

Memoria presentada por

Alessandra Borgognone

para optar al Grado de Doctora por la Universidad Pública de Navarra

**Characterization of transposon activity and genome-wide epigenetic regulation
throughout the life cycle of *Pleurotus ostreatus***

Tesis dirigida por:

Dra. Lucía Ramírez Nasto

Catedrática de Genética y Mejora Vegetal
Grupo de Genética y Microbiología
Departamento de Producción Agraria
Universidad Pública de Navarra

Dr. Raúl Castanera Andrés

Investigador contratado
Grupo de Genética y Microbiología
Departamento de Producción Agraria
Universidad Pública de Navarra

UNIVERSIDAD PÚBLICA DE NAVARRA

Pamplona, 2017

EVALUATION COMMITTEE

President

Dr. Michael Donald Thon

Centro de Hispanoluso de Investigaciones Agrarias
Universidad de Salamanca, Spain

Secretary

Dr. Javier Pozueta Romero

Instituto de Agrobiotecnología (IDAB)
Consejo Superior de Investigaciones Científicas (CSIC)
Univesidad Pública de Navarra, Spain

Vocal

Dr. Claude Murat-Furminieux

Institute National de la recherché agronomique (INRA)
Unité Interactions Arbres/Micro-organismes
Centre INRA Nancy-Lorraine, France

Substitute

Dra. Annegret Kohler

Institute National de la recherché agronomique (INRA)
Unité Interactions Arbres/Micro-organismes
Centre INRA Nancy-Lorraine, France

External reviewers

Dra. Barbara Montanini

Dipartimento di Scienze Chimiche, della Vita
e della Sostenibilità Ambientale
Università degli Studi di Parma, Italy

Dr. David Casero

Department of Pathology and Laboratory Medicine
University of California, Los Angeles, CA, USA

External reviewer substitute

Dr. Israel Ausin Pascua

Research Center of FAFU
Fujian Agriculture and Forestry University, China

Dra. LUCÍA RAMÍREZ NASTO, Catedrática de Genética y Mejora Vegetal del Departamento de Producción Agraria de la Universidad Pública de Navarra, y

Dr. RAÚL CASTANERA ANDRÉS, Investigador contratado del Departamento de Producción Agraria de la Universidad Pública de Navarra,

INFORMAN:

que la presente memoria de Tesis Doctoral titulada “**Characterization of transposon activity and genome-wide epigenetic regulation throughout the life cycle of *Pleurotus ostreatus***” elaborada por **Dña. Alessandra Borgognone**, ha sido realizada bajo su dirección, y que cumple las condiciones exigidas por la legislación vigente para optar al grado de Doctor.

Y para que así conste, firma la presente en Pamplona, a 30 de Mayo de 2017



Fdo. Dra. Lucía Ramírez Nasto



Fdo. Dr. Raúl Castanera Andrés

Acknowledgements

Undertaking this PhD has been a truly life-changing experience for me and it would not have been possible to do without the support and guidance that I received from many people.

I would like to first thank my supervisor Professor Lucía Ramírez Nasto for her dedication and guidance over the course of my PhD project. She introduced me to the fascinating fields of genetic and epigenetic, which made my research experience productive and stimulating.

I am also grateful to Professor Antonio Gerardo Pisabarro for all the support (in both research and administrative issues), unreserved advice and encouragement.

I wish to give my appreciation for the opportunity and trust they gave me to join the Genetics and Microbiology research group some years ago and pursue this PhD.

A very special thanks to Raúl, who has been by my side throughout this PhD, living every single minute of it. He brought his invaluable enthusiasm, willingness and constructive criticisms that helped shape this research work, supporting me in hard moments and understanding my stress moaning attitude.

His contributions will have a lasting impact upon my personal and professional life and for that I am immensely grateful.

Furthermore, I want to thank past and current GENMIC group members, Gumer, Elaia, Alejandra, Manuel, Martha, Leticia, Fran and Iñaki, for their presence, which brightened the atmosphere in the lab and contributed a lot to my time at UPNA and in Pamplona. Much of the work discussed in this thesis was done in collaboration with them all. I am especially thankful to Elaia and Alejandra for our enjoyable time outside the lab.

I extend my thanks to many people from the Departments of Agrarian Production and Environmental Science at UPNA for their moral and logistic support throughout this research period.

The epigenetic studies discussed in this dissertation would not have been possible without the outstanding support of Professor Matteo Pellegrini from the University of Los Angeles, California, who showed great interest in collaborative projects and who kindly welcomed me in his lab.

Special words of thanks go to Marco Morselli, who taught me how to carry out the sequencing data analysis and the handling of bioinformatics tools. I very much appreciated his patience, enthusiasm and hardworking spirit as well as our great time at UCLA. I am obliged to Mila Rubbi for her useful suggestions regarding the experiments and her kind help in facilitating all things during my stay in Los Angeles. I would also like to thank Alice, a great scientist and friend, and other people for being as a big family in LA.

During my stay at UPNA, I had the opportunity to join the synchronized swimming team, where I met amazing people, enjoying a wonderful training. A special acknowledgement goes to Alicia and Patricia, for their unyielding support and for every great moment during these years.

I am grateful for amazing time spent with roommates in Calle Amaya apt. My stay in Pamplona was made enjoyable in large part due to people that became a part of my life. I cannot forget Bea and Chechu for being great friends and amazing roommates in too many ways. We have shared countless moments of laughter and sorrow and their friendship has made all the difference during my stay. I will cherish our friendship for years to come.

Thanks also to the members of "Mountains en komunak" group not mentioned above, Libertad, Maria and David, for encouraging me enormously during the final part of this PhD and for many good moments together.

I take this opportunity to thank Mari Carmen and Antonio for the warmth they extended to me during these last years and for always making me feel so welcome.

I am indebted to all my friends in Italy and elsewhere in the world, in particular Noemi, Claudia, Giulia, Simona Solange, Isabella and Marta, for always helping me in whatever way they could during this challenging period and for their unconditional friendship that has filled my life with much happiness.

I would also like to say a heartfelt thank you to my uncles, aunts and cousins, for always believing in me and encouraging me in every single step, even in the distance.

I own my deepest gratitude to my mum, my dad, my brother and little Maya, for always being there for me and encouraging me in every endeavor. My thesis came to reality with the strength they gave me and without whom all my milestones could not have been practically possible.

Lastly, I dedicate this accomplishment to my grandmas, Rosa and Anna, who have been a source of inspiration throughout my life. They raised me with infinite love, giving me the best example of sacrifice, persistence, work dedication and family devotions, always encouraging me to find and realize my potential.

TABLE OF CONTENTS

Resumen	11
Summary	15
CHAPTER I – general introduction	17
Evolution of genome sequencing	19
Uncovering the “non-coding” fraction of genomes	20
Classification of transposable elements.....	21
Impact of transposable elements on genome architecture and functionality.....	22
Epigenetic regulation of transposon activity.....	24
DNA methylation	24
Transcriptional and post-transcriptional silencing mediated by srnas	27
The kingdom of fungi.....	30
Fungal genomics: a picture from basidiomycota	33
Lifestyle and genomics of the basidiomycete <i>pleurotus ostreatus</i>	33
References.....	37
Scope of research	43
CHAPTER II – Somatic transposition and meiotically-driven elimination of an active helitron family in <i>Pleurotus ostreatus</i>	45
Introduction	47
Materials and methods.....	51
Results and discussion	55
References.....	73
Supplementary information	74
CHAPTER III – Epigenetic and transcriptomic profiles throughout the life cycle of <i>Pleurotus ostreatus</i>	96
Introduction	97
Materials and methods.....	103
Results	109
Discussion.....	125
References.....	131
Supplementary information	135

Final remarks and General discussion	153
Conclusions	166
List of publications	169
Funding	171

List of abbreviations

5mC	5-methylcytosine
6mA	6-methyladenine
aDNA	aberrant DNA
AGO	Argonaute
bp	base pair
BS-seq	Bisulfite Sequencing
DCL	Dicer-like
DNA	Deoxyribonucleic acid
DNMT	DNA methyltransferase
dNTPs	deoxynucleotide Triphosphates
dsRNA	double-stranded RNA
Fig.	Figure
gDNA	genomic Deoxyribonucleic Acid
JGI	Joint Genome Institute
mg	milligrams
min	minutes
ml	millilitres
MIP	Methylation induced premeiotically
miRNA	microRNA
MSUD	Meiotic silencing by unpaired DNA
ncRNA	noncoding RNA
NGS	Next-Generation Sequencing
nt	nucleotide
<i>P. ostreatus</i>	<i>Pleurotus ostreatus</i>
PCR	Polymerase chain reaction
PTGS	Post-transcriptional gene silencing
qPCR	quantitative Polymerase chain reaction
rasiRNA	Repeat associated small interfering RNA
RdDM	RNA-directed DNA methylation
RIP	Repeat-Induced point mutation
RISC	RNA-induced silencing complex
RNA	Ribonucleic acid
RNAi	RNA interference
RNA-seq	RNA sequencing
rpm	revolutions per minute
RPKM	Reads per kilobase per million
RPM	Reads per million
rRNA	ribosomal RNA

siRNA	short interfering RNA
sRNA	small RNA
sRNA-seq	small RNA sequencing
ssRNA	single-stranded RNA
TE	Transposable element
TGS	Transcriptional gene silencing
tRNA	transfer RNA
WGS	Whole Genome Sequencing
w/v	weight to volume ratio

Resumen

La mayoría de organismos procariotas y eucariotas contiene en su genoma secuencias de ADN repetitivas, definidas como elementos transponibles (ETs), que pueden movilizarse e insertarse aleatoriamente generando inestabilidad genómica que en definitiva afecta la función de secuencias codificantes. Durante mucho tiempo, los ETs se describieron como fragmentos de ADN 'independientes', debido a su proliferación dentro del genoma del huésped sin aportar ningún beneficio o induciendo efectos deletéreos cuando se insertan en genes. Sin embargo, con el tiempo, esta visión de los ETs fue cambiando al conocerse la contribución de los mismos en el mantenimiento de la estabilidad y evolución del genoma. En los últimos años, la disponibilidad de genomas procariotas y eucariotas secuenciados ha permitido el estudio de la distribución de estos elementos móviles y el efecto de sus copias activas entre genomas relacionados.

Es conveniente señalar que debido a la capacidad que tienen de generar mutaciones, los genomas del huésped han desarrollado mecanismos endógenos para limitar la movilización y reprimir la actividad transcripcional. En eucariotas, las secuencias repetidas pueden silenciarse a través de modificaciones epigenéticas de la secuencia de nucleótidos, que modifican la expresión génica sin producir cambios permanentes. Los mecanismos epigenéticos, que incluyen la metilación del ADN y las estrategias de silenciamiento por ARN de interferencia (RNAi), pueden regularse a diferentes niveles, determinando el silenciamiento de regiones específicas genómicas. Por lo tanto, el desarrollo de las llamadas tecnologías de nueva generación (Next-Generation) ha permitido obtener perfiles epigenéticos completos del genoma y ampliar la visión general de los ETs en la evolución de los genomas. Las modificaciones epigenéticas en hongos filamentosos se han estudiado con gran detalle en *Neurospora crassa*, aunque también se han descrito en otros eucariotas superiores. Sin embargo, el efecto de los cambios epigenéticos en el ADN de basidiomicetes permanece aún casi sin explorar.

Estudios recientes realizados en *Pleurotus ostreatus* mostraron que el genoma de este basidiomicete modelo está poblado por numerosas familias de ETs, entre las que se encuentran los helitrones, un grupo de transposones de ADN que se movilizan a través de un mecanismo de transposición de círculo rodante. En esta tesis nos hemos propuesto investigar la herencia de los helitrones y caracterizar los perfiles epigenéticos y transcriptómicos del hongo *P. ostreatus* en diferentes estadios de desarrollo. Ambos objetivos se han llevado a cabo mediante el uso de técnicas moleculares integradas con el análisis de datos de secuenciación Sanger o Next-Generation.

Esta tesis doctoral se compone de cuatro capítulos principales.

En el **Capítulo I** se describe el estado actual de la técnica de secuenciación del genoma, identificación del transposones y descripción de los mecanismos epigenéticos (metilación del ADN y silenciamiento por ARN de interferencia). Se incluye también, una introducción general sobre el Reino de los hongos y del organismo modelo *P. ostreatus*.

En el **Capítulo II** se determinan los perfiles de transmisión de las dos familias de helitrones (HELPO1 y HELPO2) en la progenie derivada meióticamente en *P. ostreatus*. Los resultados describen unos patrones de segregación distorsionada de los helitrones pertenecientes a la familia HELPO2 que demuestran una evidente sub-representación de estos elementos en la progenie. Análisis de las regiones flanqueantes de las secuencias de los elementos HELPO2 indican que el proceso meiótico de conversión génica podría estar relacionado con la eliminación de estos elementos repetitivos, favoreciendo la presencia de loci vacíos resultantes de la transposición de elementos HELPO2. El análisis del contenido de elementos HELPO2 en una genealogía de subclones mantenidos en diferentes condiciones de cultivo demostró la existencia de transposición somática de un helitron HELPO2. Análisis adicionales de la secuencia genómica y transcriptómica revelaron que *P. ostreatus* posee una maquinaria de RNAi activa que podría estar implicada en el control de los elementos transponibles. Estos hallazgos demuestran, por primera vez, la movilización de un helitron en el reino de los hongos y muestran la puesta en marcha de mecanismos de defensa del genoma contra ADN invasivo usando *P. ostreatus* como modelo.

En el **Capítulo III** se describen las estrategias de defensa epigenética en *P. ostreatus* mediante el análisis integrado de metilación del ADN, transcripción de sRNAs y mRNA en distintos estadios de su ciclo de vida. El análisis realizado utilizando tecnologías de nueva generación aplicada a distintos contextos genómicos (BS-seq, sRNA-seq y mRNA-seq) se llevó a cabo en cepas monocarióticas y dicarióticas.

El estudio realizado reveló que el perfil de metilación de ADN específico en cada cepa, así como la producción de ARN pequeños de 21-22 nt, están involucrados en la represión de la actividad de los transposones. Además, estos análisis revelan que la metilación de los transposones—pero no la producción de sRNA—está directamente involucrada en el silenciamiento de genes localizados cerca de transposones. Finalmente, estos resultados describen los genes directamente involucrados en el proceso de fructificación a través del análisis comparativo de transcriptomas.

En el **Capítulo IV**, se presenta una discusión general sobre los resultados de esta tesis. Se discute sobre los perfiles epigenéticos y transcripcionales en las dos familias de helitrones de las cepas de *P. ostreatus*, así como la variabilidad de los perfiles de metilación de cepas dicarióticas que comparten idéntico complemento genético pero que se han mantenido en diferentes condiciones de subcultivo durante varios años. Finalmente, se discute la presencia de los elementos necesarios para la activación de la maquinaria de RNAi en las distintas cepas de *P. ostreatus*.

Summary

Most prokaryotic and eukaryotic life forms must deal with the presence of repetitive DNA sequences called transposable elements (TEs), whose ability to mobilize through the genome and insert at a random position has an impact on genome stability and functionality. For a long time, TEs were described as ‘selfish’ DNA fragments, owing to their proliferation within the host genome without conferring any benefit or inducing detrimental effects when inserted in gene-coding regions. Over time, however, this view was changed by the discovery of the contribution of transposons in genome integrity and evolution.

Recent genome-wide characterizations have focused on the distribution of these mobile elements and the effects of active transposon copies within closely related genomes. Given their mutagenic potential, host genomes have evolved endogenous mechanisms to limit the mobilization and repress the transcriptional activity of transposons. In higher organisms, repeat sequences are transcriptionally silenced through epigenetic modifications, which modulate gene expression without producing permanent modifications along the nucleotide sequence. The integrated and dynamic nature of epigenetic pathways, including DNA methylation and RNA-silencing systems, can be regulated at different levels, leading to the targeted silencing of specific genomic regions. Thus, the revolutionary advent of high-throughput sequencing led to an unprecedented opportunity to generate genome-wide epigenetic profiles and extend the understanding of the contribution of TEs in genome evolution.

The filamentous fungi—in particular, *Neurospora crassa* and other ascomycetes—have provided fundamental advances in many of the aforementioned areas. These organisms possess complex epigenetic pathways that are also conserved in other higher eukaryotes to efficiently shut down transposon activity. Despite their importance, the occurrence of epigenetic events as well as the impact of transposon activity in basidiomycetes have been poorly analyzed so far.

Recent studies in *Pleurotus ostreatus* uncovered that the genome of this basidiomycete model is populated by a diverse set of TE families, providing a detailed picture of the distribution and importance of helitrons, which are a group of DNA transposons that mobilize through a rolling-circle mechanism.

Therefore, the principal aim of this work is to investigate the inheritance of helitron transposons and profile the epigenetic and transcriptomic landscape of *P. ostreatus* at different growing stages. Both objectives have been performed through the use of molecular techniques integrated with subsequent Sanger or Next-Generation sequencing data analysis.

This PhD dissertation is comprised of four main chapters.

Chapter I describes the current state-of-the-art of genome sequencing, transposon identification and epigenetic mechanisms (DNA methylation and RNA silencing pathways); it also provides an introductory overview on the fungal kingdom and the model system *P. ostreatus*.

Chapter II presents the inheritance patterns of HELPO1 and HELPO2 helitron families in the meiotically-derived progeny of *P. ostreatus*. Our results report the distorted segregation patterns of HELPO2 helitrons that lead to a strong under-representation of these elements in the progeny. Further analyses of the HELPO2 flanking sites showed that the meiotic process of gene conversion may contribute to the elimination of such repetitive elements, favoring the presence of HELPO2 vacant loci. Moreover, the analysis of HELPO2 content in a reconstructed pedigree of subclones maintained under different culture conditions revealed an event of helitron somatic transposition. Additional investigations of genome and transcriptome data indicated that *P. ostreatus* carries active RNAi machinery that could be involved in the control of transposable element proliferation. These findings provide the first evidence of helitron mobilization in the fungal kingdom and highlight the interaction between genome defense mechanisms and invasive DNA using *P. ostreatus* as a model.

Chapter III describes the occurrence of epigenetic defense strategies in this fungal model. The analyses report a picture of genome-wide epigenetic (DNA methylation and small RNAs) and transcriptional (mRNA) patterns in *P. ostreatus* throughout its life cycle. High-throughput sequencing analyses (BS-seq, sRNA-seq and mRNA-seq) performed in monokaryon and dikaryon samples revealed epigenetic differences among strains but not within developmental stages. Our results uncovered strain-specific DNA methylation profiles (~ 2 to 7 % global methylation levels) and 21-22nt small RNA production primarily involved in the repression of transposon activity. Furthermore, our findings provide evidence that TE-associated DNA methylation—but not small RNA production—is directly involved in the silencing of genes surrounded by transposons. Finally, these findings also report key genes that are activated in the fruiting process through a comparative analysis of transcriptomes.

A general discussion on findings presented in this thesis is given in **Chapter IV**. This last part reports a critical outlook on the epigenetic and transcriptional state of the two helitron families of the *P. ostreatus* strains that are differentially subcultured during several years as well as nucleus-specific methylation variations in dikaryotic strains that share an identical genetic complement but different subculture conditions. Lastly, it is given a general overview on the strain-specific RNAi machinery in *P. ostreatus*.

Chapter I – General Introduction

Evolution of genome sequencing

The development and constant improvement of DNA sequencing, which refers to the determination of the precise nucleotide sequence within a DNA molecule, has made it possible to obtain exhaustive information about the genomic landscapes, ranging from short fragments (i.e., gene regulatory regions) up to the characterization of entire genomes. The advent of the first DNA sequences was documented during the early 1970s. Around that time, a short sequence of the *lac* operator from bacteria was obtained through the application of time consuming and laborious techniques (Gilbert and Maxam 1973). The first revolution in this field took place in the late 1970s after the development of innovative methods proposed simultaneously by the researchers Maxam and Gilbert (Maxam and Gilbert 1977) and Sanger (Sanger et al. 1977), which extraordinarily increased the sequencing throughput. These techniques, defined as Whole Genome Shotgun (WGS), enabled the sequencing of longer DNA sequences with higher efficiency compared to the initial strategies. By then, Sanger sequencing was predominant due to its improved accuracy and reduced utilization of toxic compounds. According to these procedures, the DNA chain is broken up into smaller fragments, which are cloned and sequenced individually. Subsequently, the short sequences are reassembled into their original order based on overlaps, ultimately yielding the complete sequence through the use of sophisticated computer programs (Staden 1979).

In the last few decades, the extensive use of WGS techniques for DNA sequencing and the progressive technological optimization led to the successful determination of the first set of eukaryotic genomes, including human (Consortium 2001; Venter et al. 2001) and mouse (Waterston 2002) genomes. Despite its reliability and constant improvement (Collins et al. 2003), sequencing based on the Sanger method displayed some disadvantages in terms of reagents, intensive working time and monetary cost. Given the importance of DNA sequencing in different fields of biological research, considerable efforts were made to develop more efficient methods, reducing their cost.

This encouraged the production of alternative technologies and considerable breakthroughs that led to the advent of the so-called high-throughput technologies (Next-Generation Sequencing [NGS]) (Mardis 2008). So far, three NGS platforms are the most prevalent and commercially available: i) 454 array-based pyrosequencing (Margulies et al. 2005), ii) Illumina sequencing by synthesis (Bentley 2006) and iii) SOLiD sequencing by ligation (Valouev et al. 2008). Although second-generation sequencing has enabled the discovery of many new genomes (Schatz et al. 2010), the techniques experienced some difficulties in terms of small size of reads (i.e. 25 – 250 bp) and lack of sufficient robustness to study overall chromosome architecture (Jia et al. 2013).

The need for technologies that would produce longer reads has resulted in the advent of new sequencing, called Third-Generation Sequencing (TGS), which facilitated the development of long-read sequencing (3 – 15 kb). PacBioRS is the most successful TGS technology so far (Eid et al. 2009; Korlach et al. 2010). Furthermore, the release of another TGS technology, the so-called Nanopore DNA (Stoddart et al. 2009), ushered in a new era of cheap and portable long-read sequencing techniques. The increasing improvement of these technologies as well as the proliferation of genome databases and bioinformatics software led to extraordinary advances in many fields of genomics through the applications of specific techniques, such as RNA-seq to analyze comparative transcriptomes (Mortazavi et al. 2008), Methyl-seq to identify the methylation state across the genome (Brunner et al. 2009), ChIP-seq to detect DNA-protein binding (Johnson et al. 2007) and whole-exome sequencing to investigate the variations of all protein-coding regions (Rabbani et al. 2014).

In addition, highly accurate *de novo* genome assemblies and resequencing analysis provided a better picture of the genome fraction occupied by non-coding DNA (Costa 2008). Overall, the improvement of NGS technology and the deluge of sequenced data provide a chance to explore unsolved questions in many fields of biology and genetics.

Uncovering the “non-coding” fraction of genomes

Several genomic studies documented that the diverse and integrated regulation of each organism relies on a complex interaction of “non-coding” regions rather than just on protein-coding genes (Doolittle and Sapienza 1980). Initially defined as “junk DNA”, non-genic DNA sequences such as transposable elements (TEs) have been progressively recognized as important players in genome plasticity. Their study has impressively risen during the past decade with the advent of comparative genomics and bioinformatics tools dedicated to their detection. Before that, transposable elements gained extraordinary interest through the revolutionary discoveries performed by the Nobel Prize winner Barbara McClintock in the late 1940s. Her experiments on maize chromosome breakage elucidated the mechanism by which these jumping DNA units modulated the expression of certain genes, generating phenotypic changes (McClintock 1984).

Since then, several studies uncovered that these segments of DNA are ubiquitous across living organisms and are able to induce mutations and chromosomal rearrangements as a consequence of their mobilization. In the last several decades, the proliferation of whole-genome sequencing projects provided increasing information on the transposon landscapes, particularly on their function and distribution.

The surprising impact of TEs on the genome structure contribute to a turnaround from considering them as ‘selfish’ parasitic sequence (Orgel and Crick 1980) to potent drivers of genome evolution. Given their ability to transpose and multiply within the DNA sequence, these mobile genetic sequences can increase in copy number and occupy variable genome fractions, ranging from a few copies to millions of copies (Szmulewicz et al. 1998). In higher eukaryotes, TEs represent one of the most significant contributors to genome size, and they account for up to approximately 80% in some plants (Wicker et al. 2011) and up to 50% in mammals (Zamudio and Bourchis 2010).

Classification of transposable elements

The main ability of transposable elements is to mobilize within the host genome. According to their mechanism of transposition, TEs are generally classified into two main categories: Class I and Class II (Fig. 1). Class I elements, also known as retrotransposons, transpose by a copy-paste mechanism through reverse transcription of an RNA intermediate encoded by the element. Based on their structure and transposition mechanism, retrotransposons can be divided into five orders: LTR (Long terminal repeat) retrotransposons, DIRS (*Dictyostelium* intermediate repeat sequence), PLE (*Penelope*-like element), LINEs (Long interspersed nuclear element) and SINEs (Short interspersed nuclear element) (Wicker et al. 2007). The principal distinction in this group concerns LTR-retrotransposons, which are flanked by specific long-terminal-repeats (LTRs) at the end of both sides and non-LTR retrotransposons (LINEs and SINEs). Class II transposons proliferate directly from DNA to DNA. This class is divided into two subclasses that are comprised of four orders: TIR (terminal inverted repeats), Crypton, Helitrons and Mavericks. During their transposition, DNA transposons can excise from the original position and integrate into a new target site through a “cut and paste” mechanism or mobilize through rolling-circle replication, as described for Helitron elements, similarly to some bacterial mobile elements (Kapitonov and Jurka 2001). Based on specific traits (e.g., target site duplication, expression of specific proteins), elements belonging to both classes can be further categorized into superfamilies and families. Both retrotransposons and DNA-transposon families comprise non-autonomous elements that require the retrotransposase/transposase of an autonomous element for their mobilization across the genome.

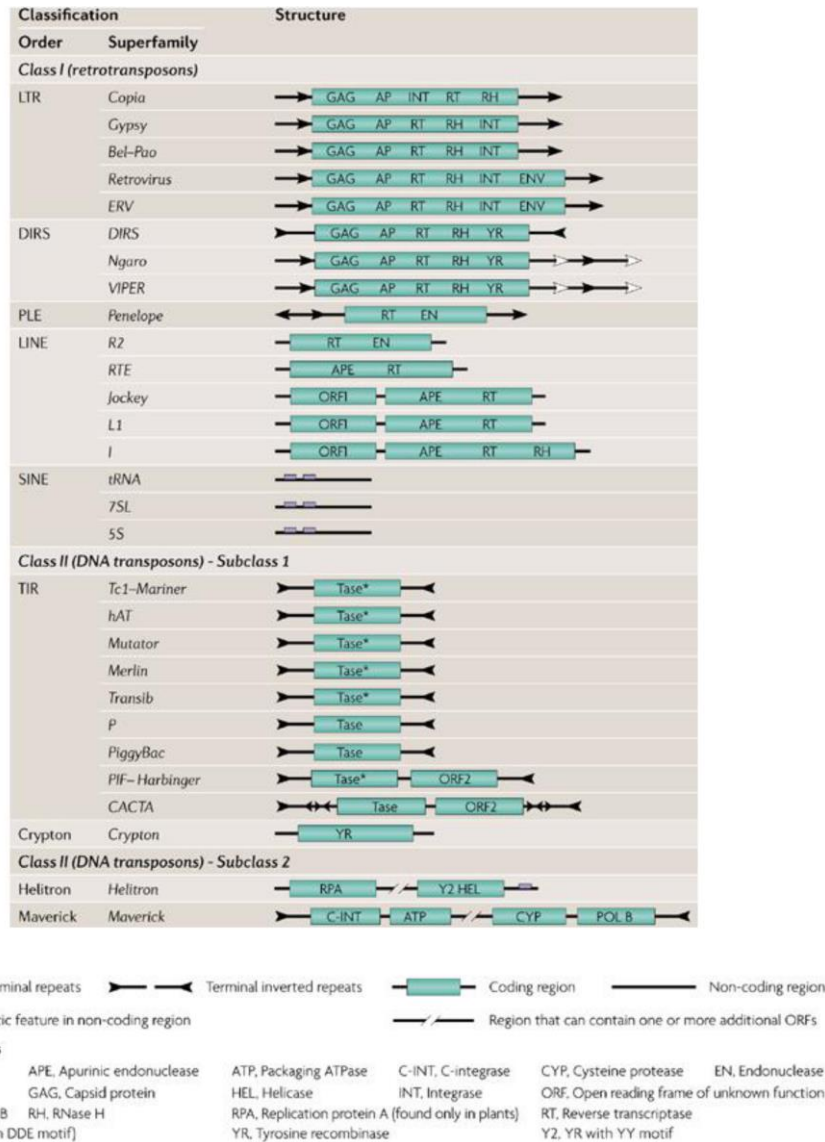


Figure 1. Hierarchical classification of transposable elements. Obtained from Wicker et al. (2007).

Impact of transposable elements on genome architecture and functionality

Transposon mobilization can be more than simple events of genomic deletions and insertions. Beside their potential detrimental impact, it has been postulated that transposons play an essential role in promoting the plasticity within the host genomes and exert both positive and negative regulation over the proximity of genes (Raffaele and Kamoun 2012). Given their ability to multiply across the genome, the insertion of a few or higher number of transposon copies contributes to generate variability between genomes at inter- and intraspecific levels. On a structural scale, transposon bursts can lead to an increase of the genome size and generate polymorphisms between closely related lineages as is extensively reported in plants (Naito et al. 2006; El Baidouri and Panaud 2013). Additionally, these repetitive sequences play an essential

role in chromatin organization and the maintenance of genome integrity. In fact, high amounts of repeated elements and TEs have been detected in chromosomal regions involved in processes of segregation and replication in eukaryotic cells (Gao et al. 2015). For example, transposon blocks and other repeats are associated with heterochromatin formation in the centromeric regions of *S. pombe* (Almeida and Allshire 2005), and they also constitute the telomeric regions of *D. melanogaster* chromosomes (Biessmann et al. 1992). Although TEs can positively contribute to genome stability and variability, they are primarily known for their potential to act as insertional mutagens and cause chromosome breakage (McClintock 1947).

