CENTRO DE INVESTIGACIÓN Y DE ESTUDIOS AVANZADOS DEL INSTITUTO POLITÉCNICO NACIONAL

Unidad Irapuato

Análisis transcripcional de *Trichoderma atroviride* revela cuatro etapas funcionales durante el proceso de fotoconidiación

Tesis que presenta

LCG José Fabricio López Hernández

Para obtener el grado de

Maestro en Ciencias

En especialidad de

Biología Integrativa

Director de Tesis:

Dr. Alfredo Heriberto Herrera Estrella

Irapuato, Guanajuato Agosto, 2015

Acknowledgements

To Consejo Nacional de Ciencia y Tecnología (CONACyT), for the scholarship (Number 356376) I received for two years.

To Cinvestav Unidad Irapuato and LANGEBIO-Unidad de Genómica Avanzada.

To Dr. Alfredo Heriberto Herrera Estrella, for accepting me in his lab once again, for his guidance, discussions, patience but mostly for his trust in my work and vision of science.

To Dr. Robert Winkler, for his mentoring in the first semester of the masters and accepting me in his lab as a rotation student. Also, for all the comments, suggestions and observation made in this project to make it more clear and concise.

To Dr. Luis Rafael Herrera Estrella, for all the comments made with particular shrewdness along this project in order to present this work in a more tangible way.

To the members of Lab 7, for all the great moments we had at the lab as well as outside. Without exception all of them were great colleagues and great friends. In especial, Martha Rendón, Soledad Saburido, Miguel Hernández, Pedro Martínez and Edmundo Santillán, for creating the mezzanine the best place to work in the lab.

To my classmates and friends Beto, Joel, Alex, Carol, Jhonatan and Felipe. For being such incredible persons to talk, enjoy and share with. I am going to miss each and every one of you. Thank you for those outstanding moments in classes as well as in parties.

To the people in Ensenada, Dr. Rosa Mouriño for accepting me in her lab. To Raquel, Maria Jose and Leonora for being good friends and lab partners.

To the people who played every Tuesday rugby, for teaching me to be a best person every week. To the people that were part of the volleyball teams were I played as captain, for trusting me and teaching me to be a better human every day.

To Roberto Álvarez for being always open to talk about science, life, humanity, society and every interesting phenomena in nature.

To Marcos Paolineli for being a good friend in Irapuato as well as in Ensenada, for being always curious and teaching me how to push myself forward.

To people from “Casa del Migrante”, Alejandro Sánchez for accepting me in the house; Miguel Hernández for teaching the value of responsibility, Edgar Balcázar for such good meals and moments and Roberto Bermúdez for being a good friend and a true brother.

To my friends in León that always made me feel at home, Erny, Gavo, Maldonado, Geraz and Tato. To my *Toporos* friends Ricardo, Melvin, Turing, Abiel and Jorge for the support during these years and the good time we could have together.

To Karla for being the motivation and best support I could find through these years… thank you.

To my parents Luis López and Elsa Hernández, they have supported all my decisions, adventures and desires. To my sisters, Luisa and Ana López for all the concerns expressed and help given every time they saw me stuck. To my brother in law Adán Espinoza for being always helpful and attentive.

To my family who have helped me in the most precise manners, with compassion, wishes, kind and encouraging words, but mostly with love.

To all the people that I have met these years and before that are and were part of my life during the masters. To all of you, thank you.

Abbreviations

BAM Binary Alignment Map

BL Blue Light

BLP Blue Light Pulse

BLR Blue Light Receptor (Proteins)

Bp Base pair

bZIP basic Leucine Zipper

CCD Charge Coupled Device

cpm Count Per Million

DEG Differentially Expressed Genes

DEPC Di-Ethyl-Pyro-Carbonate

DGE Differential Gene Expression

DIM-2 DNA Methylation Defective 2. (Protein)

disiRNA Dicer Independent Small Interference RNA

D Diversity

DNA Deoxyribonucleic Acid

DNMT DNA Methyl Transferases (Protein)

FAD Flavin Adenine Mononucleotide

FDR False Discovery Rate

GB Gigabytes

GCF Gene Cluster Finding

GO Gene Ontology

GOEA Gene Ontology Enrichment Analysis

HP1 Herterchromatine Protein 1

IGV Integrative Genome Viewer

Kbp Kilo base pair

LFC Log Fold Change

LOV Light- Oxygen- Voltage- (Domain)

MI Mechanical Injury

PDA Potato Dextrose Agar

RNA Ribonucleic Acid

SAM Sequence Alignment Map

SMRT Single Molecule Real-Time

S Specialization

SNP Single Nucleotide Polimorfism

TF Transcriptional Factor

TIFIIA-γ Transcription Initiation Factor IIA gamma subunit

TIFIIE Transcription Initiation Factor IIE

TIFIID Transcription Initiation Factor IID

WC White Collar (Protein)

5mC 5 Methyl Cytosine

Glossary

Contig: A sequence *contig* is a contiguous, overlapping *sequence* resulting from the reassembly of the small DNA fragments.

General Index

Tables Index

Figure Index

1. Resumen

Los hongos pueden reproducirse de forma sexual y asexual, uno de los mecanismos para reproducirse asexualmente es la conidiación. Este mecanismo juega un papel fundamental en la supervivencia y dispersión de los hongos. En el género *Trichoderma* este fenómeno puede ser desencadenado por diferentes estímulos como la percepción de luz, daño mecánico y falta de nutrientes.

Trichoderma atroviride requiere 24 horas para producir conidios maduros después de un estímulo luminoso. Trabajos previos se han enfocado en las respuestas transcripcionales tempranas justo después del estímulo con luz azul.

En el presente trabajo se analizaron y cuantificaron los estados y cambios transcripcionales globales en puntos considerados críticos para este programa de desarrollo. Las muestras analizadas fueron tomadas 30 minutos, tres, seis, doce y veinticuatro horas después del estímulo por luz azul. Estos tiempos abarcan las primeras respuestas transcripcionales así como diferentes estadíos de desarrollo hasta la formación de conidios maduros.

Se observó un alto nivel de cambios en los patrones transcripcionales durante el desarrollo de los conidios, cerca del 21 % de los genes en el genoma son activados o reprimidos, y que la población de transcritos afectada es diferente entre etapas tempranas y tardías.

Se encontró una represión incrementada de los genes involucrados en la traducción a partir de las seis horas después del estímulo y represión del metabolismo secundario. Finalmente, estos cambios pudieron ser clasificados en cuatro fases transcripcionales asociadas a procesos celulares, i) respuesta a daño de DNA, ii) síntesis de proteínas, iii) metabolismos carboxílico y organonitrogenado y iv) encendido de genes necesarios para la germinación.

A su vez, desarrollé una herramienta bayesiana para encontrar grupos de genes encendidos o apagados en un segmento de DNA. Se encontró un incremento en el número de grupos y de genes asociados a ellos conforme se desarrolla el conidióforo. Esto ocurrió tanto en genes que se encienden como los que se apagan.

Para entender el estado de la cromatina de las células durante la fotoconidiación determinamos los niveles de metilación de DNA, observándose niveles bajos de metilación y un incremento de la conidiación cuando se inhibe la metilación de DNA con 5-Azacitidina. De esto podemos concluir que existe metilación en el *T. atroviride* a bajos niveles y que la falta del mismo tiene consecuencias en la producción de conidios.

Con base en los resultados, se concluyó que existen diferentes poblaciones de transcritos a lo largo proceso de fotoconidiación en *Trichoderma atroviride.* A su vez se mostró que seis horas después del estímulo, es un punto crucial durante el desarrollo. Antes de dicho punto la transcripción está asociada a la respuesta a luz azul y la síntesis de proteínas necesarias para conidiación; después de las seis horas hay una represión de la traducción pero un alto nivel transcripcional asociado a la maduración de los conidios y la germinación.

2. Abstract

Fungi can reproduce both, sexually and asexually. One of the mechanism to reproduce asexually is conidiation. This mechanism plays a fundamental role in their survival and dispersal. In the fungal genus *Trichoderma* this phenomenon can be triggered by different stimuli, such as blue light perception, mechanical injury and nutrient deprivation.

Upon exposure to light, *Trichoderma atroviride* requires 24 hours to produce mature conidia. Previous studies have focused in the early transcriptional responses after the blue light stimulus. In this study I analyzed and quantified the global transcriptional changes at time points considered critical for this developmental program, unto generation of mature conidia.

It was observed a high level of changes in transcriptional patterns during the development of conidia. Nearly 21% of the genes in the genome are activated or repressed, and that the population of transcripts affected differs between the early and late stages.

There was an increasing repression of genes associated with translation that starts six hours after the stimulus and repression of genes associated with secondary metabolism at later stages. These changes in four transcriptional phases associated with cellular processes, i) response to DNA damage, ii) protein synthesis, iii) carboxylic and organonitrogen metabolism and iv) expression of genes necessary for germination.

I developed a Bayesian tool to find clusters of genes in a DNA segment that are activated and repressed.

In order to understand the chromatin state of the cells during photoconidiation, we determined the levels of DNA methylation, which showed low levels of methylation with clear increment in the production of conidia when the DNA methylation inhibitor 5-Azacytidine was applied.

Based on the results, was concluded that exist different transcripts populations during the photoconidiation process in *Trichoderma atroviride*. It was shown that after stimulus, six hour is a crucial point during the development. Before six hours

the transcription is associated with light response and protein synthesis to condiate; after six hours, there is a repression of translation but high transcriptional levels related with conidia maturation and germination.

3. Introduction

Sunlight is the main energy source in our planet. Living organisms use sunlight as energy source and as indicator of environmental conditions. In order to perceive light, organisms have evolved mechanism for sensing and responding to the presence or absence of sunlight.

In *Trichoderma atroviride* sunlight induces the production of asexual reproduction structures called conidia. This phenomenon is a cellular differentiation process triggered by the exposure to a specific range of wavelength, the blue light (400-480 nm).

Asexual reproduction triggered by light is not exclusive for the *Trichoderma* genus, also in Aspergillus nidulans light regulates conidiation. It is known that blue and red light play together a role in conidiation inhibiting sexual reproduction. For instance, when blue (460nm) and red (680nm) light are present there is a maximum level of conidiation. Each light separately cannot induce conidiation at the same level that combined. (Purschwitz *et al*., 2008).

Asexual reproduction by light is not only achieved through conidiation. In the zygomycete *Phycomyces blakesleeanus,* different wavelengths ranging from blue to green (480-540 nm) can also induce the formation of sporangiophores, asexual reproductive structures with enclosed cells (Bergman, 1972).

Blue and green light has also a morphologic effect in other zygomycete called *Mucor circinelloides*. This fungus has tree copies of the homolog White-Collar1 (WC-1) named as mcwc-1a, mcwc-1b and mcwc-1c. The gene mcwc-1c was described as the responsible of the high production of carotenoids after blue or green light (525-540nm) exposure. The mcwc-1a was described as the responsible for positive phototropism, i.e. sporangiophore bends towards light. Positive phototropism were not observed with red light or darkness. The third gene (mcwc-1b) has no effect in both phenomena when deleted (Torres-Martínez & Garre, 2006).

As occurs in other fungi, asexual reproduction in *M. circinelloides* can be induced by blue and white light. However, none of the WC homologues play an essential

role sensing this process. It has been suggested that their function might be redundant in *M. cricinelloides* (Nicolás *et al*., 2008).

Regarding asexual reproduction, in *T. atroviride* the cell has to make a decision based on the environmental conditions in order to continue with the differentiation process or arrest it (Carreras-Villaseñor *et al*., 2012). Once the decision has been made conidiation can start.

Differentiation during conidiation in fungi has been widely studied from a morphological point of view, but the transcriptional steps that produce and prepare conidiation have been studied only at the early stages of this developmental process. Moreover, the lack of studies in the final stages is necessary in order to understand the whole process.

3.1. Cellular differentiation and chromatin

The capacity of a genome to establish a new cell type depends on the interaction of different molecular elements. These elements work together in the control gene activation and repression. The differentiation process is activated by environmental conditions, which are sensed and interpreted by the cell. To achieve the correct differentiation pattern, many cellular elements work together at different levels. For instance, transcription factors (TF) are involved in the activation or repression of genes, chromatin structure allows or restrain transcription, non-coding RNAs that could affect gene expression by transcriptional activation or repression as well as in post-transcriptional steps. and cis-elements on the genome, such as promoters (Higgs, 2007; Fig.1).

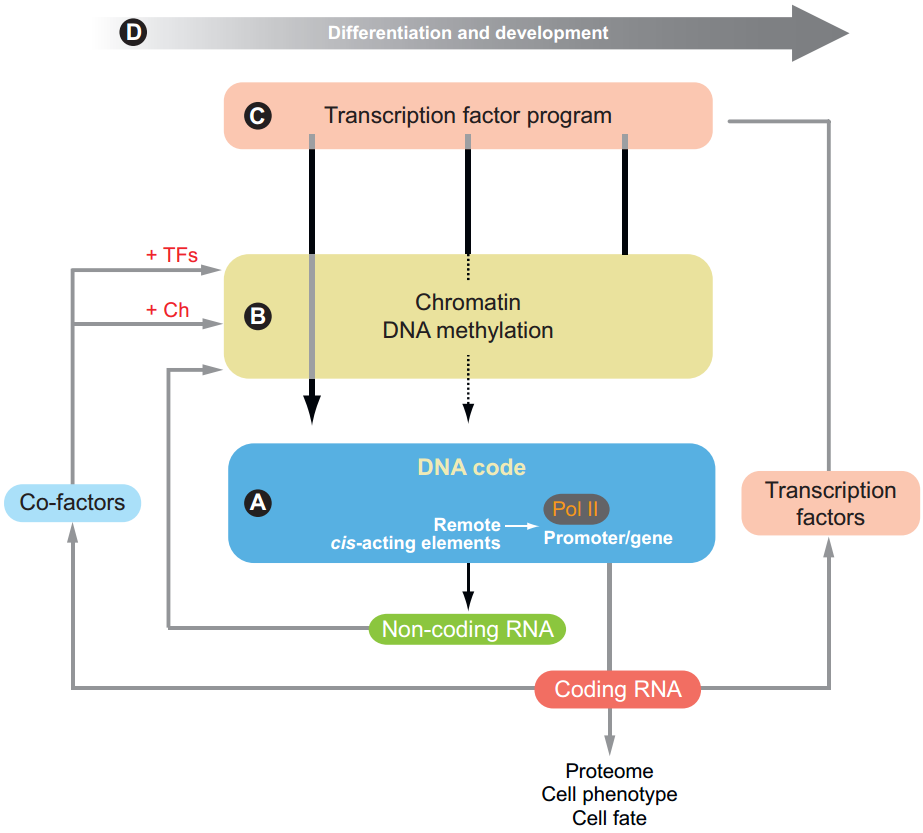

Figure 1. Molecular elements that lead differentiation. A) DNA elements that control gene expression. B) Elements that regulate DNA in chromatin level. C) Proteins that activate and repress genes. D) Differentiation process that is the result of interaction between these elements in a time lapse. Figure taken from (Higgs *et al*., 2007)

Analyses with transcriptomes of differentiated cells have shown distinct transcriptional profiles, either based in the level of activation of the genome or the level of transcription specificity of the active genes (Reyes-valde, 2008).They indicated a transcriptional commitment when cells are differentiated.

The aim of this study was to elucidate the transcriptional consequences of an environmental factor that induces a cellular differentiation process. The organism have to interpret and respond with the initial stimulus and subsequently react based in its current physiology and transcriptional plasticity. In such regard, *Trichoderma atorviride* shows up as model to study asexual differentiation owed to the morphological response by given environmental stimulus.

In order to do that, it was necessary to study the transcriptional state of the cell during the differentiation process from a transcriptomic perspective based in

RNA-Seq. This methodology allowed to take cellular snapshots of the transcriptional components transcribed as mRNA. Furthermore, it allowed me to follow the transcriptional transitions between the different states along the development of conidiophores.

The capacity of a genome to establish a new cell type when induced by an environmental stimulus allowed me to better understand how this interaction has consequences in the initial response as well as the destiny of a cell in order to survive and disperse. In other words, this considers the environment as well as the genomic component of the cell.

3.2. Sequencing technologies

The Genomic era has enabled us to obtain vast information of certain molecular elements related with nucleic acids, DNA and RNA. In such regard, new methodologies have been developed in order to increase the depth as well as the quality of the data generated.

After the elucidation of the human genome, next generation sequencing platforms were created and improved. There are multiple technologies available such as Roche/454 pyrosequencing, Illumina/Solexa, Ion Torrent and SOLiD.

Third generation platforms are referred as single molecule sequencing. Some of them are under development and trial. They are HeliScope form Helicos BioSciences, PacBio/SMRT, Oxford Nanopore with two products GridION and MinION, and a still promising single-molecule sequencing platform called VisiGen's Starlight. There is one single third generation commercially available PacBio/SMRT.

Each one of these technologies takes advantage of different biophysical and chemical properties of nucleic acids. In principle, the technology to use depends on the kind of study to make; for instance, for transcriptomic analyses Illumina/Solexa platforms have been considered as the best technology owed to sequencing depth and reads size. It is used not only based in cost-benefit, it is also because it avoids certain bias generated in other technologies such as

incorrect detection of single nucleotide polymorphisms (SNPs) and low coverage (Quail *et al*., 2012).

Illumina technology is the current ideal platform to sequence RNA based data. This is owed to the costs, time of data generation and sequencing depth (Tarazona, 2011).

4. Background

4.1. Asexual reproduction in fungi

Fungi have been used to study reproduction mechanisms due to the fact that they can reproduce both sexually and asexually. The conidium is a structure for asexual reproduction used by different fungi (Cole, 1986), which is used as a mechanism for survival and dispersal. Besides their natural role, these structures are important biotechnological applications, such as food maturation, biological control (Steyaert *et al*., 2010; Pascual *et al*., 1997), and inoculum for antibiotic production (Smith & Calam, 1980).

The formation of conidia is not expected to be the same among the fungal clade. However the cellular ontogeny of these structures can be classified in different groups such as ampullate, sympodial, determinate, percurrent and phialidic among others (Cole, 1986). Some of the fungal *genera* with phialidic conidia development are *Aspergillus*, *Penicillum*, *Neurospora* and *Trichoderma*, among others. In this work I focused in the transcriptional changes of phialidic conidiation in *Trichoderma atroviride*.

The genus *Trichoderma* is used mostly for enzyme production in the industry and as biocontrol agent in agriculture. Conidiation can be induced by different stimuli such as blue light (BL) exposure (Gressel & Galun, 1967), mechanical injury (Casas-Flores *et al*., 2004) and nutrient deprivation (Steyaert *et al*., 2010). More recently, it has been proven that extracellular Ca^{2+} (Simkovic *et al*., 2008) and native volatile compounds present in conidia (Nemčovič *et al*.,, 2008) may induce conidiation too.

At first sight, it seems that the cells that conidiate after the stimulus are different depending in the kind of stimulation (Fig.2). However, it has been observed that in *T. atroviride* the induction of conidiation by MI, the cell adjacent to the disrupted one blocks the drain of cytoplasmic content, this one regenerates and produces a conidiophore with conidia (Hernández-Oñate *et al*., 2012).On the other hand, the cells that respond to light are the apical cells in the radial zone, this results in the production of a dark green ring of conidia where it once was the apical zone

(Fig. 2). Even though the cells stimulated are not the same between the injured and Blue Light Pulse (BLP) induced), the cell that leads conidiation is an apical one in both processes.

