., Marton, A., Vizler, C., Tóth, L., Vágvölgyi, C., Marx, F., & Galgóczy, L. (2015). Insight into the antifungal mechanism of *Neosartorya fischeri* antifungal protein. *Protein and Cell*, *6*(7), 518–528. <https://doi.org/10.1007/s13238-015-0167-z>
- Yount, N. Y., & Yeaman, M. R. (2004). Multidimensional signatures in antimicrobial peptides. *Proceedings of the National Academy of Sciences*, *101*(19), 7363–7368. <https://doi.org/10.1073/pnas.0401567101>