In this sense, the activity of some transposons has been described as a continual rather than occasional event in several eukaryotic genomes (Huang et al. 2012). In the human fungal pathogen *Candida albicans*, the transcriptional activity of LTR and non-LTR retrotransposons has been shown to induce random mutations across the genome (Goodwin et al. 2001; Holton et al. 2001). Transposons can also produce mutations after their excision by either leaving or not leaving “footprints” that can point to their previous presence or lack thereof (Britten 1996). Some deleterious effects of TEs can also result from illegitimate chromosomal rearrangements as well as the insertion into or near protein-coding genes (Wright et al. 2003). Indeed, transposons can negatively impact host fitness by extending their epigenetic marks on flanking regions (Wessler 2006) and downregulating the expression of coding genes (Lankenau and Volff 2009).

Taken together, genetic variations associated to the activity of mobile elements may result in dramatic alterations at somatic levels. An example is the color variegation in maize kernels and leaves documented by McClintock (McClintock 1951), but somatic variations caused by TEs have also been documented in a wide range of eukaryotic organisms (Kidwell and Lisch 1997). In *Drosophila melanogaster*, massive chromosome breakage in larval cells resulting from the transposition of *P* elements showed a direct impact on pupal lethality and hybrid dysgenesis in progeny (Bregliano and Kidwell 1983). Another paradigmatic example reports that LINE-mediated transposition in the mammalian brain can induce a broad spectrum of phenotypic modification and ultimately generate variations in neuronal genomes (Singer et al. 2010).

Epigenetic regulation of transposon activity

Transposable elements have been described as important drivers of genome plasticity and evolution; nonetheless, they remain mutagenic that agents able to produce detrimental effects. Therefore, host-genomes have evolved endogenous mechanisms to protect themselves from transposon proliferation and limit their expansion. In fact, plant and animal transposable elements have been described as the principal target of epigenetic silencing mechanisms that consist of integrated and dynamic pathways that are regulated at different levels (Slotkin and Martienssen 2007; Fedoroff 2012).

The term 'epigenetics' refers to modifications that can modulate the transcription of genes in the absence of changes along the DNA sequence. Initially described as heritable modification, the term is now applied to changes that may occur in terminally differentiated cells as well (Popova et al. 2006). The best characterized epigenetic strategies include DNA methylation (correlated with chromatin remodeling and histone modification) and mechanisms of transcriptional and post-transcriptional silencing mediated by the RNA interference (RNAi) complex. It is becoming clear that these epigenetic modifications can be regulated by various silencing effector complexes in response to the activity of invading transposons.

DNA methylation

DNA methylation was the first epigenetic mark that was discovered to impact gene expression (Holliday and Pugh 1975). It is involved in the regulation of several cellular processes ranging from development to gene silencing in a variety of eukaryotic genomes (Suzuki and Bird 2008; Law and Jacobsen 2010a).

Such modification occurs by the covalent transfer of methyl groups (CH₃) to the DNA, typically at position 5 of the cytosine ring (mC). The addition of methyl groups is controlled at different levels in cells and orchestrated by a family of enzymes known as DNA methyltransferases (DNMTs) (Bird 2002) (Fig. 2).

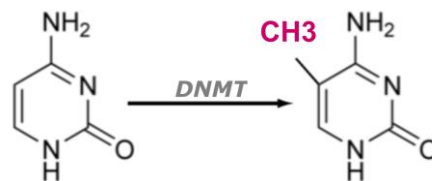


Figure 2. Mechanism of cytosine methylation mediated by DNA methyltransferases.

This modification can modulate the activity of a DNA segment without involving permanent changes in the nucleotide sequence of the organism. An exhaustive picture of molecular mechanisms leading the establishment and maintenance of DNA methylation is still unknown. However, several efforts have been made to characterize the molecular pathways that are involved in this process in the last two decades (Law and Jacobsen 2010b). The state-of-the-art method to profile DNA methylation in all cytosine contexts (CG, CHG and CHH) based on bisulfite (BS) conversion was discovered in 1980 (Wang et al. 1980).

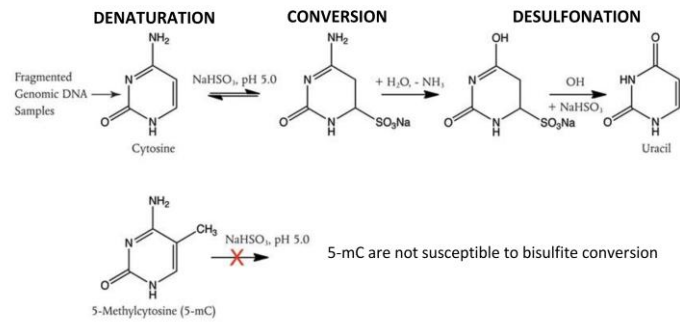
After chemical treatment of DNA molecules with bisulfite under denaturing conditions, the procedure converts unmethylated cytosines to uracil, whereas the methylated cytosines remain unchanged.

During subsequent PCR amplifications and sequencing, uracil residues are converted to thymine (Fig. 3). The technique was initially applied to a few genomic regions with individual Sanger sequencing (Frommer et al. 1992), but it can now be performed at the whole-genome scale using Next-Generation Sequencing (NGS) (Lister and Ecker 2009; Laird 2010). The complete conversion and high sequencing coverage obtained allow to generate comprehensive DNA methylome and detect methylated cytosines at single-base resolution along the whole genome.

According to a recent comparative study (Harris et al. 2010), BS-Seq is considered to be the most reliable method to achieve a large-scale single-bp resolution of the methylation profile. Thus, the increasing improvement of tools for genome-wide profiling of DNA methylation enabled researchers to identify these genomic modifications and gain a comprehensive picture of epigenetic regulation in different organisms. In mammalian stem cells (Lister and Ecker 2009) plants (Lister et al. 2008) and fungi (Selker et al. 2003; Zemach et al. 2010), methylated cytosines have been detected in symmetrical CG/CHG or asymmetrical CHH (where H stands for A, T or C) contexts (Dyachenko et al. 2010). In differentiated mammals, however, 5mC occurs nearly exclusively in CG dinucleotides. Further studies extended this view by reporting that CG and CHG are found in similar locations, whereas CHH methylation (as a non-symmetric site and requiring active *de novo* methylation for its presence) is often rare and transient (Cokus et al. 2008; Schmitz and Ecker 2012).

In fungi, the occurrence and function of this epigenetic modification is still unclear. DNA methylation has been uncovered in *Neurospora crassa* and *Ascobolus immersus* and some other filamentous fungi, while in others it seems to be lacking (Freitag and Selker 2005; Galagan et al. 2005). In *N. crassa*, it was found to be primarily associated with relics of the genome defense system repeated-induced point (RIP) mutation (Selker et al. 1987). Also, in the transposons-rich ascomycete *Tuber melanosporum*, 5mC was mainly detected in transposable elements (Montanini et al. 2014).

Bisulfite conversion



PCR amplification

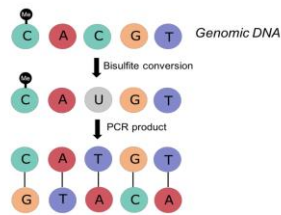


Figure 3. Schematic representation of bisulfite conversion in unmethylated and methylated cytosine residues and subsequent PCR amplification.

Generally, genomic regions possessing high methylation levels display a tightly packaged chromatin state referred to as heterochromatin, whereas regions with low methylation and loose chromatin are defined as euchromatin. Heavily methylated genomic regions are often considered a mark of silenced heterochromatin. Nevertheless, comparisons between genome-wide DNA methylation and transcriptional profiles suggest that this association is not totally repressive (Zilberman et al. 2007; Vaughn et al. 2007).

A recent analysis of the genome-wide distribution of 5mC in plants shows that intergenic DNA regions display higher methylation levels when compared to gene bodies due to the preferential location of repeated elements within intergenic regions (Regulski et al. 2013). However, the highest methylation levels are generally detected in the pericentromeric and centromeric regions of the chromosomes (Lippman et al. 2004; Eichten et al. 2013), which are enriched in repeated elements and depleted of coding genes.

It is considered that methylation is principally involved in genome defense and the regulation of gene expression. The first function concerns the inactivation of non-self fragments, such as potential insertion of viral and retroviral sequences, and the silencing of repeated sequences, such as TEs, pseudogenes and repetitive sequences (Zhang et al. 2006; Slotkin and Martienssen 2007). In eukaryotes, DNA methylation maintains the silent state of transposable elements and regulates the chromatin status through the coordinate interaction of DNMTs with components of additional epigenetic complexes.

According to this, when the activity of transposons is not fully suppressed, cryptic promoters located in the transposon sequence can induce the transcription of noncoding RNAs (ncRNAs). Specific classes of ncRNAs participate as targeting factors in epigenetic pathways, such as RNA-directed DNA methylation (RdDM) involved in the silencing of transposable elements (Law and Jacobsen 2010).

Transcriptional and post-transcriptional silencing mediated by sRNAs

Beyond their crucial role in the pathway from DNA into proteins, RNA molecules are involved in other functional processes, such as the establishment of DNA methylation and histone modifications. RNAs are conventionally distinguished into coding (mRNA) and non-coding (ncRNAs) molecules (Szymanski et al. 2003). This latter group comprises functional RNA molecules transcribed from a DNA molecule but not translated into proteins. Among the ncRNAs, only ribosomal RNA (rRNA) and transfer RNA (tRNA) have been characterized in depth. However, recent studies uncovered a novel class of ncRNAs called small RNAs (sRNAs), which are involved in gene regulation at both transcriptional (TGS) and post-transcriptional (PTGS) levels (Moazed 2009).

TGS refers to the blocking of transcription, which prevents the accumulation of mRNA transcripts, and the role of small RNAs in TGS mediated by heterochromatin formation was described in *Schizosacharomyces pombe* (Volpe et al. 2002). On the other hand, PTGS involves the degradation of mRNA products in the cytoplasm, which prevent their translation into proteins. The occurrence of post-transcriptional gene silencing mediated by small RNAs was first discovered in plants in 1998 (Waterhouse et al. 1998), and subsequent experiments described that this mechanism was led by sRNAs that were 21-25 nt long (Hamilton and Baulcombe 1999). RNA-sequencing (RNA-Seq) is a highly sensitive and accurate tool for measuring expression across the transcriptome using deep-sequencing technologies. Studies using this method enable researchers to view the complexity of eukaryotic transcriptomes and uncover the functional elements during development or different conditions. Small RNA sequencing, which is a type of RNA-Seq, is a technique to isolate and sequence small RNA species. This method can query thousands of small RNAs and examine the differential expression at single-base resolution (Shore et al. 2016). The sRNAs described to date vary in terms of size and are classified depending on their biogenesis and function (Kim 2005; Kim et al. 2009) (Fig. 4).

The main small RNA categories are piwi interacting RNAs (piRNAs), microRNAs (miRNAs) and small interfering RNAs (siRNAs). piRNAs originate from transposons, repeats and piRNA clusters, their mechanism of action is RNA cleavage (PTGS) and they have been exclusively identified in animals. miRNAs are processed from single-stranded hairpin precursors of variable length and produce translational repression and mRNA degradation in plants and animals.

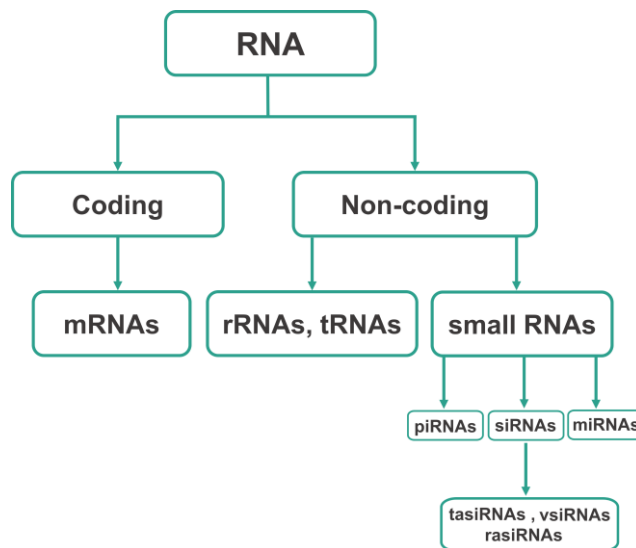


Figure 4. Classification of coding and non-coding RNA (ncRNAs). Coding RNA comprises messenger RNAs (mRNAs). Non-coding RNAs are divided into ribosomal (RNAs), transfer (tRNAs) and small (sRNAs) RNAs. SmallRNA include three main types: short interfering RNAs (siRNAs), microRNAs (miRNAs) and piwi interacting RNAs (piRNAs). Short interfering RNAs have been recently grouped into different classes, including virus-derived RNAs (vsRNAs), trans-acting RNAs (tasiRNAs) or repeat associated RNAs (rasiRNAs). Adapted from Guleria et al. (2012).

Regarding siRNAs, they are present in plants, animals and fungi and originate from repeats, transposons or virus. siRNAs can be involved in PTGS, which starts with the production of aberrant double-stranded RNAs (dsRNAs). The dsRNAs are cleaved by the Dicer RNase III complex to produce short fragments of 21-25 nucleotides and transported to the cytoplasm. These siRNAs are loaded in the RNA-induced silencing complex (RISC) and guide it towards complementary mRNAs. Targeted mRNAs are subsequently degraded by the Argonaute slicer activity, leading to the cleavage or suppression of translation of complementary mRNA sequences (Valencia-Sanchez et al. 2006; Carthew and Sontheimer 2009) (Fig. 5).

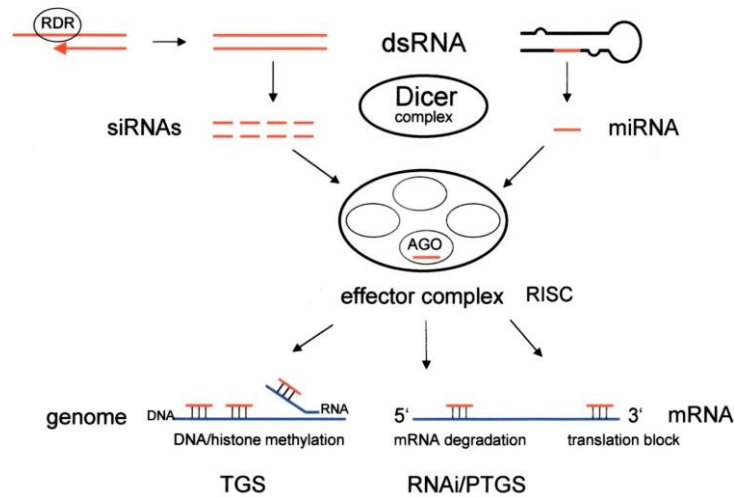


Figure 5. Schematic diagram of transcriptional (TGS) and post-transcriptional (PTGS) gene silencing. Adapted from M Matzke MA, Matzke AJM (2004).

In addition, siRNAs can produce TGS by promoting histone methylation (in animals) and DNA methylation (in plants). Recent studies in plants and mammals have stressed such a link between DNA methylation and the RNAi pathways (Huang et al. 2014; Zhao et al. 2016) in transposon silencing, indicating that both mechanisms are tightly related (Zilberman et al. 2003). It has been postulated that *de novo* methylation in plants as well as the maintenance of DNA methylation in asymmetric sequence contexts are mediated by RNA-directed DNA methylation (RdDM), which is an epigenetic process directed by siRNA or rasiRNA (repeat associated siRNA) (Wassenegger et al. 1994; Mette et al. 2000; Law and Jacobsen 2010). RdDM guided by siRNAs has been well characterized in plants, where *de novo* methylation can occur in all cytosine contexts (CG, CHG and CHH) in contrast to animals (He et al. 2011). According to this mechanism, the precursors of endogenous siRNA are synthesized by a specific nuclear RNA polymerases (Pol IV and Pol V) (Herr 2005; Wierzbicki et al. 2009) and processed by a RNA-dependent RNA polymerase (RdRP) homolog (RDR2) (Xie et al. 2004) and Dicer enzyme DCL3 (Matzke and Mosher 2014). These siRNAs bind to an Argonaute protein AGO4 and facilitate the targeting of complementary loci. Since AGO4 is physically associated with the *de novo* DNA methyltransferase DRM2 (Zhong et al. 2014), an interaction establishes the direct link between siRNAs and transcriptional silencing (Fig. 5).

In *Arabidopsis*, high-resolution, genome-wide methylation revealed that ~37% of methylated regions are associated with high levels of siRNAs (Zhang et al. 2006b), suggesting that the integrated interaction within the molecular pathways associated with RdDM mechanisms contribute to the maintenance of

transcriptional silencing (Law et al. 2013; Johnson et al. 2014). These observations highlight the function of methylation regulated by the TGS mechanism in the dynamic control of transposon activity.

Despite that most RNAi studies have been performed on plants and animals, this pathway has also been described in fungi, even though the precise molecular pathway has not yet been fully understood. Nevertheless, this mechanism is thought to play a crucial role in the control of development and genome integrity (Janbon et al. 2010). In the fungal model *N. crassa*, two different silencing pathways that are mediated by siRNAs have been described: quelling (Romano and Macino 1992) and meiotic silencing by unpaired DNA (MSUD) (Shiu et al. 2001). Both mechanisms seem to participate in genome defense against invasive nucleic acids, such as transposons and viruses (Dang et al. 2011).

The kingdom of Fungi

The kingdom of fungi includes some of the most important eukaryotic organisms in terms of ecological roles. Fungi are not able to photosynthesize and, similarly to animals, are classified as heterotrophic organisms. They acquire nutrients directly from their living hosts or by absorbing non-living organic matter from the external environment through an efficient enzymatic system. Some fungi digest nutrients by the release of enzymes that break down large organic molecules such as polysaccharides, proteins and lipids into smaller molecules, which are then absorbed into the fungal cells.

Recent studies based on molecular approaches suggest that they may have appeared more than 1000 million years ago (Parfrey et al. 2011), adding strength to the first dating based on fossil records (Redecker et al. 2000). Initially considered as primitive plant forms, following taxonomic classifications suggested that fungi were more closely related to animalia than to plant kingdom and that they diverged from the former about 1500 million years ago (Hedges et al. 2004). A reasonable estimation reports 1.5 million fungal species, though only about 5% have been formally classified (Hawksworth 2001).

The fungal kingdom encompasses an extraordinary diversity of taxa. The group *Eumycota* comprises organisms traditionally studied as ‘fungi’ and classified in four main phyla: Ascomycota, Basidiomycota, Chytridiomycota and Zygomycota (Alexopoulos et al. 1996; Webster and Weber 2007). However, recent taxonomic studies provide support for the recognition of additional fungal phyla, such as Glomeromycota and Microsporidia, which were initially identified as plant and animal parasites (Hibbett et al. 2007) (Fig. 6). A subkingdom Dikarya has been proposed for Ascomycota and Basidiomycota phyla (Hibbett et al. 2007), which include most of the so-called “higher fungi” and account for the vast majority of fungal species described up to now (James et al. 2006).

Most fungi are formed by multibranching cellular structures called hyphae that enable the colonization of various habitats. Abundantly distributed worldwide, they can grow in different ecological niches, principally as decomposers, pathogens or saprotrophs, and play essential roles in several processes, from nutrient cycling to production of active molecules of great biotechnological interest. Besides their use as a nutritional source in the form of mushrooms and truffles, they are involved in a number of food industrial processes (i.e., leavening agent for bread and fermentation of wine and beer). Given their ability to produce antibiotics and active metabolites, fungi are also a prominent source for pharmaceutical and medical applications.

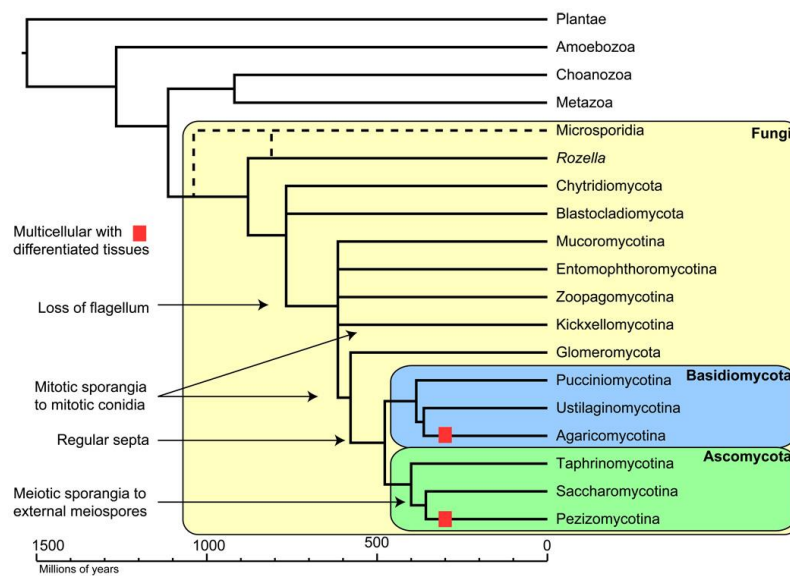


Figure 6. Phylogenetic tree and classification of fungi. Obtained from Stajich et al. (2009).

Fungal reproduction is complex and reflects genetic and ecological differences between phyla. Fungi can reproduce both sexually and asexually by formation of spores and can grow as unicellular or multicellular mycelia. During the life cycle, the asexual cycle produces mitospores, while the sexual cycle produces meiospores (Fig. 7). The development of haploid nuclei, the cellular (plasmogamy) and nuclear (karyogamy) fusions before meiosis, represent the key steps of sexual reproduction. In *Dikarya*, the life cycle of a sexually reproducing fungus alternates between monokaryotic and dikaryotic stages containing haploid and diploid nuclei. The haploid phase ends with nuclear fusion, and the diploid phase begins with the formation of the zygote (the diploid cell resulting from fusion of two compatible haploid nuclei). The meiotic event restores the haploid number of chromosomes and initiates the haploid phase through the

production of meiospores. In both Ascomycota and Basidiomycota, the fusion of nuclei (karyogamy) can occur much later than the fusion of cytoplasm (plasmogamy).

During this stage, which is defined as the dikaryotic phase, the two compatible haploid nuclei coexist and remain separate in one hyphal segment. The nuclear fusion and formation of diploid cells occurs only immediately before meiosis, where uninucleate spores are produced and dispersed.

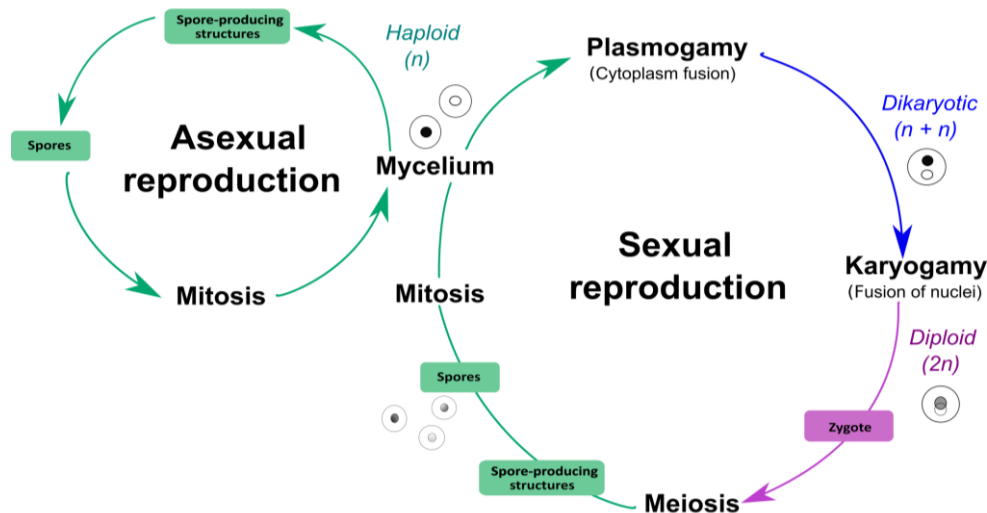


Figure 7. Generalized sexual and asexual life cycle in fungi.

The presence of specialized structures where spores are contained and the process involved in the formation of meiospores reflect the evolution of specific features considered for taxonomic classification. In Ascomycota, sexual spores are produced inside structures of the fruiting bodies, known as ascus, whereas members of Basidiomycota produce the basidiospores externally in a structure called basidia. Unlike basidiomycetes, the asexual reproduction is very common in members of the Ascomycota phylum, where spores known as conidia are produced by mitosis or budding (like in yeasts). Despite that the Ascomycota phylum comprises the largest number of members and the most relevant fungal models, many basidiomycete exemplars are commonly known for the presence of unique structures—e.g., mushrooms, which are a type of fruiting body of relevant interest for their nutritional values.

Fungal genomics: a picture from Basidiomycota

The phylum Basidiomycota comprises approximately 32% of the taxonomically categorized fungal species (Kirk et al. 2008), and it is extraordinarily important in agriculture, forestry and medicine.

This heterogeneous phylum includes plant (Johansson and Stenlid 1985) and human (Brown et al. 2007) pathogens, ectomycorrhizal symbionts (Martin et al. 2008) as well as most of the wood-decaying fungi (Floudas et al. 2012). Three main subphyla are identified within the Basidiomycota (Blackwell et al. 2006), depending on their lifestyle and the basidium structure. *Ustilaginomycotina* and *Pucciniomycotina* encompass mainly plant parasitic and crop pathogen lineages, known as smut and rust fungi. The third subphylum, *Agaricomycotina*, includes members that display morphological differences and produce basidia in various types of fruiting bodies. Some edible species are generally known for their excellent nutritional and organoleptic characteristics and are industrially cultivated for food production, such as the shiitake (*Lentinula edodes*), common button (*Agaricus bisporus*) and oyster (*Pleurotus ostreatus*) mushrooms. Besides these properties, basidiomycetes displaying ligninolytic activity have fueled several fields of research for their potential biotechnological applications (Morita et al. 2009; Riley et al. 2014).

The extraordinary ecological diversity reflected in a number of industrial applications make the characterization of basidiomycete genomes of increasing interest.

The recent throughput of phylogenetic comparative methods and applications of NGS technologies provided new insights from a large number of basidiomycete genomes, leading to a better understanding of their genetics and biology. According to these studies, basidiomycetes show an extreme variability in genome size, ranging from approximately 8 up to 190 Mb (Nemri et al. 2014) and a number of protein-coding genes commonly correlated with their lifestyle (Grigoriev et al. 2014). Interestingly, it has also been estimated that a variable fraction of sequenced genomes is occupied by transposable elements (reviewed by Castanera et al. 2017), which cover from small (0.1%) to very large genome fractions (~ 45%).

Lifestyle and genomics of the basidiomycete *Pleurotus ostreatus*

Pleurotus ostreatus, which is commonly referred to as the oyster mushroom, is an edible basidiomycete that belongs to the monophyletic *Pleurotacea* family within the order Agaricales (Thorn et al. 2000). It produces fleshy fruiting bodies that range in color from white to gray or even dark-brown and has shell-shaped, semicircular caps that span 5-25 cm, although some variability can originate from their widespread distribution (Woller 2007) (Fig. 8). Its biological plasticity and the fact that it can grow relatively easy on several natural and artificial substrates compared to other gourmet mushrooms make this species

particularly suitable for mushroom cultivation. In the natural environment, this white-rot fungus lives as a saprophyte on dead or decaying wood and plays a key role as an active lignin degrader (Eugenio and Anderson 1968; Sánchez 2010).



Figure 8. *Pleurotus ostreatus* fruiting bodies. Obtained from Genetics and Microbiology Research Group, Public University of Navarre.

Similarly to other wood decomposers fungi, *P. ostreatus* possess an efficient ligninolytic enzymatic system (Langmead and Salzberg 2012; Riley et al. 2014), which is described as an ancestral toolkit within members of Agaricomycota (Riley et al. 2014; Alfaro et al. 2016). Enzymes involved in lignin degradation (e.g. laccases, monooxygenases, peroxidases) produced by *P. ostreatus* and other ligninolytic fungi have gained considerable attention for several biotechnological applications, such as agro-waste recycling (Aggelis et al. 2003), bioremediation (Purnomo et al. 2010) or pulp bleaching (Dwivedi et al. 2010). Additionally, the simplicity of its life cycle and its ability to grow and fructify under laboratory conditions make it an interesting model for fungal genetics. Since the early 1990s, a commercial strain identified as *P. ostreatus* N001 has been used as a model organism for cultivated and lignin degrading basidiomycetes in the Genetics and Microbiology research group at the Public University of Navarre.

Its life cycle alternates between monokaryotic (haploid cells) and dikaryotic (dihaploid cells) phases (Eugenio and Anderson 1968). Two compatible monokaryons can mate to generate a dikaryon (sexually competent form) in which the two haploid nuclei remain independent throughout the vegetative growth and fruit body development until the occurrence of karyogamy. The karyogamy (diploid condition) occurs only in the basidia at the end of the life cycle, immediately before the meiotic division that produces four haploid basidiospores.

In 1999, the two nuclei harbored in the dikaryotic strain N001 were isolated by a de-dikaryotization process based on an enzymatic treatment of N001 protoplast, leading to the obtainment and the subsequent identification of the parental monokaryons containing either of the two nuclei (Larraya et al. 1999). This process yielded two compatible protoclones (monokaryons PC15 and PC9) bearing each one of the non-recombined haploid nuclei. The availability of these strains as well as a meiotic progeny derived from the N001 dikaryotic strain enabled further genetic studies of *P. ostreatus*.

The molecular karyotype determined the presence of 11 chromosomes (Larraya et al. 1999) that fit with 11 linkage groups that were subsequently described in the genetic map of *P. ostreatus* (Larraya et al. 2000; Park et al. 2006). Using this framework, it has been possible to map quantitative linkage loci that control the mycelial growth rate in both monokaryotic and dikaryotic strains (Larraya et al. 2002) and characterize production and quality traits (Larraya et al. 2003) and the activity of ligninolytic enzymes (Santoyo et al. 2008). These data supported the full sequencing of the PC15 and PC9 genomes in collaboration with the DOE Joint Genome Institute (JGI, <http://jgi.doe.gov/>), and their genome assemblies and annotations are publicly available in the MycoCosm Database (Riley et al. 2014b; Alfaro et al. 2016; Castanera et al. 2016).