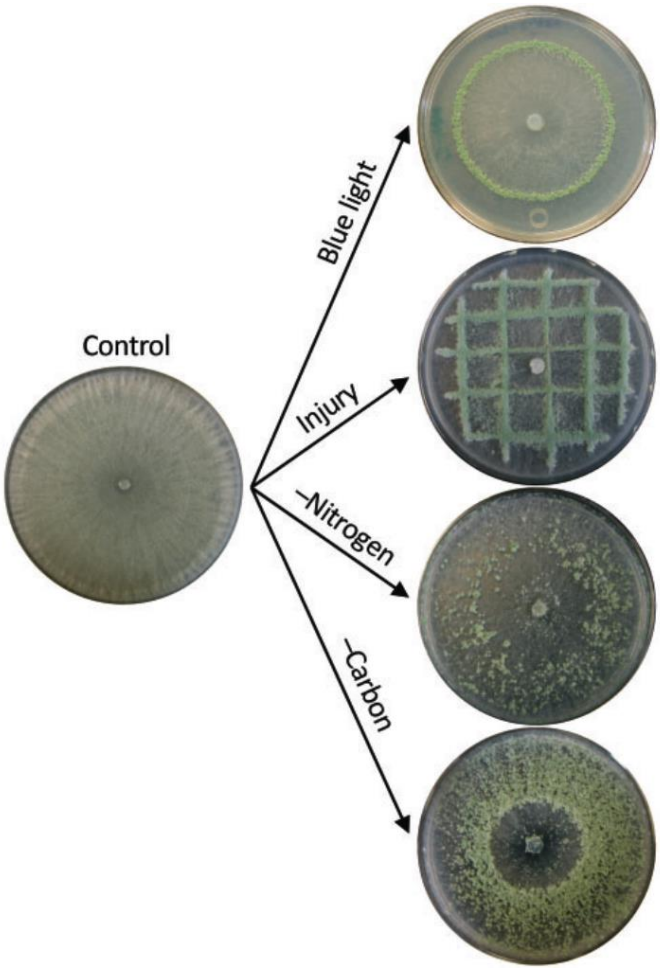

Figure 2. Conidiation induced in wild type *T. atroviride* by blue light pulse, mechanical injury, nitrogen starvation and carbon deprivation. Photographs taken 36 h after treatments. Figure taken from (Carreras-Villaseñor, *et al*., 2012)

Under optimal conditions, the morphological differentiation induced by BL starts two hours after the stimulus. The asexual reproduction structure starts to develop when the responding cells go upward and form aerial hyphae, while a new hypha is branched to keep growing radially at four hours. Subsequently aerial hyphae form branches at eight hours. Then branches generate structures called phialides 12 h. The phialides are coat-like structures where conidia are maintained and

maturated. This process requires 24 h from the direct stimulus to the formation of mature conidia induced by BLP or MI (Casas-Flores *et al*., 2004). (Fig.3)

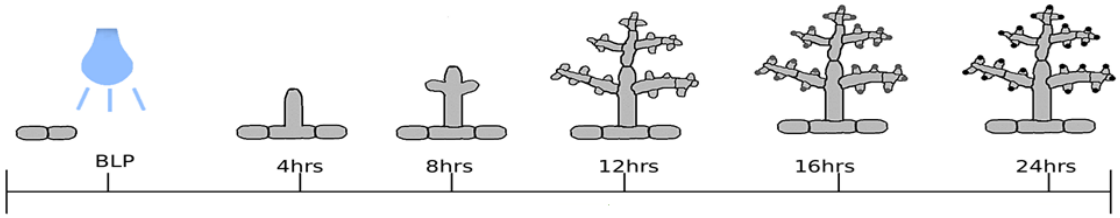

Figure 3. Conidia development after BLP. At four hours after stimulus the responding cell goes upward, at eight hours start to form branches, at 12 h phialides are observed, at 16 h unmaturated conidia are present and finally at 24 h dark pigmented conidia can be observed.

4.2. Light perception and transcriptomic response

The interest of this work is focused in BL induced conidiation. It is known that the perception of light can be made 16 h after inoculum, but it has been suggested that does not depend on age nor in colony size but in the metabolic state. When cells are capable of producing conidia in response to a BLP are photocompetent (Gressel & Galun, 1967).

Conidiation induced by BL in *T. atroviride* is initiated by the transcriptional factors Blue Light Receptor 1 and 2 (BLR-1 and BLR-2), homologues of the *Neurospora crassa* WC-1 and WC-2, respectively. BLR-1 has a domain found in proteins that perceive changes in Light, Oxygen and Voltage (LOV), which associates with FAD chromophore. When the chromophore is excited, it forms a flavin-cysteinyl adduct, that leads to protein activation (Crosson & Moffat, 2001). Once activated BLR-1, it is assumed to form a complex with BLR-2 to start a transcriptional cascade necessary for photoconidiation. The presence of these receptor is fundamental to blue-light perception, it has been proven that the lack of any of these proteins blocks photoconidiation in *T. atroviride* (Casas-Flores *et al*., 2004) and *T. reesei* (Castellanos *et al*, 2010) .

It has also been demonstrated that the inhibition of transcription before seven hours after photoinduction suppresses conidiation (Gressel & Galun, 1967).

Interestingly but not surprisingly, the inhibition of translation before 7 hours also suppresses conidiation (Betina & Zajacová, 1978).

A previous transcriptomic study in *T. atroviride* based on microarray technologies showed a strong transcriptional response after BLP at 30 min. They used 1,428 probes and they found that 2.8 % of the analyzed genes respond to blue light, 2% (30 genes) being upregulated and 0.8% (10) downregulated. The up-regulated genes were named *blu* and the down-regulated genes *bld* (Rosales-Saavedra *et al.*, 2006).

4.3. Cellular differentiation and chromatin

Differentiation of conidia is guided by early transcriptional activation and repression of genes in *T. atroviride* (Carreras-Villaseñor *et al*., 2012) and *A. nidulans* (Garzia *et al.*, 2013). It is known that transcriptional regulation in other organisms during cellular differentiation is controlled by chemical modification of the chromatin (Lister *et al.,* 2009). One of the main mechanisms of transcriptomic regulation via chromatin modification is DNA methylation, which is a covalent modification of the DNA with a methyl group in the 5th carbon of cytosines (5mC).

5mC is associated with activation or repression of genes depending on the genomic context and the organism where it is present. In mammals a high level of methylation in the promoter regions it is associated with gene repression; but when located in the gene body it is associated with gene activation. In plants it is highly associated with gene repression regardless of the context (Martienssen & Colot, 2001).

In fungi there are few reports with evidence of methylation in DNA (Ikeda *et al*., 2013; Reyna-López *et al.*, 1997). The associated functions are related with gene repression in functional genes (Dang *et al.*, 2013) and repression of repeated regions (Selker *et al*., 2003).

The effects of inhibition of DNA methylation with 5-Azacytidine have been reported in conidiation in *Aspergillus flavus*, where conidiation is lost and instead it creates a “fluffy” phenotype in different culture media (Lin *et al*., 2012). The

transcriptional consequences in differentiation produced by the lack of methylation are unknown in *A. flavus*. However, they found transcriptional deregulation of aflatoxin genes; showing higher expression than in cultures without 5-Azacytidine.

However one of the limitations of using this inhibitor is the toxicity effect; for instance, in different *Aspergilli* it was determined that an increaset in the concentration of 5-Azacytidine in the culture media is highly toxic for *A. flavus* and *A. nidulans*(Lin *et al*., 2012).

The genes responsible for DNA methylation in fungi belong to the DNA methylation transferase family proteins, DNMT1. This protein was reported in *N. crassa* and named as DIM-2 because its absence results in defective DNA methylation in *N. crassa*. Deletion of this gene disrupts the detection of methylation in DNA. Interestingly, deletion of the homologue of DIM-2 in *Ascolobus immersus* does not reduce methylation level in that organism.

DNA methylation in *N. crassa* is associated with gene repression in functional genes (Dang *et al*., 2013) and repression of DNA repatitive regions (Selker *et a*l., 2003).

To establish DNA methylation in *N. crassa* in repeated regions, it is necessary to recruit a protein called heterochromatin protein 1 (HP1), which is highly conserved in the clade eukarya (Honda & Selker, 2008). HP1 recognizes the trimethylated lysine 9 of histone 3, which is established by the DNA methylase DIM-5. DIM-2 forms a complex with HP1 and methylates the regions that are close to the histone marks. In case of the deletion of any of the three proteins, methylation cannot be carried out in these regions; Fig. 4). On the other hand, DIM-5 is not required to methylate DNA in gene regions where Dicer Independent smallRNAs (disiRNAs) are present but DNA methylation by DIM-2 is necessary to establish histone methylation by DIM-5 (Dang *et al*., 2013).

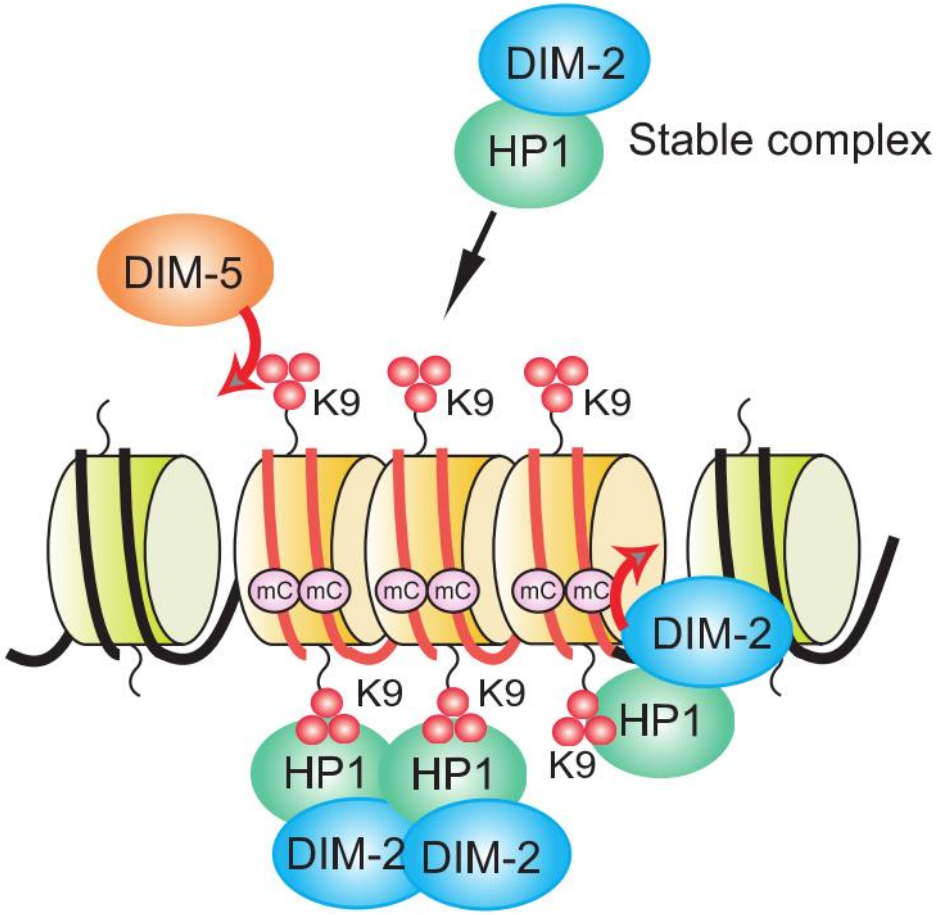

Figure 4. DNA methylation pathway in repeated regions in *N. crassa*. DIM-5 trimethylates K9 of histone 3. This mark is recognized by HP1 in complex with DIM-2 which methylates DNA.

4.4. Sequencing tools and pipelines

To study RNA it is necessary to take advantage of the molecular features of each kind of RNA. For instance, some studies consider RNA transcribed by RNA-Pol II, they use as a tag the Poly-A tail which is generated in eukaryotes; other use RNA traps to catch small RNAs, etc. Transcriptomic studies take as much as possible for a specific RNA and analyze all the information generated as a global understanding. Coding RNA is analyzed with insights to understand the genes responsive to a certain condition or treatment.

Genomic tools and platforms have enable us to detect high number of mRNA and the relative amounts of each gene in the cell. This approach uses a methodology

that sequences nucleic acids in a high-throughput fashion. The Illumina® HiSeq 2500 platform in format 2X100 generates 33Gb approx. of data and 168 million reads by lane when they are not filtered. This sequencing platform enables to get 10 million reads per sample, which is enough considering genome size. (Tarazona *et al*., 2011)

After transcriptomic data are generated different filters are used in order to get sequences of interest aligned to a reference and the number of sequencing reads associated to genes to perform multiple bioinformatics analyses such as differential expression of genes.

5. Hypothesis

If cellular differentiation controls development during photoconidiation process in *Trichoderma atroviride*, morphological and functional changes are reflected in transcriptional changes as well as in changes of DNA methylation levels after blue light induction.

6. Main Objective

Quantify the transcriptional changes and the transcriptional activation of the genome as well as the amount of methylated DNA in different stages during photoconidiation process in *Trichoderma atroviride*.

7. Particular Objectives

- Establish the number of differentially expressed gene during photoconidiation.
- Assign functional annotation to activated or repressed genes during photoconidiation.
- Quantify the levels of transcriptional activation of the genome and the degree of transcriptional specialization of the genome among the sequenced samples.
- Identify genomic context patterns associated with gene activation or repression.
- Measure the quantity of DNA methylation during photoconidiation.

8. Materials and Methods

8.1. RNA sequencing

8.1.1. Sample taking

Samples were collected at six different times one set before exposure to the BLP and five sets after. The first one was taken after 36 h of growth of the fungus in darkness, and the other five were taken at 30 min when transcriptional activity has been reported previously after BLP (Rosales-Saavedra et al., 2006), three hours as an intermediate point between response to light and the sample of six hours when transcripts before this time are necessary to conidiation, 12 h when phialides are formed and 24 h after the BLP when the conidia are mature. In order to enrich the samples with tissue that responds to light and participate in photoconidiation, we collected only there are corresponding to the photoconidiation ring.

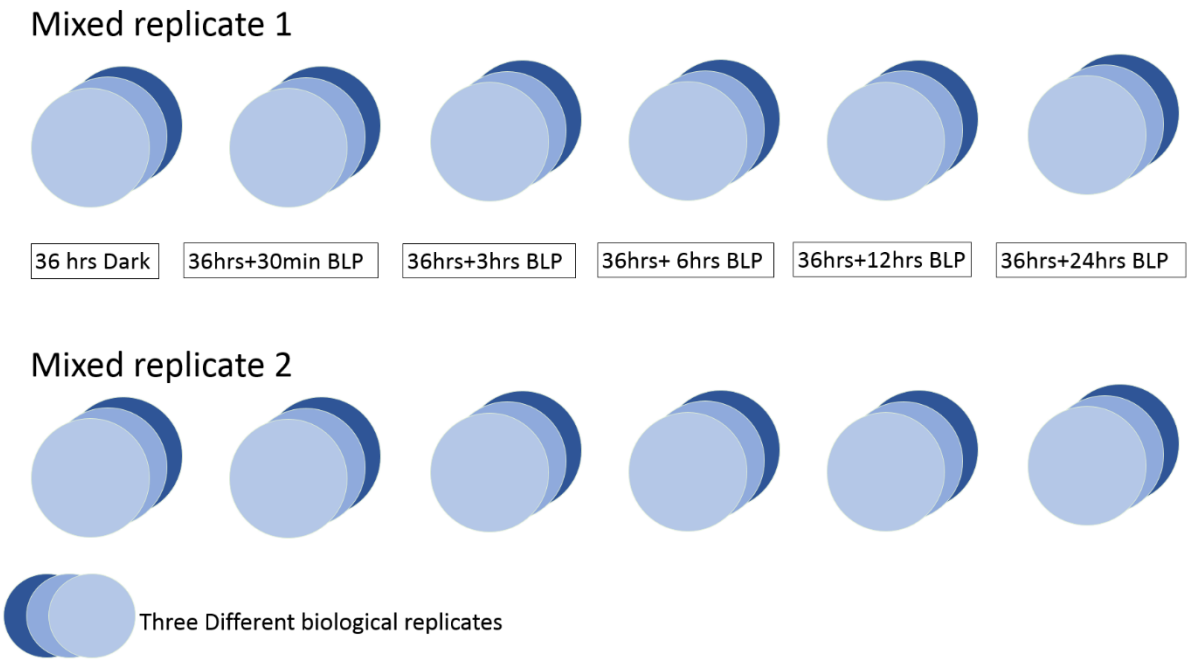

Figure 5. Each mixed replicate is composed by three biological replicates. We made that pool for each treatment. 36 h in darkness, 30min, three, six, 12 and 24 hours after BLP.

There were created two replicates for sequencing, each of them composed by three biological replicates that were mixed in a single sample. By biological replicates is referred as collections from the same treatment, in different days. Then, two pools were formed by three biological replicates that were sequenced separately.

Protocol of sample taking.

1. Cut rings of sweet cellophane (cellulose composed) with an intern diameter of 4 cm and an external diameter of 6 cm. This is the approximated size of *T. atroviride* at 36 h in darkness in PDA media. Cut a spare number of cellophane rings.
2. Put the cellophane rings in a recipient with distilled water, heat them in the microwave and wash softly, drop the water and repeat this step three times.
3. Cut a useful size of foil and put a wet paper towel over the foil. This paper bed contained and separated sets of cellophane rings in order to use them separately. Once the wet paper bed was full with an arbitrary number of cellophane rings, cover them with another wet paper towel. Repeat the process until the cellophane rings are finished, then cover them with a new wet paper towel. Finally cover the wet paper towels and make a package ready to sterilize.
4. We prepare 14 Petri dishes (10 cm diameter) with PDA. Twelve dishes for sample taking and one as positive and the other as positive and negative controls of conidiation respectively.
5. Set right in the center of Petri dishes the sterile cellophane rings.
6. In a dark room (we used a safe red light to illuminate), we placed an inoculum in the center of the Petri dish and allowed the fungus to grow for 36 h in darkness at 26 ºC in a diode chamber.
7. To collect the control sample at 36 h in darkness use sterile laboratory tweezers and place cellophane rings with sample in foil sheets. Close the foil sheet carefully and softly, then drop it in liquid nitrogen for each collection (two by day).

8. To deliver the BLP, use a blue light fluency of 1200 µMol/m^2 distributed in 57 sec. Induce the 12 remaining dishes. Waited 30min, three hours, six hours, 12 h and 24 h to collect and freeze the two samples separately as made in step 7.
9. Repeat the process two more times in order to get three different biological replicates at different day.
10. Positive and negative controls of conidiation were considered for each biological replicate. The first is an induced culture without taking the sample, the second is a not induced sample.

8.1.2. Extraction of nucleic acids method

There were extracted from the same sample DNA and RNA. This allowed to get the most related result between the DNA and RNA analyses.

The sample pools by replicate were obtained by maceration in liquid nitrogen. The macerate was transferred to a pre-cooled Eppendorf tube. The tubes were filled until they reached 500 µl of macerate. This was done for all samples. The tubes were frozen in liquid nitrogen and store at -80 ºC.

RNA extraction protocol:

1. Add 1mL of TRizol ®.
2. Vortex tubes softly for 10 min.
3. Centrifuge samples 10 min at 12,000 rpm or 11,000 g at 4 ºC.
4. Recover supernatant and put in a new Eppendorf tube.
5. Keep samples at room temperature 5 min.
6. Add 0.5 volumes of phenol-chloroform-isoamyl alcohol (25:24:1)
7. Vortex tubes 5 min.
8. Centrifuge samples 12,000 rpm or 8,000 g by 15 min at 4ºC.
9. Repeat steps 7-8 on more time.
10. Recover supernatant.
11. Added 0.2 volumes of chloroform- isoamyl alcohol (24:1)
12. Vortex samples 2 min.
13. Keep samples at room temperature 2min..

14. Centrifuge tubes at 12,000 rpm or 8,000 g by 15 min at 4ºC.
15. Recover aqueous phase.
16. Add 0.25 volumes of isopropanol and 0.25 volumes of concentrated salt solution (Sodium citrated 0.8M and NaCl 1.2 M).
17. Mix by inversion and incubate in room temperature 10 min.
18. Tubes were centrifuged at 12,000 rpm or 11,000 g by10 min at 4ºC.
19. Remover water were and we wash the pellet with 1 volume of ethanol 75% (DEPC diluted)
20. Vortex 30 s.
21. Centrifuge tubes at 7,500 rpm or 4,500 g by 5 min at 4ºC.
22. Remove water and let the pellet dry in room temperature.
23. Suspend pellet were in H_20 DEPC overnight at 4 ºC.

DNA extraction protocol:

1. Add 600 µL of urea buffer to each Eppendorf tube.
2. Vortex until sample is completely mixed and keep 30min at room temperature.
3. Adde 300 µL of phenol and 300 µL of chloroform.
4. Vortex for 10 min.
5. Centrifuge samples 15min at room temperature.
6. Take çand repeated the 1-5 step one more time.
7. Added a volume of isopropanol and mixed by inversion.
8. Collected the precipitated DNA with a glass tip.
9. Wash precipitated DNA strands with ethanol 70%.
10. Finally, dissolve the strand in 50 µL of H_2O MiliQ.