The genome sequencing yielded 34.3 and 35.6 Mbp assemblies for PC15 and PC9, respectively, with a similar number of annotated genes (Table 1).

Table 1. Summary of PC15 and PC9 genome sequencing and annotation

Nuclear genome Assembly	PC15 v2.0	PC9 v1.0
Nuclear genome size (MbP)	34.3	35.6
Number of nuclear scaffolds	11	572
Number of gene models	12,330	12,206
Sequencing method	Sanger	454 + Sanger
Total repeats*	1051 (204)	873 (65)
Repeats content (%)	6.20	2.50
Genome assembly gaps (%)	0	9.72

* RepeatMasker reconstructed copies (Full-length copies are shown in parentheses).

The PC15 assembly contains 12 scaffolds that match with the 12 chromosomes of *P. ostreatus* that were previously described (11 nuclear + 1 mitochondrial), while the assembly of the PC9 genome led to 572 scaffolds that account for a total of 476 gaps and cover 9.72 % of the assembly length. Both genomes showed an overall conserved macrosynteny, although comparative genomic analyses uncovered an important amount of non-homologous regions that are reminiscent of genomic rearrangements promoted by the mobilization of repeated elements. In a recent study, 80 TE families were identified, and they accounted for 6.2 and 2.5 % of PC15 and PC9 genome size, respectively (Castanera et al. 2016).

The results of this work showed that both Class I and Class II TEs populate the *P. ostreatus* genome (Fig. 9), and there is a clear abundance of LTR-retrotransposons and a strong preference to accumulate within TE-rich clusters.

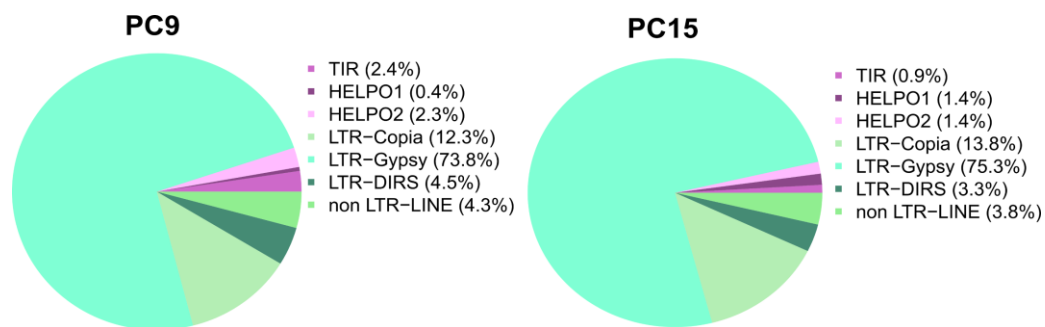


Figure 9. Class I and Class II transposon coverage of the genome of PC9 and PC15 *P. ostreatus* reference genomes.

The revolutionary advent of high-throughput sequencing technologies has triggered the development of a whole range of methods for genome-wide mapping of DNA methylation and comparative transcriptomic profiling, which have been separately used to uncover specific genomic contexts.

In the last few years, however, the integrated application of these methods enabled the combined analysis of epigenetic events and transcriptional profiles, highlighting their complex interaction during the development and transposons inactivation in many eukaryotes at multicellular (Mishra et al. 2011; Chodavarapu et al. 2012) and single-cell levels (Angermueller et al. 2016; Hu et al. 2016).

References

- Aggelis G, Iconomou D, Christou M, Bokas D, Kotzailias S, Christou G, Tsagou V, Papanikolaou S (2003) Phenolic removal in a model olive oil mill wastewater using *Pleurotus ostreatus* in bioreactor cultures and biological evaluation of the process. *Water Res* 37:3897–3904. doi: 10.1016/S0043-1354(03)00313
- Alexopoulos CJ, Mims CW, Blackwell M (1996) *Introductory Mycology*, 4th Edition. Wiley, New York 868. doi:10.2105/AJPH.43.6_Pt_1.781-a
- Alfaro M, Castanera R, Lavín JL, Grigoriev I V, Oguiza JA, Ramírez L, Pisabarro AG (2016) Comparative and transcriptional analysis of the predicted secretome in the lignocellulose-degrading basidiomycete fungus *Pleurotus ostreatus*. *Environ Microbiol*. doi: 10.1111/1462-2920.13360
- Almeida R, Allshire RC (2005) RNA silencing and genome regulation. *Trends Cell Biol*. 15:251–258.
- Bentley DR (2006) Whole-genome re-sequencing. *Curr Opin Genet Dev* 16:545–552. doi: 10.1016/j.gde.2006.10.009
- Biessmann H, Valgeirsdottir K, Lofsky A, Chin C, Ginther B, Levis RW, Pardue ML (1992) HeT-A, a transposable element specifically involved in “healing” broken chromosome ends in *Drosophila melanogaster*. *Mol Cell Biol* 12:3910–3918. doi: 10.1128/MCB.12.9.3910.Updated
- Bird A (2002) DNA methylation patterns and epigenetic memory. *Genes Dev*. 16:6–21.
- Blackwell M, Hibbett DS, Taylor JW, Spatafora JW (2006) Research Coordination Networks: a phylogeny for kingdom Fungi (Deep Hypha). *Mycologia* 98:829–37.
- Bregliano JC, Kidwell MG (1983) Hybrid dysgenesis determinants.
- Britten RJ (1996) DNA sequence insertion and evolutionary variation in gene regulation. *Proc Natl Acad Sci U S A* 93:9374–9377. doi: 10.1073/pnas.93.18.9374
- Brown SM, Campbell LT, Lodge JK (2007) *Cryptococcus neoformans*, a fungus under stress. *Curr. Opin. Microbiol*. 10:320–325.
- Brunner AL, Johnson DS, Kim SW, Valouev A, Reddy TE, Neff NF, Anton E, Medina C, Nguyen L, Chiao E, Oyolu CB, Schroth GP, Absher DM, Baker JC, Myers RM (2009) Distinct DNA methylation patterns characterize differentiated human embryonic stem cells and developing human fetal liver. *Genome Res* 19:1044–1056. doi: 10.1101/gr.088773.108.
- Carthew RW, Sontheimer EJ (2009) Origins and Mechanisms of miRNAs and siRNAs. *Cell* 136:642–55. doi: 10.1016/j.cell.2009.01.035
- Castanera R, Borgognone A, Pisabarro AG, Ramírez L (2017) Biology, dynamics, and applications of transposable elements in basidiomycete fungi. *Appl Microbiol Biotechnol* 101:1337–1350. doi: 10.1007/s00253-017-8097-8
- Castanera R, López-Varas L, Borgognone A, LaButti K, Lapidus A, Schmutz J, Grimwood J, Pérez G, Pisabarro AG, Grigoriev I V, Stajich JE, Ramírez L (2016) Transposable Elements versus the Fungal Genome: Impact on Whole-Genome Architecture and Transcriptional Profiles. *PLoS Genet* 12:e1006108. doi: 10.1371/journal.pgen.1006108
- Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE (2008) Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. *Nature* 452:215–219. doi: 10.1038/nature06745
- Collins FS, Morgan M, Patrinos A (2003) The Human Genome Project: Lessons from Large-Scale Biology. *Science* (80-) 300:286–290. doi: 10.1126/science.1084564
- Consortium IHGS (2001) Initial sequencing and analysis of the human genome. *Nature* 409:860–921. doi: 10.1038/35057062
- Costa FF (2008) Non-coding RNAs, epigenetics and complexity. *Gene* 410:9–17.
- Dang Y, Yang Q, Xue Z, Liu Y (2011) RNA interference in fungi: pathways, functions, and applications. *Eukaryot Cell* 10:1148–55. doi: 10.1128/EC.05109-11
- Doolittle WF, Sapienza C (1980) Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284:601–603. doi: 10.1038/284601a0
- Dwivedi P, Vivekanand V, Pareek N, Sharma A, Singh RP (2010) Bleach Enhancement of Mixed Wood Pulp by Xylanase–Laccase Concoction Derived Through Co-culture Strategy. *Appl Biochem Biotechnol* 160:255–268. doi: 10.1007/s12010-009-8654-4
- Dyachenko O V, Schevchuk T V, Kretzner L, Buryanov YI, Smith SS (2010) Human non-CG methylation: are human stem cells plant-like? *Epigenetics* 5:569–72. doi: 10.4161/epi.5.7.12702
- Eichten SR, Briskine R, Song J, Li Q, Swanson-Wagner R, Hermanson PJ, Waters AJ, Starr E, West PT, Tiffin P, Myers CL, Vaughn MW, Springer NM (2013) Epigenetic and Genetic Influences on DNA Methylation Variation in Maize Populations. *Plant Cell* 25:2783–2797. doi: 10.1105/tpc.113.114793
- El Baidouri M, Panaud O (2013) Comparative Genomic Paleontology across Plant Kingdom Reveals the Dynamics of TE-Driven Genome Evolution. *Genome Biol Evol* 5:954–965. doi: 10.1093/gbe/evt025
- Eugenio CP, Anderson NA (1968) The Genetics and Cultivation of *Pleurotus ostreatus*. *Mycologia* 60:627. doi: 10.2307/3757430
- Fagard M, Vaucheret H (2000) (Trans)gene silencing in plants: How Many Mechanisms? *Annu Rev Plant Physiol Plant Mol Biol* 51:167–94. doi: 10.1146/annurev.arplant.51.1.167
- Floudas D, Binder M, Riley R, Barry K, Blanchette RA, Henrissat B, Martinez AT, Otillar R, Spatafora JW, Yadav JS, Aerts A, Benoit I, Boyd A, Carlson A, Findley K, Foster B, Gaskell J, Glotzer D, Gorecki P, Heitman J, Hesse C, Hori C, Igarashi K, Jurgens JA, Kallen N, Kersten P, Kohler A, Kues U, Kumar TKA, Kuo A, LaButti K, Larrondo LF, Lindquist E, Ling A, Lombard V, Lucas S, Lundell T, Martin R, McLaughlin DJ, Morgenstern I, Morin E, Murat C, Nagy LG, Nolan M, Ohm RA, Patyshakuliyeva A, Rokas A, Ruiz-Duenas FJ, Sabat G, Salamov A, Samejima M, Schmutz J, Slot JC, St. John F, Stenlid J, Sun H, Sun S, Syed K, Tsang A, Wiebenga A, Young D, Pisabarro A, Eastwood DC, Martin F, Cullen F, Grigoriev I V., Hibbett DS (2012) The Paleozoic Origin of Enzymatic Lignin Decomposition Reconstructed from 31 Fungal Genomes. *Science* (80-) 336:1715–1719. doi: 10.1126/science.1221748
- Freitag M, Selker EU (2005) Controlling DNA methylation: Many roads to one modification. *Curr. Opin. Genet. Dev*. 15:191–199.
- Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A* 89:1827–31. doi: 10.1073/PNAS.89.5.1827
- Galagan JE, Calvo SE, Cuomo C, Ma L-J, Wortman JR, Batzoglou S, Lee S-I, Baştürkmen M, Spevak CC, Clutterbuck J, Kapitonov V, Jurka J, Scaccocchio C, Bell-Pedersen D, Griffiths-Jones S, Doonan JH, Yu J, Vienken K, Pain A, Freitag M, Selker EU, Archer DB, Peñalva MÁ, Oakley BR, Momany M, Tanaka T, Kumagai T, Asai K, Machida M, Nierman WC, Denning DW, Caddick M, Hynes M, Paoletti M, Fischer R, Miller B, Dyer P, Sachs MS, Osmani SA, Birren BW (2005) Sequencing of *Aspergillus nidulans* and comparative analysis with *A. fumigatus* and *A. oryzae*. *Nature* 438:1105–1115. doi: 10.1038/nature04341

- Gao D, Jiang N, Wing R a, Jiang J, Jackson S a (2015) Transposons play an important role in the evolution and diversification of centromeres among closely related species. *Front Plant Sci* 6:216. doi: 10.3389/fpls.2015.00216
- Gilbert W, Maxam A (1973) The nucleotide sequence of the lac operator. *Proc Natl Acad Sci U S A* 70:3581–4.
- Goodwin TJD, Ormandy JE, Poulter RTM (2001) L1-like non-LTR retrotransposons in the yeast *Candida albicans*. *Curr Genet* 39:83–91. doi: 10.1007/s002940000181
- Grigoriev I V., Nikitin R, Haridas S, Kuo A, Ohm R, Otilar R, Riley R, Salamov A, Zhao X, Korzeniewski F, Smirnova T, Nordberg H, Dubchak I, Shabalov I (2014) MycoCosm portal: Gearing up for 1000 fungal genomes. *Nucleic Acids Res.* doi: 10.1093/nar/gkt1183
- Hamilton AJ, Baulcombe DC (1999) A species of small antisense RNA in posttranscriptional gene silencing in plants. *Science* 286:950–2.
- Harris RA, Wang T, Coarfa C, Nagarajan RP, Hong C, Downey SL, Johnson BE, Fouse SD, Delaney A, Zhao Y, Olshen A, Ballinger T, Zhou X, Forsberg KJ, Gu J, Echipare L, O'Geen H, Lister R, Pelizzola M, Xi Y, Epstein CB, Bernstein BE, Hawkins RD, Ren B, Chung W-Y, Gu H, Bock C, Gnirke A, Zhang MQ, Haussler D, Ecker JR, Li W, Farnham PJ, Waterland RA, Meissner A, Marra MA, Hirst M, Milosavljevic A, Costello JF (2010) Comparison of sequencing-based methods to profile DNA methylation and identification of monoallelic epigenetic modifications. *Nat Biotechnol* 28:1097–1105. doi: 10.1038/nbt.1682
- Hawksworth DL (2001) The magnitude of fungal diversity: the 1.5 million species estimate revisited. *Mycol Res* 105:1422–1432. doi: 10.1017/S0953756201004725
- He X-J, Chen T, Zhu J-K (2011) Regulation and function of DNA methylation in plants and animals. *Cell Res* 21:442–465. doi: 10.1038/cr.2011.23
- Hedges S, Blair JE, Venturi ML, Shoe JL (2004) A molecular timescale of eukaryote evolution and the rise of complex multicellular life. *BMC Evol Biol* 4:2. doi: 10.1186/1471-2148-4-2
- Herr AJ (2005) RNA Polymerase IV Directs Silencing of Endogenous DNA. *Science* (80-) 308:118–120. doi: 10.1126/science.1106910
- Hibbett DS, Binder M, Bischoff JF, Blackwell M, Cannon PF, Eriksson OE, Huhndorf S, James T, Kirk PM, Lücking R, Thorsten Lumbsch H, Lutzoni F, Matheny PB, McLaughlin DJ, Powell MJ, Redhead S, Schoch CL, Spatafora JW, Stalpers JA, Vilgalys R, Aime MC, Aptroot A, Bauer R, Begerow D, Benny GL, Castlebury LA, Crous PW, Dai YC, Gams W, Geiser DM, Griffith GW, Gueidan C, Hawksworth DL, Hestmark G, Hosaka Kurtzman CP, Larsson KH, Lichtwardt R, Longcore J, Miadlikowska J, Miller A, Moncalvo JM, Mozley-Standridge S, Oberwinkler F, Parmasto E, Reeb V, Rogers JD, Roux C, Ryvarden L, Sampaio JP, Schüßler A, Sugiyama J, Thorn RG, Tibell L, Untereiner WA, Walker C, Wang Z, Weir A, Weiss M, White MM, Winka K, Yao YJ, Zhang N (2007) A higher-level phylogenetic classification of the Fungi. *Mycol Res* 111:509–547. doi: 10.1016/j.mycres.2007.03.004
- Holliday R, Pugh JE (1975) DNA modification mechanisms and gene activity during development. *Science* 187:226–232. doi: 10.1126/science.1111098
- Holton NJ, Goodwin TJ, Butler MI, Poulter RT (2001) An active retrotransposon in *Candida albicans*. *Nucleic Acids Res* 29:4014–4024.
- Huang CRL, Burns KH, Boeke JD (2012) Active transposition in genomes. *Annu Rev Genet* 46:651–75. doi: 10.1146/annurev-genet-110711-155616
- Huang Y-Z, Sun J-J, Zhang L-Z, Li C-J, Womack JE, Li Z-J, Lan X-Y, Lei C-Z, Zhang C-L, Zhao X, Chen H (2014) Genome-wide DNA Methylation Profiles and Their Relationships with mRNA and the microRNA Transcriptome in Bovine Muscle Tissue (*Bos taurine*). *Sci Rep* 4:6546. doi: 10.1038/srep06546
- James TY, Kauff F, Schoch CL, Matheny PB, Hofstetter V, Cox CJ, Celio G, Gueidan C, Fraker E, Miadlikowska J, Lumbsch HT, Rauhut A, Reeb V, Arnold AE, Amtoft A, Stajich JE, Hosaka K, Sung G-H, Johnson D, O'Rourke B, Crockett M, Binder M, Curtis JM, Slot JC, Wang Z, Wilson AW, Schüßler A, Longcore JE, O'Donnell K, Mozley-Standridge S, Porter D, Letcher PM, Powell MJ, Taylor JW, White MM, Griffith GW, Davies DR, Humber RA, Morton JB, Sugiyama J, Rossman AY, Rogers JD, Pfister DH, Hewitt D, Hansen K, Hambleton S, Shoemaker RA, Kohlmeyer J, Volkmann-Kohlmeyer B, Spotts RA, Serdani M, Crous PW, Hughes KW, Matsuura K, Langer E, Langer G, Untereiner WA, Lücking R, Büdel B, Geiser DM, Aptroot A, Diederich P, Schmitt I, Schultz M, Yahr R, Hibbett DS, Lutzoni F, McLaughlin DJ, Spatafora JW, Vilgalys R (2006) Reconstructing the early evolution of Fungi using a six-gene phylogeny. *Nature* 443:818–822. doi: 10.1038/nature05110
- Janbon G, Maeng S, Yang DH, Ko YJ, Jung KW, Moyrand F, Floyd A, Heitman J, Bahn YS (2010) Characterizing the role of RNA silencing components in *Cryptococcus neoformans*. *Fungal Genet Biol* 47:1070–1080. doi: 10.1016/j.fgb.2010.10.005
- Johansson M, Stenlid J (1985) Infection of roots of Norway spruce (*Picea abies*) by *Heterobasidion annosum*. *For Pathol* 15:32–45. doi: 10.1111/j.1439-0329.1985.tb01040.x
- Johnson DS, Mortazavi A, Myers RM, Wold B (2007) Genome-wide mapping of in vivo protein-DNA interactions. *Science* 316:1497–502. doi: 10.1126/science.1141319
- Johnson LM, Du J, Hale CJ, Bischof S, Feng S, Chodavarapu RK, Zhong X, Marson G, Pellegrini M, Segal DJ, Patel DJ, Jacobsen SE (2014) SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation. *Nature* 507:124–128. doi: 10.1038/nature12931
- Kapitonov V V, Jurka J (2001) Rolling-circle transposons in eukaryotes. *Proc Natl Acad Sci U S A* 98:8714–9. doi: 10.1073/pnas.151269298
- Kidwell MG, Lisch D (1997) Transposable elements as sources of variation in animals and plants. *Proc Natl Acad Sci USA* 94:7704–7711. doi: 10.1073/pnas.94.15.7704
- Kim VN (2005) Small RNAs: classification, biogenesis, and function. *Mol Cells* 19:1–15. doi: 806 [pii]
- Kirk, P.M.; Cannon, P.F.; Minter, D.W.; Stalpers JA (2008) *Dictionary of the Fungi*, 10th edn. CAB International, Wallingford, UK
- Laird PW (2010) Principles and challenges of genomewide DNA methylation analysis. *Nat Rev Genet* 11:191–203. doi: 10.1038/nrg2732
- Lankenau D, Volf J-N (2009) Transposons and the Dynamic Genome.
- Larraya LM, Alfonso M, Pisabarro AG, Ramírez L (2003) Mapping of genomic regions (quantitative trait loci) controlling production and quality in industrial cultures of the edible basidiomycete *Pleurotus ostreatus*. *Appl Environ Microbiol* 69:3617–3625. doi: 10.1128/AEM.69.6.3617-3625.2003
- Larraya LM, Idareta E, Arana D, Ritter E, Pisabarro AG, Ramírez L (2002) Quantitative trait loci controlling vegetative growth rate in the edible basidiomycete *Pleurotus ostreatus*. *Appl Environ Microbiol* 68:1109–14.
- Larraya LM, Pérez G, Peñas MM, Baars JJ, Mikosch TS, Pisabarro AG, Ramírez L (1999) Molecular karyotype of the white rot fungus *Pleurotus ostreatus*. *Appl Environ Microbiol* 65:3413–7.
- Larraya LM, Pérez G, Ritter E, Pisabarro AG, Ramírez L (2000) Genetic linkage map of the edible basidiomycete *Pleurotus ostreatus*. *Appl Environ Microbiol* 66:5290–300.

- Law JA, Du J, Hale CJ, Feng S, Krajewski K, Palanca AMS, Strahl BD, Patel DJ, Jacobsen SE (2013) Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. *Nature* 498:385–389. doi: 10.1038/nature12178
- Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* 11:204–220. doi: 10.1038/nrg2719
- Lippman Z, Gendrel A-V, Black M, Vaughn MW, Dedhia N, Richard McCombie W, Lavine K, Mittal V, May B, Kasschau KD, Carrington JC, Doerge RW, Colot V, Martienssen R (2004) Role of transposable elements in heterochromatin and epigenetic control. *Nature* 430:471–476. doi: 10.1038/nature02651
- Lister R, Ecker JR (2009) Finding the fifth base: Genome-wide sequencing of cytosine methylation. *Genome Res.* 19:959–966.
- Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR (2008) Highly Integrated Single-Base Resolution Maps of the Epigenome in *Arabidopsis*. *Cell* 133:523–536. doi: 10.1016/j.cell.2008.03.029
- Mardis ER (2008) Next-Generation DNA Sequencing Methods. *Annu Rev Genomics Hum Genet* 9:387–402. doi: 10.1146/annurev.genom.9.081307.164359
- Margulies M, Egholm M, Altman WE, Attiya S, Bader JS, Bemben LA, Berka J, Braverman MS, Chen Y-J, Chen Z, Dewell SB, Du L, Fierro JM, Gomes X V., Godwin BC, He W, Helgesen S, Ho CH, Irzyk GP, Jando SC, Alenquer MLI, Jarvie TP, Jirage KB, Kim J-B, Knight JR, Lanza JR, Leamon JH, Lefkowitz SM, Lei M, Li J, Lohman KL, Lu H, Makhijani VB, McDade KE, McKenna MP, Myers EW, Nickerson E, Nobile JR, Plant R, Puc BP, Ronan MT, Roth GT, Sarkis GJ, Simons JF, Simpson JW, Srinivasan M, Tartaro KR, Tomasz A, Vogt KA, Volkmer GA, Wang SH, Wang Y, Weiner MP, Yu P, Begley RF, Rothberg JM, Rothberg JM (2005) Genome sequencing in microfabricated high-density picolitre reactors. *Nature* 437:376–80. doi: 10.1038/nature03959
- Martin F, Aerts a, Ahrén D, Brun a, Danchin EGJ, Duchaussoy F, Gibon J, Kohler a, Lindquist E, Pereda V, Salamov a, Shapiro HJ, Wuyts J, Blaudez D, Buée M, Brokstein P, Canbäck B, Cohen D, Courty PE, Coutinho PM, Delaruelle C, Detter JC, Deveau a, DiFazio S, Duplessis S, Fraissinet-Tachet L, Lucic E, Frey-Klett P, Fourrey C, Feussner I, Gay G, Grimwood J, Hoegger PJ, Jain P, Kilaru S, Labbé J, Lin YC, Legué V, Le Tacon F, Marmeisse R, Melayah D, Montanini B, Muratet M, Nehls U, Niculita-Hirzel H, Oudot-Le Secq MP, Peter M, Quesneville H, Rajashekar B, Reich M, Rouhier N, Schmutz J, Yin T, Chalot M, Henrissat B, Kúes U, Lucas S, Van de Peer Y, Podila GK, Polle a, Pukkila PJ, Richardson PM, Rouzé P, Sanders IR, Stajich JE, Tunlid a, Tuskan G, Grigoriev I V (2008) The genome of *Laccaria bicolor* provides insights into mycorrhizal symbiosis. *Nature* 452:88–92. doi: 10.1038/nature06556
- Matzke MA, Mosher RA (2014) RNA-directed DNA methylation: an epigenetic pathway of increasing complexity. *Nat Rev Genet* 15:394–408. doi: 10.1038/nrg3683
- Maxam AM, Gilbert W (1977) A new method for sequencing DNA. *Proc Natl Acad Sci U S A* 74:560–4.
- McClintock B (1984) The significance of responses of the genome to challenge. *Science* 226:792–801.
- McCLINTOCK B (1951) Chromosome organization and genic expression. *Cold Spring Harb Symp Quant Biol* 16:13–47. doi: 10.1101/SQB.1951.016.01.004
- McClintock B (1947) Cytogenetic studies of maize and *Neurospora*. *Carnegie Inst Wash Yearb* 46:146–152
- ST – Cytogenetic studies of maize and *Neu*.
- Mette MF, Aufsatz W, van der Winden J, Matzke MA, Matzke AJ (2000) Transcriptional silencing and promoter methylation triggered by double-stranded RNA. *EMBO J* 19:5194–201. doi: 10.1093/emboj/19.19.5194
- Montanini B, Chen P-Y, Morselli M, Jaroszewicz A, Lopez D, Martin F, Ottonello S, Pellegrini M (2014) Non-exhaustive DNA methylation-mediated transposon silencing in the black truffle genome, a complex fungal genome with massive repeat element content. *Genome Biol* 15:411. doi: 10.1186/s13059-014-0411-5
- Morita T, Fukuoka T, Imura T, Kitamoto D (2009) Production of glycolipid biosurfactants by basidiomycetous yeasts. *Biotechnol Appl Biochem* 53:39–49. doi: 10.1042/BA20090033
- Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nat Methods* 5:621–8. doi: 10.1038/nmeth.1226
- Naito K, Cho E, Yang G, Campbell M a, Yano K, Okumoto Y, Tanisaka T, Wessler SR (2006) Dramatic amplification of a rice transposable element during recent domestication. *Proc Natl Acad Sci U S A* 103:17620–17625. doi: 10.1073/pnas.0605421103
- Nemri A, Saunders DGO, Anderson C, Upadhyaya NM, Win J, Lawrence GJ, Jones DA, Kamoun S, Ellis JG, Dodds PN (2014) The genome sequence and effector complement of the flax rust pathogen *Melampsora lini*. *Front Plant Sci* 5:98. doi: 10.3389/fpls.2014.00098
- Orgel LE, Crick FH (1980) Selfish DNA: the ultimate parasite. *Nature* 284:604–607. doi: 10.1038/284604a0
- Parfrey LW, Lahr DJG, Knoll AH, Katz LA (2011) Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc Natl Acad Sci U S A* 108:13624–9. doi: 10.1073/pnas.1110633108
- Park S-K, Peñas MM, Ramírez L, Pisabarro AG (2006) Genetic linkage map and expression analysis of genes expressed in the lamellae of the edible basidiomycete *Pleurotus ostreatus*. *Fungal Genet Biol* 43:376–387. doi: 10.1016/j.fgb.2006.01.008
- Popova EY, Claxton DF, Lukasova E, Bird PI, Grigoryev SA (2006) Epigenetic heterochromatin markers distinguish terminally differentiated leukocytes from incompletely differentiated leukemia cells in human blood. *Exp Hematol* 34:453–462. doi: 10.1016/j.exphem.2006.01.003
- Purnomo AS, Mori T, Kamei I, Nishii T, Kondo R (2010) Application of mushroom waste medium from *Pleurotus ostreatus* for bioremediation of DDT-contaminated soil. *Int Biodeterior Biodegradation* 64:397–402. doi: 10.1016/j.ibiod.2010.04.007
- Rabbani B, Tekin M, Mahdieh N (2014) The promise of whole-exome sequencing in medical genetics. *J Hum Genet* 59:5–15. doi: 10.1038/jhg.2013.114
- Redecker D, Kodner R, Graham LE (2000) Glomalean fungi from the Ordovician. *Science* 289:1920–1921. doi: 10.1126/science.289.5486.1920
- Regulski M, Lu Z, Kendall J, Donoghue MTA, Reinders J, Llaca V, Deschamps S, Smith A, Levy D, McCombie WR, Tingey S, Rafalski A, Hicks J, Ware D, Martienssen RA (2013) The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. *Genome Res* 23:1651–62. doi: 10.1101/gr.153510.112
- Riley R, Salamov AA, Brown DW, Nagy LG, Floudas D, Held BW, Levasseur A, Lombard V, Morin E, Otiillar R, Lindquist EA, Sun H, LaButti KM, Schmutz J, Jabbour D, Luo H, Baker SE, Pisabarro AG, Walton JD, Blanchette RA, Henrissat B, Martin F, Cullen D, Hibbett DS, Grigoriev I V (2014b) Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. *Proc Natl Acad Sci U S A* 111:9923–8. doi: 10.1073/pnas.1400592111
- Romano N, Macino G (1992) Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. *Mol Microbiol* 6:3343–53.