8.1.3. RNA sequencing specifications

14 RNA-Seq libraries type TrueSeq were prepared at the Genomic Services Laboratory of Advanced Genomics Unit in Irapuato, Guanajuato. Each library was labeled and sequenced using the *1X100 high performance* format of the Hiseq2500 Illumina platform.

8.1.4. DNA digestion for analytical chemistry analysis of DNA methylation

The DNA digestion must produce nucleosides i.e. no reactant phosphate linked. In order to do that, three reactions were performed. The first reaction is a digestion with DNAse I, which produces several cuts in DNA but leaves some uncut nucleotides; the second digestion uses nuclease P1, which produces nucleotides; and finally we used Alkaline Phosphatase CIP.

DNAse I reaction.

Reactants	**Procedure**
10 µl of DNA-H_2O Tris-HCl	1. Reaction 1 h in 37 ºC. 2. Denature reaction at 100 ºC 3 minutes. 3. Put in ice until next step.
1 µl of DNAse I	
1.2 µl of Reaction Buffer	

Nuclease P1 reaction:

Add to the previous reaction

REACTANTS	**PROCEDURE**
1 µl of Nuclease P1= 5 µg of enzyme	1. Reaction 18 h in 37 ºC.
2 µl of Sodium Acetate .5 M and pH 5.2	
2 µl $ZnSO_4$ 1 M	

Alkaline Phosphatase CIP reaction:

From the same reaction

REACTANTS	**PROCEDURE**
3 µl Tris-HCl 1 M pH 8	1. Reaction 2 h in 37 ºC.
1.5 µl Alkaline Phosphatase CIP	
2 µl Reaction Buffer	

8.2. Differential Gene Expression

In order to perform a DGE analysis the sequences were filtered, mapped to a reference genome and the associated reads were counted to each gene.

8.2.1. Reads filtering

To filter reads it was used FastQC (S, Andrews. Version 0.10.1) a tool that shows the reads quality based in the quality of the *fastq* files generated and the length of the read. All libraries had a good quality (bigger than 20), this avoids the need of using the trimming methodology in the mapping reads.

8.2.2. Mapping reads

The reads mapping or reads alignment is the location of each read generated by the sequencer to a reference. The reference could be the organism's genome, a near one phylogenetically or a specific set of nucleotides to align. We used the genome of *T. atroviride* IMI 206040 Version 2, available at http://genome.jgi-psf.org/pages/search-for-genes.jsf?organism=Triat2. To align each library to the genome we used a Perl script that called Bowtie2 then the output is directed to Samtools (Li *et al.*, 2009).

We used the default parameters and very sensitive alignment criteria, where only the match between the read and the reference is 100 %. If a read aligned in two different locations in the genome only, the first one mapped was used.

Once the mapping was done and Samtools converted the SAM file to BAM we sorted the reads. The sorting is the readjustment of the mapped reads in the genome with an index reference, in this case the genome. This process left the unmapped reads at the end of the sorted file. The sorting of BAM files allowed us to see in the Integrative Genome Viewer (IGV) tools (Thorvaldsdóttir, Robinson, & Mesirov, 2013) browser the alignment of reads in each section of the genome.

A Perl script **(cd provided)** describes in a more detailed way the procedure and instructions of the mapping part.

8.2.3. Gene counting

There are multiple ways to count reads mapped to a location of the reference. We used R/Bioconductor package called RSamtools (Morgan, Obenchain, & Hayden, 2015). This package takes all the BAM files and matches the reads to a specific locus.

8.2.4. Statistical analysis of differential gene expression

A generalized linear model (GLM) method was selected to get differentially expressed genes (DEGs). This methodology uses a relation variance-mean to stablish a log-linear model. The method has been implemented in R/Bioconductor package EdgeR (McCarthy, Chen, & Smyth, 2012).

An R Script takes counts associated to genes, then changes counts to count per million (cpm). Then indexes the comparisons with its respective replicate. The Script also calculates normalization factors, and *common*, *trended* and *tagwise* dispersions.

The maximum value of false discovery rate (FDR) was 0.01, this could be considered as strict. The genes found in that FDR had a lower fold change of 1.3 for up-regulated genes and -1.3 for down-regulated. We then compared BLP libraries against the control kept darkness. To determine DEG in transitions we used the BLP induced libraries against its previous stage, i.e. tree hours using as control 30min, 6h uses as control 3h and so on. These comparisons made up 9 lists of DEG in total.

8.2.5. Gene Ontology Analysis of differentially expressed genes

To perform Gene Ontology Analysis of DEG enrichment. José Manuel Villalobos made the annotation of the *T. atroviride* V2 genome. It was developed at BLAST2GO (Götz *et al*., 2008). In BLAST2GO we performed Gen Ontology Enrichment Analyses (GOEAs) with DEGs by stage as well as transition. We selected the enrichment with an FDR < .05. (Conesa *et al*., 2005)

GO terms are divided in three main ontologies, cellular components, cellular functions and cellular processes. We obtained all the enrichments for the three groups. In order to reduce the amount of GO terms enrichments we used a tool to reduce GO terms to the broadest and most specific term enriched, implemented in the BLAST2GO platform, and used a cut off in FDR < .05.

8.3. Gene clusters by genomic context

8.3.1. Bayesian inference of gene clustering

We implemented a strategy based in Bayesian probability in order to find gene clusters with the same transcriptional profile, upregulated or downregulated.

The method draws a null distribution based in a hypergeometric distribution of a K number of genes activated/repressed from an N number of total genes.

$$P(X = k|j, N, K) = \frac{\binom{K}{ki}\binom{N-K}{nj-ki}}{\binom{N}{nj}}$$

Where k_i is the chance to get *i*th number of genes in a certain *n* window with *j*th number of genes. Where $0 \leq i \leq j$ and $j > 1$.

The drawn distribution shows the probability to find certain number of genes with the feature requested within a certain number of contiguous genes in the genome by chance.

Then the algorithm searches for the probabilities lower < .01 to find *ki* genes activated/repressed in a window with *j* number of genes given a null model.

After the null distribution is drawn, the algorithm takes the real DEG data (observed data) and check the number of genes activated/repressed given a *j* number of genes and compares with null distribution generated previously. We used *a posteriori* probability lower than .01 to detect odd clusters. This is summarized in the next equation.

$$P(P(k_i|n_j)|H_0)$$

In case of clusters found as odd clusters, the algorithm concatenates overlapped clusters and asks the probability to find that concatenated cluster with that size. If the concatenated cluster has *a posteriori* probability <0.01 it is taken as a true cluster that now has the most accurate cluster size.

8.4. Transcriptional status during photoconidiation

8.4.1. Transcriptional Diversity and Specialization

It is not possible to understand the transcriptional status of the cell using DEG analysis. In 2008 an extrapolation of Shannon`s entropy was developed (Reyes-valde, 2008), this measurement was called Diversity, which is defined by how expressed is the genome and how well distributed are all the relative frequencies of the reads in the treatment or in this case the time of collection. This Diversity equation is expressed as:

$$H_j = -\sum_{i=1}^{g} p_{ij} log_2(p_{ij})$$;

Where H_j is the Diversity of the *j* library using the relative frequencies *p* of the gene *i*th in library *j*th. H_j will vary from zero when only one gene is transcribed up to $\log_2(g)$, where all *g* genes are transcribed at the same frequency: $1/_g$. If we consider the average frequency of the *i*th gene among tissues, say,

$$p_i = \frac{1}{t}\sum_{j=1}^{t} p_{ij}$$

,

where *t* is the total number of times of collection.

With the average frequency of the gene among times of collection we can obtain the gene specificity that is provided by

$$S_i = \frac{1}{t}(\sum_{i=1}^{t} \frac{p_{ij}}{p_i} log_2 \frac{p_{ij}}{p_i})$$

,

and gene specificity is defined as the information that its expression provides about the identity of the source tissues. S_i will take a value of zero if the gene is transcribed at the same relative frequency in all collection times and a maximum value of $\log_2(t)$ if the gene is exclusively expressed in a single time.

S_i will give a value of zero if the gene is transcribed at the same frequency in all tissues and a maximum value of $\log_2(t)$ if the gene is exclusively expressed in a single tissue. If we consider S_i in the Shannon´s entropy equation we can obtain the tissue specialization for each *j*th time. Specialization might be defined as how specific and expressed are the whole genes in that time. The Specialization value

$$\delta_j = \sum_{i=1}^{g} p_{ij} S_i$$

,

varies from zero if all genes expressed in the tissue are completely unspecific i.e. $S_i = 0$ for all *g*, up to a maximum of $\log_2(t)$, when all genes expressed in the tissue are not expressed.

8.5. DNA methylation in *Trichoderma atroviride*

8.5.1. Inhibition of DNA methylation in photoconidiation

The inhibition of DNA methylation was possible using 5-Azacytidine, a competitive inhibitor of DNMT. We prepared Petri dishes containing PDA amended with the inhibitor with varying concentrations, 25 µM, 50 µM, 100 µM and 200 µM. We used 5-Azacytidine (≥98 % (HPLC)) provided by Sigma-Aldrich ®.

The protocol to prepare media in petri dishes with PDA and 5-Azacytidine was:

1. Sterilize 5 bottles with 150, 148, 146, 144 mL of PDA.
2. Weight the required quantity of 5-Azacytidine and place it in a sterile falcon tube.
   In this experiment we used a total of 5.49 mg.

5-Azacytidine mg	**Final concentrations in 150 mL PDA**
0.9	25 µM
1.8	50 µM
3.7	100 µM
7.3	200 µM
Total= 13.7 mg	

3. Cool a falcon tube with H_2O MiliQ.
4. Add 7 mL of cold water in the falcon tube with 5-Aza.
5. Vortex the sample and put it in ice repeatedly to keep it cold until 5-Aza is completely solved.
6. Filter the solution with a syringe coupled with a filter HVLP .45 µm and let drop the sterile solution in a new sterile falcon tube. Make up the solution up to 11mL.
7. Take 1 bottle of 150 mL of PDA and prepare 6 petri dishes 5-Aza.
8. Take 1 bottle of 149 mL of PDA and let it cold until reaches temperatures near solidification and drop 1 mL of 5-Azacitydine solution, mix gently. Prepare 6 Petri dishes.

9. Take 1 bottle of 148 mL of PDA and let it cold until reaches temperatures near solidification and drop 2 ml of 5-Azacitydine solution, mix gently. Prepare 6 petri dishes.
10. Take 1 bottle of 146 mL of PDA and let it cold until reaches temperatures near solidification and drop 2 mL of 5-Azacitydine solution, mix gently. Prepare 6 Petri dishes.
11. Take 1 bottle of 142 mL of PDA and let it cold until reaches temperatures near solidification and drop 2 mL of 5-Azacitydine solution, mix gently. Prepare 6 Petri dishes.
12. Keep media Petri dishes in darkness.

Room conditions are room temperature and low light exposition. Petri dishes must be inoculated the same day as prepared.

Cultures were allowed to grow 36 h in darkness at 26 ºC in a diode equipped growth chamber in darkness. From each set, 3 inocula were left in darkness and 3 exposed to light (500 µMol/m^2, in 27 sec).

Conidia were collected for every culture in 10ml and quantified in a Neubauer chamber.

Two biological replicates were analyzed.

8.5.2. Analytical measurement of DNA methylation

The analytical measurement of DNA was performed by our collaborator Kazimierz Wrobel from Universidad de Guanajuato with a method previously published and developed by his group (Magaña *et al*., 2008).

The methodology is based in a HPLC separation of nucleosides with three mobile phases, deionized water, 50 mM phosphate buffer and methanol. The compound detection is provided by a mass spectrometer and the quantity of methylated DNA is inferred with the area under the curve generated by the mass spectrometer.

Four replicates in total were made by each stage, two for each mixed DNA sample.

9. Results

9.1. General strategy

This work uses a strategy to elucidate the transcriptional and DNA methylation states in a global manner. Two strategies were usedin this study, one to determine the transcriptional states and changes based in RNA-Seq and other for the measurement of DNA methylation using an analytical chemistry approach and the inhibition of DNA methylation.

Different approaches were used, from detailed descriptions of differential gene expression analyses (DEAs) to a numerically holistic measurement of transcription with Diversity (*D*) and Specialization (*S*) analyses. These approaches have been used previously in the analysis of the chili pepper transcriptome (Martínez-López *et al*., 2014). Both sets of information were needed to interpret the functional and transcriptional cellular status during the photoconidiation process.

The usual strategy is to find Differentially Expressed Genes (DEGs) using statistical methodologies previously reported and proven. Once DEGs are found, they are described in order to explain what are their potential roles depending on the expression values. Another common way of analyzing this information is to search the statistical enrichment of GO terms based on DEG.

A powerful, but no common strategy, is the Diversification (*D*) and Specialization (*S*) analyses (Reyes-Valde, 2008), based in Shannon's entropy information theory. Transcriptomes are interpreted as a whole in *D* and *S* as the interaction of relative frequencies of genes within and between samples.

Particularly, it was assumed that there might be regulation of genes with common expression patterns when they are close in the genome as previously reported (Gibbons *et al*., 2012) and for that it was developed a tool based in Bayesian Inference to find gene clusters.

It is known that during differentiation in mammals the amount of DNA methylation changes. Thus, it was necessary to measure the amount of methylated DNA through the photoconidiation process using HPLC.

I finally interpreted all the results to infer the transcriptional states during photoconidiation as well as the role of DEGs by genomic context and cellular role. The complete strategy is summarized in (Figure 6).

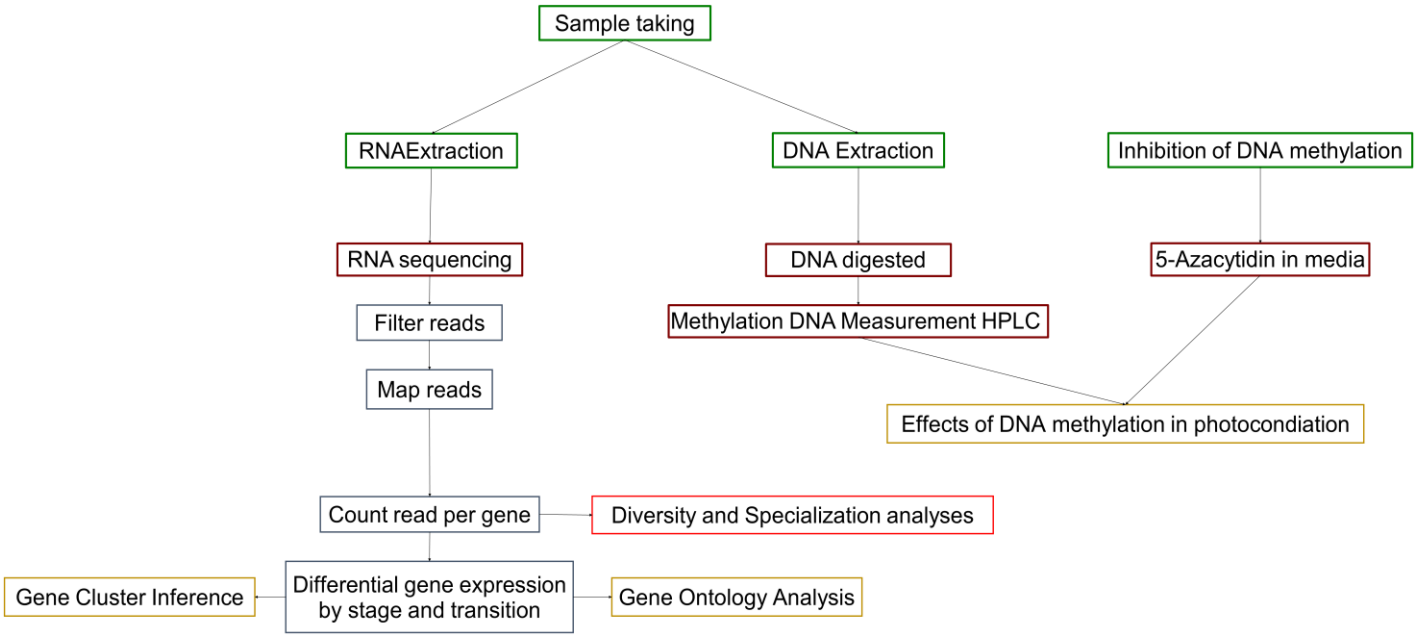

Figure 6.Two mixed strategies were developed to understand the transcriptomic and chromatic state of cell during photoconidiation.

9.2. RNA sequencing summary

The Illumina/Solexa HiSeq2500 platform was selected to sequence 12 RNA libraries from independent samples collected at 6 different times 36 h in darkness, 30min, three, six, 12 and 24 hours after BLP.

Each time was taken based in different assumptions. Previously in our laboratory strong transcriptional changes had been observed at 30 minutes after a BLP (Rosales-Saavedra et al., 2006). Three hours were selected by two reasons, i) cells are start to differentiate to form aereal hypha, which could indicate the initiation of the differentiation process and ii) to have a checkpoint before six hours, because it has been reported as a time were inhibition of transcription still

blocks production of conidia. Samples at 12 h already have phialides and also work as a checkpoint before 24 h where conidia are mature.

The sequencer produced a total of 169,140,730 reads. When aligned to the genome and quantified to gene regions without UTRs, we observed that nearly 50 % of the reads are associated with coding sequences in the *T. atroviride* genome V2 (Table1)**.**

Because library preparation is based in a conversion from RNA to cDNA, at this step the directionality of the mRNAs is lost. That was an issue when we tried to quantify the number of reads associated with genes. For that reason we left aside genes that overlap. Overlapped genes were only 60.

Table 1. Sequenced samples Illumina/Solexa HiSeq2500 platform. Format 1X100. 14 libraries were sequenced. The sample name is given by the time followed by the technical replicate and the condition.

Sample	ID	Total number of Reads obtained	Reads Aligning to Genes
36 h.1 Dark	AH1TL1SS61.11	11,056,114	6,209,232
30 min.1 BLP	AH1TL1SS62.12	13,448,399	5,837,113
3 h.1 BLP	AH1TL1SS63.13	10,284,070	4,396,168
6 h.1 BLP	AH1TL1SS64.14	14,321,886	6,166,557
12 h.1 BLP	AH1TL1SS65.15	14,124,116	6,037,106
24 h.1 BLP	AH1TL1SS66.16	13,810,979	5,811,974
36 h.2 Dark	AH1TL1SS680.19	13,745,055	5,941,906
30min.2 BLP	AH1TL1SS690.20	12.426.585	5.465.898
3 h.2 BLP	AH1TL1SS700.21	11.932.158	5.287.834
6 h.2 BLP	AH1TL1SS710.22	10.679.611	4.574.171
12 h.2 BLP	AH1TL1SS720.23	8.042.670	3.412.173
24 h.2 BLP	AH1TL1SS73.25	11.174.038	4.692.952

9.3. Differential Gene Expression

9.3.1. Genes differentially expressed by stage

I performed a DEA using the samples after BLP against the control in darkness, using R/Bioconductor software with the EdgeR package. This group of algorithms can measure the dispersion among samples and works better when experiments that have replicates with the same treatment. After statistical analyses, and given that our samples had low variability by treatment, or time of collection, we considered a differentially expressed gene (DEG)when false discovery rate (FDR) is lower than0.01 and a Log Fold Change (LFC) greater than 0.29 for up-regulated and lower than -0.29 for down-regulated genes. The values are given by the FDR restriction and its relation with the variability of the samples; it was observed high dispersion in counts values in genes with a LFC lower than .3.

Table 2. Number of genes differentially expressed upregulated as well as downregulated. The third column indicates the ratio of genes upregulated over downregulated.

Time after BLP	Upregulated	Downregulated	Upregulated/Downregulated
30 min	284	497	0,57
3 h	615	927	0,66
6 h	895	1127	0,79
12 h	1381	1521	0,91
24 h	1910	2020	0,95

A greater number of repressed genes was observed in comparison with activated genes in early stages. Surprisingly the ratio between activated and repressed genes reaches almost 1 at later stages (Table 2).

Along the photoconidiation process there are genes that are kept repressed or activated in all libraries (Fig. 7). I found 59 activated (Fig. 7b) and 280 repressed (Fig. 7a) genes in all the developmental stages. Surprisingly, the number of repressed genes in all libraries is greater than that of activated genes. On the other hand, the number of genes activated and unique for each developmental

stage is greater. For instance, the number of activated genes when conidia are mature is larger than the repressed ones; a similar situation is found at 12 h stage when phialides are present.

a) b)

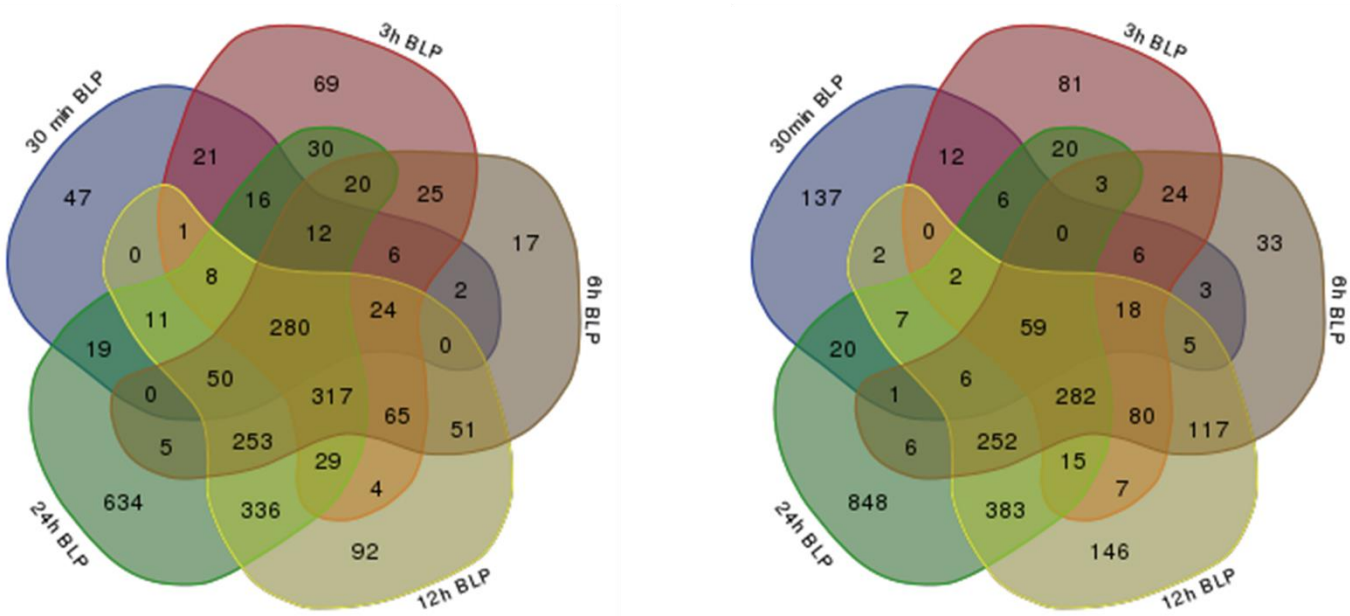

Figure 7. Venn Diagram of DEG during developmental stages. a) Repressed genes during photoconidiation. b) Activated genes during development.

The total number of genes that modify their pattern of expression during the photoconidiation process was 2,581 out of 11,863 in *T. atroviride* genome V2. Among the overall DEG, there were genes that gradually increase their level of activation or repression. In order to detect them, only those genes with an FDR lower to 0.01 in all stages where selected. From 59 genes that are constantly activated, only eight where increasingly activated (Fig. Supp. 1). From 280 genes that are constantly repressed 40 were increasingly repressed (Fig. Supp. 2).

The changes in LFC are bigger as time advances. It is noteworthy that in the comparison between 24 h after BLP and the control in darkness we found the genes that are more strongly repressed or activated than in any other comparison are present in that final stage. On the other hand, the sample with the fewest DEGs corresponds to 30min after the BLP, with low LFC values associated (Fig. 8).

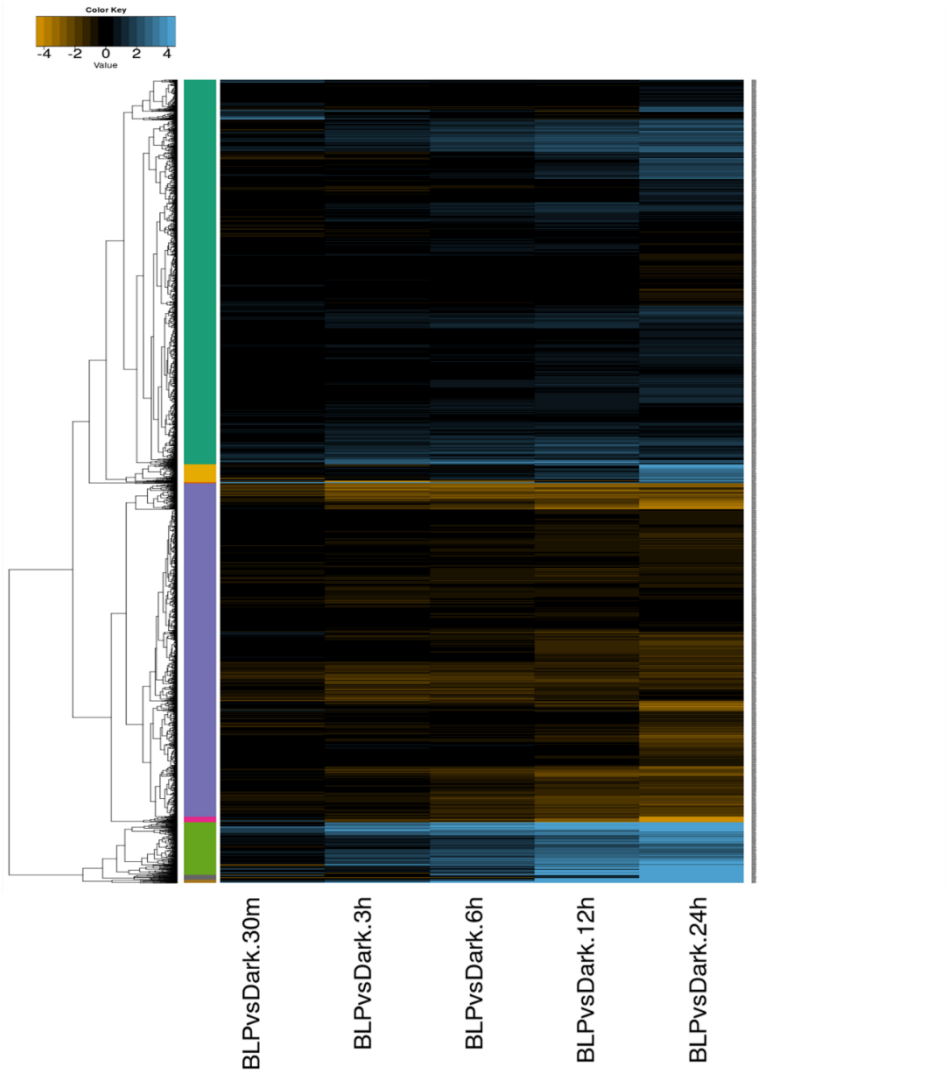

Figure 8. Heatmap of differentially expressed genes. Five comparisons were performed against 36 h in darkness. LFC limit is 4 and -4, genes over or under those LFC have saturated colors. Complete UPGMA (Unweighted Pair Group Method with Arithmetic Mean) clustering algorithm. The x axis shows the comparisons between samples at different times, e.g. BLPvsDark.30 is the 30 min sample using as a control the sample in darkness. The y axis shows the different clusters formed, the dendogram is composed by each differentially expressed gene.

9.3.2. Gene Ontology of Genes differentially expressed by stage

In order to understand the functional changes associated with the differentially expressed genes in each comparison, GO terms enrichments were perfomed GO terms related to cellular processes with the BLAST2GO tool with an FDR of 0.05 were selected. It was used a reduction tool of the dataset that enabled to see the

most specific and broadest GO term associated with each group of genes as possible.

There was observed a repression of translation that starts at six hours and is maintained until the mature conidiophore stage. It was also observed repression of genes associated with ion transport, catabolic processes and cytoskeleton organization at 24 h. The repression of translation is present in GO terms associated with cellular components in genes related with ribosome. On the other hand, there was an enrichment of up-regulated genes at 30 min associated with DNA metabolic process, this stage is the first response to blue light. At three hours there is a single GO term enrichment, related with protein metabolic process, more specifically, related to endoplasmic reticulum, which is in charge of protein maturation (Fig. Supp. 3). At six and 12 hours, when phialides are present, terms related with carboxylic acid metabolism and organonitrogen compounds (use of nitrogen in amide, amines, nitrates, and assimilated nitrogen with structural function) are activated. Finally in conidiophores at 24h there is activation of genes related with response to stimulus, single-organism cellular and DNA metabolic process (Figure 9).

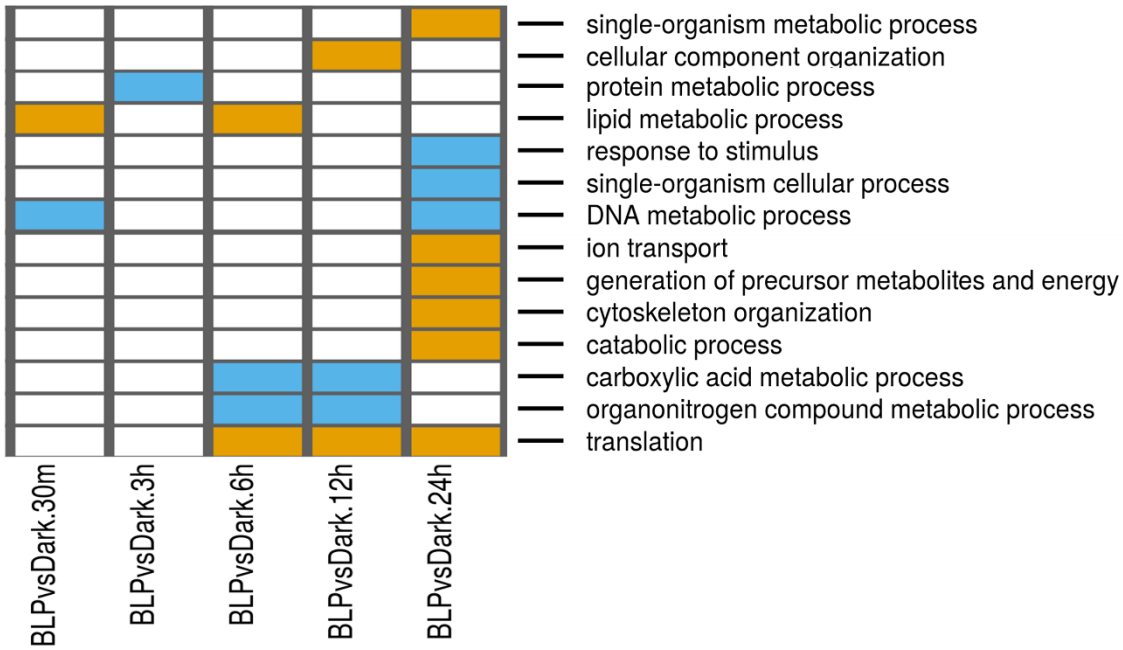

Figure 9. GO terms enriched by stage. Go terms enriched by upregulated genes are represented in blue and GO terms enriched by downregulated genes are represented in yellow. GO terms are written on the right. Enriched terms were considered when p-value was lower than .01 and an FDR lower than .05.

It is observed that there are 4 phases based in similarity, one at 30 min, another at three hours, one more from six to 12 hours and a final phase at 24 h (Figure 9).

9.3.3. Differentially expressed genes by transition

DEA was performed using the samples after BLP against its previous stage. The comparisons were between three hours using as control 30 min after BLP, six hours against three hours and so on. I considered a gene differentially expressed when its LFC was bigger than 0.3 for up-regulated and lower than -0.3 for down-regulated genes, and a false discovery rate (FDR) lower than 0.01.

The lowest ratios between activated and repressed genes were present at early stages, but at later stages it changes. The rations are greater in the transition between six and 12 hours and almost the same in the last transition (Table 3).

Table 3. Number of genes differentially expressed upregulated as well as downregulated by transition. The third column indicates the ratio of genes upregulated over downregulated.

Time after BLP	Upregulated	Downregulated	Upregulated/Downregulated
3 h vs 30 min	426	578	0,74
6 h vs 3 h	153	190	0,81
12 h vs 6 h	284	228	1,25
24 h vs 12 h	1055	1052	1

There were 2,888 DEG among transitions, those changes have contrasting patterns. There were interesting patterns of expression, some genes that were activated or repressed at certain transition do not necessarily change in the next transition (Fig. 10).

There was greater transcriptional changes in the last transition by number of genes (Table 1) and magnitude (Fig. 10). Also the transition from 30 min to three hours showed great changes in LFC.

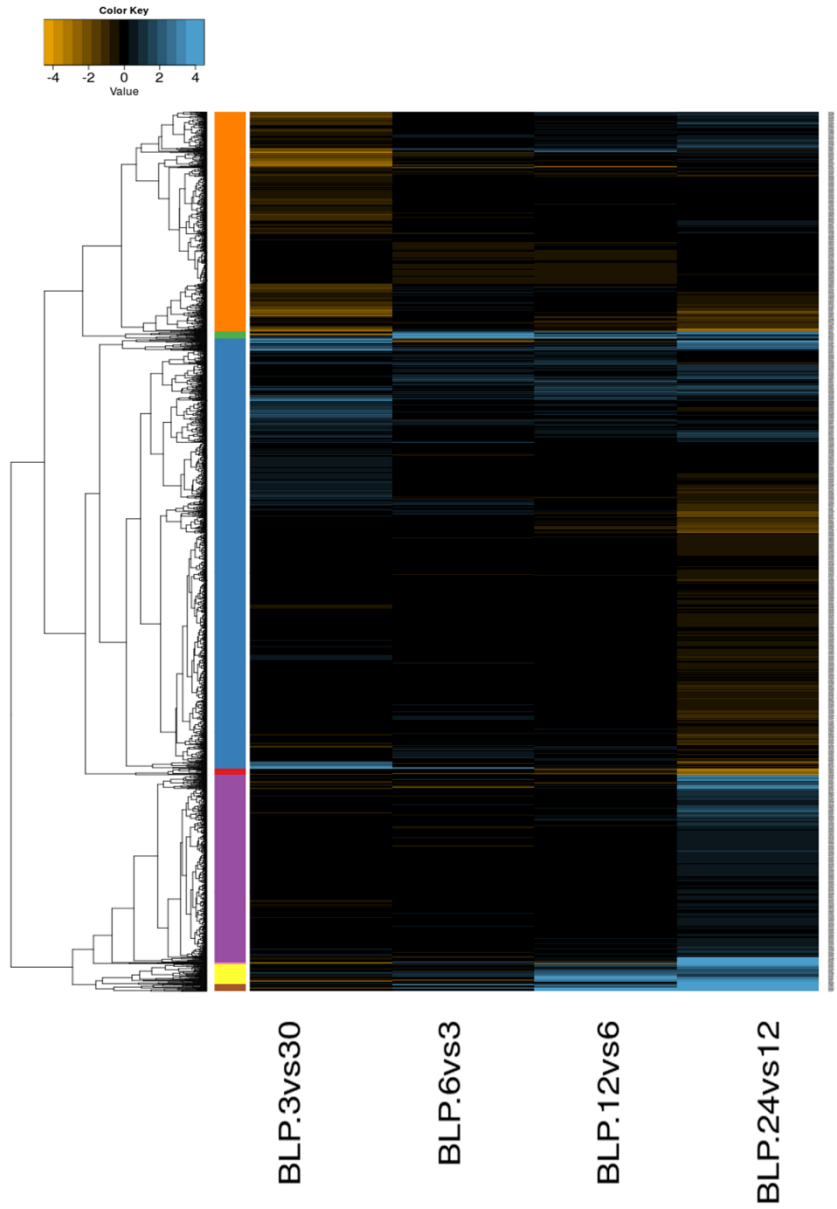

Figure 10. Heatmap of differentially expressed genes. Five comparisons were performed against 36hrs in darkness. LFC limit is 4 and -4, genes over or under those LFC have saturated colors. Complete UPGMA (Unweighted Pair Group Method with Arithmetic Mean) clustering algorithm. The x axis shows the comparisons between samples at different times, e.g. BLP.3Vs30 is the 3 h sample using as a control the sample at 30 min after BLP. The y axis shows the different clusters formed, the dendogram is composed by each differentially expressed gene.

Surprisingly, when compared the first, the second and third transitions they have approximately three hours of collection difference but the number of genes is much greater at the first one (30 min vs 3 h). On the other hand greater number of genes is activated in the fourth transition. This indicates that the number of genes that modify their pattern of expression does not only depend on the time elapsed during the process, but on the changes in a specific stage.

9.3.4. Gene Ontology of Genes differentially expressed by transitions

The analysis of transitions produced interesting GO terms enrichments. There was observed a repression of translation related genes starting at six hours, using three hours as control; this pattern is also present in each transition after six hours. This means that repression occurs from six hours to 24 hours and it increases at each stage.

Interestingly there is a repression of terms associated with secondary metabolism, carbohydrate metabolic and catabolic processes, ion transport, and cytoskeleton organization. I also observed a repression at the final transition of carboxylic acid metabolic and organonitrogen compounds metabolic processes at 24 h.

The up-regulated genes also had enrichments of GO terms. The first transition had an enrichment of terms related with cellular homeostasis, carboxylic and organonitrogen compounds metabolic processes. The second transition (three to 6 hours), had enrichments associated with lipid metabolic, carbohydrate metabolic and catabolic processes. Interestingly, at the last transition we found GO terms of organelle organization, response to stress, cell cycle, termination of signal transduction G-protein (proteins related with GCPRs inactivation found) and an odd term of symbiosis were enriched (Fig.11)**.**

Interestingly the cellular component GO term of ribosome is also enriched in by repressed genes at the later transition starting at six hours. These genes involved are associated with proteins associated with the ribosome (Fig. Supp. 3).

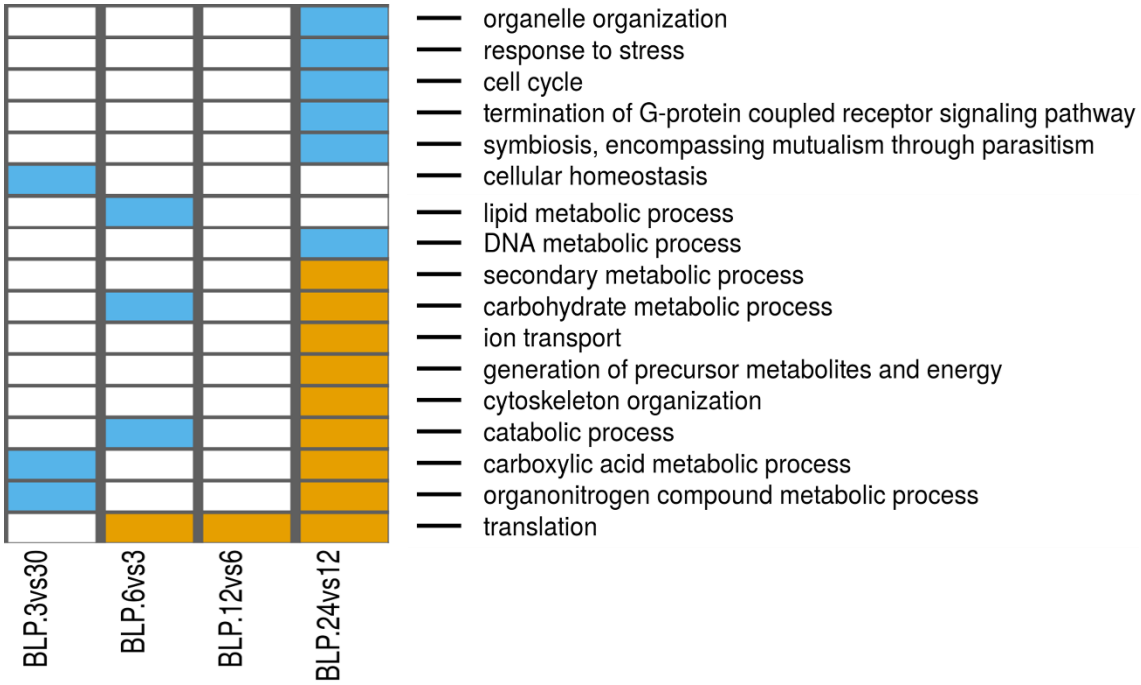

Figure 11. GO terms enriched by stage. Go terms enriched by upregulated genes are represented in blue and GO terms enriched by downregulated genes are represented in yellow. GO terms are written on the right. Enriched terms were considered when p-value was lower than .01 and an FDR lower than .05.