- Sánchez C (2010) Cultivation of *Pleurotus ostreatus* and other edible mushrooms. *Appl Microbiol Biotechnol* 85:1321–37. doi: 10.1007/s00253-009-2343-7
- Sanger F, Nicklen S, Coulson AR (1977) DNA sequencing with chain-terminating inhibitors. *Proc Natl Acad Sci U S A* 74:5463–7.
- Santoyo F, González AE, Terrón MC, Ramírez L, Pisabarro AG (2008) Quantitative linkage mapping of lignin-degrading enzymatic activities in *Pleurotus ostreatus*. *Enzyme Microb Technol* 43:137–143. doi: 10.1016/j.enzymictec.2007.11.007
- Schmitz RJ, Ecker JR (2012) Epigenetic and epigenomic variation in *Arabidopsis thaliana*. *Trends Plant Sci* 17:149–154.
- Selker EU, Cambareli EB, Jensen BC, Haack KR (1987) Rearrangement of duplicated DNA in specialized cells of *Neurospora*. *Cell* 51:741–752. doi: 10.1016/0092-8674(87)90097-3
- Selker EU, Tountas NA, Cross SH, Margolin BS, Murphy JG, Bird AP, Freitag M (2003) The methylated component of the *Neurospora crassa* genome. *Nature* 422:893–897. doi: 10.1038/nature01564
- Shiu PK, Raju NB, Zickler D, Metzberg RL (2001) Meiotic silencing by unpaired DNA. *Cell* 107:905–16.
- Singer T, McConnell MJ, Marchetto MCN, Coufal NG, Gage FH (2010) LINE-1 retrotransposons: Mediators of somatic variation in neuronal genomes? *Trends Neurosci* 33:345–354. doi: 10.1016/j.tins.2010.04.001
- Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* 8:272–85. doi: 10.1038/nrg2072
- Staden R (1979) A strategy of DNA sequencing employing computer programs. *Nucleic Acids Res* 6:2601–2610. doi: 10.1093/nar/6.7.2601
- Suzuki MM, Bird A (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* 9:465–76. doi: 10.1038/nrg2341
- Szymanski M, Barciszewska MZ, Zywicki M, Barciszewski J (2003) Noncoding RNA transcripts. *J Appl Genet* 44:1–19.
- Thorn RG, Moncalvo J-M, Reddy CA, Vilgalys R (2000) Phylogenetic Analyses and the Distribution of Nematophagy Support a Monophyletic Pleurotaceae within the Polyphyletic Pleurotoid-Lentinoid Fungi. *Mycologia* 92:241. doi: 10.2307/3761557
- Valencia-Sanchez MA, Liu J, Hannon GJ, Parker R (2006) Control of translation and mRNA degradation by miRNAs and siRNAs. *Genes Dev* 20:515–24. doi: 10.1101/gad.1399806
- Valouev A, Ichikawa J, Tonthat T, Stuart J, Ranade S, Peckham H, Zeng K, Malek JA, Costa G, McKernan K, Sidow A, Fire A, Johnson SM (2008) A high-resolution, nucleosome position map of *C. elegans* reveals a lack of universal sequence-dictated positioning. *Genome Res* 18:1051–1063. doi: 10.1101/gr.076463.108
- Vaughn MW, Tanurdžić M, Lippman Z, Jiang H, Carrasquillo R, Rabinowicz PD, Dedhia N, McCombie WR, Agier N, Bulski A, Colot V, Doerge R., Martienssen RA (2007) Epigenetic Natural Variation in *Arabidopsis thaliana*. *PLoS Biol* 5:e174. doi: 10.1371/journal.pbio.0050174
- Venter JC, Adams MD, Myers EW, Li PW, Mural RJ, Sutton GG, Smith HO, Yandell M, Evans CA, Holt RA, Gocayne JD, Amanatides P, Ballew RM, Huson DH, Wortman JR, Zhang Q, Kodira CD, Zheng XH, Chen L, Skupski M, Subramanian G, Thomas PD, Zhang J, Gabor Miklos GL, Nelson C, Broder S, Clark AG, Nadeau J, McKusick VA, Zinder N, Levine AJ, Roberts RJ, Simon M, Slayman C, Hunkapiller M, Bolanos R, Delcher A, Dew I, Fasulo D, Flanigan M, Florea L, Halpern A, Hanchenalli S, Kravitz S, Levy S, Mobarry C, Reinert K, Remington K, Abu-Threideh J, Beasley E, Biddick K, Bonazzi V, Brandon R, Cargill M, Chandramouliswaran I, Charlab R, Chaturvedi K, Deng Z, Di Francesco V, Dunn P, Eilbeck K, Evangelista C, Gabrielian AE, Gan W, Ge W, Gong F, Gu Z, Guan P, Heiman TJ, Higgins ME, Ji RR, Ke Z, Ketchum KA, Lai Z, Lei Y, Li Z, Li J, Liang Y, Lin X, Lu F, Merkulov G V, Milshina N, Moore HM, Naik AK, Narayan VA, Neelam B, Nusskern D, Rusch DB, Salzberg S, Shao W, Shue B, Sun J, Wang Z, Wang A, Wang X, Wang J, Wei M, Wides R, Xiao C, Yan C, Yao A, Ye J, Zhan M, Zhang W, Zhang H, Zhao Q, Zheng L, Zhong F, Zhong W, Zhu S, Zhao S, Gilbert D, Baumhueter S, Spier G, Carter C, Cravchik A, Woodage T, Ali F, An H, Awe A, Baldwin D, Baden H, Barnstead M, Barrow I, Beeson K, Busam D, Carver A, Center A, Cheng ML, Curry L, Danaher S, Davenport L, Desilets R, Dietz S, Dodson K, Doup L, Ferreira S, Garg N, Gluecksmann A, Hart B, Haynes J, Haynes C, Heiner C, Hladun S, Hostin D, Houck J, Howland T, Ibegwam C, Johnson J, Kalush F, Kline L, Koduru S, Love A, Mann F, May D, McCawley S, McIntosh T, McMullen I, Moy M, Moy L, Murphy B, Nelson K, Pfannkoch C, Pratts E, Puri V, Qureshi H, Reardon M, Rodriguez R, Rogers YH, Romblad D, Ruhfel B, Scott R, Sitter C, Smallwood M, Stewart E, Strong R, Suh E, Thomas R, Tint NN, Tse S, Vech C, Wang G, Wetter J, Williams S, Williams M, Windsor S, Winn-Deen E, Wolfe K, Zaveri J, Zaveri K, Abril JF, Guigó R, Campbell MJ, Sjolander K V, Karlak B, Kejariwal A, Mi H, Lazareva B, Hatton T, Narechania A, Diemer K, Muruganujan A, Guo N, Sato S, Bafna V, Istrail S, Lippert R, Schwartz R, Walenz B, Yooseph S, Allen D, Basu A, Baxendale J, Blick L, Caminha M, Carnes-Stine J, Caulk P, Chiang YH, Coyne M, Dahlke C, Mays A, Dombroski M, Donnelly M, Ely D, Esparham S, Fosler C, Gire H, Glanowski S, Glasser K, Glodek A, Gorokhov M, Graham K, Gropman B, Harris M, Heil J, Henderson S, Hoover J, Jennings D, Jordan C, Jordan J, Kasha J, Kagan L, Kraft C, Levitsky A, Lewis M, Liu X, Lopez J, Ma D, Majoros W, McDaniel J, Murphy S, Newman M, Nguyen T, Nguyen N, Nodell M, Pan S, Peck J, Peterson M, Rowe W, Sanders R, Scott J, Simpson M, Smith T, Sprague A, Stockwell T, Turner R, Venter E, Wang M, Wen M, Wu D, Wu M, Xia A, Zandieh A, Zhu X, Sinsheimer RL, Sanger F, Seeburg PH, Strauss EC, Kobori JA, Siu G, Hood LE, Gocayne J, Martin-Gallardo A, McCombie WR, Jensen MA, Adams MD, Adams MD, Adams MD, Kerlavage AR, Fields C, Venter JC, Adams MD, Soares MB, Kerlavage AR, Fields C, Venter JC, Polymeropoulos MH, Marra M, Adams MD, White O, Sanger F, Coulson AR, Hong GF, Hill DF, Petersen GB, Mahy BWJ, Esposito JJ, Venter JC, Fleischmann RD, Fraser CM, Bult CJ, Tomb JF, Klenk HP, Venter JC, Smith HO, Hood L, Schmitt H, Zhao S, Lin X, Weber JL, Myers EW, Green P, Pennisi E, Venter JC, Adams MD, Marshall E, Pennisi E, Adams MD, Rubin GM, Myers EW, Collins FS, Sanger F, Nicklen S, Coulson AR, Prober JM, Myers G, Selznick S, Zhang Z, Miller W, Hattori M, Dunham I, Carvalho AB, Lazzaro BP, Clark AG, Schuler GD, Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ, Olivier M, Chaudhari N, Hahn WE, Milner RJ, Sutcliffe JG, Dickson D, Ewing B, Green P, Crollius HR, Pruitt KD, Katz KS, Sicotte H, Maglott DR, Uberbacher EC, Xu Y, Mural RJ, Burge C, Karlin S, Mural RJ, Salamov AA, Solovyev V V., Miklos GL, John B, Francke U, Horvath JE, Schwartz S, Eichler EE, Bickmore WA, Sumner AT, Holmquist GP, Bernardi G, Zoubak S, Clay O, Bernardi G, Ohno S, Broman KW, Murray JC, Sheffield VC, White RL, Weber JL, McEachern MJ, Krauskopf A, Blackburn EH, Bird A, Gardiner-Garden M, Frommer M, Larsen F, Gundersen G, Lopez R, Prydz H, Cross SH, Bird A, Grunau C, Hindermann W, Rosenthal A, Antequera F, Bird A, Cross SH, Slavov D, Smit AF, Riggs AD, Elliott DJ, Makeyev A V., Chkheidze AN, Lievhaber SA, Pan Y, Decker WK, Huq AHM, Craig WJ, Nouvel P, Goncalves I, Duret L, Mouchiroud D, Smith TF, Waterman MS, Delcher AL, Trask BJ, Sharon D, Barbazuk WB, McLysaght A, Enright AJ, Skrabanek L, Wolfe KH, Burk DW, Skrabanek L, Wolfe KH, Taillon-Miller P, Gu Z, Li Q, Hillier L, Kwok PY, Taillon-Miller P, Piernot EE, Kwok

- PY, Altshuler D, Marth GT, Cargill M, Halushka MK, Zhang J, Madden TL, Nachman MW, Bauer VL, Crowell SL, Aquadro CF, Nickerson DA, Jorde L, Wang DG, Przeworski M, Hudson RR, Rienzo A Di, Tavare S, Clark AG, Kaessmann H, Heissig F, Haeseler A von, Paabo S, Sonnhammer EL, Eddy SR, Durbin R, Bateman A, Ponting CP, Schultz J, Milpetz F, Bork P, Goodenough DA, Goliger JA, Paul DL, Wilkinson DG, Nakamura F, Kalb RG, Strittmatter SM, Horner PJ, Gage FH, Casaccia-Bonnel P, Gu C, Chao M V., Wang S, Barres BA, Geppert M, Sudhof TC, Littleton JT, Bellen HJ, Maximov A, Sudhof TC, Bezprozvanny I, Lemke G, Perrimon N, Bernfield M, Lindahl U, Kusche-Gullberg M, Kjellen L, Hurskainen TL, Hirohata S, Seldin MF, Apte SS, Black RA, White JM, Aravind L, Dixit VM, Koonin E V., Garcia-Meunier P, Etienne-Julan M, Fort P, Piechaczyk M, Bonhomme F, Mansur NR, Meyer-Siegler K, Wurzer JC, Sirover MA, Tatton NA, Kenmochi N, Chen FW, Ioannou YA, Madsen HO, Poulsen K, Dahl O, Clark BF, Hjorth JP, Chambers DM, Peters J, Abbott CM, Khalyfa A, Carlson BM, Carlson JA, Wang E, Aeschlimann D, Thomazy V, Munroe P, Wu SM, Cheung WF, Frazier D, Stafford DW, Furie B, Kehoe JW, Bertozzi CR, Pawson T, Nash P, Velden AW van der, Thomas AA, Fraser CM, Tettelin H, Brett D, Muller HJ, Kern H, Feinberg AP, Collins CA, Guthrie C, Eddy SR, Wang Q, Khillan J, Gadue P, Nishikura K, Holcik M, Sonenberg N, Korneluk RG, McKinsey TA, Zhang CL, Lu J, Olson EN, Capanna E, Romanini MGM, Smith JM, Charlesworth D, Charlesworth B, Morgan MT, Bailey JE, Maleszka R, Couet HG de, Miklos GL, Miklos GL, Crutchfield JP, Young K, Gell-Mann M, Lloyd S, Barabasi AL, Albert R, Colucci-Guyon E, Ewing B, Green P, Ewing B, Hillier L, Wendl MC, Green P, Lander ES, Waterman MS, Krogh A, Sjölander K, Sjölander K, Bairoch A, Apweiler R, Tatusov RL, Galperin MY, Natale DA, Koonin E V. (2001) The sequence of the human genome. *Science* 291:1304–51. doi: 10.1126/science.1058040
- Voinnet O (2009) Origin, Biogenesis, and Activity of Plant MicroRNAs. *Cell* 136:669–687. doi: 10.1016/j.cell.2009.01.046
- Wang RYH, Gehrke CW, Ehrlich M (1980) Comparison of bisulfite modification of 5-methyldeoxycytidine and deoxycytidine residues. *Nucleic Acids Res* 8:4777–4790. doi: 10.1093/nar/8.20.4777
- Wassenegger M, Heimes S, Riedel L, Sängler HL (1994) RNA-directed de novo methylation of genomic sequences in plants. *Cell* 76:567–76.
- Waterhouse PM, Graham MW, Wang MB (1998) Virus resistance and gene silencing in plants can be induced by simultaneous expression of sense and antisense RNA. *Proc Natl Acad Sci U S A* 95:13959–64.
- Waterston RH (2002) Initial sequencing and comparative analysis of the mouse genome. *Nature* 420:520–562.
- Webster J, Weber R (2007) Introduction to fungi. Cambridge University Press
- Wessler SR (2006) Transposable elements and the evolution of eukaryotic genomes. *Proc Natl Acad Sci* 103:17600–17601. doi: 10.1073/pnas.0607612103
- Wicker T, Mayer KFX, Gundlach H, Martis M, Steuernagel B, Scholz U, Simková H, Kubaláková M, Choulet F, Taudien S, Platzer M, Feuillet C, Fahima T, Budak H, Dolezel J, Keller B, Stein N (2011) Frequent gene movement and pseudogene evolution is common to the large and complex genomes of wheat, barley, and their relatives. *Plant Cell* 23:1706–18. doi: 10.1105/tpc.111.086629
- Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH (2007) A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* 8:973–982. doi: 10.1038/nrg2165
- Wierzbicki AT, Ream TS, Haag JR, Pikaard CS (2009) RNA polymerase V transcription guides ARGONAUTE4 to chromatin. *Nat Genet* 41:630–634. doi: 10.1038/ng.365
- Woller R (2007) The Pearl Oyster Mushroom. http://bioweb.uwlax.edu/bio203/2011/woller_ryan/index.htm. Accessed 31 Aug 2014
- Wright SI, Agrawal N, Bureau TE (2003) Effects of recombination rate and gene density on transposable element distributions in *Arabidopsis thaliana*. *Genome Res.* 13:1897–1903.
- Zamudio N, Bourcuhis D (2010) Transposable elements in the mammalian germline: a comfortable niche or a deadly trap? *Heredity (Edinb)* 105:92–104. doi: 10.1038/hdy.2010.53
- Zemach A, McDaniel IE, Silva P, Zilberman D (2010) Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* 328:916–919. doi: 10.1126/science.1186366
- Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW-L, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, Ecker JR (2006) Genome-wide High-Resolution Mapping and Functional Analysis of DNA Methylation in *Arabidopsis*.
- Zhao J-H, Fang Y-Y, Duan C-G, Fang R-X, Ding S-W, Guo H-S (2016) Genome-wide identification of endogenous RNA-directed DNA methylation loci associated with abundant 21-nucleotide siRNAs in *Arabidopsis*. *Sci Rep* 6:36247. doi: 10.1038/srep36247
- Zhong X, Du J, Hale CJ, Gallego-Bartolome J, Feng S, Vashisht AA, Chory J, Wohlschlegel JA, Patel DJ, Jacobsen SE (2014) Molecular mechanism of action of plant DRM de novo DNA methyltransferases. *Cell* 157:1050–1060. doi: 10.1016/j.cell.2014.03.056
- Zilberman D, Cao X, Jacobsen SE (2003) ARGONAUTE4 Control of Locus-Specific siRNA Accumulation and DNA and Histone Methylation. *Science* (80-) 299:716–719. doi: 10.1126/science.1079695
- Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S (2007) Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. *Nat Genet* 39:61–69. doi: 10.1038/ng1929

Scope of research

The objectives of this work can be divided into two groups as follows:

I

- Analysis of the inheritance of two helitron families in a meiotically-derived progeny and comparison with parental strains.
- Evaluation of helitron content in *P. ostreatus* sub-clones and experimental demonstration of their transposition.

II

- Analysis and functional characterization of DNA methylation and small RNA production at different growing stages of *P. ostreatus*.
- Identification of possible epigenetic origin of TE-mediated gene silencing.
- Comparative transcriptome analysis throughout the life cycle of *P. ostreatus*.

Chapter II – Somatic transposition and meiotically-driven elimination of an active helitron family in *Pleurotus ostreatus*

Introduction

Transposable elements (TEs) constitute an important fraction of the eukaryotic and prokaryotic genomes. According to their transposition mechanism, the vast majority of TEs produce target site duplications (TSDs) at their insertion sites, with some exceptions as Helitrons, a subclass of DNA transposons discovered in 2001 by computational analysis in *Arabidopsis thaliana*, *Oryza sativa*, and *Caenorhabditis elegans* genomes (Kapitonov and Jurka 2001). Helitrons are present in a wide range of eukaryotic genomes (Kapitonov and Jurka 2001; Poulter et al. 2003; Pritham and Feschotte 2007; Roffler et al. 2015) and show several structural and enzymatic features that differentiate them from the rest of the TEs.

They do not display Terminal Inverted Repeats (TIRs) as other DNA transposons do, do not generate TSDs at the insertion site, and carry conserved 5'-TC and CTRR-3' ends and a 16 bp to 20 bp palindromic hairpin located approximately 10-20 nucleotides upstream of the 3' terminus (Kapitonov and Jurka 2001; Kapitonov and Jurka 2007). Additionally, putative autonomous helitrons encode a RepHel protein containing a replication initiation (Rep) and a helicase (Hel) domain. The conservation of the Rep catalytic motifs with the replication initiators of plasmids and ssDNA viruses led to the hypothesis that these elements transpose with a rolling-circle mechanism (RC), which is consistent with the absence of TSDs (Fig. 1A). Nevertheless, a study carried out in maize haplotypes identified helitron somatic excision events based on the presence of footprints in polymorphic helitron loci (occupied vs vacant loci occurrence in the genome) (Li and Dooner 2009), suggesting that this TE subclass may show both excision and replication based transposition mechanisms.

Similar to other transposons, it has been described that helitrons can promote rearrangements in plant genomes (Lai et al. 2005; Choi et al. 2007) and generate intra-specific variability by breaking the genetic co-linearity among haplotypes (Fu and Dooner 2002; Lai et al. 2005; Choi et al. 2007). In addition, during the transposition process, helitrons can capture and disperse gene fragments, producing chimeric transcripts (Lai et al. 2005). After their discovery, *in silico* studies have been carried out to characterize the nature of these eukaryotic DNA TEs. Despite several helitron-related findings being described and explained, the mechanism of helitron mobilization remained unclear because, until very recently, there was no experimental evidence for their transposition.

Based on the proposed "rolling-circle" mechanism, helitrons transpose through the cleavage of a single-stranded DNA intermediate and subsequent insertion in another chromosomal location. It has been postulated that during mobilization, two transposases cleave at the donor and target sites and bind to the

resulting 5' ends (Fig. 1B). The free 3' OH in the target DNA forms a bond with the donor strand resulting in strand transfer. Replication at the cleaved donor site starts at the free 3' OH where the donor strand serves as a primer for DNA synthesis by host DNA polymerase and replication initiates to displace one strand of Helitron. If the palindrome sequence and 3' end of the element are recognized, cleavage occurs after the CTRR sequence and the one Helitron strand is transferred to the donor site where DNA replication resolves the heteroduplex (Fig. 1B – bottom left panel). When the inserted fragment of donor DNA is flanked at one end by IRR and at the other end by the CTTG or GTTC sequence present in the donor in a way that usually results in multiple tandem insertions of the donor plasmid or capture of flanking sequence in the target site. Failure to recognize the termination signal for Helitron transposition may result in the transfer of DNA sequence flanking the 3' end of the Helitron border to the donor site and capture of this flanking sequence in the target site (Fig. 1B – bottom right panel). This last mechanism has been proposed to explain the acquisition of additional coding sequences along the Helitron sequence. Interestingly, a novel study reports the first experimental evidence of helitron insertions in cell culture, providing insights into the transposition and gene capture mechanisms (Grabundzija et al. 2016). In addition, the finding of circular DNA intermediates containing head-to-tail junctions of helitron ends strongly supports the originally proposed RC transposition mechanism (Pritham and Feschotte 2007; Grabundzija et al. 2016).

Previous bioinformatics analyses have reported that helitrons account for ~3% of the mammalian genome size (Zhou et al. 2006; Pritham and Feschotte 2007) and 1-5% of the genome size in insects (Poulter et al. 2003; Kapitonov and Jurka 2007). In plants, these DNA transposons constitute a variable portion, contributing between 0.01% to 6% among different species (Kapitonov and Jurka 2001; Lal 2003; Morgante et al. 2005; Xiong et al. 2014).

In fungi, the presence of helitrons has been reported often as part of *in silico* comprehensive TE annotations of several ascomycetes and basidiomycetes sequenced genomes (Kapitonov and Jurka 2007; Cultrone et al. 2007; Santoyo et al. 2008; Martin et al. 2010; Labbé et al. 2012; Castanera et al. 2014). In addition, a study in *Aspergillus nidulans* described the capture and duplication of a gene promoter during the generation of a non-autonomous helitron (Cultrone et al. 2007).

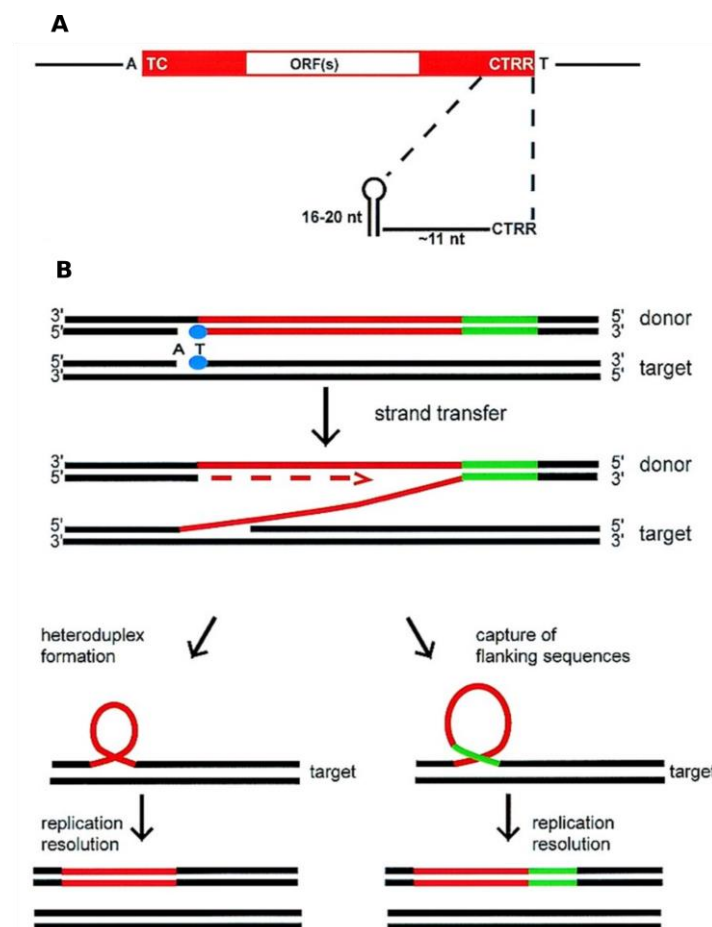


Figure 1. Generic Helitron structure (A) and transposition mechanism (B). Red lines represent Helitron element (autonomous or nonautonomous element), blue ellipses indicate transposase molecules and green lines show DNA sequence flanking the 3' end helitron extreme. Obtained from Feschotte and Wessler (2001).

In *P. ostreatus*, a recent study identified and characterized two helitron families, HELPO1 and HELPO2, which showed differential patterns of expression and distribution (Castanera et al. 2014). Helitron elements were detected by a structure-based approach using HelSearch (Yang and Bennetzen 2009) and classified into families according to the conservation of their 3' ends (30 bp with at least 80% identity). Subsequent manual inspection of both 5' and 3' boundaries of each helitron led to the identification of subfamilies (80% or higher identity in the 30 bp 3' end but lower than 80% in the 5' end). Thus, based on the similarity of the 5' and 3' boundaries helitrons of the HELPO1 family were further classified into three subfamilies: HELPO1.1, HELPO1.2 and helpo1.3, with elements varying from 1.5 to 13.7 kb length. HELPO2 contained

elements ranging from 3.9 to 10 kb in length. Both the HELPO1 and HELPO2 families include putative autonomous elements, containing genes encoding a RepHel protein with a rolling-circle replication initiator (Rep) and a helicase (Hel) domain, thought to be essential for transposition (Fig. 2). *P. ostreatus* helitrons carry most canonical features such as the 3' subterminal hairpin and the RepHel helicase. Nevertheless, in contrast to plants and animal helitrons, their encoded helicase does not carry RPA or zinc fingers domains. They occupy 0.35% and 0.05% in PC15 and PC9 genomes respectively and are mainly located in polymorphic loci.

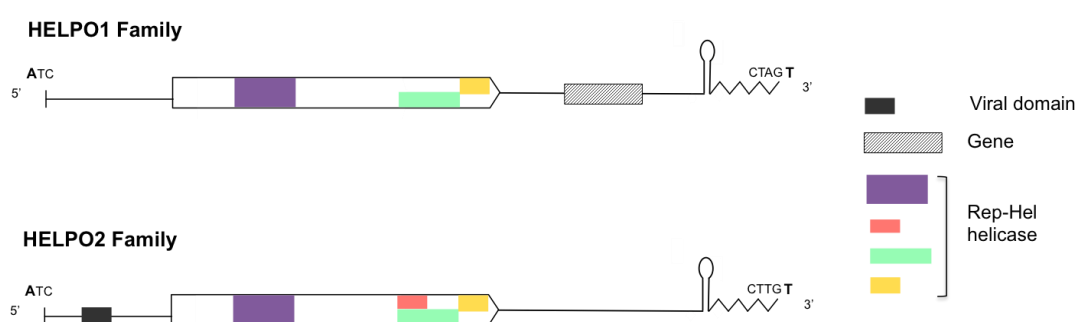


Figure 2. Structural representation of HELPO1 and HELPO2 helitron families in *P. ostreatus*.

In addition, comprehensive analysis of the transposons landscape in *P. ostreatus* genome unravel that helitrons were the most abundant DNA transposons, with HELPO1 family showing among the highest transcriptional levels when compared to other Class I and Class II TE families (Castanera et al. 2016). In the present study, we used experimental approaches to uncover the differential amplification dynamics of helitrons in *P. ostreatus* subclones maintained on solid culture under high and low subculture frequencies. In addition, we analysed the pattern of inheritance of helitron TEs in the offspring of 68 monokaryotic strains meiotically obtained from the dikaryotic strain N001. We provide evidence of somatic HELPO2 transposition in one of the analysed subclones and show that the HELPO2 family is under-represented in the progeny. A detailed analysis of two HELPO2 loci suggests that helitrons could be eliminated during meiosis by gene conversion.

Materials and Methods

Origins of the *P. ostreatus* monokaryotic and dikaryotic strains

A total of seventy-four strains and sub-clones of *P. ostreatus* with different nuclear allelic compositions were used in this study. All of them were derived from N001, the dikaryotic strain used by our group as a model for mushroom breeding, genetics and genomics since 1994 (Larraya et al. 2003). The N001 strain was de-dikaryotized in 1999, and the two corresponding haploid-monokaryotic protoclones were recovered and identified as PC9 and PC15 (Larraya et al. 1999). For each of the three strains, a subclone was deposited in the Spanish Type Culture Collection (CECT) on the dates indicated in Table 1 and maintained under low subculture frequency (N001-03, PC9-99 and PC15-99, two subcultures per year); another subclone was kept under high subculture frequency (approximately 8 subcultures per year) for routine laboratory work up to the present day (N001-14, PC9-14 and PC15-14). In addition, we used a haploid-monokaryotic progeny of 68 monokaryotic strains meiotically derived from N001 in 1994 (68mK, Table 1). This collection has been stored in our laboratory for 20 years under low subculture frequency.

Table 1. Description of *P. ostreatus* strains and subclones used in this stud

Subclone	Strain	Nuclear condition (ploidy)	Subculture regime (frequency)	Use/storage	Culture period
N001-03	N001	Dikaryon (n+n)	High/Low *	Culture Collection (CECT20600)	1994-2003/2003-2014
N001-14	N001	Dikaryon (n+n)	High (8/year)	Laboratory work	1994-2014
PC9-99	PC9	Monokaryon (n)	Low (2/year)	Culture Collection (CECT20311)	1999-2014
PC9-14	PC9	Monokaryon (n)	High (8/year)	Laboratory work	1999-2014
PC15-99	PC15	Monokaryon (n)	Low (2/year)	Culture Collection (CECT20312)	1999-2014
PC15-14	PC15	Monokaryon (n)	High (8/year)	Laboratory work	1999-2014
68mK #	N001 single spore isolates	Monokaryon (n)	Low (2/year)	Culture Collection (Laboratory)	1994-2014

* This subclone was maintained under high subculture frequency from 1994 to 2003. Then, it was deposited in the CECT and maintained under low subculture frequency on solid medium until 2014.

Collection of 68 monokaryotic strains (mk = monokaryotic).

Culture conditions and nucleic acid extraction

All of the strains and subclones used in this work were cultured in liquid SMY submerged fermentation (10 g of sucrose, 10 g of malt extract, 4 g of yeast extract, 1 litre H₂O [pH 5.6]) in the dark at 24°C under orbital shaking (130 rpm). The cultures were kept for eight growing days, and the mycelia were collected using vacuum filtration, ground in a sterile mortar in the presence of liquid nitrogen and stored at -80°C. Genomic DNA extractions were carried out using an E.Z.N.A. Fungal DNA Mini Kit (Omega Bio-Tek, Norcross, GA), following the manufacturer's instructions. After additional incubation with 4 µl of RNase A (10 mg/ml) for 30 minutes at 37°C, the DNA solutions were treated twice with phenol-chloroform (3:1), and the pellet was resuspended in 40 µl of nuclease-free water. DNA concentrations were determined using a Qubit® 2.0 fluorometer (Life Technology, Carlsbad, CA). The evaluation of total DNA purity was based on the Nanodrop™ 2000 A₂₆₀/A₂₈₀ ratio (Thermo Scientific, Wilmington, DE). DNA preparations served as a template for both real-time quantitative PCR (qPCR) and conventional PCR approaches.

Design and validation of primer sequences

Primers were designed based on the *P. ostreatus* reference genome sequences PC15 v2.0 and PC9 v1.0, using Primer3 software (Untergasser et al. 2012). Specificity of primer pairs was verified *in silico* by two approaches: i) BLASTN searches (Altschul et al. 1990) against *P. ostreatus* genomes using primer sequences as a query, and ii) manual verification of perfect matches between oligonucleotides and binding sites in PC15-PC9 alignments of the target loci. The validated primers were used to perform experimental analyses based on qPCR and conventional PCR.

Real-time quantitative PCR analysis

The helitron content was estimated by qPCR using a relative quantification approach (Pfaffl 2001). Reactions were carried out in a CFX96 thermal cycler (Bio-Rad Laboratories, S.A.). Amplification products were monitored at each cycle of PCR using SYBR green fluorescent dye. Each amplification mixture (20 µl) contained 1X IQ SYBR green Supermix (10 µl), 6 µM forward and reverse primers and 0.5 ng of genomic DNA in nuclease-free water. Primers were designed targeting family-specific regions to selectively amplify the HELPO1 and HELPO2 elements. The list of primers designed for the amplification of helitrons and reference single-copy genes is shown in Supplementary Table S1.

Reactions were run according to the following cycling conditions: DNA templates were denatured at 95 °C for 5 min, followed by 40 cycles of amplification (95 °C for 15 sec, 60 °C for 30 sec, and 95 °C for 1 min). The final melting curve was set by increasing 0.5 °C every 5 sec from 65 °C to 95 °C. Each assay was performed in triplicate in 96-well plates, and a non-template control (NTC) for each primer pair was included. Relative fluorescence units and threshold cycle values were processed using Bio-Rad CFX Manager Software, and the results were exported to Microsoft Excel for further analysis. The helitron copy number was estimated as a relative value in relation to the average signal of two single-copy reference genes, according to formula 1:

$$[1] \quad \mathbf{RCN} = 2^{-(Ct_{hel} - Ct_{ref})}$$

where RCN stands for “Relative copy number”, Ct_{hel} is the threshold cycle value of the target region (helitron), and Ct_{ref} is the average threshold cycle value of *sar1* and *lacc3* single-copy reference genes (reference).

Conventional PCR reactions

Several PCR protocols were designed according to the amplicon size and were run on a thermal cycler PCT-200 (MJ Research, MN). Biotaq DNA Polymerase (Bioline, Luckenwalde, Germany) was used to amplify products lower than 3.5 kb. Each PCR mix was performed in a final volume of 25 µL containing 1 unit of Biotaq polymerase, 50 mM MgCl₂, dNTP mix at 100 mM, and forward and reverse primers at 6 µM. Template DNA was firstly denatured at 95 °C for 5 min followed by 30 cycles of amplification (95 °C for 40 sec, 58 °C for 60 sec, 72 °C for 1 min to 3 min, depending on the expected product size) and a final extension cycle at 72 °C for 10 min. RANGER DNA Polymerase (Bioline, Luckenwalde, Germany) was used to amplify products longer than 3.5 kb. PCR reactions were prepared following the manufacturer’s recommendations in a final volume of 50 µl containing 4 units of RANGER DNA polymerase, 20 µM of specific primers and 20 ng of total DNA template suspended in nuclease-free water. PCR conditions comprised initial denaturation at 95 °C for 3 min followed by 30 cycles of 98 °C for 10 sec, 57 °C for 60 sec and 68 °C for 5 min to 10 min (depending on the expected product size), with a final extension at 68 °C for 10 min. For both protocols, negative controls were included using nuclease-free water instead of DNA template. Amplicons were resolved by electrophoresis on 0.8% w/v TAE agarose gels (1X Tris-acetate-EDTA), stained with 1:20000 Red Safe™ (iNtRON Biotechnology) and visualized under UV illumination.

PCR products were isolated from the agarose gels, purified with an E.Z.N.A Gel Extraction Kit (Omega Bio-Tek, Norcross, GA) and sequenced using Sanger technology (Sistemas Genómicos S.L., Valencia, Spain). Primer pairs used for the conventional PCR approach are listed in Supplementary Table S2-S5.

Southern blot hybridization

An internal fragment (PIF-1 like helicase domain) from HELPO2 helitrons was used as a probe on Southern blots hybridization according to standard procedures (Sambrook and Russell 2001). Probe fragments were obtained by PCR, using 10 ng of DNA template in a total reaction volume of 25 μ L. The oligonucleotides used as primers were designed by PRIMER3 (Untergasser et al. 2012) software, and used at 6 μ M concentration. Genomic DNA (10 μ g) of PC15 and PC9 strains was digested with 20 units of *HpaII* and *MspI* restriction enzymes (Promega) for 14 h at 37 °C. These isoschizomers selectively target CCGG restriction sites, although they display differential sensitivity to C5-methylcytosine. Fragmented DNA was then recovered by phenol-chloroform (1:1) extraction, followed by ethanol precipitation. Each aliquot was resuspended in a volume of 30 μ L and resolved in a 0,8 % agarose gel stained with Red Safe at 50 V for 16 hours. After running, the gel was blotted onto nitrocellulose membrane. Restriction enzyme-digested products were hybridized with a digoxigenin-labeled (Roche Diagnostic GmbH, Mannheim, Germany) PIF-1 like probe.

Detection of RNAi pathway genes in *P. ostreatus*

Key genes involved in Quelling and Meiotic Silencing of Unpaired DNA (MSUD) pathways described in *Neurospora crassa* were used as a query for BLASTP searches (cutoff e value = e^{-5}) against *P. ostreatus* PC15 v2.0 and PC9 v1.0 annotations. *GenBank* accessions of *N. crassa* proteins were the following: *qde-1*: EAA29811.1, *qde-2*: ESA42123.1, *qde-3*: AAF31695.1, *sad-1*: AAK31733.1, *sms-2*: EAA29350, *dcl-1*: EAA32662.1, *dcl-2*: EAA34302.3, and *qip*: XP_011393741.1. Conserved domains of *P. ostreatus* proteins with significant hits were identified using HMMER3 software (Eddy 2011) with PFAM-A target database (Finn et al. 2014), as well as NCBI Conserved Domain Database (Marchler-Bauer et al. 2015). In parallel, proteins were clustered using all by all BLASTP (cutoff e value = e^{-10}) followed by mcl (Enright et al. 2002) (inflation value =2). Clusters of proteins carrying the same conserved domains as the query proteins were retained. Orthology between PC15 and PC9 proteins was inferred by reciprocal best blast hit.

Results and discussion

HELPO1 and HELPO2 families show different inheritance patterns in the N001 meiotic progeny

To analyse the inheritance patterns of the two helitron families present in *P. ostreatus*, we studied helitron segregation in a progeny of 68 monokaryons meiotically derived from the N001 strain (population 68mK) by means of qPCR. The relative copy number of HELPO1 elements showed a normal distribution in the progeny (Shapiro-Wilks test, $W=0.978$, $p\text{-value}=0.271$), with monokaryons displaying RCN values ranging from 2 to 13 and a population peak showing 6-7 RCNs (Fig. 3A). Surprisingly, results corresponding to the HELPO2 family showed a trend that did not fit with a normal distribution ($W=0.727$, $p\text{-value}=5.135 \times 10^{-10}$, Fig. 3B). In this case, although RCNs ranged from 0 to 9, up to 49 monokaryons exhibited only 0 or 1 helitrons, uncovering a strong bias towards the lack of HELPO2 elements in the N001 progeny.

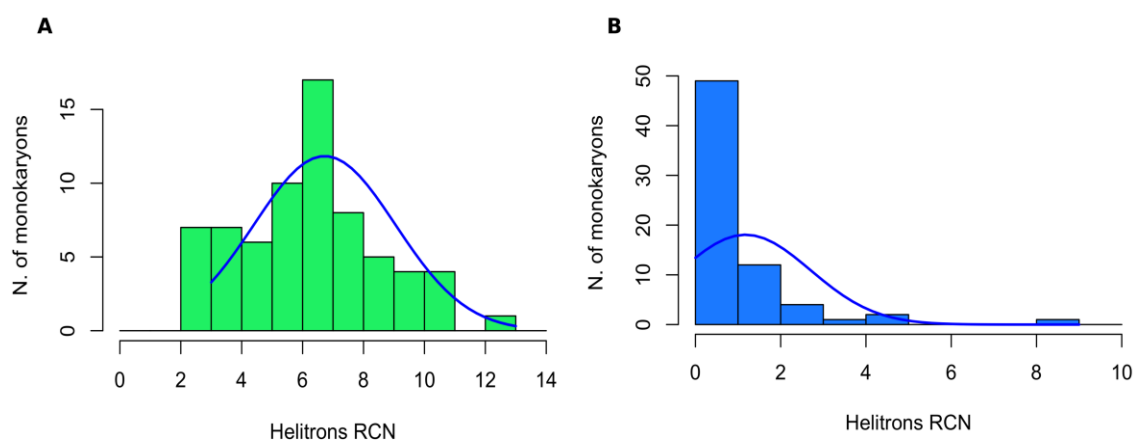


Figure 3. Distribution of HELPO1 (A) and HELPO2 (B) relative copy number in the 68-mK meiotic population. Relative copy numbers are presented as histograms with density lines.

Previous studies carried out by our group indicated that the HELPO2 family amplified very recently in the protoclone PC15, whereas it is absent in the compatible strain PC9 (Castanera et al. 2014). The distorted segregation of HELPO2 elements might be the consequence of a specific genome defence mechanism against the invasion of repetitive DNA. In this sense, the HELPO2 family shows elements with high similarity among them and low transcription levels (Castanera et al. 2014; Castanera et al. 2016), indicating that it might be targeted by transposon-silencing mechanisms. In contrast, HELPO1 elements are more divergent and showed very high transcription levels (Castanera et al. 2014; Castanera et al. 2016).

This fact suggests that this family is not targeted by genome defense mechanisms. A possible explanation for the unique profile of HELPO2 might be the very recent amplification of RepHel-coding elements in this family, evidenced by the presence of five identical full-length copies. These elements could be detected as foreign, invasive DNA and targeted by the genome defense machinery. To understand the striking differences found in the inheritance patterns of HELPO1 and HELPO2 families in the meiotic progeny, we reconstructed the pedigree of the monokaryotic collection along with their parental strains N001, PC15 and PC9 (Fig. 4), which have been used as a genetic model of study since 1994.

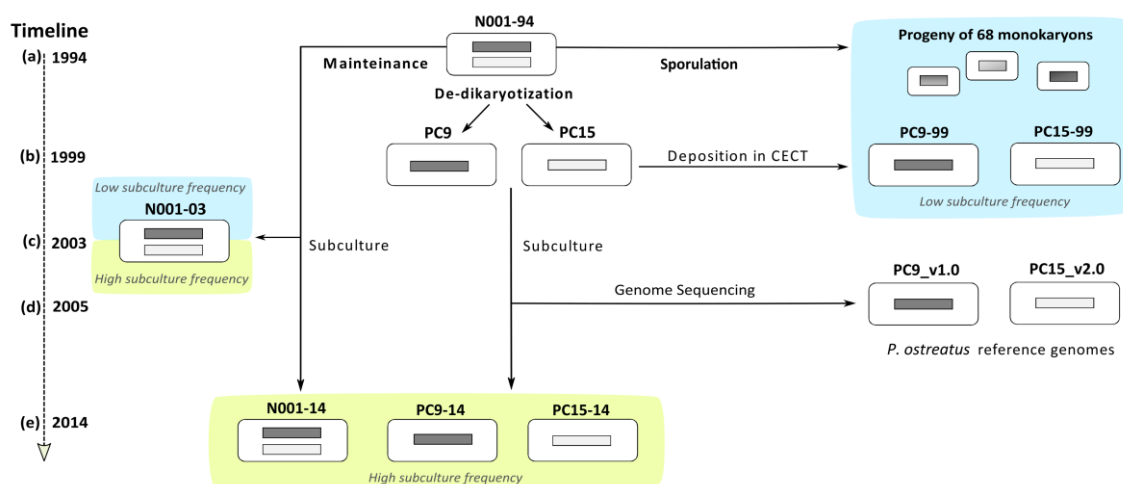


Figure 4. Pedigree and related time-scale events of *P. ostreatus* strains and subclones used in this study. (a) (1994) The N001-94 strain was sporulated leading to the monokaryotic population 68 mK. (b) (1999) PC9 and PC15 protocloned were obtained from N001 by de-dikaryotization and deposited in CECT (PC9-99 and PC15-99). (c) (2003) N001 was deposited in CECT (N001-03). (d) (2005) PC9 and PC15 protocloned were sequenced at JGI. (e) (2014) N001-03, PC9-99 and PC15-99 were retrieved from CECT and used in this study, along with N001-14, PC9-14 and PC15-14 subclones used for laboratory routine work up to 2014.

HELPO2 dynamics diverge in subclones maintained under different subculture conditions

The abundance of HELPO1 and HELPO2 elements was estimated in several subclones of the monokaryons and dikaryon parental strains. These subclones have been stored under different subculture frequencies for up to 20 years, as shown in Table 1 and Figure 4. Using the previously described qPCR

approach, we observed that HELPO1 RCNs were largely conserved and consistently identical among the subclones maintained under low subculture frequency (PC9-99, PC15-99, N001-03) and the corresponding subclones maintained in the laboratory at a high subculture frequency (PC9-14, P15-14 and N001-14; $p < 0.05$, two-tailed Student's t-test, Fig. 5A).

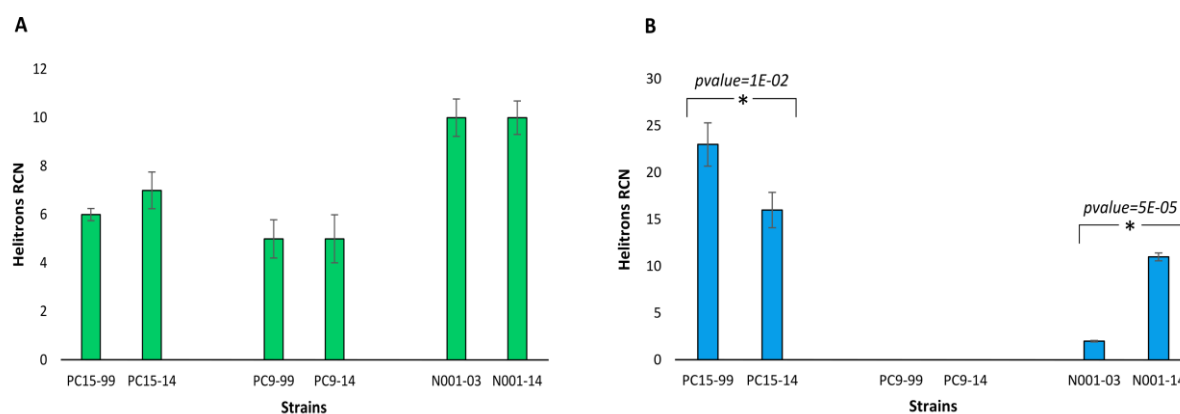


Figure 5. HELPO1 (A) and HELPO2 (B) relative copy number in parental subclones maintained under different subculture frequencies. Relative copy number of elements belonging to HELPO1 and HELPO2 families examined in the dikaryotic (N001) and monokaryotic (PC15 and PC9) parental strains. Data are reported as the mean \pm SD.

Furthermore, the RCNs observed in N001 dikaryon subclones fit with the sum of PC9 and PC15 RCNs, as expected for an additive model (N001 is a dikaryotic strain that carries the PC15 and PC9 nuclei). In contrast, the estimation of HELPO2 abundance in the monokaryotic and dikaryotic subclones revealed an unexpected scenario. Significant differences were found in HELPO2 content within PC15 and N001 subclone pairs (Fig. 5B). More specifically, PC15-99 showed a RCN value higher than PC15-14 (23 vs 15, respectively), whereas N001-14 showed a striking five-fold increment in comparison to N001-03 (11 vs 2 RCNs). In the case of PC9, both subclones lacked HELPO2 elements. Previous bioinformatics analysis reported a widespread distribution of HELPO2 copies in the chromosomes of PC15, whereas no intact HELPO2 helitrons were detected in the genome of the PC9 strain (Castanera et al. 2014). In this sense, qPCR results confirmed the lack of HELPO2 elements in both PC9-99 and PC9-14 strains. Additionally, these observations were confirmed, by performing Southern blot hybridization using the PIF-1 helicase domain of HELPO2 elements, as a probe (Supplementary Figure S1).

Nevertheless, the differential HELPO2 content in the PC15 and N001 subclones suggested that these elements have been mobilized in the mycelia (somatic transposition) and accumulated in the genome as a result of a continuous subculture. The fact that the addition of PC15 and PC9 HELPO2 RCNs surpasses that of any of the N001 subclones reinforces the hypothesis of an amplification burst in both PC15 subclones. Previous studies have described an amplification burst of DNA transposons (*IS30* and *IS5*) in bacterial subclones maintained for decades in stab culture, giving rise to genetic and phenotypic diversity (Naas et al. 1994). In addition, it is known that environmental stresses increase transposition rates (Capy et al. 2000; Tittel-Elmer et al. 2010). In this sense, the de-dikaryotization process underwent by the strain N001 in 1999 might have played a critical role in helitron increases, as protoplast isolation requires an enzymatic digestion protocol that could have triggered a stressful scenario leading to HELPO2 elements activation. We hypothesize that high subculturing rates may increase the chance of helitron transposition and accumulation. The five-fold increment of the HELPO2 element content in N001-14 vs N001-03 supports this hypothesis, but does not explain the results observed in PC15-99 and PC15-14. In the latter case, the subclone maintained under low-subculturing frequency (PC15-99) shared a higher HELPO2 RCNs than its counterpart maintained under higher subculturing rates.

This finding suggest that TE insertions without strong deleterious effects have a random chance of being passed to successive generations by clonal subculturing, thus bringing to the equation a randomness parameter. On the contrary, HELPO1 family elements did not show any sign of amplification. This finding, along with the uniform patterns of its inheritance found in the N001 mk68 meiotic progeny suggested that transposition of HELPO1 elements is infrequent in the *P. ostreatus* genome.

Molecular analysis of HELPO2 polymorphisms uncovers HELPO2 transposition events

The results obtained by the qPCR approach strongly suggest that HELPO2 is an active family, and the RCN increments observed in PC15 and N001 suggest that new elements of this family could be populating unknown loci. Previous bioinformatics analyses identified seven HELPO2 copies encoding RepHel helicases, distributed in 5 of 11 nuclear chromosomes of PC15 (Table 2). This genome-wide helitron annotation was carried out with *P. ostreatus* PC9 v1.0 and PC15 v2.0 genome assemblies (Castanera et al. 2014) produced by the JGI in 2005. Using conventional PCR, we re-analysed the parental subclone pairs to track the presence or absence of such HELPO2 loci.

The objective of this approach was to confirm the bioinformatics predictions and to identify putative helitron polymorphisms between and within subclone pairs that could be the result of helitron mobilizations.

Table 2. Helitrons described *in silico* in *P. ostreatus* PC15 genome assembly v2.0

Locus	Localization (chromosome*)	Start (bp)	End (bp)	Length (bp)
<i>Helpo2_I</i>	I	619761	626150	6388
<i>Helpo2_V</i>	V	387607	398218	10611
<i>Helpo2_VI</i>	VI	1150050	1156438	6388
<i>Helpo2_VII</i>	VII	1635256	1641644	6388
<i>Helpo2_VIII_a</i>	VIII	1367660	1374048	6388
<i>Helpo2_VIII_b</i>	VIII	1722302	1726240	3938
<i>Helpo2_VIII_c</i>	VIII	2234922	2241310	6388

*Chromosomes are equivalent to scaffolds in PC15 v2.0 assembly

A PCR strategy was designed to differentiate the two possible allelic states corresponding to the presence or absence of a helitron element in a given locus. We designed two primer pairs per locus flanking the left and right helitron ends. In each pair, one primer was designed inside the element, and the other was designed outside the helitron boundary, as shown in Figure 6A and 6B. This strategy (outer/inner) allowed us to identify loci carrying HELPO2 insertions (“occupied loci”) with positive PCR products, but no amplification should occur in loci with no HELPO2 insertion (“vacant loci”).

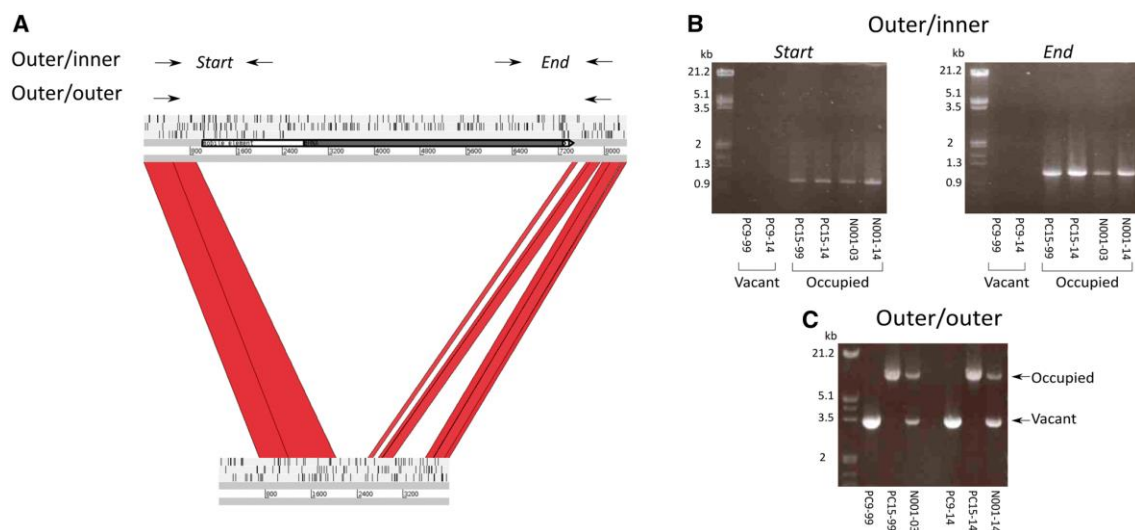


Figure 6. Detection of HELPO2 polymorphisms by the outer/inner and outer/outer strategies. (A) Example of a HELPO2 (*Helpo2_VIII_a*) polymorphic site in PC15 and PC9 haplotypes shown as an ACT (Artemis Comparison Tool) comparison. Blocks indicate conserved regions (>85% similarity). The HELPO2 element is shown as a white arrow with a grey rectangle inside the arrow representing its encoded RepHel helicase. Vertical black lines show stop codons predicted in the three forward reading frames. (B) and (C) represent the amplification profiles of a polymorphic HELPO2 locus in PC9 (vacant), PC15 (occupied) and N001 (vacant + occupied) subclones using the outer/inner and outer/outer strategies.

As expected, none of the PC9 subclones showed amplification in any of the HELPO2 loci (Table 3, Supplementary Figure S2). Interestingly, a striking difference was observed between PC15-99 and PC15-14 subclones: although the first one displayed HELPO2 insertions in the seven loci described in Table 2, the latter showed no amplification in *Helpo2_I* and in *Helpo2_VIII_c*. Intriguingly, both N001 subclones (N001-03 and N001-14) showed an identical profile to PC15-14 (Table 3, Supplementary Figures. S2 A and S2 G). This result lead us to hypothesize that *Helpo2_I* and *Helpo2_VIII_c* helitrons were transposed and inserted exclusively in PC15-99 subclone, after the de-dikaryotization of N001 in 1999. The experimental evidence of helitron transposition has been recently demonstrated in vitro using a reconstructed element from the bat genome (Grabundzija et al. 2016).

Table 3. Molecular validation of HELPO2 polymorphisms using the PCR outer/inner strategy. Each term represents an amplification product. Primers and expected product sizes are shown in Supplementary Table S2.

Strain \ Locus	PC9_99	PC9_14	PC15_99	PC15_14	N001-03	N001-14
<i>Helpo2_I</i>	vacant	vacant	occupied	vacant	vacant	vacant
<i>Helpo2_V</i>	vacant	vacant	occupied	occupied	occupied	occupied
<i>Helpo2_VI</i>	vacant	vacant	occupied	occupied	occupied	occupied
<i>Helpo2_VII</i>	vacant	vacant	occupied	occupied	occupied	occupied
<i>Helpo2_VIII_a</i>	vacant	vacant	occupied	occupied	occupied	occupied
<i>Helpo2_VIII_b</i>	vacant	vacant	occupied	occupied	occupied	occupied
<i>Helpo2_VIII_c</i>	vacant	vacant	occupied &	vacant	vacant	vacant

& The right flank displayed a higher size than the bioinformatics prediction (2 kb instead of 1.38 kb).

Our finding represents the first evidence of a fungal helitron transposition in its host genome under laboratory conditions. PCR products corresponding to *Helpo2_I* and *Helpo2_VIII_c* were purified, sequenced by Sanger technology and aligned to PC15 v2.0 and PC9 v1.0 genome sequences. The results obtained confirmed the helitron presence in the *Helpo2_I* locus, with the alignments spanning both external and internal regions of the HELPO2 (Fig. 7).

Regarding *Helpo2_VIII_c*, it is noteworthy that the PCR product of the right flank displayed a size 0.6 kb larger than expected. In this case sequenced PCR products could be aligned to the internal helitron regions but not to the external ones due to the low quality of the sequence. To discard that this phenomenon as a false negative, we designed a second strategy to validate vacant sites with a positive PCR product. Using the PC15 sequence as a reference, primers were designed for regions adjacent to HELPO2 boundaries and homologous to the PC9 regions flanking the vacant site (outer/outer, Fig. 6A). The complexity of such regions (often carrying other transposons, especially *Gypsy* LTR-retrotransposons) made it difficult to find homologous regions between PC15 and PC9, yielding long expected PCR fragment sizes for vacant sites, ranging from 2,603 bp to 8,578 bp (Supplementary Table S3).

To overcome this problem, we used a long-range DNA polymerase that allowed us to amplify vacant and occupied sites including PCR product size up to 10-15 kb (i.e., Fig. 6C).

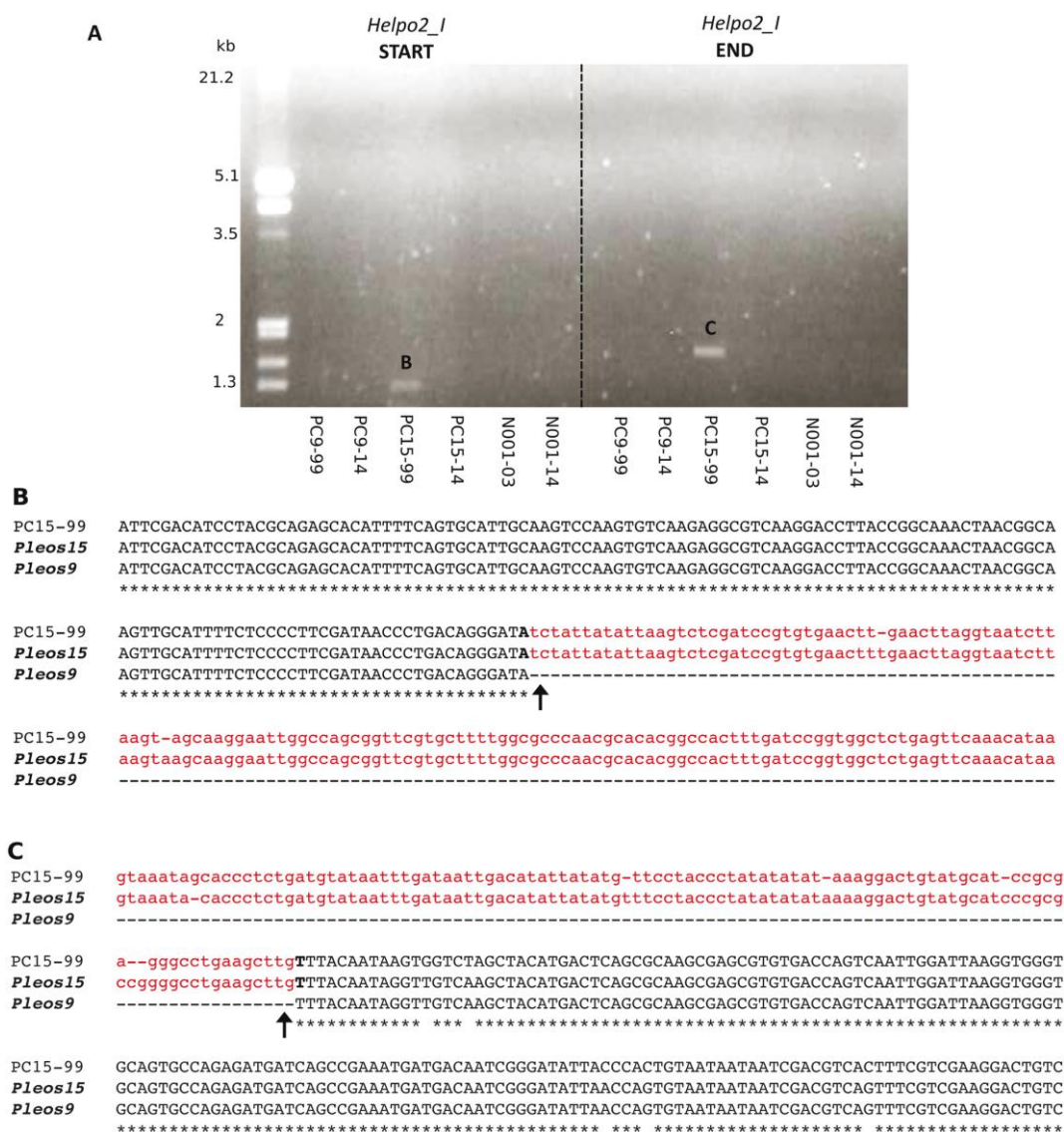


Figure 7. Evidence of HELPO2 insertion in the *Helpo2_I* locus of the PC15-99 subclone. Left (START) and right (END) *Helpo2_I* boundaries amplified exclusively in PC15-99 subclone (A). Helitron insertion in the PC15-99 subclone is reported by aligning both sequenced PCR products (B and C) to the occupied and empty sites of PC15 v2.0 (*Pleos15*) and PC9 v1.0 (*Pleos9*) reference genomes. Lowercase typeface indicates helitron sequences.