9.4. Transcriptional states during photoconidiation

Differential expression analysis quantifies and defines in a categorical way the activation or repression of genes but cannot explain the whole transcriptional state of samples. A methodology to accomplish this matter was published with an approach based in information theory, Shannon's entropy. The measure developed was called Diversity which can be defined as the level of transcriptional activation in the genome based in the relative gene expression in each sample. The measure gives information about the transcriptional state of the cells under study. Moreover, a derivation based in the transcriptional specificity of the genes explains the Specialization of the sample among all samples tested.

9.4.1. Transcriptomic diversity and specialization by sample.

To perform a *D* and S analysis we sum the replicates in a single set to measure transcriptional relativity. As more replicates are included, more robust the result is.

The *D* values observed among the samples where different. In the samples analyzed it is necessary to notice which one is the upper and lower limit of *S* and *D*. The less diverse sample was the 3h sample with a low specialization. The less specialized sample was that of 6 h after the BLP, which has an intermediate D value. Surprisingly, the most diverse and specialized sample was that of 24 h, this can only happen when the genes expressed are highly specific for that sample and there is a big number of genes being expressed evenly in the sample (Fig.12).

The relation between *D* and *S* could be expected as inversely proportional assuming that if few genes are activated hence the sample must be highly specialized. Instead we observed samples that were highly specific and had low diversity.

As previously described, time allows differentiation as(Higgs *et al*., 2007). Differentiation has consequences in trancriptional changes that results in different transcriptional profiles where different values of *S* and *D* can be observed (Reyes-valde, 2008). It could be expected that the largest geometric disctance could be between undifferentiated cells and mature conidiophores. However, the largest distance is between the three and 24 h samples (Fig. 12). This is owed to the transcriptional state at 3 h where cells are in homeostasis and have high translational activity while in mature conidiophores they are completely differentiated and prepared with transcripts for germination and reduced translation (Fig. 9).

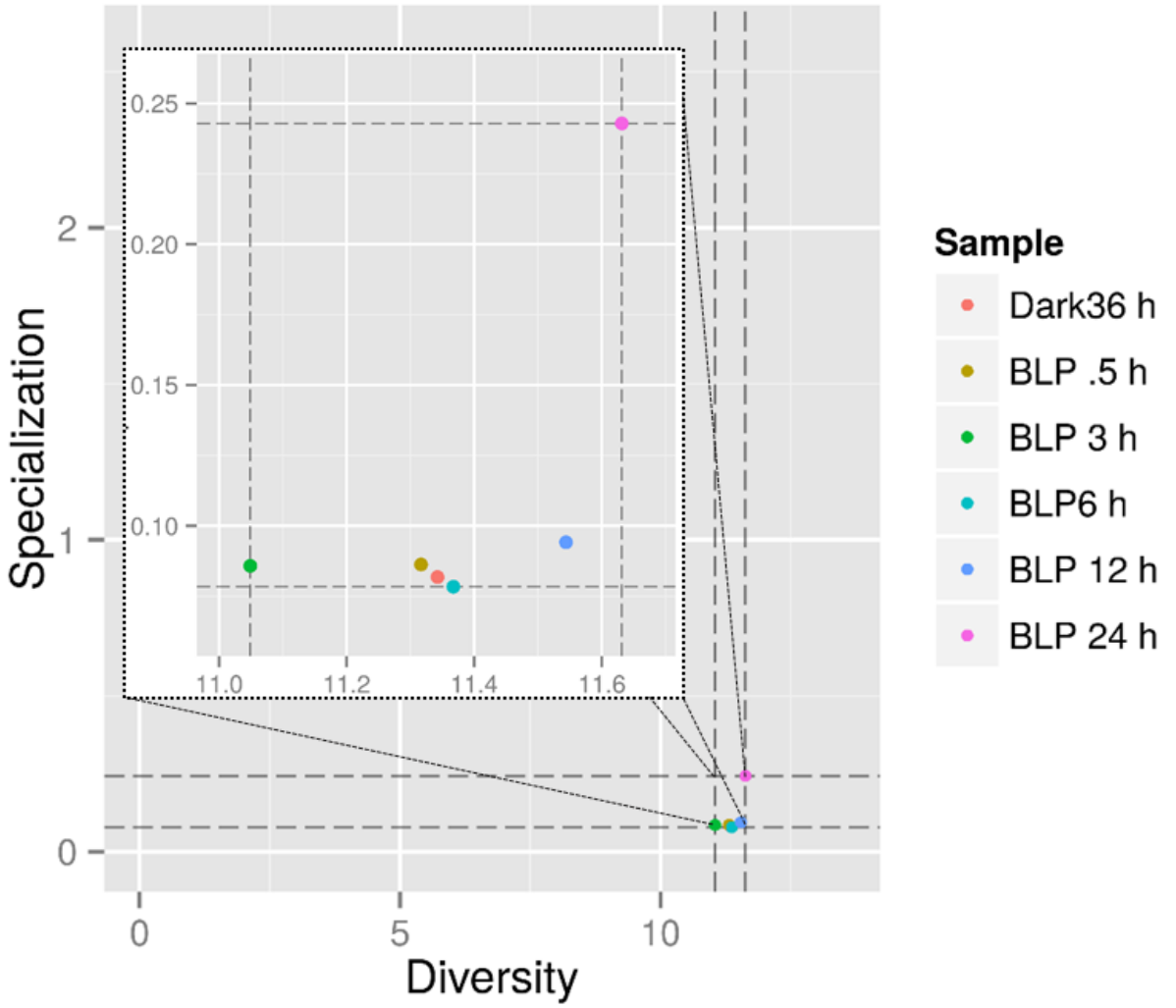

Figure 12. Diversity and specialization values. Diversity is in X axis and Specialization in Y axis. The lower *D* value is near 11.0 and the maximum is near 11.6. The max value that *D* can have is 13.53. *S* goes to values near to ~.07 to almost 0.25.

9.4.2. Differentially Expressed Genes and Gene Specificity values intersection

It has been suggested that the large changes of expression in DEG might be associated with the importance of a gene in each stage. On the other hand, the specificity of a gene gives information about the exclusivity of expression among certain conditions. Those hints are not strong enough separately. In such regard, I combined the strength of both methodologies looking for genes with high LFC and high specificity. This methodology allowed me to find genes with high transcriptional changes that might be highly specific for that change (Figure 12).

The points that are not upper outliers might owe their specificity to the increment of expression in LFC. When a high transcriptional change occurs it could be reflected in a higher specificity. On the other hand when a LFC is low but the gene has high specificity it occurs only because the activation of the gene is specific for that stage. In order to select genes that are highly specific but had a low LFC, I fitted a linear model between LFC values and specificity for up-regulated genes, then selected the upper outliers.

A total of 18 outliers were found in different stages. There were 3 at three hours, 3 at six hours, 10 at 12h and 2 at 24h. At the blue light response, at 30 min, there were no highly specific genes with low LFC. Among those genes only 4 are annotated.

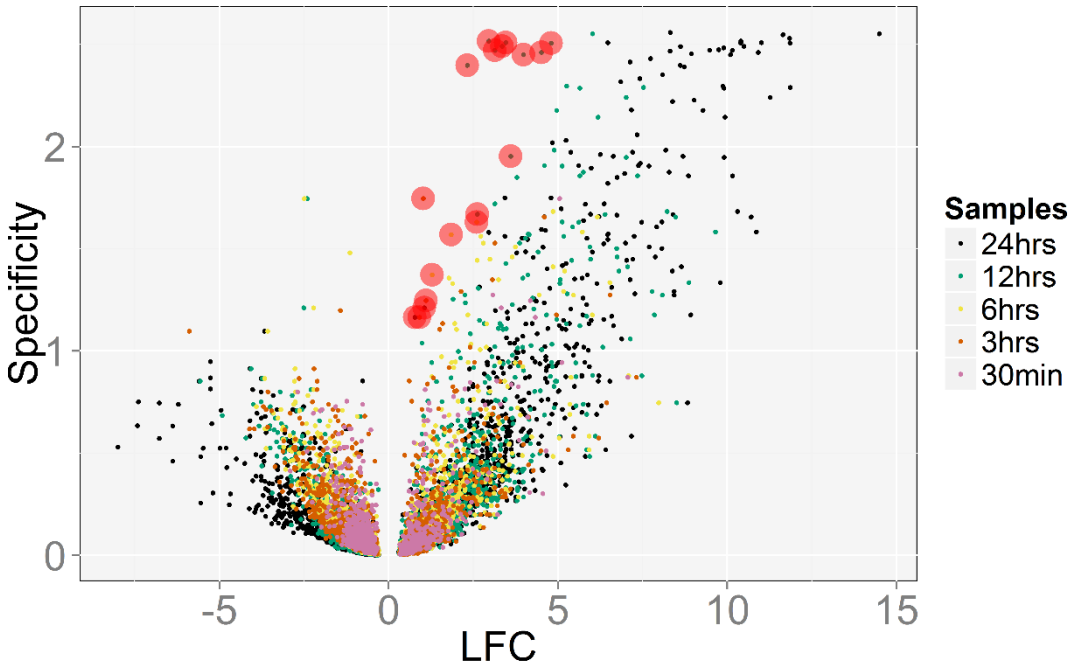

Figure 13. DEGs scattered using LFC and specificity values. A linear model fitted to upregulated genes. Slop= .322. Outliers selected under Bonferroni corrected $p < .05$. Outliers are marked in red transparent dots.

The limits of LFC between activated and repressed genes are different. The maximum LFC for activated genes was 14.2, while repressed genes had a minimum LFC of -7. This is owed to a low relative repression of genes after BLP, that repression allows the activation of other genes that were repressed in darkness, thus when they are activated the LFC was high. This is reflected by the

high specificity of genes activated at 12 and 24 hours. This effect explains the lack of repressed DEGs when microarrays where used (Rosales-Saavedra *et al*., 2006).

9.5. Gene clusters by genomic context

9.5.1. Gene clusters associated with differential expression

Cellular differentiation is directed by big changes in transcriptional patterns. In order to identify patterns of genomic regulation several reports in fungi have shown the presence of genes that are close to each other and have the same transcriptional profile (Gibbons *et al*., 2012; Garzia *et al*., 2013; Dhingra *et al*., 2013). To find those clusters they used the negative binomial probability for a certain number of genes activated or repressed within a set of contiguous genes.

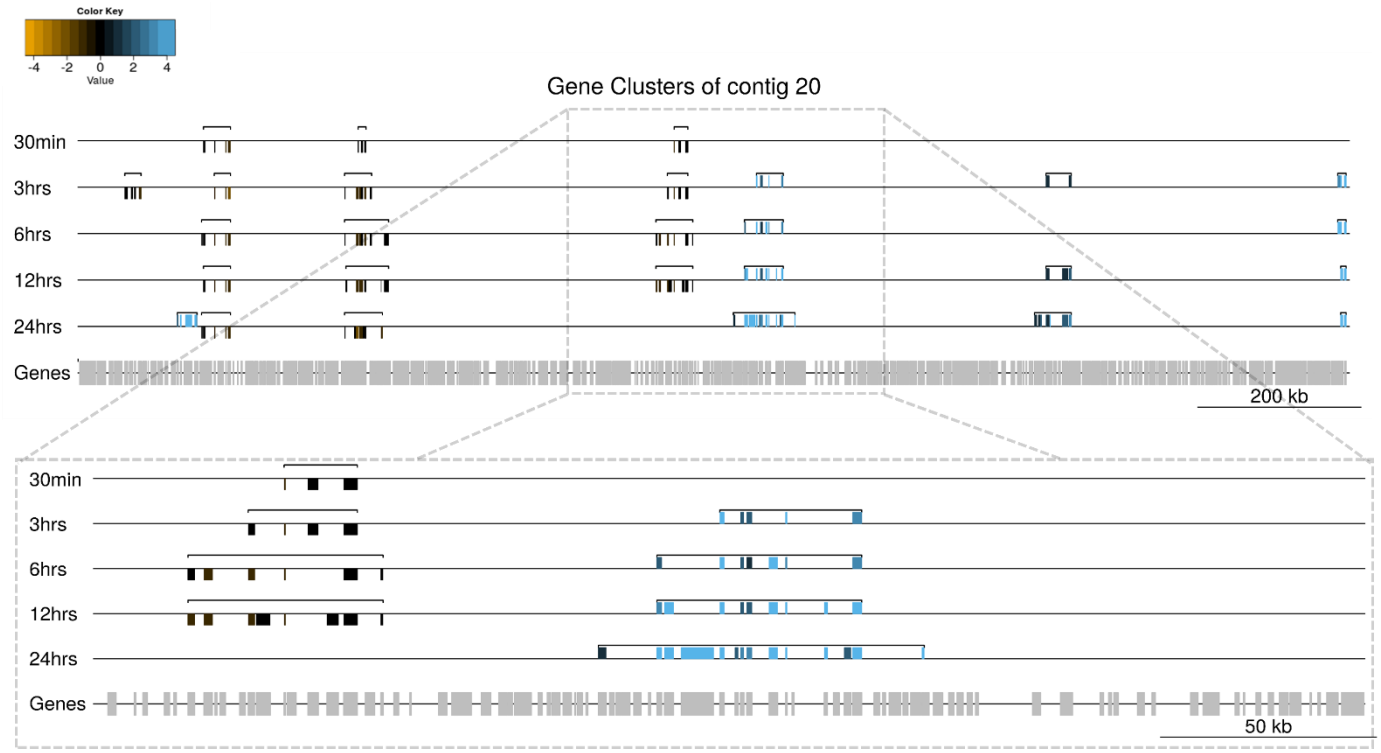

Figure 14. Gene clusters representation in the genome. Contig 20 of the *T. atroviride* genome was selected. A zoom is shown in the lower panel of the figure. LFC is associated with color saturation, dark ones have low LFC. Activated genes are above lines and are colored in blue. Repressed genes are under the line and are colored in yellow. This contig has clusters of differentially expressed genes in every stage of development.

I decided to develop a new methodology based in the Bayesian inference of clusters **(see materials and methods. Secc. 9.5.1)**. A window size seed of 35 Kb, which is nearly 0.1% of the genome; a shift size across the genome of 1 Kb were used. A cluster was consider significant if the probability to find that cluster was <0.01 in the null model. Then I collected all clusters found and concatenated them if they were overlapped. After that, is asked again for the significance of new concatenated clusters. There were searched odd clusters in every stage (Fig. 14)**.**

The Bayesian probability allows to determinate the probability to get an event based in the data itself. So, *a posteriori* probabilities are not the same for all window. A representation of how we can determinate odd clusters is presented in Fig.15

Genes found by window of 35 Kbps

Genes activated/repressed

	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14
0	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
1	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
2	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
3	FALSE	FALSE	FALSE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
4	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE
5	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE	FALSE	FALSE	FALSE
6	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	FALSE	TRUE	TRUE	TRUE	TRUE	FALSE
7	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
8	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
9	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE	TRUE
10	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE	TRUE
11	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE	TRUE
12	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE	TRUE
13	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE	TRUE
14	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	FALSE	TRUE

Figure 15. Example of a table of significant clusters of genes activated/repressed found in different gene number of contiguous genes. In the X axis are the number of contiguous genes in the genome (i). In Y axis the number of genes in cluster (j) that are considered odd when they are within i number of genes.

Multiple clusters are found at different stages, before concatenation the number of clusters is bigger due to the reducing effect of concatenating two elements and also the drop of gene clusters that are not significant when the size of the window has changed. This is observed in all the stages and also in clusters of up- and down-regulated genes. (Table 4)

Table 4.Number of clusters found for up-regulated and down-regulated genes in all stages. Odd clusters before and after concatenated are shown.

STAGE/CLUSTERS FOUND	30 MIN	3 HRS	6 HRS	12 HRS	24 HRS
ODD CLUSTERS DOWN-REGULATED	46	87	97	107	143
ODD CONCATENATED CLUSTERS DOWN-REGULATED	15	29	36	51	61
ODD CLUSTERS UP-REGULATED	25	51	86	138	150
ODD CONCATENATED CLUSTERS UP-REGULATED	30	46	46	49	59

9.5.2. Comparison between binomial method and Bayesian inference of gene clusters

I decided to compare both methodologies in order to understand which one will be a better to find gene clusters with the same feature within a group of contiguous genes.

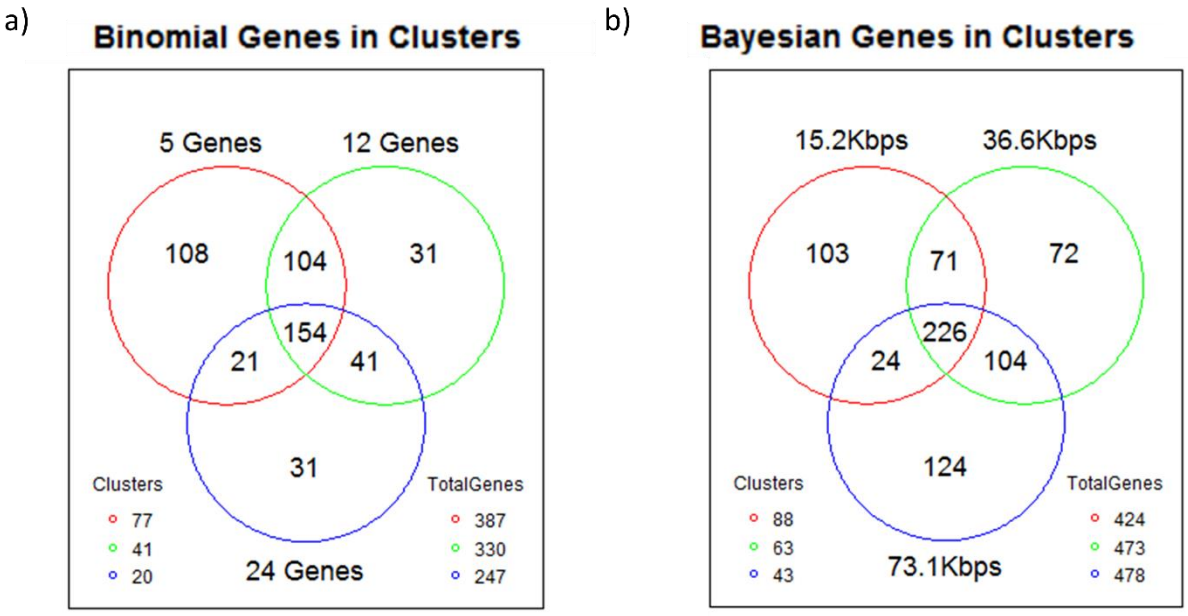

Figure 16. Comparison between binomial and Bayesian inference of clusters. A) Binomial inference of clusters with 5, 12 and 24 contiguous genes. B) Bayesian inference of clusters with expected number of contiguous genes, 5 genes in 15.2 Kbps, 12 genes in 36.6 Kbps and 24 genes in 73.1 Kbps.

Both methodologies are by philosophical principles different, frequentist and Bayesian probabilities have different assumptions. The first lies in the previously known data or known distributions, the second one takes what is in the data and obtains the probability to find certain values. I couldn't directly compare the results obtained with the two methodologies. Nevertheless, I used both Gene Cluster Finding (GCF) algorithms and compared them as closest as possible. I selected the set of the activated genes at 24hrs to test both algorithms.

To compare them equally, I made a slightly different inference of clusters for binomial algorithm. For binomial GCF I used the significance of genes within windows of 5, 12 and 24 genes. Then I concatenated overlapped clusters and calculated again the negative binomial probability to get the new cluster size. I found a total of 387, 330, and 247 genes in clusters, respectively. Among the genes in clusters, 154 were found in the three comparisons (Fig 16a).

There is no way to use a specific number of genes to perform GCF with Bayesian inference, so there was selected a window size with an expected number of genes. The sizes selected sizes were 15.2 Kb with an expected value to find 5 genes, 36.6 Kb to find 12 genes and 73.1 Kb to find 24 of genes. In these analyses there were found 424, 473 and 476 genes. 226 genes where present in all the GCF analyses (Fig 16b). While increasing the number of genes by window the number of clusters got reduced from 77 clusters to 20 using binomial method. On the other hand while increasing the window size from 15.2 to 73.1 Kbp with the Bayesian method got reduced to just half the number of clusters.

The frequentist methodology starts to lose clusters when the number of genes by cluster increases and hence loses overlapped genes with Bayesian method. On the other hand, the Bayesian method does not diminish the number of genes and still keeps many genes that cover great part of the genes found by both methodologies (Fig. 17). However neither of them maintains the same number of genes or clusters (Fig. 16).

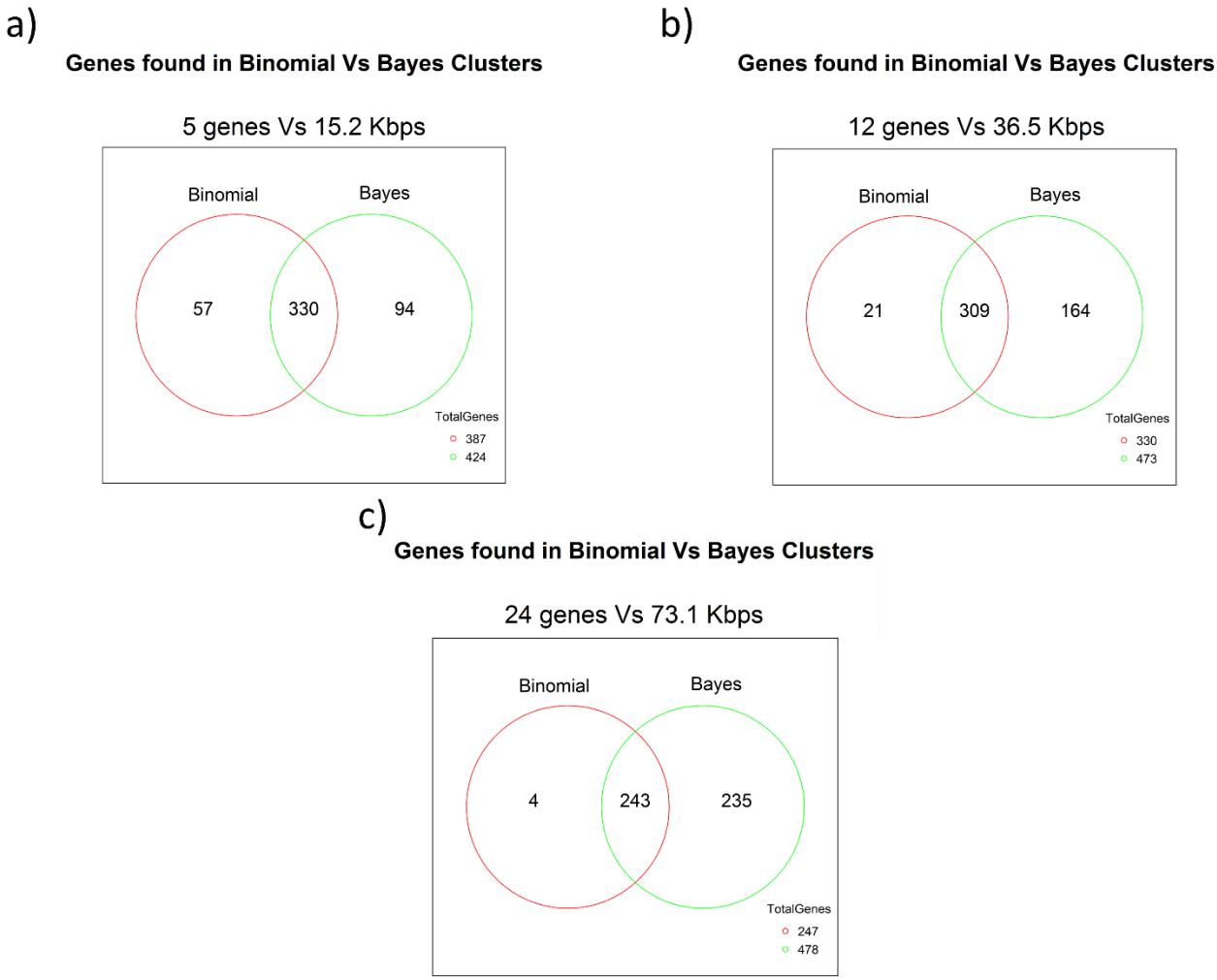

Figure 17. Overlapped genes in clusters between Binomial and Bayesian model. a) 5 genes *vs* 15.2 Kbps (5 genes expected). b) 12 genes *vs* 36.5 Kbps (12 genes expected). c) 24 genes *vs* 73.1Kbps (24 genes expected).

9.6. DNA methylation during photoconidiation process

9.6.1. Inhibition of DNA methylation

Genomic regulation is controlled by different molecular elements, one of them is the chromatin. Chromatin is a structure composed by multiple elements such as histone modifications, histone deposition and covalent changes to DNA. One of the most studied is DNA methylation.

I hypothesized that DNA methylation could play a role during the development of photoconidiation in *T. atroviride*. To test it, I performed an essay to inhibit DNA methylation using 5-Azacitydine.

I used different increasing concentrations of 5-Azacytidine to study its effect in the number of conidia produced (Fig.17). I observed increasing production of conidia with increasing concentrations of the inhibitor, being highest at 100µM. Concentrations higher than 200µM turned out to be toxic for *T. atroviride* (Data not shown).

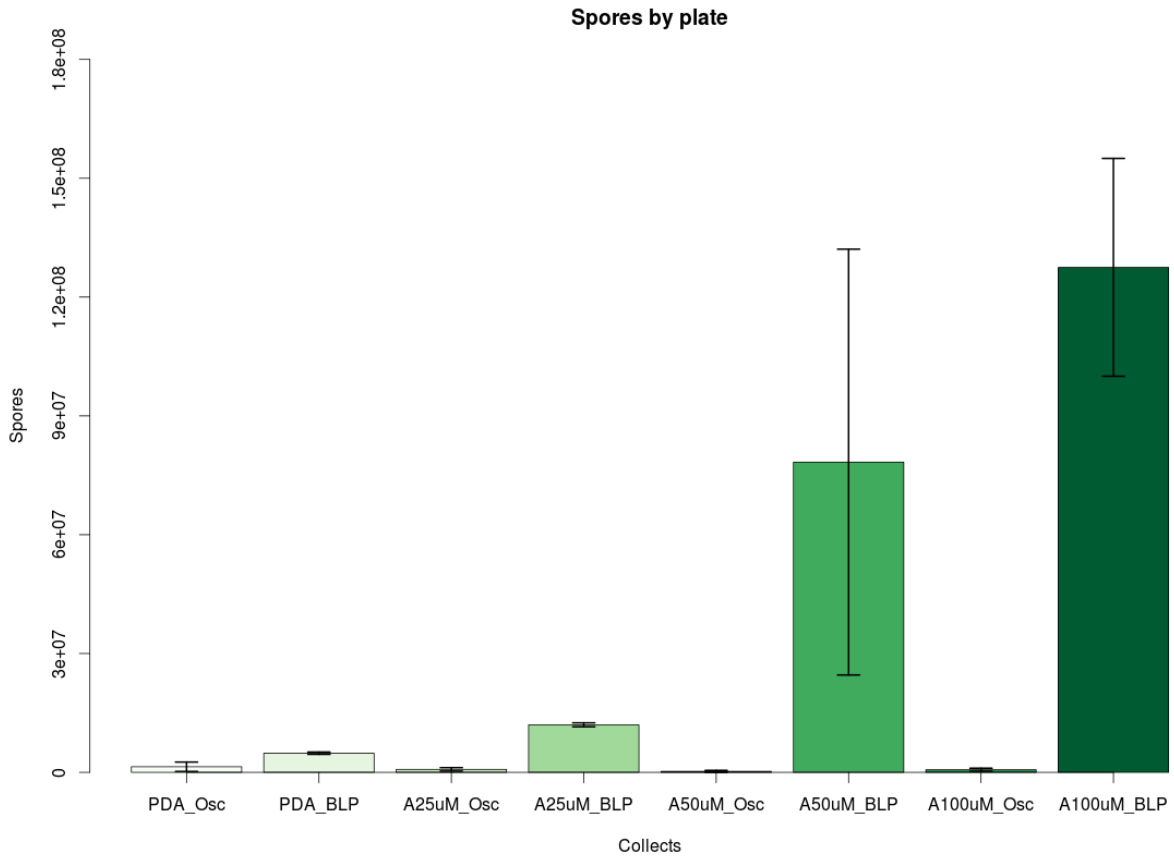

Figure 18. Effect of DNA methylation inhibition. Conidiation is increased in relation with inhibition of DNA methylation. Sample with 100µM with BLP is significantly different from samples in PDA/5-Aza-Free with BLP. All the samples in darkness were significantly different from BLP samples. Tukey test. Confidence level 95%. p = .05. Two biological replicated were made with the same result.

9.6.2. Analytical quantification of DNA methylation

With the help of Dr. Kazimiers Wrobel, we detected DNA methylation in *T. atroviride* in the samples between ranging from 0.5 to 1 percent. Quantification of DNA methylation of samples in darkness and after BLP showed no overall changes in DNA methylation but there was a significant difference between mean samples of 6 h after BLP and 12 h after BLP (Fig. 18).

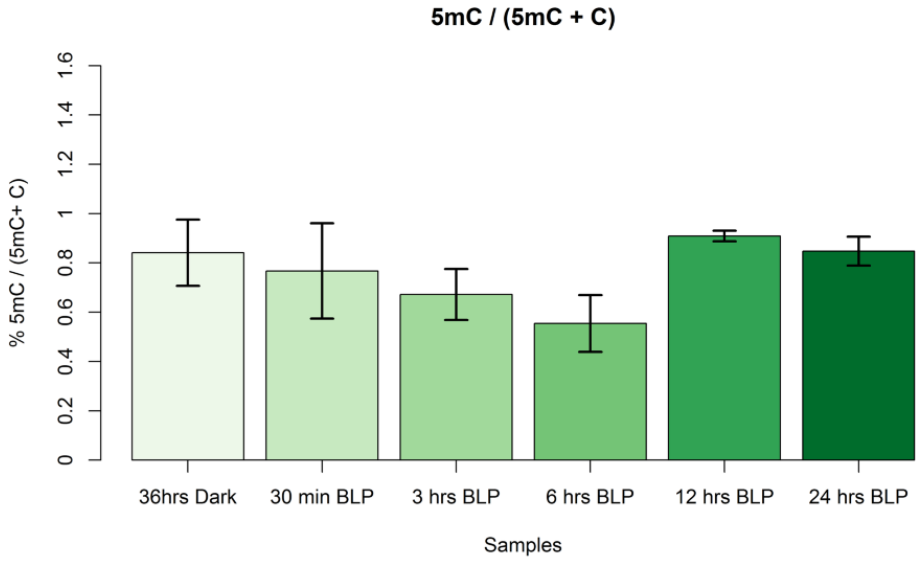

Figure 19. % of 5mC/ (5mC +C). Significant changes were observed between six and 12 hours. Tukey test. Confidence level 95%. p = .05

10. Discussion

Cellular differentiation is a complex phenomenon where different kinds of cells that have the same genome information are generated. This is achieved by the interaction of molecular elements in the cell and the response to environment.

Transcriptional analyses help understanding cellular phenomena as a compendium of a myriad of events. I decided to start with a numerical and quantitative approach, which is based on the assumption that statistically significant measures can be classified in order to make categorical comparisons as well as quantity inferences. This methodology allows to get a general interpretation of the transcriptional state of the cells as well as the changes during a certain process.

Asexual reproduction has been studied for several years given the capacity of the spores to resist and disperse. Spores are used in the food industry, as biocontrol agents, inoculum for antibiotics production among other activities. *Trichoderma atroviride* has biotechnological interest as a biocontrol agent which application method is by spores dispersion. Understanding spore production could help to improve the quality of spores production as well as increase the yield.

Different transcriptional analyses have been made in order to study the early stages of asexual reproduction in different fungi, including *Aspergillus nidulans* (Gibbons *et al*., 2012) and *Trichoderma atroviride* (Hernández-Oñate *et al*., 2012). Similar analyses have been carried out but considered only the final stage of asexual development (Judelson *et al*., 2008) and there is one that specifically compared vegetative cells and conidia in *Magnaporthe orzyae* (Kim & Lee, 2012). In this study the apical zone was taken before BLP this is a stage when cells are highly proliferative; thus, germination could be considered the opposite process, because it goes from conidia to cellular proliferation with mostly apical zones. In such regard, some studies have been carried out in *Aspergillus fumigatus* (Lamarre *et al*., 2008)*, Aspergillus niger* (Novodvorska *et al*., 2013) and *Fusarium graminearum* (Seong, 2008). This work, however, to the best of my knowledge, represents the first transcriptomic analysis through the whole conidiation process from the induction to the production of mature conidia. Additionally, the

combinatorial strategy used in this study is novel in terms of the informational power as well as the global understanding of this developmental program.

In order to obtain a deeper understanding and obtain *bona fide* information, I decided to enrich the biological samples. It was collected the conidiating ring formed after BLP, this kind of sampling enriches the biological material that is actually responding to light and hence it enriches the RNA sample associated with conidiation. The selection and mixing of three biological replicates reduces noise of spurious transcriptional values or outliers.. This method adds robustness to differential expression analysis. This scenario can be imagined where an outlier might be present in a single cell but the other two reduce that average value; moreover a second replicate with the same principle helps us identify the variance and accomplish DEGs. On the other hand, still persists the possibility to have a sample that may not be well treated hence might not represent the light response, however this is overcome by the dispersion analyses made in the Differential Expression Analysis.

It was observed a high transcriptional rearrangement based on the DEA. Nearly 21 % of the genome is differentially expressed during development for genes that were either repressed or activated. The number of genes with transcriptional changes increases as the developmental process advances, suggesting the establishment of a new cell type.

Interestingly, there was observed a higher number of genes repressed in samples at early stages than activated ones. The ratio between activated and repressed genes reaches almost 1 at final stages. In a previous work (Rosales-Saavedra *et al*., 2006) there were found more genes activated (30 genes) than repressed (10 genes) at 30 minutes after BLP. This may be explained either for the low data depth derived from the microarray or the used probes which included 3’ UTR regions and not CDSs, as the transcriptome does.

The comparison between theactivated genes at 30 min and the previous work, showed that only 14 can be matched to the current genome and from those, five were up-regulated too; one of them is *phr1* which has been reported as a photolyase that repairs the DNA damage by light, by the photoreactivation of pyrimidine cyclobutane dimers.

Among the repressed genes, one was not possible to match with the genome. From the nine genes remaining we observed the repression of three. The genes found have not function reported in conidiation.

Sadly, most of the possible comparison between Rosales-Saavedra and this work are limited by microarrays constrictions such as design (probe selection of 3' UTRs) and background noise.

The depth of transcriptomic data allowed to see a more accurate relation between transcriptional repression and activation. The distribution of LFC is more biased to higher values in activated genes. This indicates that the genome is being highly activated in the conidiation process (Fig. 12).

In transitional DEA we observed DEG that change from one stage after BLP to the next one. In the transition between 30 min to three hours there is a greater proportion of genes being repressed. This is also observed from 3hrs to six hours, but after those transition the ratio changes and the genome appears to start an activation process. This is in agreement with the DEA made by stage, where I observed that the ratio between activated/repressed genes begins to approach 1. This suggests a cellular response after BLP in order to contend with blue light and prepare the cells to start a conidiation program.

Among the genes found to be increasingly activated there were five annotated genes. A C_2H_2Zn TF (Gene ID. 310560), two peptidases, an ankyrin repeat protein and a protein that belongs to the major facilitator superfamily.

Regarding increasingly repressed genes, 29 were annotated out of 40 with different function. Among the genes annotated there was a protein of the Calcipressin family, a Phospholipase-C like protein, different enzymes and transporters with no reported functions associated with conidiation.

No single GO term was found in the activating pattern that had 8 genes nor in the repressing pattern that had 51 genes.

The DEAs suggest two transcriptional phases based on the transcripts population before and after six hours. However the DEAs cannot tell the importance of those changes in biological terms.

GO enrichment analyses are restricted by the number of genes annotated and associated with GO terms. This means that the enrichment of terms is biased by the number and kind of genes previously described in literature and databases. BLAST2GO controls the first issue because it uses Fischer's exact test which uses annotated and not annotated genes in each numerical comparison (Blüthgen *et al*., 2005). This methodology allows to compare from a global reference background (Huang *et al*., 2009).

Three main GO categories are available, cellular component, function and process. We performed three of them and used the cellular process category based on the clarity of results.

GO terms enriched for activated genes where found at every stage. At 30 min DNA metabolic process was enriched, some of the genes found are a DNA Mismatch Repair Protein msh6 (ID.252385), a Base Excision DNA Repair Protein (ID. 26345), *phr1* a photolyase encoding gene (ID. 26345) (Berrocal-Tito *et al*., 1999) and a DNA ligase (ID. 133844). It is known that one of the first responses to blue light is DNA repair (Rosales-Saavedra *et al*., 2006; Berrocal-Tito *et al*., 2007) Thus, I conclude that the main transcriptional response at 30 min is associated with DNA repair.

A single GO term was enriched at 3 h, related with protein metabolic process. Among the genes activated in such GO term we found 16 tRNA synthases. Also two heat Shock Proteins, HSP90 (ID. 297563) and HSP60 (ID. 297734) were found and several genes related with the synthesis of aminoacids**.** Also when analyzed the cellular component GO terms enriched for activated genes the endoplasmatic reticulum where enriched. This evidence indicates high production of proteins, probably the proteins necessary to conidiate before the 6-8h threshold, time after which transcription and translation of *de novo* proteins are no longer needed to conidiation.

Carboxylic acid metabolic and organonitrogen compound metabolic processes were found enriched at six and 12 hours that are also associated with synthesis of amino acids and other related pathways. 12 hours is one of the morphologically described stages during photoconidiation. At this time the conidiophore is almost complete but determinated, phialides are already present but conidia are not.

Finally at 24 h, DNA metabolic process is enriched but with different genes for DNA repair from those found at 30 min. For instance, a DNA lyase which cleaves apurinc/apyrimidic sites and removes 3'-blocking groups at single strand breaks of damaged DNA is activated (ID. 225032); this one is not the photolyase found activated at 30 min. There are two ATP-DNA dependent DNA helicase (IDs. 284954 & 234113). At this stage we found a total of 46 genes compared with 13 genes found at 30min. This number is increased not only by the DNA damage response genes, but also by the genes that start replication such as subunits of the origin recognition complex among others. Several genes found in DNA metabolic process are involved in response to stimulus and single-organisms cellular processes. Surprisingly, *blr*1 is slightly activated at 24h after BLP.

For repressed genes we found GO terms enriched in all stages, except at 3hrs.

At 30 min lipid metabolic process is repressed. Among the genes found at 30 min, there is a hypothetical Phospholipase C (ID. 157590) with homologue in *N. crassa* (ID. NCU02175), which has been associated with Ca+ signaling (Zelter *et al*., 2004). There are also two subunits of synthases of fatty acids (IDs. 226146 & 85662). It can be observed that at six hours lipid metabolic process is also enriched, both times share 21 genes, out of 37 at 30 min and 24 at six hours. There is no enrichment at 3 h associated with lipid metabolic process. We can hypothesize that transcription is compromised to contend with DNA damage at 30 min and to keep the development of conidiophore at six hours, while at three hour transcription is compromised with protein synthesis. The enrichment of terms in separated stages suggests that the differentiation process requires the activation or repression of genes at certain stages without a progressive continuum transcriptional profile.

At 12 h when conidiophore is determinated after the BLP, cellular component organization is repressed. In summary, this term represents the major components of the cell for replication, transcription, translation and cellular transport. This term is not associated with metabolism. To mention a few of this kind of genes, in this list, there are ATPases, cytoskeleton components, histones and ribosomal proteins; a wide range of protein classes fall in this GO term. This suggests that cells have started to stop the processes associated cell division and components necessary to produce proteins.

At the 24 h genes associated with cytoskeleton organization is repressed. This set shares seven out of 14 genes with cellular component organization at the previous stage.

One of the most interesting GO terms repressed in DEA by stage is the translation process GO term. This is present at three stages, 6, 12 and 24 h with 53, 118 and 144 genes respectively. Six and 12 hours shared 51 genes, 12 and 24 hours shared 101 and all of them shared 47 genes. This genes are mainly ribosomal proteins and translation initiation factors.

The apparent repression of translation that starts six hours after stimulation might reflect the preparation of entry into a dormant state (Lamarre *et al*., 2008). The last stage also indicates the lack of metabolic activity in the cell. This effect is reinforced with the repression of ion transport (a highly energetic process), generation of precursor metabolites and energy, cytoskeleton organization and catabolic process GO terms at 24 h.