The presence of vacant sites was tested in all subclones (Table 4, Supplementary Figure S4) including the discordant *Helpo2_I* and *Helpo2_VIII_c* loci in the PC15-99 subclone, despite the outer/inner strategy confirming the presence of HELPO2 in both of them. Using this approach, we validated all vacant sites except PC9 *Helpo2_VII*. This particular locus was further analysed with additional primer pairs and could

not be amplified in any of the PC9 subclones, probably due to the inefficacy of long-range RANGER polymerase. Remarkably, the amplification and sequencing of *Helpo2_I* and *Helpo2_VIII_c* in PC15-99 revealed products corresponding to vacant sites, even though we had previously validated the presence of HELPO2 helitrons in both loci by the outer-inner strategy. Validation of these two vacant sites was performed by sequencing and subsequent alignment to the PC15 v2.0 reference (Supplementary Figure S3). These results could suggest that cells with and without HELPO2 elements (HELPO2+ / HELPO2-) coexist in hyphae of the PC15-99 subclone, giving rise to a mosaic mycelium. This fact reinforces the hypothesis of a very recent transposition of both elements, which were not entirely fixed or eliminated from the PC15-99 subclone in the successive subculturing performed after their insertion. In addition, these results provide indirect evidence that helitron was mobilized by replicative or semi-replicative transposition, as HELPO2 elements increased in PC15_99 without the excision of any of the other five occupied loci.

Table 4. Molecular validation of HELPO2 polymorphisms using the PCR outer/outer strategy. Each term represents an amplification product. White cells show results in agreement with the outer/inner strategy, and black cells result in disagreement. Grey cells represent occupied loci previously validated by the outer/inner strategy. Primers and expected product sizes are shown in Supplementary Table S3.

Strain \ Locus	PC9_99	PC9_14	PC15_99	PC15_14	N001-03	N001-14
<i>Helpo2_I</i>	vacant	vacant	vacant	vacant	vacant	vacant
<i>Helpo2_V</i>	vacant	vacant				
<i>Helpo2_VI</i>	vacant	vacant				
<i>Helpo2_VII</i>	N/A	N/A				
<i>Helpo2_VIII_a</i>	vacant	vacant				
<i>Helpo2_VIII_b</i>	vacant	vacant				
<i>Helpo2_VIII_c</i>	vacant	vacant	vacant	vacant	vacant	vacant

N/A no amplification

Distorted segregation patterns of HELPO2 in the N001 meiotic-derived progeny

We investigated the segregation patterns of the seven HELPO2 loci in 68 monokaryons using the previously described outer/inner strategy. As N001 is a dikaryon formed by the mating of PC15 and PC9 monokaryons and due to the seven loci being polymorphic (the presence of helitrons in PC15 vs absence of helitrons in PC9), it should be expected that there is a proportion of 1:1 (presence vs absence of helitrons) in the progeny. In contrast, the results revealed that the inheritance of HELPO2 elements was clearly distorted (Fig. 8). First of all, three of the seven loci carrying a HELPO2 in the PC15 strain (*Helpo2_I*, *Helpo2_VII* and *Helpo2_VIII_c*) were completely absent in the progeny.

The fact that *Helpo2_I* and *Helpo2_VIII_c* were present in the PC15-99 subclone but absent in PC15-14, N001-03, N001-14 and the whole N001 progeny (68mK) reinforces the hypothesis of its unique (and recent) transposition in PC15-99, which according to this result, would have taken place after N001 de-dikaryotization (1999) but before PC15 genome sequencing (2005), as both loci carry HELPO2 elements in the PC15 v2.0 assembly.

This observation demonstrated that these loci were vacant in the original PC15 and PC9 strains that mated in the wild to form N001. Following similar reasoning, a HELPO2 element must have been inserted in *Helpo2_VII* in the PC15 nucleus of the N001 dikaryon in the time that elapsed between the isolation of the meiotic progeny (in 1994 all monokaryons carry a vacant locus) and N001 de-dikaryotization (1999, PC15 carries an occupied locus; Fig. 4). Furthermore, PC9 helitron (vacant) alleles were consistently overrepresented in a ratio close to 3:1 in the remaining four loci. These results are in concordance with the biased distribution of HELPO2 elements previously found in the monokaryotic collection by qPCR (Fig. 3B). This scenario suggests the occurrence of a preferential selection of PC9 helitron (vacant) alleles during meiosis, or, in other words, an elimination of helitrons. To rule out the possibility that such a finding was derived from an artificially skewed progeny, we revisited the *P. ostreatus* genetic linkage map.

This map was constructed based on the segregation of 214 molecular markers in a population of 80 sibling monokaryons meiotically derived from N001 (Park et al. 2006), and 68 of them were used in this study. According to the genetic map, 86% of the markers followed the expected 1:1 ratio. Additionally, the segregation of genetic markers surrounding HELPO2 loci did not show distorted segregation. Taking together, these data reinforced the fact that the segregation of HELPO2 elements is selectively skewed with respect to the expected Mendelian segregation.

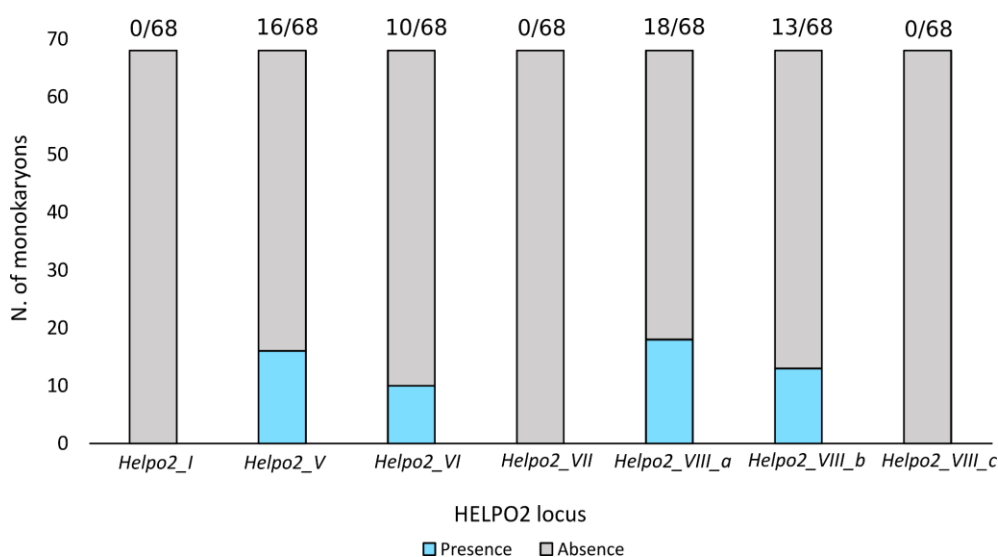


Figure 8. HELPO2 segregation rates in the 68-mK monokaryotic progeny. Distribution of HELPO2 helitrons in a meiotic progeny of 68 individuals is reported as stacked histograms. The seven HELPO2 loci are shown on the X-axis. The Y-axis indicates the number of monokaryons. Black and grey bars show the number of monokaryons presenting occupied and vacant loci, respectively. The frequency of helitron presence per locus is shown above each bar.

Somatic transposition of HELPO2 generates mosaicism in *P. ostreatus*

The transposition of an autonomous HELPO2 element to a new chromosomal location in PC15-99 (i.e., *Helpo2_I*, chromosome I, position 619,761-626,150) was validated by two PCR strategies and reinforced by the presence of a vacant locus in the other strains/subclones, including the monokaryotic progeny. We integrated these results in the reconstruction of the pedigree of *P. ostreatus*' strains maintained under different subculture conditions to propose a time-scale model that could explain the HELPO2 transposition in PC15-99 as well as the current HELPO2 profiles obtained in all the strains under study (Fig. 9).

After 1999, a somatic transposition of the helitron HELPO2 took place in the nucleus of a cell of a PC15 culture plate and was integrated in the *Helpo2_I* locus by a copy and paste mechanism that could be triggered by the stress of the de-dikaryotization process. As a result, the transposition event originated a new HELPO2 copy in such a cell, which after mitotic replication led to the occurrence of mosaicism in the PC15 mycelia, displaying *Helpo2_I+* (occupied) and *Helpo2_I-* (vacant) cells. Thus, subsequent subclones derived from the PC15 strain might have been originated from agar plugs located in different sections of the mycelium plate, thus presenting polymorphic loci.

We hypothesize that the PC15-14 sample maintained during 20 years under high subculture frequency and used for routine laboratory work could have derived from a section of the PC15 *Helpo2_{-/-}* cells. Alternatively, another likely explanation could be that *Helpo2_{+/+}* cells were also sampled but lost in the PC15-14 subclone after multiple subculturing replications due to lower fitness in comparison to *Helpo2_{-/-}* or simply by the result of random drift. In contrast, PC15-99 still maintains the phenotype of a genetic mosaic. In this sense, the minimal subculturing frequency of this subclone likely favoured the preservation of both *Helpo2_{+/+}* and *Helpo2_{-/-}* cells in the same mycelium net.

The presence of mosaicism in basidiomycetes has been widely studied in *Armillaria gallica*, where variable numbers of nuclei genotypes have been detected in single fruiting bodies collected for decades (Peabody et al. 2000). In this sense, it has been described that among-cell-line genetic variation in the haploid mycelium of *A. gallica* may confer plasticity to adapt to changing environmental conditions (Peabody et al. 2005).

Allelic gene conversion may contribute to the elimination of HELPO2 helitrons from *P. ostreatus* progeny

As was previously described, the PCR strategies leading to an understanding of the inheritance patterns of the HELPO2 family in the progeny of N001 uncovered an over-transmission of the HELPO2- vacant alleles from the parental PC9. To understand if such phenomena were derived from a preferential inheritance or by helitron excisions during meiosis, we proceeded as follows: *Helpo2_{VI}* and *Helpo2_{VIII_b}* loci were chosen on the basis of their clear distortion segregation ratios. All the monokaryons carrying vacant sites according to the lack of amplification by the outer/inner strategy (58 of 68 for *Helpo2_{VI}* and 55 of 68 for *Helpo2_{VIII_b}*) were subjected to additional amplification reactions using the primers shown in Figure 10 A and B. The objectives were as follows: i) to validate the presence of vacant sites with a positive PCR product, and ii) to determine the parental (PC9 or PC15) origin of such vacant sites as well as that of the adjacent regions (i.e., if they were inherited from the parental PC15 or PC9). For this purpose, primers were designed to amplify homologous regions of the parental strains carrying small indels and single nucleotide polymorphisms (SNPs), which allowed us to identify their PC9 or PC15 origin (primer sequences shown in Supplementary Table S4 and S5).

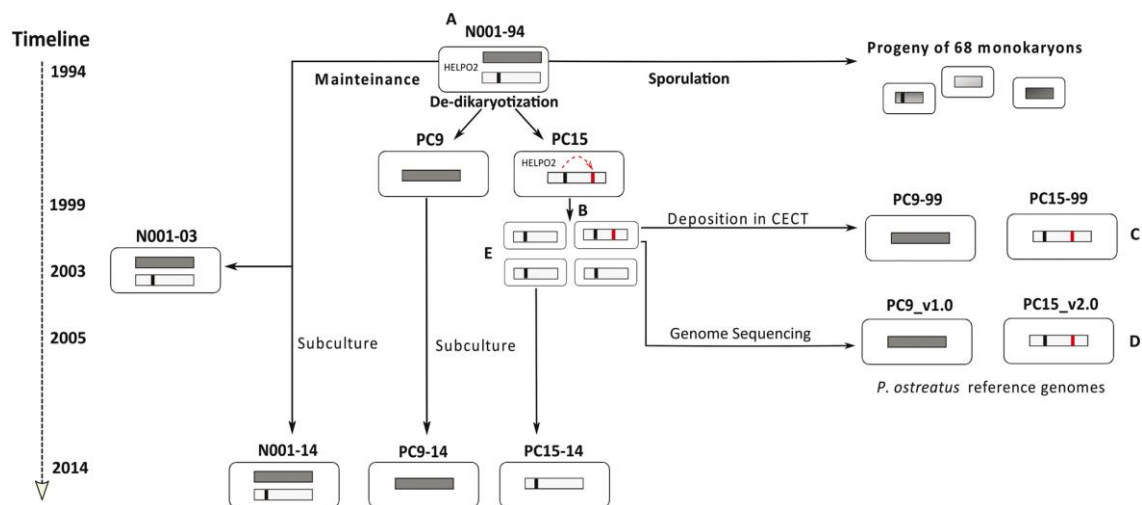


Figure 9. Chronological model explaining HELPO2 somatic insertions in the PC15-99 subclone. (a) (1994) N001-94 parental strain displays 4 helitron loci (*Helpo2_V*, *Helpo2_VI*, *Helpo2_VIII_a* and *Helpo2_VIII_b*), which are transmitted to the meiotic progeny 68 mK. (b) (1999). After de-dikaryotization, a HELPO2 from one of the four loci mobilizes in a cell of PC15 and integrates into a new locus (i.e. *Helpo2_I*), producing a mosaic mycelium displaying both occupied (*Helpo2_I+*) and vacant (*Helpo2_I-*) cells (dashed arrow indicates HELPO2 helitron transposition and subsequent integration). (c) (1999–2014) A PC15 subclone presenting mosaicism is deposited in the CECT and maintained under low subculture frequency up to 2014 (PC15-99). (d) (2005) A PC15 subclone carrying *Helpo2_I+* cells is used for whole genome sequencing by JGI. (e) (1999–2014) PC15 subclone used for routine laboratory work up to 2014 (PC15-14) is derived from a section of PC15 *Helpo2_I-* cells. Alternatively, *Helpo2_I+* cells were lost by random drift as a consequence of the high subculture frequency.

For each locus, the four amplification products shown in Figure 10 A and 10 B were sequenced and aligned to both PC15 v2.0 and PC9 v1.0 assemblies. SNPs were manually identified in the alignments and used to assign their parental origin. In 94% of the monokaryons analysed, the four regions studied were inherited from PC9, pointing to a preferential inheritance of the vacant site instead of a HELPO2 excision during meiosis. Nevertheless, a surprising profile was found in 7 monokaryons (4 in *Helpo2_VI* locus and 3 in *Helpo2_VIII_b* locus). Although the three flanking sites were inherited from the PC15 parental strain, SNPs of the amplification product corresponding to the vacant site (1,489 bp in *Helpo2_VI* and 907 bp in *Helpo2_VIII_b*) indicated that they were transmitted from the parental PC9 (Fig. 10A and B).

These results might be explained by a mechanism of allelic gene conversion, which results in the loss of HELPO2 helitrons. To verify this hypothesis, the presence of gene conversion junctions were validated in the 7 monokaryons by PCR as shown in Figure 10C. To ensure the linearity of the short regions located between the amplified regions (shown in Fig. 10A and B as white rectangles), primers were designed to amplify longer regions, spanning *Helpo2_VI* and *Helpo2_VIII_b* loci as well as their flanking sites (yellow arrows in Fig 10C, Supplementary Table S6).

Our results showed the presence of bands corresponding to the expected size, confirming the linearity of such regions and discarding the presence of rearrangements (Supplementary Figure S5). After sequencing, PCR products were aligned to both PC15 and PC9 parental strains. Interestingly, the pattern of SNPs/indels showed that gene conversion occurred precisely in the same sites in all monokaryons analysed for each locus (Supplementary Figure S6), suggesting that the mechanism leading to this recombination might have specific targets. The absence of HELPO2 helitrons in the monokaryons with profiles consistent with allelic gene conversion events suggest that this mechanism might be used to eliminate unpaired repetitive sequences (such TEs) during meiosis. Allelic gene conversion is a molecular mechanism associated with recombination in which a genomic fragment is “copied/pasted” onto another homologous fragment. Little is known about gene conversion in fungi, although it is associated with non-Mendelian segregation ratios in yeast and mammals (Lamb 1998; Galtier 2001).

In fungi, previous studies have shown that gene conversion leads to 3:1 ratios in heteroallelic crosses (Holliday 2007). A recent analysis of the recombination landscape in the basidiomycete *Agaricus bisporus* revealed a skewed segregation ratio from the expected Mendelian 1:1 in certain loci of the haploid offspring of 139 isolated single spores. In addition, results showed that crossover events (COs) occurred almost exclusively at chromosome ends including gene conversion in the reciprocal CO context (Sonnenberg et al. 2016). In 1982, it was proposed that gene conversion could be involved in the elimination of selfish DNA during meiosis when hybrid DNA structures are homologous but one of them displays a DNA insert present in one parent (Holliday 1982), as is the case with HELPO2 loci in PC15 and PC9. In such a case, a single-strand loop could be formed, being targeted by endonucleases and resulting in gene conversion with the elimination of the insert. Events of a model of gene conversion (homing) mediated by site-specific endonucleases have also been detected in different strains of *Botrytis cinerea*.

The analysis of meiotic products obtained by crossing strains carrying polymorphisms permitted determination of gene conversion tracts that corresponded to unpaired regions previously identified in the parental strains (Bokor et al. 2010).

Helitron elimination in *P. ostreatus* progeny: genome defence against invasive DNA?

Eukaryotes have developed defence mechanisms to control the proliferation of transposable elements and avoid the potentially harmful effects of their insertions. One of the most impressive mechanisms of genome defence is called chromatin diminution, which has been described in several eukaryotes from distinct lineages (Kataoka and Mochizuki 2011; Sun et al. 2014). This mechanism selectively eliminates DNA and targets mainly young TEs displaying high frequencies and similarities (Sun et al. 2014). The molecular mechanism underlying chromatin diminution bears similarity to piRNA-directed transposon silencing in metazoans and requires an active RNAi machinery. In fact, it has been proposed that small RNAs could play an essential role directing such DNA elimination in the protozoan *Tetrahymena thermophile* (Kataoka and Mochizuki 2011). No study has described the presence of such mechanisms in fungi, but a handful of studies have shown the presence and activity of the main components of the RNAi pathway in filamentous fungi (Dicer, Argonaute, and RNA-dependent RNA polymerase) reviewed in (Dang et al. 2011). From the diversity of the fungal RNAi pathway, two mechanisms have evolved to control TEs at the vegetative (Quelling) and sexual (MSUD) states. The latter is of great relevance to our study as it involves the identification of unpaired DNA during meiosis by a *trans*-sensing mechanism. The origin of such unpaired regions are likely produced by TE insertions in only one of the two parental strains, such as in the case of the HELPO2 family in *P. ostreatus*. We screened PC15 and PC9 genomes for proteins homologous to those involved in *N. crassa* MSUD and *Quelling* and found a total of 18 genes in both strains: 5 argonaute proteins, 5 RNA-dependent RNA polymerases (RdRPs), 4 dicers, 3 RecQ helicases and 1 exonuclease. *P. ostreatus* RNAi proteins carry the *corresponding* conserved domains (Supplementary Table S7) and are actively expressed in vegetative mycelia (Table 5). Further experimental work is needed to understand whether the transcriptional repression and meiotic elimination of HELPO2 is somehow linked to the activity of RNAi pathway genes. Due to the ability of fungi to diversify and develop “epigenetic solutions” to control TE proliferation, such a possibility seems to be an interesting field for future studies.

Table 5. Identification of proteins belonging to the RNAi pathway in *P. ostreatus*. Transcription of PC15 and PC9 orthologs is shown in RPKM (Reads Per Kilobase per Million mapped reads)

PC15-ID *	PC9-ID *	RPKM	RPKM	Description
25449	90477	10.8	11.1	RNA-dependent RNA polymerase
1053861	85609	16.34	30.2	RNA-dependent RNA polymerase
1076136	83385	2.74	0.6	RNA-dependent RNA polymerase
1041769	52011	36.06	22.6	RNA-dependent RNA polymerase
154946	85193	53.19	21.8	RNA-dependent RNA polymerase
1060211	89407	87.2	233.8	Argonaute
1110274	43687	116.78	59.7	Argonaute
173501	122655	85.31	124.2	Argonaute
21640	91748	12.43	20.5	Argonaute
44554	114784	27.76	39.9	Argonaute
33722	67302	15.34	30.7	RecQ Helicase
11210	24796	12.58	3.3	RecQ Helicase
160303	87316	7.47	4.8	RecQ Helicase
1064031	83509	12.96	6.6	DICER
1033048	81629	9.4	7.6	DICER
1093523	50672	27.64	16.8	DICER
1112895	84917	34.41	47.1	DICER
1039826	83395	41.63	54.4	DnaQ-like exonuclease

* Protein identifiers of PC15 v2.0 and PC9 v1.0 MycoCosm databases

References

- Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ (1990) Basic local alignment search tool. *J Mol Biol* 215:403–10. doi: 10.1016/S0022-2836(05)80360-2
- Bokor AAM, van Kan JAL, Poulter RTM (2010) Sexual mating of *Botrytis cinerea* illustrates PRP8 intein HEG activity. *Fungal Genet Biol* 47:392–8. doi: 10.1016/j.fgb.2010.01.003
- Capy P, Gasperi G, Biéumont C, Bazin C (2000) Stress and transposable elements: co-evolution or useful parasites? *Heredity (Edinb)* 85:101–106. doi: 10.1046/j.1365-2540.2000.00751.x
- Castanera R, López-Varas L, Borgognone A, LaButti K, Lapidus A, Schmutz J, Grimwood J, Pérez G, Pisabarro AG, Grigoriev I V, Stajich JE, Ramírez L (2016) Transposable Elements versus the Fungal Genome: Impact on Whole-Genome Architecture and Transcriptional Profiles. *PLoS Genet* 12:e1006108. doi:10.1371/journal.pgen.1006108
- Castanera R, Perez G, Lopez L, Sancho R, Santoyo F, Alfaro M, Gabaldon T, Pisabarro AG, Oguiza JA, Ramirez L (2014) Highly expressed captured genes and cross-kingdom domains present in Helitrons create novel diversity in *Pleurotus ostreatus* and other fungi. *BMC Genomics* 15:1071.
- Choi JD, Hoshino A, Park K II, Park IS, Iida S (2007) Spontaneous mutations caused by a Helitron transposon, Hel-It1, in morning glory, *Ipomoea tricolor*. *Plant J* 49:924–934. doi: 10.1111/j.1365-313X.2006.03007.x
- Cultrone A, Domínguez YR, Drevet C, Scazzocchio C, Fernández-Martín R (2007) The tightly regulated promoter of the xanA gene of *Aspergillus nidulans* is included in a helitron. *Mol Microbiol* 63:1577–87. doi: 10.1111/j.1365-2958.2007.05609.x
- Dang Y, Yang Q, Xue Z, Liu Y (2011) RNA interference in fungi: Pathways, functions, and applications. *Eukaryot. Cell* 10:1148–1155.
- Eddy SR (2011) Accelerated Profile HMM Searches. *PLoS Comput Biol* 7:e1002195. doi: 10.1371/journal.pcbi.1002195
- Enright AJ, Dongen S Van, Ouzounis CA (2002) An efficient algorithm for large-scale detection of protein families. *Nucleic Acids Res* 30:1575–1584. doi: 10.1093/nar/30.7.1575
- Finn RD, Bateman A, Clements J, Coggill P, Eberhardt RY, Eddy SR, Heger A, Hetherington K, Holm L, Mistry J, Sonnhammer ELL, Tate J, Punta M (2014) Pfam: the protein families database. *Nucleic Acids Res* 42:D222–D230. doi: 10.1093/nar/gkt1223
- Fu H, Dooner HK (2002) Intraspecific violation of genetic colinearity and its implications in maize. *Proc Natl Acad Sci U S A* 99:9573–9578. doi: 10.1073/pnas.132259199
- Galtier N (2001) GC-Content Evolution in Mammalian Genomes: The Biased Gene Conversion Hypothesis. *Dev Psychobiol* 39:251–6. doi: 10.3138/physio.61.1.51
- Grabundzija I, Messing SA, Thomas J, Cosby RL, Bilic I, Miskey C, Gogol-Döring A, Kapitonov V, Diem T, Dalda A, Jurka J, Pritham EJ, Dyda F, Izsvák Z, Ivics Z (2016) A Helitron transposon reconstructed from bats reveals a novel mechanism of genome shuffling in eukaryotes. *Nat Commun* 7:10716. doi: 10.1038/ncomms10716
- Holliday R (2007) A mechanism for gene conversion in fungi. *Genet Res* 89:285–307. doi: 10.1017/S0016672308009476
- Holliday R (1982) Gene Conversion: A Possible Mechanism for Eliminating Selfish DNA. In: *Molecular and Cellular Mechanisms of Mutagenesis*. Springer US, Boston, MA, pp 259–264
- Kapitonov V V., Jurka J (2007) Helitrons on a roll: eukaryotic rolling-circle transposons. *Trends Genet.* 23:521–529.
- Kapitonov V V, Jurka J (2001) Rolling-circle transposons in eukaryotes. *Proc Natl Acad Sci U S A* 98:8714–9. doi: 10.1073/pnas.151269298
- Kataoka K, Mochizuki K (2011) Programmed DNA elimination in *Tetrahymena*: A small RNA-mediated genome surveillance mechanism. *Adv Exp Med Biol* 722:156–173. doi: 10.1007/978-1-4614-0332-6_10
- Labbé J, Murat C, Morin E, Tuskan GA, Le Tacon F, Martin F (2012) Characterization of transposable elements in the ectomycorrhizal fungus *Laccaria bicolor*. *PLoS One*. doi: 10.1371/journal.pone.0040197
- Lai J, Li Y, Messing J, Dooner HK (2005) Gene movement by Helitron transposons contributes to the haplotype variability of maize. *Proc Natl Acad Sci U S A* 102:9068–73. doi: 10.1073/pnas.0502923102
- Lal SK (2003) The Maize Genome Contains a Helitron Insertion. *PLANT CELL ONLINE* 15:381–391. doi: 10.1105/tpc.008375
- Lamb BC (1998) Gene conversion disparity in yeast: its extent, multiple origins, and effects on allele frequencies. *Heredity (Edinb)* 80 (Pt 5):538–52.
- Larraya LM, Alfonso M, Pisabarro AG, Ramirez L (2003) Mapping of Genomic Regions (Quantitative Trait Loci) Controlling Production and Quality in Industrial Cultures of the Edible Basidiomycete *Pleurotus ostreatus*. *Appl Environ Microbiol* 69:3617–3625. doi: 10.1128/AEM.69.6.3617-3625.2003
- Larraya LM, Perez G, Penas MM, Baars JJP, Mikosch TSP, Pisabarro AG, Ramirez L (1999) Molecular Karyotype of the White Rot Fungus *Pleurotus ostreatus*. *Appl Environ Microbiol* 65:3413–3417.
- Li Y, Dooner HK (2009) Excision of Helitron transposons in maize. *Genetics* 182:399–402. doi: 10.1534/genetics.109.101527
- Marchler-Bauer A, Derbyshire MK, Gonzales NR, Lu S, Chitsaz F, Geer LY, Geer RC, He J, Gwadz M, Hurwitz DI, Lanczycki CJ, Lu F, Marchler GH, Song JS, Thanki N, Wang Z, Yamashita RA, Zhang D, Zheng C, Bryant SH (2015) CDD: NCBI's conserved domain database. *Nucleic Acids Res* 43:D222–D226. doi: 10.1093/nar/gku1221
- Martin F, Kohler A, Murat C, Balestrini R, Coutinho PM, Jaillon O, Montanini B, Morin E, Noel B, Percudani R, Porcel B, Rubini A, Amicucci A, Amselem J, Anthouard V, Arcioni S, Artiguenave F, Aury J-M, Ballario P, Bolchi A, Brenna A, Brun A, Buée M, Cantarel B, Chevalier G, Couloux A, Da Silva C, Denoeud F, Duplessis S, Ghignone S, Hilselberger B, Iotti M, Marçais B, Mello A, Miranda M, Pacioni G, Quesneville H, Riccioni C, Ruotolo R, Splivallo R, Stocchi V, Tisserant E, Viscomi AR, Zambonelli A, Zampieri E, Henrissat B, Lebrun M-H, Paolucci F, Bonfante P, Ottonello S, Wincker P (2010) Périgord black truffle genome uncovers evolutionary origins and mechanisms of symbiosis. *Nature* 464:1033–8. doi: 10.1038/nature08867
- Morgante M, Brunner S, Pea G, Fengler K, Zuccolo A, Rafalski A (2005) Gene duplication and exon shuffling by helitron-like transposons generate intraspecific diversity in maize. *Nat Genet* 37:997–1002. doi: 10.1038/ng1615
- Naas T, Blot M, Fitch WM, Arber W (1994) Insertion Sequence-Related Genetic-Variation In Resting *Escherichia- Coli* K-12. *Genetics* 136:721–730.
- Peabody RB, Peabody DC, Sicard KM (2000) A genetic mosaic in the fruiting stage of *Armillaria gallica*. *Fungal Genet Biol* 29:72–80. doi: 10.1006/fgbi.2000.1187