It is noteworthy that we observed the repression of genes that work in the transcription pre-initiation complex at different times. TFIIA-γ is repressed at three, six and twelve hours. In addition, TFIIE is repressed at six hours and TFIID subunit 14 at 12 hours. TFIIA-γ is the small subunit of the complex. In this regard, the function of TFIIA is conserved among the domain eukarya and is responsible for the activation of transcription (Ozer *et al*., 1994). TFIIE binds directly with RNA-Pol II and recruits other components of the RNA-Pol II holoenzyme (Roeder, 1996). Finally the function of subunit 14 of TFIID is not clear, but mutation of the corresponding gene in *Schizosaccharomyces pombe* indicates that it is responsible of contending with certain stresses and works with the TIFFD complex (Kimura & Ishihama, 2004) and changes in their expression influence the formation of RNA-Pol II holoenzyme, and hence affect gene expression (Colgan & Manley, 1992). The regulation of the subunits of these complexes suggests a fine-tuning in the repression of genes at three, six and twelve hours. When we compared the relation of GOEAs by transition and stage we observed very concordant results between them. The use of transitions allowed us to increase our knowledge of the stages. In such regard, the enrichment of translation term in the last three transitions indicates an increasing repression of translation as conidia reach maturity.

Carboxylic acid and organonitrogen compounds metabolic processes are enriched in transitions from 30min-3h, but not in 3-6 h nor in 6-12 h while by stages at six and 12 hours those terms were enriched. I did not observe an increment of those terms at 3 h because the number of genes associated was not enough compared with darkness control. However, metabolic process term was enriched, this is a highly related term. A similar contrasting effect is observed also when these terms are enriched in the last transition but not in the last stage.

The fact that repressed genes of the term related with metabolism at the last transition are also present at the last stage is noteworthy. This not only means that the genes associated with metabolism are silenced in the last transition also indicates a repression of metabolism compared with darkness. Silencing of metabolism is highly concordant with the other GO terms by stage and transition. However, the enrichment of organelle organization, response to stress and cell cycle by activated genes at the last transition indicates that the cells are transcribing those genes to be active when dormancy is finished and germination starts. This hypothesis is supported by the repression of genes associated with translation in those stages.

In the last stages there is substantial repression of secondary metabolism related genes. However, this is due to the number of genes associated but not all secondary metabolism genes are repressed. In fact, there were eight Polyketide Synthases activated and 2 repressed at 24hrs.

Among the GO terms enriched in different stages and transitions, is clear to see two kind of patterns, one before six hours when enrichment are non-continuous and could be intermittent and a second one after six hours where in general the transcriptional profile is sustained and increased. These patterns suggest that the functional purposes of transcription are different, the first one is compromised with the response to light and the second the establishment of a new cell type.

So far we have discussed the functional relevance of DEGs but differential expression analyses cannot explain or elucidate the transcriptional state of the cells during the photoconidiation process. For this purpose *D* and *S* where obtained.

The lowest value of *D* is at 3 h this means that the genes have a biased transcription for certain genes and its low *S* indicates that those genes are not highly specific. This could be due to the transcription of housekeeping genes or primary metabolism rather than secondary metabolism of stress genes. Inadvertently, 3 h is the stage with less GO enrichments even though they have high number of activated and repressed genes, the one term enriched is protein synthesis and in the transition between 30 min to 3 h there is an enrichment of homeostatic process. This suggest that 3 h is the moment when cells have contended with blue light stress and have started to synthesize proteins necessary to conidiation.

The 6 h sample showed the lowest *S* value. This is not so clear based in the number of genes activated and repressed. The only reason this could occur is that the genes present at that time are so not specific among the conditions analyzed. This means the transcriptional levels of the genes in that moment recall the transcriptional levels of the genes in other samples. This occurs because *S* depends on specificity value that depends on the relative frequencies among all samples.

The highest *D* and *S* values are for the 24 h sample. This can only occur when there is a high number of genes activated with more even transcriptional levels within the sample all expressed genes and the fact that activated genes are highly specific. This means the conidia have transcribed genes for dormancy and germination that are specific for either of the two conditions. The unitary transcription contribution in *S* is observed in Fig. 12 by high levels of specificity for a great number of genes. When analyzed the influence of transcriptional activation on the gene specificity (Fig. 13) Several interesting genes were found. At six hours, a cytochrome p450 (ID. 183029) and a C2H2 zinc finger TF (ID. 167723). At 12 h, an Aspartyl-tRNA Synthetase (ID. 236434) and a conidial Pigment Polyketide Synthase *alb*1 (ID. 154816). The mutant of *alb1* in *A. fumigatus* produces deformed and albino conidia (Sugui *et al*., 2007).

Is noteworthy that during photoconidiation *S* and D values do not change greatly reaching the limits, but the difference between the lowest and highest *D* value could be considered as the change of 1,000 genes that are being expressed evenly. Particularly, *S* values do not changed greatly this could be owed by the

lack of sample beside photoconidiation, however at 24 h there is an increase of *S* owed to the sum of the specificity of the genes expressed when conidia are mature.

Figure 13 also shows that the log fold changes by repressed genes are in general smaller compared with activated genes. This is reflected in the distribution of genes between the limits of LFC. This effect is not produced by the number of DEGs because at 24h the number of genes is almost the same and the absolute upper and lower limits of LFC are not the same. This shows a fine tuning of gene repression while high transcriptional activation occurs.

High transcriptional activation changes are observed for genes that are also highly specific. This is owed to the sudden expression of genes that were highly repressed previously with low relative frequency of reads.

In all the analyses performed, six hours turns out to be a developmental stage of transcriptional transition between early and late activated genes. This can be observed in the level of *S* and the GO profile at six hours which changes drastically from three to six hours, where translation is repressed and organonitrogen and carboxylic metabolic processes are activated. In agreement with the observations made for the six hours sample, it has been proven that inhibition of transcription after BLP near 6 hours inhibits conidiation but beyond that time, inhibition of transcription and translation has no effect in conidia formation. After that time the transcriptional activation is associated. This leaves for later stages the expression of genes for conidia maturation, dormancy and germination.

This led us to infer two major transcriptional phases and four functional transcriptional phases based in information theory and GOEAs of DEGs respectively.

The transcriptional phases are defined before and after six hours. Functional transcriptional phases are at 30 min the light response which involves direct DNA damage repair, a second phase for protein synthesis at 3 h. The protein synthesis phase is present at six and 12 hours with almost the same GO enrichment profiles and a final phase for dormancy which involves the reduction of metabolism as well as the activation of genes for germination.

Previously, a transcriptome analysis of the transition from vegetative growth to the early stages of asexual development in *A. nidulans* was published (five hours after induction). They found genes related with the terms enriched at six hours in this study such as carbon and nitrogen metabolism. On the other hand, lipid metabolism is activated in *A. nidulans* while it is repressed at six hours by stage but by transition is activated from three to six hours in *T. atroviride*. This might be owed to i) the stage comparison is with the apical zone where lipid production is highly active while by transition cells had to respond to the stress provoked by exposure to blue light and ii) because they have taken their sample from whole mycelia and not the apical zone where transcripts of lipid metabolism might be masked and when during conidiation those are strong enough to be detected.

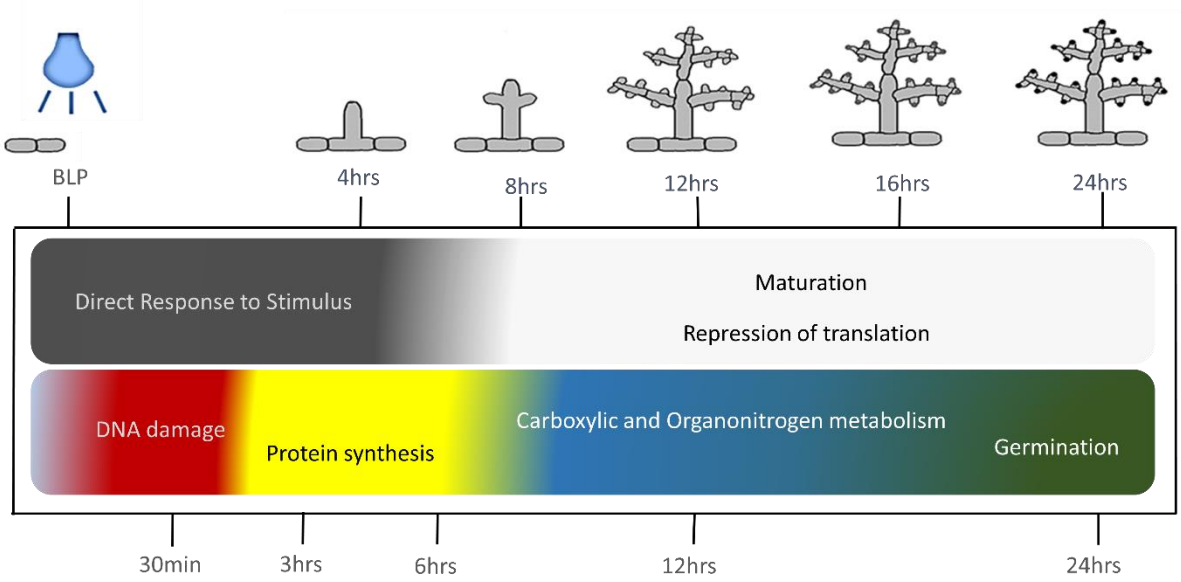

Figure 20. Transcriptional phases during photoconidiation. On the upper panel morphogenesis is described along the maturation process. The gray and white panels are the transition of the transcripts population before and after six hours. The lower panel represents the four functional phases based in the activation of genes in red color DNA damage as the main phase at 30 min; in yellow protein synthesis is the main function related with transcripts activated; in blue the genes associated with carboxylic and organonitrogen metabolisms and finally in green the genes associated with germination. The upper time line represents the reported morphologic changes at certain times and the lower one shows the samples taken in this study; both are not in scale.

Is noteworthy that they found at 5 hours a response to oxidative stress probably associated with the induction of conidiation in *A. nidulans*. This fact allows me to hypothesize that probably after some kind of stress cells have to make a decision to differentiate based in the intensity and kind of stimulus. This is observed also in *Trichoderma*, since conidiation is induced by different stresses.

The GCF algorithm found an increasing number of gene clusters while time differentiation progresses. Is noteworthy that the number of genes activated or repressed in clusters for each condition is higher than expected (Fischer Test $p < .00001$). This means that many of the genes regulated are in general near to each other. For instance, 53 out of 284 genes where found activated in 15 clusters, a great number assuming random activation of genes (3 genes by chance at clusters) in the genome. This suggest a genomic regulation by regions or sections based in the chromatin state of the cell e.g. if a section of DNA is more “open”, it is more susceptible to transcriptional changes.

Interestingly, there are seven genes in clusters that are constantly activated with unknown function and 23 genes constantly repressed. Among the genes repressed there were a Phospholipase C (ID. 157590) and a highly conserved Oxysterol Binding Protein (I.D. 155704) which has a homolog in *S. cerevisiae* named *osh3* where its overexpression induces pseudo hyphal growth, it has been suggested that is involved in hyphal development but no effect in conidiation (Park *et al*., 2002).

The aforementioned transcriptional changes are related with the establishment of a new cellular type through a differentiation process. The differentiation is led by transcriptional changes that might be regulated by chromatin.

Just to mention an example, the biggest cluster has 15 genes activated from a total of 24, from those 15 genes 9 were annotated. The genes found were two amidases, a cyclohexanone monooxygenase, a hormone sensitive lipase, an ankyrin repeat protein, a protein with Tpr domain, a transcription factor, protein with a leucine rich domain and a G protein. Apparently is difficult to associate common function in clusters despite the presence of two amidase proteins.

When analyzed the genes in repressed clusters is also not so easy to find related functions. However is noteworthy that from the 144 genes associated with translation 37 are in clusters at 24h. This makes up 25% of the genes associated with translation.

It is worth mentioning that some of the genes associated with RNAi machinery are activated**.** For instance, the genes previously reported that affect photoconidiation are *dcr2* and *rdr*3 both are differentially expressed at certain

stages in photoconidiation (Carreras-Villaseñor *et al*, 2013). Moreover, other genes that are not essential for photoconidiation are *rdr*1, *ago*-1 and *ago*-2 and they are being up-regulated at different stages. This suggests that they are could play a different role while photoconidiation occurs.

When we analyzed the changes in DNA methylation quantities we observed a significant change between six and twelve hours. The change cannot explain directly the activation nor the repression of a certain group of genes. In order to know it in a more detailed manner, it is necessary to detect at a single base resolution the methylation on DNA.

Table 5. Genes of RNAi machinery during photoconidiation. Blue color is proportional with LFC values.

Gene / Stage	30min	3h	6h	12h	24h
dcr1	**0,00**	**0,00**	**0,00**	**0,00**	**0,00**
dcr2	**0,00**	**0,00**	**0,80**	**0,93**	**0,00**
rdr1	**0,00**	**1,05**	**0,97**	**1,15**	**1,32**
rdr2	**0,00**	**0,00**	**0,00**	**0,00**	**0,00**
rdr3	**0,00**	**0,68**	**1,43**	**1,92**	**1,98**
ago-1	**0,00**	**0,00**	**0,35**	**0,40**	**0,00**
ago-2	**0,00**	**0,70**	**0,88**	**1,01**	**1,09**
ago-3	**0,00**	**0,00**	**0,00**	**0,00**	**0,00**

On the other hand, we measured the effect of the inhibition of DNA methylation, this result indicates the biological role of 5mC but that experiment cannot explain the molecular role of DNA methylation in activation or repression of genes. The data presented high standard deviation, maybe owed by the lack of replicate in each measurement, however even with those dispersions can be inferred the positive effect in spores production when methylation is inhibited.

I hypothesize that in the case that 5mC acts as repressor mark, the lack of DNA methylation could activate easier the genes necessary for conidiation. On the other hand in case that DNA methylation act as activator the lack of methylation could facilitate the repression of the genes necessary to conidiate. Furthermore, we considered the lack of 5mC might keep the DNA in an “open” conformation

susceptible to damage when exposed to BLP and this might trigger an unknown exacerbation of conidiation.

The most discordant fact is that the genes that might be responsible for DNA methylation (DIM-2 and RIP) have low expression values. They are no DEG nor even considered for DEG because their expression was considered noise. This leaves the remaining question if there is other possible DNA methylation protein in *Trichoderma atroviride.* The observations made regarding the role of DNA methylation in conidiation are not conclusive.

11. Conclusions and perspectives

I measured and quantified the transcriptional states and changes of T*. atroviride* during the photoconidiation process. This is the first study of its kind. This was possible by the sequencing depth and the selection of times during the formation of mature conidia.

In this study two approaches were performed. One is the transcriptional state of the genes expressed and repressed in a holistic manner based in S and D values and the other is the differential expression analysis with the GO terms enrichment analysis of genes differentially expressed.

In the first one, changes that indicate a global change in the population of transcripts expressed before and after six hours were observed. These changes represent the activation or the repression of the genome at certain stages. These phases do not reflect directly the biological relevance of the activated or repressed genes. Nonetheless, the numerical quantification of the present transcripts during the photoconodiation process, shows a changing population in the stages after BLP to six hours and conidia maturation

On the other hand when used the classification of genes expressed and repressed we observed the repression of translation that starts at six hours: while an increment of transcription at six hours that continues until 24 h. These results can only be coherent when dormancy is considered at the final stage of development when metabolism and primary cellular functions are turned down. However, we observed this phenomenon at the transcriptional level, but remains to be proven at the protein level.

When both approaches were combined, the relation between transcription changes and the gene specificity, revealed that activation of genes is more abrupt while repression is subjected to a fine repression (Fig. 13).

Transcriptional analyses defined four transcriptional phases based on the functional categorization of genes expressed and repressed. Each phase is defined with not strict boundaries in time. For instance, response to DNA damage is before protein synthesis, but we could not determinate the critical point when synthesis of proteins is present. In order to understand how each phase is

established it is necessary to measure the expression of genes with known molecular function during the photoconidiation development. By principle, with this approach we could observe how these genes are being activated and repressed at certain time.

Surprisingly at the final stage we observe high repression of genes necessary to translation but high transcriptional activity. I assumed that in case of repression of translation this might be occupied by genes with high affinity for the ribosome or highly expressed genes. But we couldn't determine the functional role of repressed genes associated with translation.

On the other hand, based on the expression of functional categories that help cells proliferate, respond to stimulus, and contend with the damage of DNA despite the lack of metabolic activity, it could be hypothesized that those genes have a role in germination.

When analyzed in more detail the number of genes associated to clusters during differentiation in *T. atroviride,* we observed that the activation or repression of genes occurs in a non-random fashion. The fact that we observed gene clusters indicates that transcriptional regulation might occur not only by TF but also by DNA segments influenced by chromatin.

The measurement of DNA methylation levels in *T. atroviride* during photoconidiation could not determine the relation of these levels with the global activation or repression of the genome. In order to do that is necessary to measure DNA methylation with single base resolution to study the gene and genomic context of methylation and the effect of this in gene expression.

On the other hand despite the low DNA methylation observed, the inhibition has effects in photoconidiation. Also, the lack of expression of genes that can methylate DNA suggests a new gene that is responsible of the quantified DNA methylation.

Whit the experiment related with DNA methylation is shown that DNA methylation is present and play a role in spores production, but we still do not know how it affects gene expression.

Is noteworthy the transcriptional activation of genes that are involved in development in other organisms such as components of the RNAi machinery (dcr2, ago1, ago2, rdr3, and rdr1), some of them have been proven involved in conidiation (dcr2 and rdr3).

In order to determinate the biological relevance of this observations as well as the function of genes observed is necessary to work further. In such regard many questions remain.

We do not know the functional role of genes that are expressed during the maturation of conidia. Since we hypothesized that they might be involved in germination, is necessary to disrupt translation at different times after 6 hours and analyze the germination rate.

Analysis based in biochemistry and inhibition essays suggested a possible role of DNA methylation in the control of photoconidiation. However, further work is necessary to determine if there is a functional role of DNA methylation in *Trichoderma atroviride.*

Even though this study requires more experiments to test the hypothesis generated, it could work as the basis to formulate new questions during the photoconidiation process.

This study was focused in photconidiation, nevertheless we still don´t know if the stages proposed here have analogous stages in other kind of conidiations in *T. atroviride* or other organisms.

So far with this analysis, I have grasped the holistic transcriptomic state of the cell. I encourage the reader to question this work and propose new experiments in order to elucidate the understanding of cellular differentiation induced by an environmental stimulus

12. Bibliography

Bergman, K. (1972). Blue-light control of sporangiophore initiation in Phycomyces. *Planta*, *107*(1), 53–67.

Berrocal-Tito, G. M., Esquivel-Naranjo, E. U., Horwitz, B. a, & Herrera-Estrella, A. (2007). Trichoderma atroviride PHR1, a fungal photolyase responsible for DNA repair, autoregulates its own photoinduction. *Eukaryotic Cell*, *6*(9), 1682–92.

Berrocal-Tito, G., Sametz-Baron, L., Eichenberg, K., Horwitz, B. a, & Herrera-Estrella, a. (1999). Rapid blue light regulation of a Trichoderma harzianum photolyase gene. *The Journal of Biological Chemistry*, *274*(20), 14288–94.

Betina, V., & Zajacová, J. (1978). Inhibition of photo-induced Trichoderma viride conidiation by inhibitors of RNA synthesis. *Folia Microbiologica*, *23*, 460–464.

Blüthgen, N., Brand, K., Cajavec, B., Swat, M., Herzel, H., & Beule, D. (2005). Biological profiling of gene groups utilizing Gene Ontology. *Genome Informatics. International Conference on Genome Informatics*, *16*(1), 106–115.

Carreras-Villaseñor, N., Esquivel-Naranjo, E. U., Villalobos-Escobedo, J. M., Abreu-Goodger, C., & Herrera-Estrella, A. (2013). The RNAi machinery regulates growth and development in the filamentous fungus Trichoderma atroviride. *Molecular Microbiology*, *89*, 96–112.

Carreras-Villaseñor, N., Sánchez-Arreguín, J. A., & Herrera-Estrella, A. H. (2012a). Trichoderma: sensing the environment for survival and dispersal. *Microbiology (Reading, England)*, *158*(Pt 1), 3–16.