- Peabody RB, Peabody DC, Tyrrell MG, Edenburn-MacQueen E, Howdy RP, Semelrath KM (2005) Haploid vegetative mycelia of *Armillaria gallica* show among-cell-line variation for growth and phenotypic plasticity. *Mycologia* 97:777–787. doi: 10.3852/mycologia.97.4.777
- Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 29:e45. doi: 10.1093/nar/29.9.e45
- Poulter RTM, Goodwin TJD, Butler MI (2003) Vertebrate helitrons and other novel Helitrons. *Gene* 313:201–212. doi: 10.1016/S0378-1119(03)00679-6
- Pritham EJ, Feschotte C (2007) Massive amplification of rolling-circle transposons in the lineage of the bat *Myotis lucifugus*. *Proc Natl Acad Sci U S A* 104:1895–900. doi: 10.1073/pnas.0609601104
- Roffler S, Menardo F, Wicker T (2015) The making of a genomic parasite - the *Mothra* family sheds light on the evolution of Helitrons in plants. *Mob DNA* 6:23. doi: 10.1186/s13100-015-0054-4
- Sambrook J, Russell DW (David W (2001) *Molecular cloning : a laboratory manual*. Cold Spring Harbor Laboratory Press
- Sánchez C (2010) Cultivation of *Pleurotus ostreatus* and other edible mushrooms. *Appl Microbiol Biotechnol* 85:1321–37. doi: 10.1007/s00253-009-2343-7
- Santoyo F, González AE, Terrón MC, Ramírez L, Pisabarro AG (2008) Quantitative linkage mapping of lignin-degrading enzymatic activities in *Pleurotus ostreatus*. *Enzyme Microb Technol* 43:137–143. doi: 10.1016/j.enzmictec.2007.11.007
- Sonnenberg ASM, Gao W, Lavrijssen B, Hendrickx P, Sedaghat-Tellgerd N, Foulongne-Oriol M, Kong W-S, Schijlen EGWM, Baars JJP, Visser RGF (2016) A detailed analysis of the recombination landscape of the button mushroom *Agaricus bisporus* var. *bisporus*. *Fungal Genet Biol* 93:35–45. doi: 10.1016/j.fgb.2016.06.001
- Sun C, Wyngaard G, Walton D, Wichman HA, Mueller R, Coyne R, Chalker D, Yao M, Tobler H, Beermann S, Smith J, Antonacci F, Eichler E, Amemiya C, Smith J, Baker C, Eichler E, Amemiya C, Kubota S, Ishibashi T, Kohno S, Arnaiz O, Mathy N, Baudry C, Malinsky S, Aury J-M, Wilkes CD, Garnier O, Labadie K, Lauderdale B, Mouël A Le, Marmignon A, Nowacki M, Poulain J, Prajer M, Wincker P, Meyer E, Duhaucourt S, Duret L, Betermier M, Sperling L, Rasch E, Wyngaard G, Wyngaard G, Rasch E, Connelly B, Wang J, Mitreva M, Berriman M, Thorne A, Magrini V, Koutsovoulos G, Kumar S, Mark LB, Richard ED, Etter A, Bernard V, Kenzelmann M, Tobler H, Muller F, Hunter D, Williams K, Cartinhour S, Herrick G, Chalker D, Yao M, Standiford D, Zagoskin M, Marshak T, Mukha D, Grishanin A, Drouin G, McKinnon C, Drouin G, Gregory T, Hebert P, Gregory T, Hebert P, Kolasa J, Wyngaard G, Gregory T, Wyngaard G, Rasch E, Manning N, Gasser K, Domangue R, Rasch E, Wyngaard G, Connelly B, Sun C, Shepard D, Chong R, Arriaza J, Hall K, Castoe T, Feschotte C, Pollock D, Mueller R, Novick P, Basta H, Floumanhaft M, McClure Lynch M, Bobay L-M, Catania F, Gout J-F, Rho M, Agren J, Wright S, Petrov D, Aminetzach Y, Davis J, Bensasson D, Hirsch A, Nuzhdin S, Gobbi M De, Viprakasit V, Hughes J, Fisher C, Buckle V, Ayyub H, Gibbons R, Vernimmen D, Yoshinaga Y, Jong P de, Cheng J, Rubin E, Wood W, Bowden D, Higgs D, Cavalier-Smith T, Orgel L, Kataoka K, Mochizuki K, Baudry C, Malinsky S, Restituito M, Kapusta A, Rosa S, Meyer E, Betermier M, Hosono S, Faruqi A, Dean F, Du Y, Sun Z, Wu X, Du J, Kingsmore S, Egholm M, Lasken R, Gomez-Alvarez V, Teal T, Schmidt T, Niu B, Fu L, Sun S, Li W, Price A, Jones N, Pevzner P, Benson G, Colbourne J, Pfrender M, Gilbert D, Thomas W, Tucker A, Oakley T, Tokishita S, Aerts A, Arnold G, Basu M, Bauer D, Caceres C, Carmel L, Casola C, Choi J, Detter J, Dong Q, Dusheyko S, Eads B, Frohlich T, Geiler-Samerotte K, Gerlach D, Hatcher P, Jogdeo S, Krijgsveld J, Kriventseva E, Kultz D, Laforsch C, Lindquist E, Lopez J, Casella G, Berger R (2014) Billions of basepairs of recently expanded, repetitive sequences are eliminated from the somatic genome during copepod development. *BMC Genomics* 15:186. doi: 10.1186/1471-2164-15-186
- Tittel-Elmer M, Bucher E, Broger L, Mathieu O, Paszkowski J, Vaillant I (2010) Stress-induced activation of heterochromatic transcription. *PLoS Genet* 6:1–11. doi: 10.1371/journal.pgen.1001175
- Untergasser A, Cutcutache I, Koressaar T, Ye J, Faircloth BC, Remm M, Rozen SG (2012) Primer3-new capabilities and interfaces. *Nucleic Acids Res.* doi: 10.1093/nar/gks596
- Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, Flavell A, Leroy P, Morgante M, Panaud O, Paux E, SanMiguel P, Schulman AH (2007) A unified classification system for eukaryotic transposable elements. *Nat Rev Genet* 8:973–982. doi: 10.1038/nrg2165
- Xiong W, He L, Lai J, Dooner HK, Du C (2014) HelitronScanner uncovers a large overlooked cache of Helitron transposons in many plant genomes. *Proc Natl Acad Sci* 111:10263–10268. doi: 10.1073/pnas.1410068111
- Yang L, Bennetzen JL (2009) Structure-based discovery and description of plant and animal Helitrons. *Proc Natl Acad Sci* 106:12832–12837. doi: 10.1073/pnas.0905563106
- Zhou Q, Froschauer A, Schultheis C, Schmidt C, Bienert GP, Wenning M, Dettai A, Volff J-N (2006) Helitron Transposons on the Sex Chromosomes of the Platyfish *Xiphophorus maculatus* and Their Evolution in Animal Genomes. *Zebrafish* 3:39–52. doi: 10.1089/zeb.2006.3.39

Supplementary information

Table S1. Primers sequences used for RT-qPCR protocol

Locus	Direction	Sequence (5'-3')	Product length (bp)
HELPO1.1	Fw	GGAAAGGACGCCGAGTTCACATTCAC	79
	Rv	CTTGTTGATGGTCATCGAGAACGCAA	
HELPO1.2	Fw	TCGTGCTACTGGTGCATCTC	113
	Rv	ATAGACGTCCCTCGCAGA	
helpo1.3	Fw	GCGCGCACATTCTTAATTTAC	83
	Rv	CTGCTTACAGTCACAGTGGTTATCCA	
HELPO2	Fw	AAACTGCGGACTCCTGAAGA	127
	Rv	CAGCTGTGGTGCTTCCAGTA	
Lacc3	Fw	TCGTTTCCGTCTCGTTTCTC	133
	Rv	CTGCGAAGATTTGGATGCTG	
Sar1	Fw	GGATAGTCTTCCTCGTCGATAG	134
	Rv	GGGTGCGTCAATCTTGTTAC	

Table S2. Primer sequences used for outer/inner PCR protocol

Helitron position (PC15 genome)	Scaffold	Primers identification	Sequence (5'-3')	Expected product size (bp)
619761-626150	1	Is_Fw	GCGGGCTAATAATGCACACA	1316
		Is_Rv	CCAGGTCCAGAAATCCAGGT	
		Ie_Fw	GCATGGTCGACACACAAATC	1530
		Ie_Rv	GGAAACGAAATCATGGTCGT	
387607-398218	5	Vs_Fw	CACCCCTACCAACTCTGT	1099
		Vs_Rv	CCTGTCATGATCTTGCGCAA	
		Ve_Fw	TGCCCAACCTTTACTCCAT	1061
		Ve_Rv	CGGACCACTCTTGCAAACA	
1150050-1156438	6	VIs_Fw	TCCTCACTCCCTTGTTTCCC	1202
		VIs_Rv	CCTTCAAGTGCCCACAGAAC	
		Vle_Fw	GCTTATCTGGGGCTATTTGGG	1088
		Vle_Rv	AGAGTGTATCCTCGCGAAT	
1635256-1641644	7	VIIs_Fw	CAAAGTGGTGACATCCATCG	1148
		VIIs_Rv	CTGCTCTGGCCTACCACTTC	
		VIIe_Fw	CCAGGTCCAGAAATCCAGGT	1431
		VIIe_Rv	CATCAACACTGCCGCTATCC	
1367660-1374048	8	VIII(13)s_Fw	CGCGTAAATGGTACTGACGG	1265
		VIII(13)s_Rv	CCAGGTCCAGAAATCCAGGT	
		VIII(13)e_Fw	AATGGATCTCGCGGAATCCT	1302
		VIII(13)e_Rv	CAAGTTCGCTCGTGTGTCA	
1722302-1726240	8	VIII(17)s_Fw	CGTTGTAGCCCTTAGTGTGC	3128
		VIII(17)s_Rv	GGCTGTTGACCATGAAGCAA	
		VIII(17)e_Fw	TGTTGTCTTTGGTGGGGACT	2815
		VIII(17)e_Rv	AGGTCTTTGCGAGGGATGAA	
2234922-2241310	8	VIII(22)s_Fw	CGATTGCAGGCTACTCACAA	1245
		VIII(22)s_Rv	GCATGGTCGACACACAAATC	
		VIII(22)e_Fw	CACCTCCCTAAGCCTTCAA	1381
		VIII(22)e_Rv	GCAGAGTTAGCAGCGTCTTC	

Table S3. Primer sequences used for outer/outer PCR protocol

Helitron position (PC15 genome)	Primers identification	Sequence (5'-3')	Location (scaffold)		Expected product size (bp)	
			PC9	PC15	PC9	PC15
619761-626150	I_Fw	CGAAGGTAGAACGTGGGCTA	1	1	2618	9008
	I_Rv	GATGCCCATTCGCTTTGAT				
387607-398218	V_Fw	TGGCTTGTCTAGATGCGCTA	3	5	3322	42368
	V_Rv	TTAGGCAGCAAGTCGAGGAA				
1150050-1156438	VI_Fw	TCCTTGCCGTAACCCTTCAT	8	6	3610	12080
	VI_Rv	GGACGCTTTGATCCCGTAC				
1635256-1641644	VII_Fw	TTTCTGCGACCTTTCTTGCC	6	7	8578	14966
	VII_Rv	GTGATGTTGATGGTCTGGGC				
1367660-1374048	VIII(13)_Fw	GATGCAGCGGAGATATCG	5	8	4861	13079
	VIII(13)_Rv	CTGCTATCACCATCCAGCCT				
1722302-1726240	VIII(17)_Fw	AGAGTCAGAGATGTCAGGCG	5	8	4407	8425
	VIII(17)_Rv	GTATCAATCCAGCGCACGTT				
2234922-2241310	VIII(22)_Fw	GGCAAACATCAGATCAGGCA	5	8	2603	8990
	VIII(22)_Rv	AAACAATGCATGCACTGGGG				

Table S4. Primer sequences used for PCR protocol (SNPs detection in vacant site)

Helitron position (PC15 genome)	Primers identification	Sequence (5'-3')	Location (scaffold)		Expected product size (bp)	
			PC9	PC15	PC9	PC15
1150050-1156438	P6_Fw	GCTTGATCTTCGTCCTACGC	8	6	1845	9862
	P6_Rv	AAGAGTGGAAGCCAAGTGGA				
1722302-1726240	P8_Fw	GACGGCCTTTGATTGAACCA	5	8	6876	907
	P8_Rv	GTTGCTGCGCTTCTCTAACC				

Table S5. Primer sequences used for PCR protocol (SNPs detection in flanking sites)

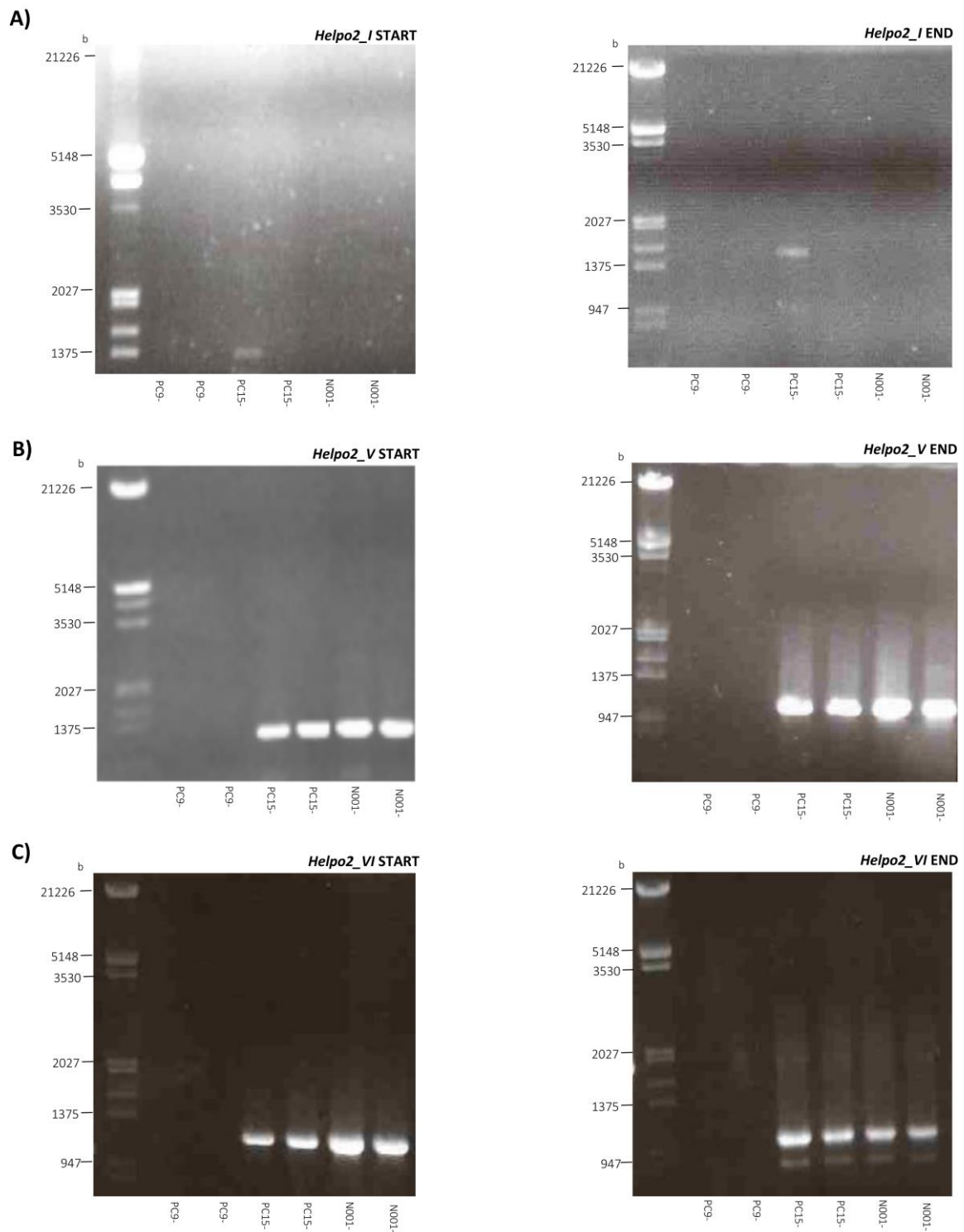
Helitron position (PC15 genome)	Scaffold	Primers identification	Sequence (5'-3')	Expected product size (bp)
1150050-1156438	6	F6s_Fw	TCAACATCTCTGCCTTCTGC	489
		F6s_Rv	CCCCCGTATCCATACCAGAT	
		F6e1_Fw	TGTTGCTGTCTGAACACGC	269
		F6e1_Rv	TCTGCATCCCACTACAACCG	
		F6e2_Fw	ATGCTCAACAATCACCAGCG	267
		F6e2_Rv	GCGTGCTATCTGGTCTTTGT	
1722302-1726240	8	F8s_Fw	CTCCCGAGCTTCCAGTTCTT	479
		F8s_Rv	GCAGTCCCTTCTAACGGAGA	
		F8e1_Fw	GCGGATTAAGAAGGCCACAG	324
		F8e1_Rv	GTCCTAACCCGCGTCCAAT	
		F6e2_Fw	TGGAACCCATCTATCCCCTTC	285
		F6e2_Rv	ACAGTGTGTGATTTGTGTACAGC	

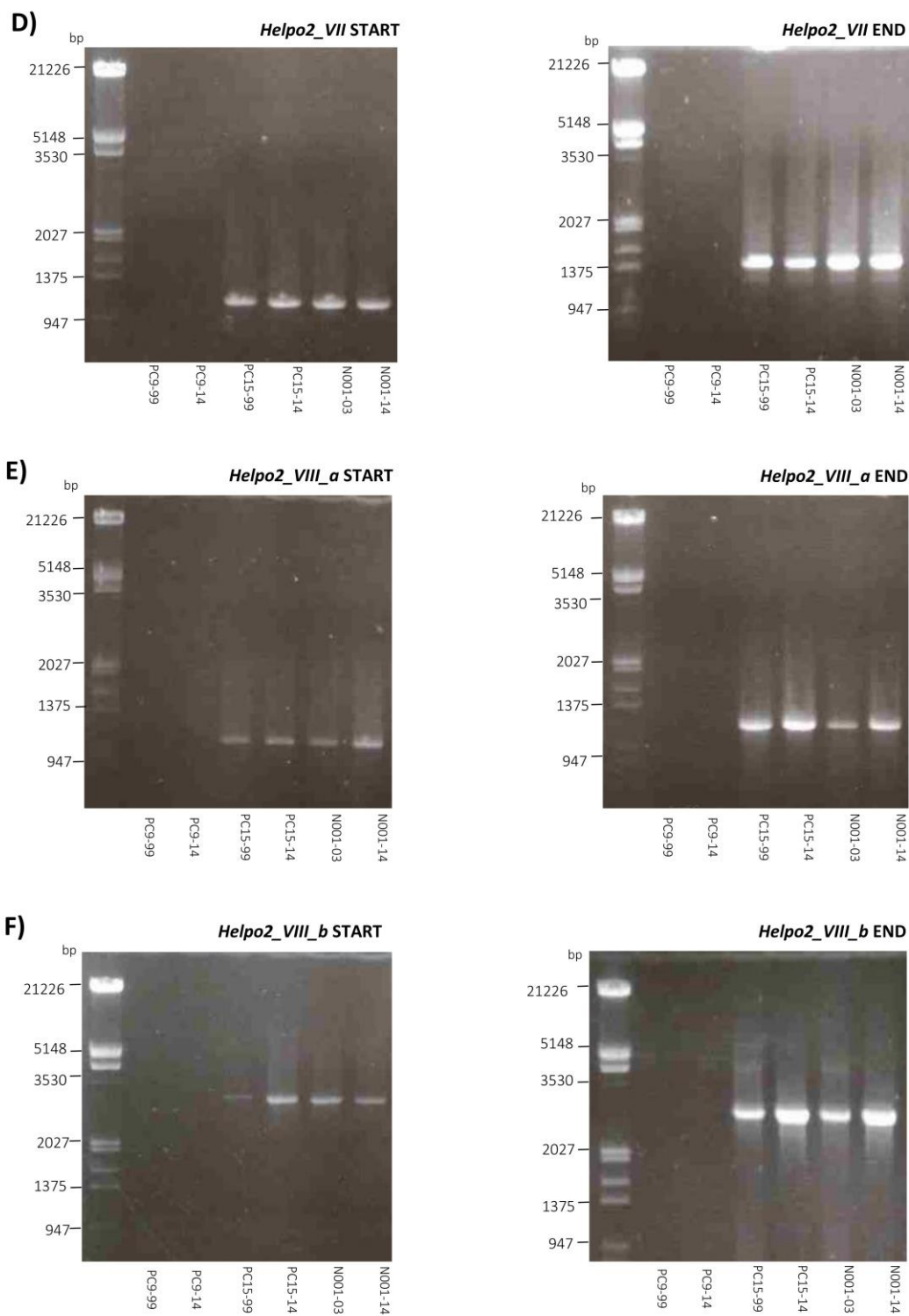
Table S6. Primer sequences used for PCR protocol (SNPs/indels validation in gene conversion region of monokaryons)

Locus	Scaffold (PC15 genome)	Primers identification	Sequence (5'-3')	Expected product size (bp)
<i>Helpo2_VI</i>	6	C6_ws_Fw	GGGAGGAATTCCATGATGAG	2971
		C6_ws_Rv	CTCGGAGATTGACCTTGAC	
<i>Helpo2_VIII_b</i>	8	C8_ws_Fw	GTCGTCTCCGCGTTTTTG	2426
		C8_ws_Rv	CTCGAGAATAGGCCATCCAG	



Figure S1 . Southern blot patterns of PC9 (lanes 1 and 2) and PC15 (lanes 3 and 4) monokaryotic strains digested with *HpaII* (lanes 1 and 3) and *MspI* (lanes 2 and 4). Digested genomic DNA products were hybridized with PIF-1 like helicase domain of HELPO2 helitrons.





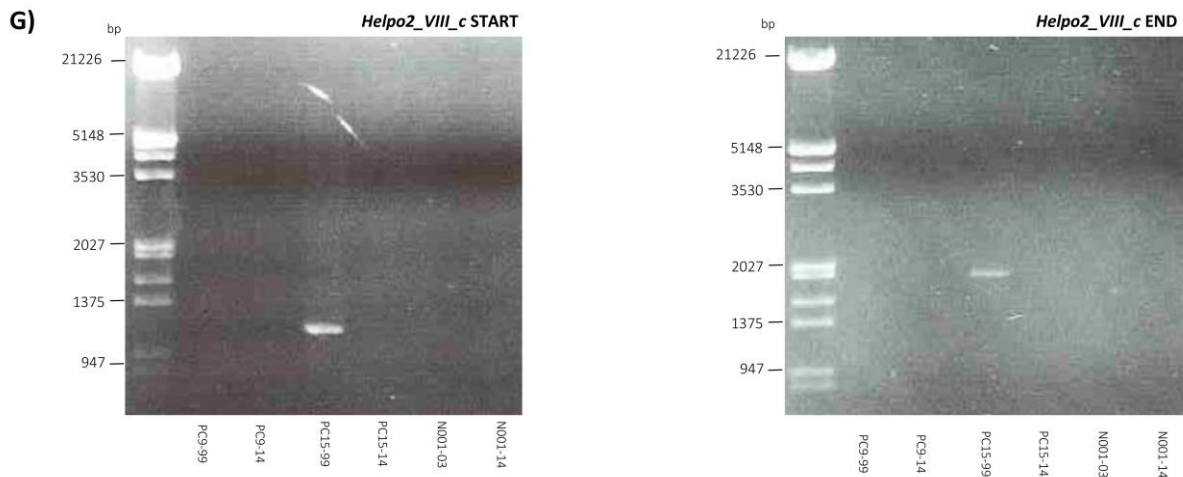


Figure S2. Analysis of full and vacant sites using the outer/inner PCR strategy

Evidence of HELPO2 polymorphic insertion in PC9, PC15 and N001 subclones maintained under different subculture conditions. Panels on left and right show respectively start and end helitrons boundaries located at different loci.

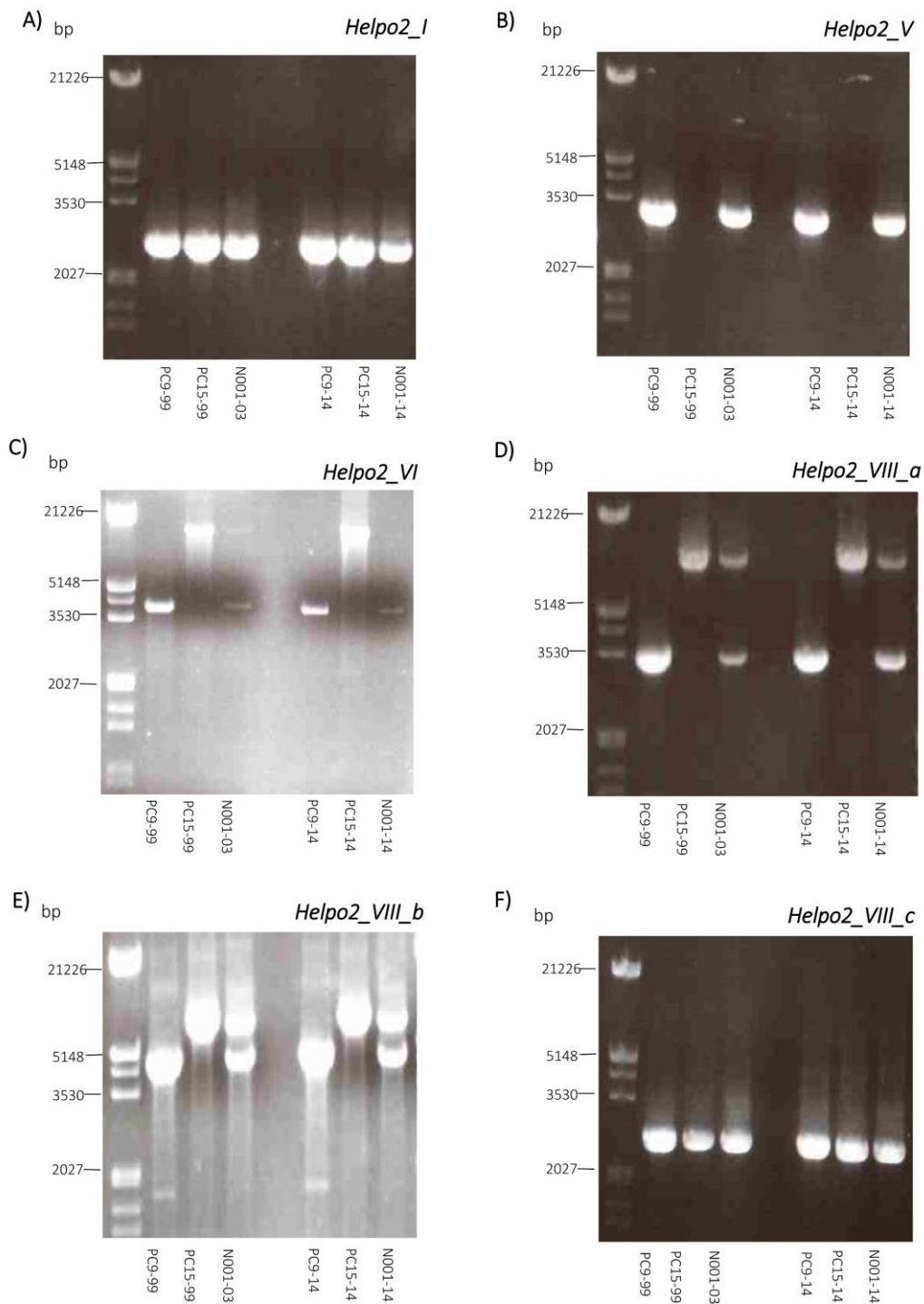


Figure S3. Analysis of full and vacant sites using the outer/outer PCR strategy

Molecular validation of HELPO2 insertion in PC9, PC15 and N001 subclones maintained under different subculture conditions. Panels show full and vacant sites corresponding to HELPO2 helitrons located at six different loci.

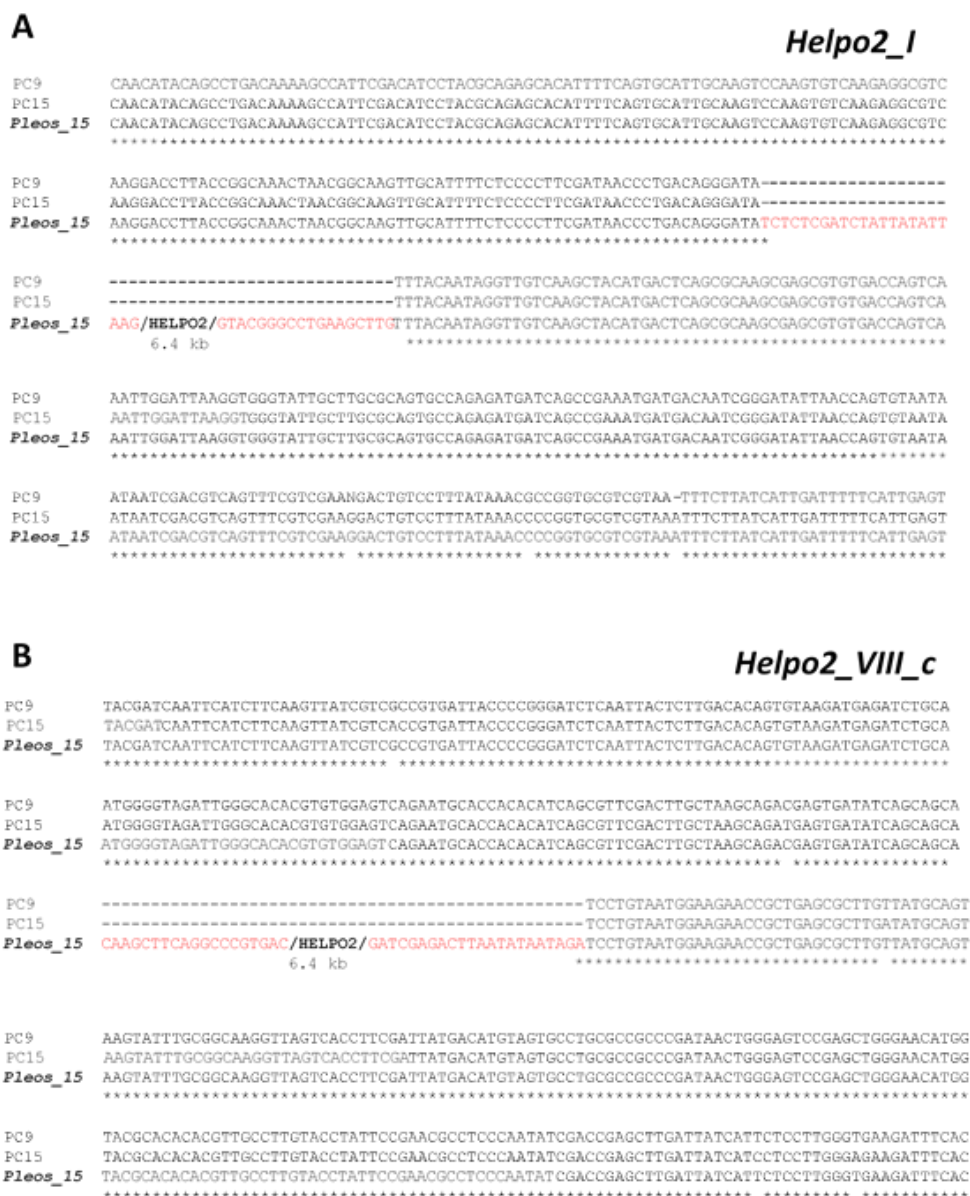


Figure S4. Alignment of *Helpo2_I* and *Helpo2_VIII_c* sequences

Alignment showing the full (*Pleos_15* : PC15 v.20 reference genome) and vacant sites (PC9 and PC15 subclones maintained under low and high subculture frequency) of two 6.4 kb HELPO2 helitrons placed at chromosome I (*Helpo2_I*) and VIII (*Helpo2_VIII_c*) of PC15 strain. Helitron insertion and flanking regions are represented in red and black type respectively.