Carreras-Villaseñor, N., Sánchez-Arreguín, J. A., & Herrera-Estrella, A. H. (2012b). Trichoderma: sensing the environment for survival and dispersal. *Microbiology (Reading, England)*, *158*(Pt 1), 3–16.

Casas-Flores, S., Rios-Momberg, M., Bibbins, M., Ponce-Noyola, P., & Herrera-Estrella, A. (2004). BLR-1 and BLR-2, key regulatory elements of photoconidiation and mycelial growth in Trichoderma atroviride. *Microbiology (Reading, England)*, *150*(Pt 11), 3561–9.

Castellanos, F., Schmoll, M., Martínez, P., Tisch, D., Kubicek, C. P., Herrera-Estrella, A., & Esquivel-Naranjo, E. U. (2010). Crucial factors of the light perception machinery and their impact on growth and cellulase gene transcription in Trichoderma reesei. *Fungal Genetics and Biology*, *47*(5), 468–476.

Cole, G. T. (1986). Models of cell differentiation in conidial fungi. *Microbiological Reviews*, *50*(2), 95–132.

Colgan, J., & Manley, J. L. (1992). TFIID can be rate limiting in vivo for TATA-containing, but not TATA-lacking, RNA polymerase II promoters. *Genes & Development*, *6*(2), 304–315.

Conesa, A., Götz, S., García-Gómez, J. M., Terol, J., Talón, M., & Robles, M. (2005). Blast2GO: A universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics*, *21*(18), 3674–3676.

Crosson, S., & Moffat, K. (2001). Structure of a flavin-binding plant photoreceptor domain: insights into light-mediated signal transduction. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(6), 2995–3000.

Dang, Y., Li, L., Guo, W., Xue, Z., & Liu, Y. (2013). Convergent Transcription Induces Dynamic DNA Methylation at disiRNA Loci. *PLoS Genetics*, *9*(9).

Dhingra, S., Lind, A. L., Lin, H. C., Tang, Y., Rokas, A., & Calvo, A. M. (2013). The Fumagillin Gene Cluster, an Example of Hundreds of Genes under veA Control in Aspergillus fumigatus. *PLoS ONE*, *8*(10), 1–16.

Garzia, A., Etxebeste, O., Rodríguez-Romero, J., Fischer, R., Espeso, E. a., & Ugalde, U. (2013). Transcriptional changes in the transition from vegetative cells to asexual development in the model fungus Aspergillus nidulans. *Eukaryotic Cell*, *12*(2), 311–321.

Gibbons, J. G., Beauvais, A., Beau, R., McGary, K. L., Latgé, J.-P., & Rokas, A. (2012). Global transcriptome changes underlying colony growth in the opportunistic human pathogen Aspergillus fumigatus. *Eukaryotic Cell*, *11*(1), 68–78.

Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., … Conesa, A. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Research*, *36*(10), 3420–3435.

Gressel, J., & Galun, E. (1967). Morphogenesis in Trichoderma : and RNA Photoinduction, *598*(1), 575–598.

Hernández-Oñate, M. a, Esquivel-Naranjo, E. U., Mendoza-Mendoza, A., Stewart, A., & Herrera-Estrella, A. H. (2012). An injury-response mechanism conserved across kingdoms determines entry of the fungus Trichoderma atroviride into development. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(37), 14918–23.

Higgs, D. R., Vernimmen, D., Hughes, J., & Gibbons, R. (2007). Using genomics to study how chromatin influences gene expression. *Annual Review of Genomics and Human Genetics*, *8*, 299–325.

Honda, S., & Selker, E. U. (2008). Direct interaction between DNA methyltransferase DIM-2 and HP1 is required for DNA methylation in Neurospora crassa. *Molecular and Cellular Biology*, *28*(19), 6044–55.

Huang, D. W., Sherman, B. T., & Lempicki, R. a. (2009). Bioinformatics enrichment tools: Paths toward the comprehensive functional analysis of large gene lists. *Nucleic Acids Research*, *37*(1), 1–13.

Ikeda, K.-I., Van Vu, B., Kadotani, N., Tanaka, M., Murata, T., Shiina, K., … Nakayashiki, H. (2013). Is the fungus magnaporthe losing DNA methylation? *Genetics*, *195*(3), 845–55.

Judelson, H. S., Ah-Fong, A. M. V, Aux, G., Avrova, A. O., Bruce, C., Cakir, C., … Windass, J. (2008). Gene expression profiling during asexual development of the late blight pathogen Phytophthora infestans reveals a highly dynamic transcriptome. *Molecular Plant-Microbe Interactions : MPMI*, *21*(4), 433–447.

Kim, K. S., & Lee, Y.-H. (2012). Gene Expression Profiling during Conidiation in the Rice Blast Pathogen Magnaporthe oryzae. *PLoS ONE*, *7*(8), e43202.

Kimura, M., & Ishihama, A. (2004). Tfg3, a subunit of the general transcription factor TFIIF in Schizosaccharomyces pombe, functions under stress conditions. *Nucleic Acids Research*, *32*(22), 6706–6715.

Lamarre, C., Sokol, S., Debeaupuis, J.-P., Henry, C., Lacroix, C., Glaser, P.,Latgé, J.-P. (2008). Transcriptomic analysis of the exit from dormancy of Aspergillus fumigatus conidia. *BMC Genomics*, *9*, 417.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., … Durbin, R. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics*, *25*(16), 2078–2079.

Lin, J.-Q., Zhao, X.-X., Wang, C.-C., Xie, Y., Li, G.-H., & He, Z.-M. (2012). 5-Azacytidine inhibits aflatoxin biosynthesis in Aspergillus flavus. *Annals of Microbiology*, *63*(2), 763–769.

Magaña, A. A., Wrobel, K., Caudillo, Y. A., Zaina, S., Lund, G., & Wrobel, K. (2008). High-performance liquid chromatography determination of 5-methyl-2’-deoxycytidine, 2'-deoxycytidine, and other deoxynucleosides and nucleosides in DNA digests. *Analytical Biochemistry*, *374*(2), 378–85.

Martienssen, R. a, & Colot, V. (2001). DNA methylation and epigenetic inheritance in plants and filamentous fungi. *Science (New York, N.Y.)*, *293*(5532), 1070–4.

Martínez-López, L. a, Ochoa-Alejo, N., & Martínez, O. (2014). Dynamics of the chili pepper transcriptome during fruit development. *BMC Genomics*, *15*, 143.

McCarthy, D. J., Chen, Y., & Smyth, G. K. (2012). Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Research*, *40*(10), 4288–4297.

Morgan, A. M., Obenchain, V., & Hayden, N. (2015). Package " Rsamtools ." http://bioconductor.org/packages/release/bioc/html/Rsamtools.html

Nemčovič, M., Jakubíková, L., Víden, I., & Farkaš, V. (2008). Induction of conidiation by endogenous volatile compounds in Trichoderma spp. *FEMS Microbiology Letters*, *284*, 231–236.

Nicolás, F. E., Calo, S., Murcia-Flores, L., Garre, V., Ruiz-Vázquez, R. M., & Torres-Martínez, S. (2008). A RING-finger photocarotenogenic repressor involved in asexual sporulation in Mucor circinelloides. *FEMS Microbiology Letters*, *280*(1), 81–88.

Novodvorska, M., Hayer, K., Pullan, S. T., Wilson, R., Blythe, M. J., Stam, H., … Archer, D. B. (2013). Transcriptional landscape of Aspergillus niger at breaking of conidial dormancy revealed by RNA-sequencing.

Ozer, J., Moore, P. a., Bolden, A. H., Lee, A., Rosen, C. a., & Lieberman, P. M. (1994). Molecular cloning of the small (γ) subunit of human TFIIA reveals functions critical for activated transcription. *Genes and Development*, *8*(19), 2324–2335.

Park, Y. U., Hwang, O., & Kim, J. (2002). Two-hybrid cloning and characterization of OSH3, a yeast oxysterol-binding protein homolog. *Biochemical and Biophysical Research Communications*, *293*(2), 733–740.

Pascual, S., Melgarejo, P., & Magan, N. (1997). Induction of submerged conidiation of the biocontrol agent Penicillium oxalicum. *Applied Microbiology and Biotechnology*, *48*(3), 389–392.

Purschwitz, J., Müller, S., Kastner, C., Schöser, M., Haas, H., Espeso, E. a., … Fischer, R. (2008). Functional and Physical Interaction of Blue- and Red-Light Sensors in Aspergillus nidulans. *Current Biology*, *18*(4), 255–259.

Quail, M., Smith, M. E., Coupland, P., Otto, T. D., Harris, S. R., Connor, T. R., … Gu, Y. (2012). A tale of three next generation sequencing platforms: comparison of Ion torrent, pacific biosciences and illumina MiSeq sequencers. *BMC Genomics*, *13*(1), 1.

Reyes-Valde, M. H. (2008).Defining diversity, specialization, and gene specificity in transcriptomes through information theory, *105*(28), 9709–9714.

Reyna-López, G. E., Simpson, J., & Ruiz-Herrera, J. (1997). Differences in DNA methylation patterns are detectable during the dimorphic transition of fungi by amplification of restriction polymorphisms. *Molecular & General Genetics : MGG*, *253*(6), 703–10. Retrieved from

Roeder, R. G. (1996). The role of general initiation factors in transcription by RNA polymerase II. *Trends in Biochemical Sciences*, *21*(9), 327–335.

Rosales-Saavedra, T., Esquivel-Naranjo, E. U., Casas-Flores, S., Martínez-Hernández, P., Ibarra-Laclette, E., Cortes-Penagos, C., & Herrera-Estrella, a. (2006). Novel light-regulated genes in Trichoderma atroviride: a dissection by cDNA microarrays. *Microbiology (Reading, England)*, *152*(Pt 11), 3305–17.

Selker, E. U., Murphy, J. G., & Freitag, M. (2003). The methylated component of the Neurospora crassa genome.*Nature*. *422*(April), 893–897.

Seong, K.-Y., Zhao, X., Xu, J.-R., Güldener, U., & Kistler, H. C. (2008). Conidial germination in the filamentous fungus Fusarium graminearum. *Fungal Genetics and Biology : FG & B*, *45*, 389–399.

Silva, F., Torres-Martínez, S., & Garre, V. (2006). Distinct white collar-1 genes control specific light responses in Mucor circinelloides. *Molecular Microbiology*, *61*(4), 1023–1037.

Simkovic, M., Ditte, P., Kurucová, A., Lakatos, B., & Varecka, L. (2008). Ca2+-dependent induction of conidiation in submerged cultures of Trichoderma viride. *Canadian Journal of Microbiology*, *54*, 291–298.

Smith, G. M., & Calam, C. T. (1980). Variations in inocula and their influence on the productivity of antibiotic fermentations. *Biotechnology Letters*, *2*(6), 261–266.

Steyaert, J. M., Weld, R. J., Mendoza-Mendoza, A., & Stewart, A. (2010). Reproduction without sex: conidiation in the filamentous fungus Trichoderma. *Microbiology (Reading, England)*, *156*(Pt 10), 2887–900.

Sugui, J. a., Pardo, J., Chang, Y. C., Müllbacher, A., Zarember, K. a., Galvez, E. M., … Kwon-Chung, K. J. (2007). Role of laeA in the regulation of alb1, gliP, conidial morphology, and virulence in aspergillus fumigatus. *Eukaryotic Cell*, *6*(9), 1552–1561.

Tarazona, S., García-Alcalde, F., Dopazo, J., Ferrer, A., & Conesa, A. (2011). Differential expression in RNA-seq: A matter of depth. *Genome Research*, *21*, 2213–2223.

Thorvaldsdóttir, H., Robinson, J. T., & Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): High-performance genomics data visualization and exploration. *Briefings in Bioinformatics*, *14*(2), 178–192.

Zelter, A., Bencina, M., Bowman, B. J., Yarden, O., & Read, N. D. (2004). A comparative genomic analysis of the calcium signaling machinery in Neurospora crassa, Magnaporthe grisea, and Saccharomyces cerevisiae. *Fungal Genetics and Biology*, *41*(9), 827–841.

13. Supplemental material

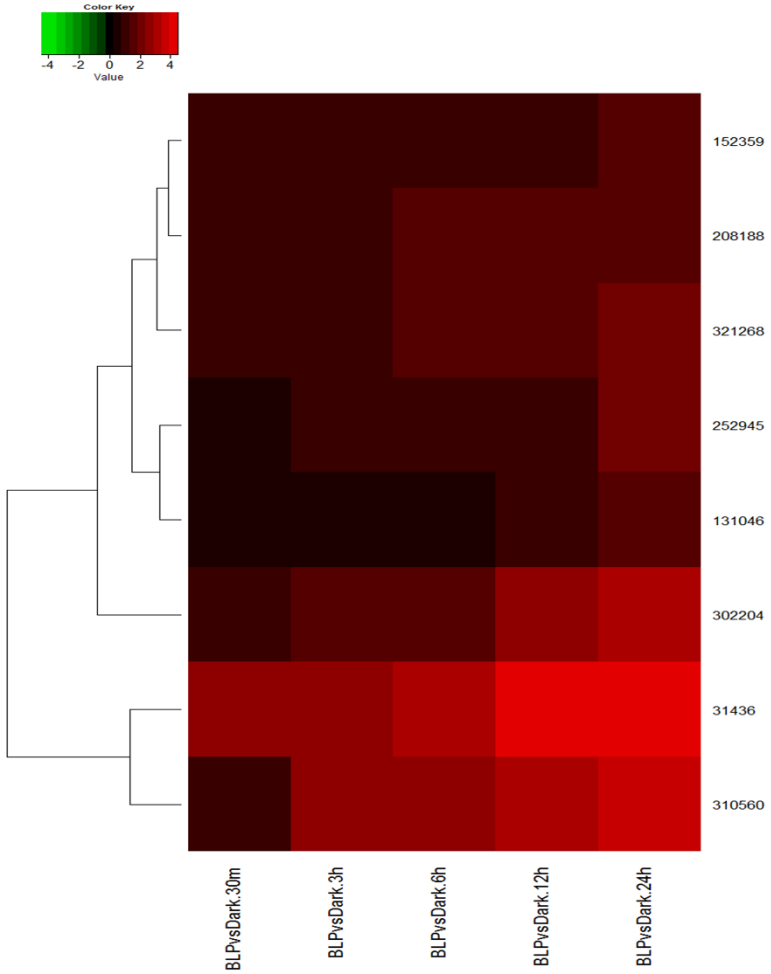

Fig. Supp.1 Progressive activation of genes during photoconidiation. Color red is proportional to activation. Complete UPGMA (Unweighted Pair Group Method with Arithmetic Mean) clustering algorithm. The x axis shows the comparisons between samples at different times, e.g. BLPvsDark.30 is the 30 min sample using as a control the sample in darkness. The y axis shows the different clusters formed, the dendogram is composed by each differentially expressed gene.

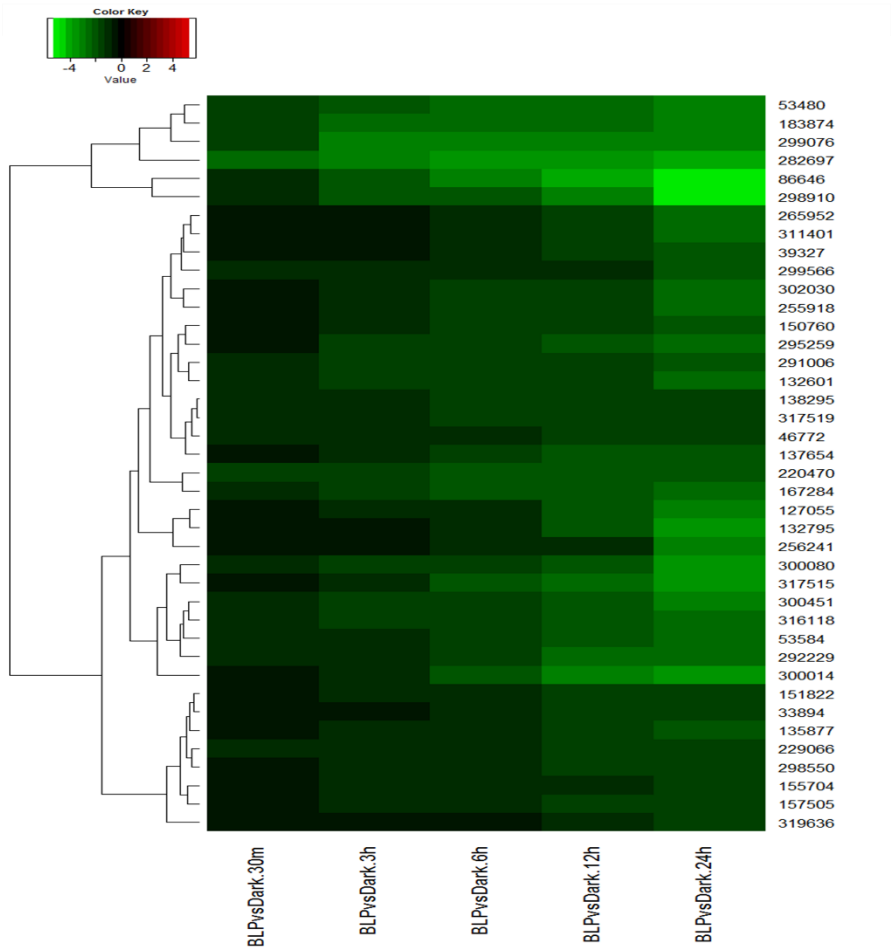

Fig. Supp.2. Progressive repression of genes during photoconidiation. Color green is proportional to repression. Complete UPGMA (Unweighted Pair Group Method with Arithmetic Mean) clustering algorithm. The x axis shows the comparisons between samples at different times, e.g. BLPvsDark.30 is the 30 min sample using as a control the sample in darkness. The y axis shows the different clusters formed, the dendogram is composed by each differentially expressed gene.

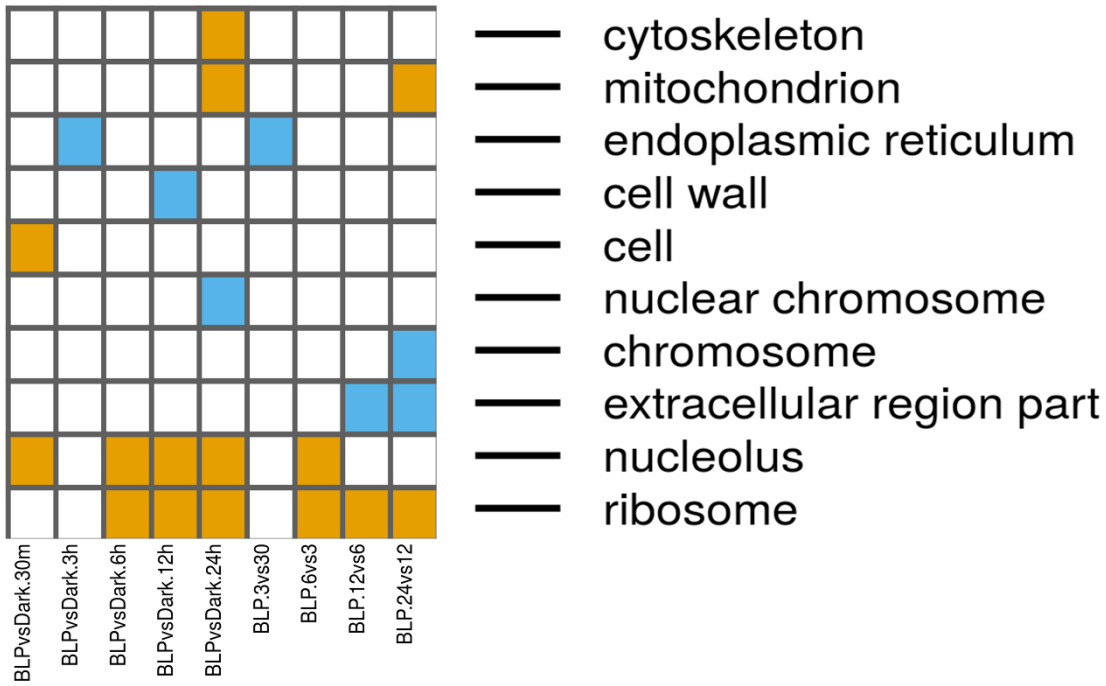

Fig. Suppl. 3. Cellular component GO terms enriched in DEA by stage and transitions. Repression is represented in yellow and activated terms in blue. GO terms selected when p-value was lower than 0.01 and FDR 0.05.