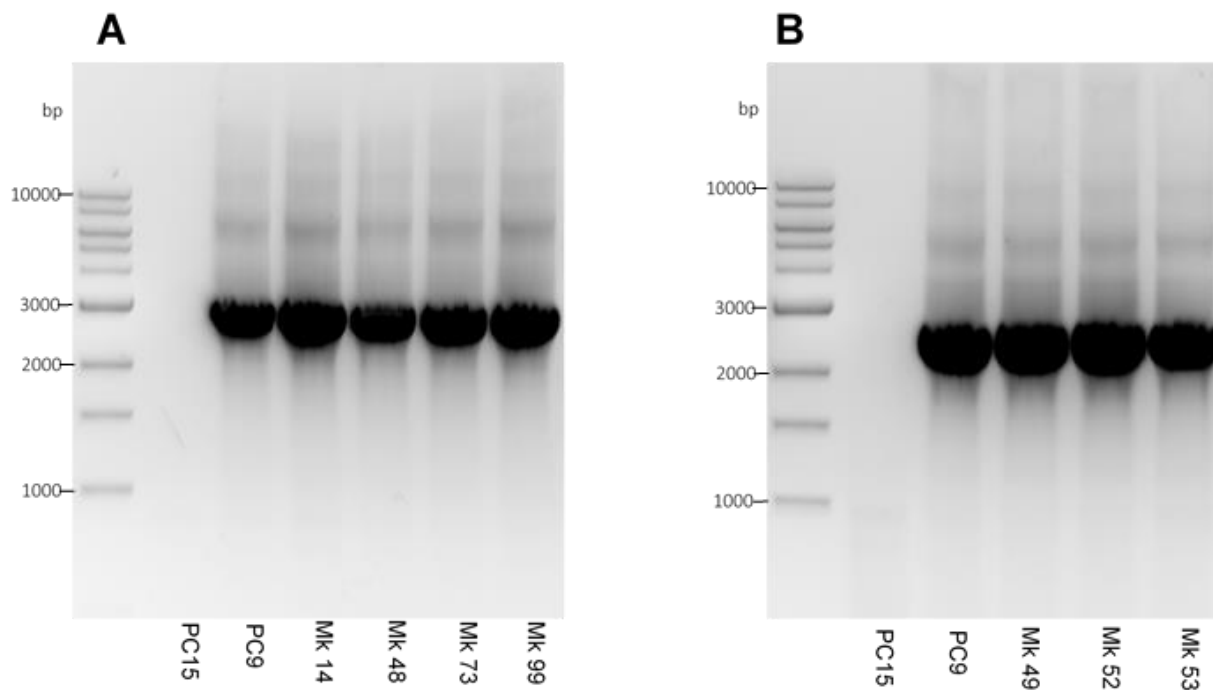


Figure S5. PCR amplification of gene conversion region in 7 monokaryons

Molecular validation of gene conversion regions in 7 individuals belonging to the monokaryotic progeny.

Panels show the 2971 bp (A) and 2426 bp (B) fragments located respectively in chromosomes VI and VIII.

A)

```

PleosPC9      TTCAACATCTCTGCCTTCTGCAGGGATTTCAATGCATACACCTTGCCCGTGTGCGACTTTC
mk14          TTCAACATCTCTGCCTTCTGCAGGGATTTCAATGCATACACCTTGCTCTTGTGCGACTTTC
mk48          TTCAACATCTAAGCCTTCTGCAGGGATTTCAATGCATACACCTTGCCCGTGTGCGACTTTC
mk73          TTCAACATCTCTGCCTTCTGCAGGGATTTCAATGCATACACCTTGCCCGTGTGCGACTTTC
mk99          TTCAACATCTCTGCCTTCTGCAGGGATTTCAATGCATACACCTTGCCCGTGTGCGACTTTC
PleosPC15     TTCAACATCTCTGCCTTCTGCAGGGATTTCAATGCATACACCTTGCCCGTGTGCGACTTTC
*****      *****      *****      *****      *****      *****      *      *****

PleosPC9      TGCACCAATCGTACCTAATGTATCGCCCATGGCAGTAAGATCAGTGGCCTAGGGTCTCT
mk14          TGCACCAATCGTACCTAATGTATCGCCCATGGCAGTAAGATCAGTGGCCTAGGGTCTCT
mk48          -GCACCAATCGTACCTAATGTATCGCCCATGGCAGTAAGATCAGTGGCCTAGGGTCTCT
mk73          TGCACNAATCGTACCTAATGTATCGCCCATGGCAGTAAGATCAGTGGCCTAGGGTCTCT
mk99          TGCACCAATCGTACCTAATGTATCGCCCATGGCAGTAAGATCAGTGGCCTAGGGTCTCT
PleosPC15     TGCACCAATCGTACCTAATGTATCGCCCATGGCAGTAAGATCAGTGGCCTAGGGTCTCT
****      *****

PleosPC9      AGAAGGTAAGTATGTACTTCGCCAAAGGCACCCTTCCCAGATGACCTTGACCGTGTGAAA
mk14          AGAATGGACGTATGTACTTCGCC---GGCACCCCTTCCCAGATGACCTTGACCGTGTGAAA
mk48          AGAATGGACGTATGTACTTCGCCAAAGGCACCCTTCCCAGATGACCTTGACCGTGTGAA-
mk73          AGAATGGACGTATGTACTTCGCCAAAGGCACCCTTCCCAGATGACCTTGACCGTGTGAAA
mk99          AGAACGGACGTATGTACTTCGCCAAAGGCACCCTTCCCAGATGACCTTGACCGTGTGAAA
PleosPC15     AGAATGGACGTATGTACTTCGCCAAAGGCACCCTTCCCAGATGACCTTGACCGTGTGAAA
****      *      *****

PleosPC9      TCGGAGAGCTTGATCTTCGTCCTACGCAGGCGCAGATGCTGCGACTCCTGCTGGCTGAAC
mk14          TCGGAGAGCTTGATCTTCGTCCTACGCAGGCGCAGATGCTGCGACTCCTGCTGGCTGAAC
mk48          TCGGAGAGCTTGATCTTCGTCCTACG-AGNNGCAGATGCTGCGACTCCTGCTGGCTGAAC
mk73          TCGGAGAGCTTGATCTTCGTCCTACGCAGGCGCAGATGCTGCGACTCCTGCTGGCTGAAC
mk99          TCGGAGAGCTTGATCTTCGTCCTACGCAGGCGCAGATGCTGCGACTCCTGCTGGCTGAAC
PleosPC15     TCGGAGAGCTTGATCTTCGTCCTACGCAGGCGCAGATGCTGCGACTCCTGCTGGCTGAAC
*****      *****

PleosPC9      GAGCGTTCAAAGTAGACGTAGTCTTCTGCGGGCTGCTGGTGTGAGATTATCTGGTATG
mk14          GAGCGTTCAAAGTAGACGTAGTCTTCTGCGGGCTGCTGGTGTGAGATTATCTGGTATG
mk48          GAGCGTTCAAAGTAGACGTAGTCTTCTGCGGGCTGCTGGTGTGAGATTATCTGGTATG
mk73          GAGCGTTCAAAGTAGACGTAGTCTTCTGCGGGCTGCTGGTGTGAGATTATCTGGTATG
mk99          GAGCGTTCAAAGTAGACGTAGTCTTCTGCGGGCTGCTGGTGTGAGATTATCTGGTATG
PleosPC15     GAGCGTTCAAAGTAGACGTAGTCTTCTGCGGGCTGCTGGTGTGAGATTATCTGGTATG
*****      *****

PleosPC9      GATACGGGGCAAAGTGC GGCTGCTGTTGCTGGTGGTGCATGGTAGTGTGCCCTGCCCTC
mk14          GATACGGGGCAAAGTGC GGCTGCTGTTGCTGGTGGTGCATGGTAGTGTGCCCTGCCCTC
mk48          GATACGGGGCAAAGNGC--CTGCTGTTGCTGGTGGTGCATGGTAGTGTGCCCTGCCCTC
mk73          GATACGGGGCAAAGTGC GGCTGCTGTTGCTGGTGGTGCATGGTAGTGTGCCCTGCCCTC
mk99          GATACGGGGCAAAGTGC GGCTGCTGTTGCTGGTGGTGCATGGTAGTGTGCCCTGCCCTC
PleosPC15     GATACGGGGCAAAGTGC GGCTGCTGTTGCTGGTGGTGCATGGTAGTGTGCCCTGCCCTC
*****      *****

PleosPC9      GTCGACTACCTTTGAGAGAACAAGGCCAAGATTTCGATCGTGAAGCCAGTCAAAGTTTTG
mk14          GTCGACTACCTTTGAGAGAACAAGGCCAAGATTTCGATCGTGAAGCCAGTCAAAGTTTTG
mk48          GTCGACTACCT-TGAGAGAACAAGGCCAAGATTTCGATCGTGAAGCCAGTCAAAGTTATG
mk73          GTCGACTACCTTTGAGAGAACAAGGCCAAGATTTCGATCGTGAAGCCAGTCAAAGTTTTG
mk99          GTCGACTACCTTTGAGAGAACAAGGCCAAGATTTCGATCGTGAAGCCAGTCAAAGTTTTG
PleosPC15     GTCGACTACCTTTGAGAGAACAAGGCCAAGATTTCGATCGTGAAGCCAGTCAAAGTTTTG
**      *****

```

```

PleosPC9      GCTGCAACACGCTGCTTTTCATATATGCCCATCAAATGCGGTAACAGTTACTATAACAAC
mk14          GCTGCAACACGCTGCTTTTCATATATGCCCATCAAATGCGGTAACAGTTACTATAACAAC
mk48          GCTGCAACACGCTGCTTTTCATA---G--CATCAAATGCGGTAACAGTTACTATAACAAC
mk73          GCTGCAACACGCTGCTTTTCATATATGCCCATCAAATGCGGTAACAGTTACTATAACAAC
mk99          GCTGCAACACGCTGCTTTTCATATATGCCCATCAAATGCGGTAACAGTTACTATAACAAC
PleosPC15    GCTGCAACACGCTGCTTTTCATATATGCCCATCAAATGCGGTAACAGTTACTATAACAAC
*****
PleosPC9      CATCGGCTAATTCATCACATGGAAATCGGCCGTACATCACAGAACACGTTTCATGGAAC
mk14          CATCGGCTAATTCATCACATGGAAATCGGCCGTACATGACAGAACACGTTTCATGGAAC
mk48          CATCGGCTAATTCATCACATGGAAATCGGCCGTACATCACAGAACACGTTTCATGGAAC
mk73          CATCGGCTAATTCATCACATGGAAATCGGCCGTACATCACAGAACACGTTTCATGGAAC
mk99          CATCGGCTAATTCATCACATGGAAATCGGCCG--CATCACAGAACACGTTTCATGGAAC
PleosPC15    CATCGGCTAATTCATCACATGGAAATCGGCCGTACATCACAGAACACGTTTCATGGAAC
*****
PleosPC9      AACAGTGATGCCTGCTCATACGTCCAAACTTCCAACATTCACGTCAGAGTCTAGAGTTGT
mk14          AACAGTGATGCCTGCTCATACGTCCAAACTTCCAACATTCACGTCAGAGTCTAGAGTTGT
mk48          AACAGTGATGCCTGCTCATACGTCCAAACTTCCAACATTCACGTCAGAGTCTAGAGTTGT
mk73          AACAGTGATGCCTGCTCATACGTCCAAACTTCCAACATTCACGTCAGAGTCTAGAGTTGT
mk99          AACAGTGATGCCTGCTCATACGTCCAAACTTCCAACATTCACGTCAGAGTCTAGAGTTGT
PleosPC15    AACAGTGATGCCTGCTCATACGTCCAAACTTCCAACATTCACGTCAGAGTCTAGAGTTGT
*****
PleosPC9      GAGT-----GAATTCTGTCTTATTCAAGTCGA-----
mk14          GAGT-----GAATTCTGTCTTATTCAAGTCGA-----
mk48          -----GAATTCTGTCTTATTCAAGTCGA-----
mk73          GAGT-----GAATTCTGTCTTATTCAAGTCGA-----
mk99          GAGT-----GAATTCTGTG-NATTCAAGTCGA-----
PleosPC15    GAGTCAAGAAACGACCAGACTTCTGTCTTATTCAAGTCGACAAGTTTCAGGCCCGTGACT
*****
PleosPC9      -----
mk14          -----
mk48          -----
mk73          -----
mk99          -----
PleosPC15    TCATCAGCATGGGACTCAGG /Helitron 6.4Kb/ GATTACCTAAGTTCAAAGTTC
*****
PleosPC9      -----TCGCTAACGCCTACAAGGTAGAGACCAATGAAG
mk14          -----TCGCTAACGCCTACAAGGTAGAGACCAATGAAG
mk48          -----TCGCTAACGCCTACAAGGTAGAGACCAATGAAG
mk73          -----TCGCTAACGCCTACAAGGTAGAGACCAATGAAG
mk99          -----TCGCTAACGCCTACAAGGTAGAGACCAATAAAG
PleosPC15    ACACGGATCGAGACTTAATATAATAGATCGCTAACGCCTACAAGGTAGAGACCAATGAAG
*****
PleosPC9      ACGGAGTGCGCCCTATGCCTGAGGGCCAACGCCCTACGATGTGACCGTCTTCAGCGC
mk14          ACGGAGTGCGCCCTATGCCTGAGGGCCAACGCCCTACGATGTGACCGTCTTCAGCGC
mk48          ACTGAGTGCGCCCTATGCCTGAGGGCCAACGCCCTACGATGTGACCGTCTTCAGCGC
mk73          ACGGAGTGCGCCCTATGCCTGAGGGCCAACNTCCTACGATGTGACCGTCTTCAGCGC
mk99          ACGGAGTGCGCCCTATGCCTGAGGGCCAACGCCCTACGATGTGACCGTCTTCAGCGC
PleosPC15    ACGGAGCGCGGCCCTATGCCTGAGGGCCAACGCCCTACGATGTGATCATGTCTTCAGCGC
*****

```

Somatic transposition and meiotically-driven elimination of an active helitron family in *Pleurotus ostreatus*

```

PleosPC9      TAAGGCAAGGGGGATGAAGTGAAGAGGAGGTACGATACGACGAAAAGACTTGATTCAATA
mk14          TACGGCAAGGGGGATGAAGTGAAGAGGAGGTACGATACGACGAAAAGACTTGATTCAATA
mk48          TAAGGCAAGGGGGATGAAGTGACCAGGAGG-ACGATACGACGAAAAGACTTGATTCTATA
mk73          TAAGGCAAGGGGGATGAAGTGAAGAGGAGGTACGATACGACGAAAAGACTTGATTCAATA
mk99          TAAGGCAAGGGGGATGAAGTGAAGAGGAGGTACGATACGACGAAAAGACTTGATTCAATA
PleosPC15    TATGGTGAGGGGGATGAAGTGAAGAGGAGGTACGATACGACGAAAAGACTTGATTCAATA
              ** ** *****
PleosPC9      TACAACATACCTAGGCTTGCGCAGCGCA-----GAGTCAGTACTAGTGTGTGTT
mk14          TACAACATACCTAGGCTTGCGCAGCGC-----GAGTCAGTACTAGTGTGTGTT
mk48          TACAACATACCTAGGCTTGCGCAGCGCA-----GAGTCAGTACTAGTGTGTGTT
mk73          TACAACATACCTAGGCTTGCGCAGCGCA-----GAGTCAGTACTAGTGTGTGTT
mk99          TACAACATACCTAGGCTTGCGCAGCGCA-----GAGTCAGTACTAGTGTGTGTT
PleosPC15    TACA-----AGGCTTGCGCAGCGCAGAGTGGCACATCGAGTCGGTACTAGTGTGTGTT
              **** *****

PleosPC9      TCAAATTCACCGACGCGATCTGGGCAGGAGATAGGAAGATAATAAGCGTGATACAGTCAG
mk14          TCAAATTCACCGACGCGATCTGGGCAGGAGATAGGAAGATAATAAGCGTGATACAGTCAG
mk48          TCAAATTCACCGACGCGATCTGGGCAGGAGATAGGAAGATAATAAGCGTGATACAGTCAG
mk73          TCAAATTCACCGACGCGATCTGGGCAGGAGATAGGAAGATAATAAGCGTGATACAGTCAG
mk99          TCAAATTCACCGACGCGATCTGGGCAGGAGATAGGAAGATAATAAGCGTGATACAGTCAG
PleosPC15    TCAAATTCACCGACGCGATCTGGGCAGGAGATAGGAAGATAATAAGCGTGATACAGTCAG
              ***** ** *****

PleosPC9      TGAATGACCCAACGTGCAGTCCGTCTGACCTCCACTTGGCTTCCACTCTTCAATCGTCGC
mk14          TGAATGACCCAACGTGCAGTCCGTCTGACCTCCACTTGGCTTCCACTCTTCAATCGTCGC
mk48          TGAATGACCCAACGTGCAGTCCGTCTGACCTCCACTTGGCTTCCACTCTTCAATCGTCGC
mk73          TGAATGACCCAACGTGCAGTCCGTCTGACCTCCACTTGGCTTCCACTCTTCAATCGTCGC
mk99          TGAATGACCCAACGTGCAGTCCGTCTGATCTCCACTTGGCTTCCACTCTTCAATCGTCGC
PleosPC15    TGAATGACCCAACGTGCAGTCCGTCTGACCTCCACTTGGCTTCCACTCTTCAATCGTCGC
              ***** ***** *****

PleosPC9      AGCACGGCTTGTTAGAAGCCAATATATCTCCTTGGAACTCCCGGCCGTTCACTTATTCC
mk14          AGCACGGCTTGTTAGAAGCCAATATATCTCCTTGGAACTCCCGGCCGTTCACTTATTCC
mk48          AGCACGGCTTGTTAGAAGCCAATATATCTCCTTGGAACTCCCGGCCGTTCACTTATTCC
mk73          AGCACGGCTTGTTAGAAGCCAATATATCTCCTTGGAACTCCCGGCCGTTCACTTATTCC
mk99          AGCACGGCTTGTTAGAAGCCAATATATCTCGTTGGAACTCCCGGCCGTTCACTTATTCC-
PleosPC15    AGCACGGCTTGTTAGAAGCCAATATATCTCCTTGGAACTCCCGGCCGTTCACTTATTCC
              ***** ***** *****

PleosPC9      GTATTGTCGAAATATACATGGCCGAAGACGACGTGCAACGACCTACCAAGAGGTCTAAA
mk14          GTATTGTCGAAATATACATGGCCGAAGACGACGTGCAACGACCTACCAAGAGGTCTAAA
mk48          GTATTGTCGAAATATACATGGCTGAAGACGACGTGCAACGACCTACCAAGAGGTCTAAA
mk73          GTATTGTCGAAATATACATGGCCGAAGATGACGTGCAA-GACC-CCCAAGAGGTCTAAA
mk99          -TATTGTCGAAATATACATGGCCGAAGACGACGTGCAACGACCTACCAAGAGGTCTAAA
PleosPC15    GTATTGTCGAAATATACATGGCCGAAGACGACGTGCAACGACCTACCAAGAGGTCTAAA
              ***** ***** *****

PleosPC9      ATAGAGCATGGATCATCTCCCACCGCATGACCAATGATGATGTTATGCAAAACGACGAC
mk14          ATAGAGCATGGATCATCTCCCACCGCATGACCAATGATGATGTTATGCAAAACGACGAC
mk48          ATAGAGCATGGATCATCTCCCACCGCATGACCAATGATGATGTTATGCAAAACGACGAC
mk73          ATAGAGCATGGATCATCTCCCACCGCATGACCAATGATGATGTTCTGCAAAACGACGAC
mk99          ATAGAGTATGGATCATCTCCCACCGCATGACCAATGATGATGTTATGCAAAACGACGAC
PleosPC15    ATAGAGCATGGATCATCTCCCACCGCATGACCAATGATGATGTTATGCAAAACGACGAC
              ***** ***** *****

```



```

PleosPC9      ACTGCCATGACGGGAGAGCCTGAAGCATCAAGCTCAAGCGCTGTTGCTGTCTGAACACGCC
mk14          ACTGCCATGACGGGAGAGCCTGAAGCATCAAGCTCAAGCGCTGTTGCTGTCTGAACACGCC
mk48          ACTGCCATGACGGGAGAGCCTGAAGCATCAAGCTCAAGCGCTGTTGCTGTCTGAACACGCC
mk73          ACTGCCATGACGGGAGAGCCTGAAGCATCAAGCTCAAGCGCTGTTGCTGTCTGAACACGCC
mk99          -CCGCCATGA-GGGGGAGCCTGAAGCTT-AACTTTAGCCCTGTTGCTGTCTGAACACCC-
PleosPC15     ACTGCCATGACGGGAGAGCCTGAAGCATCAAGCTCAAGCGCTGTTGCTGTCTGAACACGCC
               *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *   *
PleosPC9      CCAGCCGGGCAACCGGGGAAAAAAAAAGAGACGGGACGCGTCTGGGTGGGCAAATCTAGG
mk14          CCAGCCGGGCAACCGGGGAAAGGGAAAGAGACGGGACGCGTCTGGGTGGGCAAATCTAGG
mk48          CCAGCCGGGCAACCGGGGAAAGGGAAAGAGACGGGACGCGTCTGGGTGGGCAAATCTAGG
mk73          CCAGCCGGGCAACCGGGGAAAGGGAAAGAGACGGGACGCGTCTGGGTGGGCAAATCTAGG
mk99          CCAGCCGGGCAACCGGGGAAAG-GAAGAGACGGGACGCGTCTGGGTGGGCAAATCTAGG
PleosPC15     CCAGCCGGGCAACCGGGGAAAGGGAAAGAGACGGGACGCGTCTGGGTGGGCAAATCTAGG
               *****

```

B)

```

PleosPC9      AAAGAAGTTAGGAGACCAGGTGGACCCAGGCGGGTCATGGCTTACACCGTGGGCTCCCGG
Mk52          AAAGAAGTTAGGAGACCAGGTGGACCCAGGCGGGTAATGGCTTACAGCCGTGGGCTCCCGG
Mk53          TTAGAAGTTAGGAGACCAGGTGGACCCAGGCGGGTAATGGCTTACAGCCGTGGGCTCCCGG
Mk49          CAAGAAGTTAGGAGACCAGGTGGACCCAGGCGGGTAATGGCTTACAGCCGTGGGCTCCCGG
PleosPC15     AAAGAAGTTAGGAGACCAGGTGGACCCAGGCGGGTAATGGCTTACAGCCGTGGGCTCCCGG
                *****

PleosPC9      GCCTCAATCCTAGCGAAGAATACTGTCAGTGGTTGTCTTGCAGCCGCAAACATCTAAAC
Mk52          GCCGCAATCCTAGCGAAGAATACTGTCAGTGGTTGTCTCGCAGCCGCAAACATCTAAAC
Mk53          GCCTCAATCCTAGCGAAGAATACTGTCAGTGGTTGTCTCGCAGCCGCAAACATCTAAAC
Mk49          GCCTCAATCCTAGCGAAGAATACTGTCAGTGGTTGTCTCGCAGCCGCAAACATCTAAAC
PleosPC15     GCCTCAATCCTAGCGAAGAATACTGTCAGTGGTTGTCTCGCAGCCGCAAACATCTAAAC
                ***

PleosPC9      ATACAAGTTTGCTCATCACTTTGCGTCCTTGATTGGTTTCTCCGTTAGAAGGGACTGC
Mk52          ATACAAGTTTGCTCATCACTTTGCGTCCTTGATTGGTTTCTCCGTTAGAAGGGACTGC
Mk53          ATACAAGTTTGCTCATCACTTTGCGTCCTTGATTGGTTTCTCCGTTAGAAGGGACTGC
Mk49          ATACAAGTTTGCTCATCACTTTGCGTCCTTGATTGGTTTCTCCGTTAGAAGGGACTGC
PleosPC15     ATACAAGTTTGCTCATCACTTTGCGTCCTTGATTGGTTTCTCCGTTAGAAGGGACTGC
                *****

PleosPC9      GACCACTGCGGTAGCAAGGACGGCGGCCAGAATGACGGTCTTGATTGAACCATATTTGC
Mk52          GACCACTGCGGTGGCAAGGACGGCGGCCAGAATGACGGTCTTGATTGAACCATATTTGC
Mk53          GACCACTGCGGTGGCAAGGACGGCGGCCAGAATGACGGTCTTGATTGAACCATATTTGC
Mk49          GACCACTGCGGTGGCAAGGACGGCGGCCAGAATGACGGTCTTGATTGAACCATATTTGC
PleosPC15     GACCACTGCGGTGGCAAGGACGGCGGCCAGAATGACGGTCTTGATTGAACCATATTTGC
                *****

PleosPC9      AGAGTAGTAATAACTTCGATGACCGGCACCAGAAGTCGAATATGTGAGCTCAGCTGGAGA
Mk52          AGAGTAGTAATAACTTCGATGACCGGCACCAGAAGTCGAATATGTGAGCTCAGCTGGAGA
Mk53          AGAGTAGTAATAACTTCGATGACCGGCACCAGAAGTCGAATATGTGAGCTCAGCTGGAGA
Mk49          AGAGTAGTAATAACTTCGATGACCGGCACCAGAAGTCGAATATGTGAGCTCAGCTGGAGA
PleosPC15     AGAGTAGTAATAACTTCGATGACCGGCACCAGAAGTTGAATATGTGAGCTCAGCTGGAGA
                *****

PleosPC9      ACAATGCATTATAGTCGCGATTTATATACTGCGATTCCGCCAGTAACACCGTCGTTGAA
Mk52          ACAATGCATTATAGTCGCGATTTATATACTGCGATTCCGCCAGTAACACCGTCGTTGAA
Mk53          ACAATGCATTATAGTCGCGATTTATATACTGCGATTCCGCCAGTAACACCGTCGTTGAA
Mk49          ACAATGCATTATAGTCGCGATTTATATACTGCGATTCCGCCAGTAACACCGTCGTTGAA
PleosPC15     ACAATGCATTATAGTCGCGATTTATATACTGCGATTCCGCCAGTAACACCGTTGTTGAA
                *****

PleosPC9      ACGTTGTAGCCCTTAGTGTGCACGCTGCACAAGTACCGCCATCGCCGGACTTGTTCATG
Mk52          ACGTTGTAGCCCTTAGTGTGCACGCTGCACAAGTACCGCCATCGCCGGACTTGTTCATG
Mk53          ACGTTGTAGCCCTTAGTGTGCACGCTGCACAAGTACCGCCATC--CGGACTTGTTCATG
Mk49          ACGTTGTAGCCCTTAGTGTGCACGCTGCACAAGTACCGCCATCGCCGGACTTGTTCATG
PleosPC15     ACGTTGTAGCCCTTAGTGTGCACGCTGCACAAGTACCGCCATCGCCGGACTTGTTCATG
                *****

PleosPC9      CCTTGGCATGGGGCGCACCACAGAGGCTGCATGGTCCTCGTCGGGTAGGAGCGGTGAGCA
Mk52          CCTTGGCATGGGGCGCACCACAGAGGCTGCATGGTCCTCGTCGGGTAGGAGCGGTGAGCA
Mk53          CCTTGGCATGGGGCGCACCACAGAGGCTGCATGGTCCTCGTCGGGTAGGAGCGGTGAGCA
Mk49          CCTTGGCATGGGGCGCACCACAGAGGCTGCATGGTCCTCGTCGGGTAGGAGCGGTGAGCA
PleosPC15     CCTTGGCATGGGGCGCACCACAGAGGCTGCATGGTCCTCGTCGGGTAGGAGCGGTGAGCA
                **

```

```

PleosPC9      TGAACATTTCAAAAAAGCTCCCCATCGCTCCTGTTTCGAGGA-----
Mk52          TGAACATTTCAAAAAAGCTCCCCATCGCTCCTGTTTCGAGGA-----
Mk53          TGAACATTTCAAAAAAGCTCCCCATCGCTCCTGTTTCGAGGA-----
Mk49          TGAACATTTCAAAAAAGCTCCCCATCGCTCCTGTTTCGAGGA-----
PleosPC15     TGAACATTTGAAAAGCTCCCCATCGCTCCTGTTTCGAGGATCTATATTATCTATCTATC
*****
*****

PleosPC9      -----
Mk52          -----
Mk53          -----
Mk49          -----
PleosPC15     AATACATCGTGTGGGGCCTTGCTTCCAGAGCAAGGACAAGTACCACAATATGCCAACT

PleosPC9      -----
Mk52          -----
Mk53          -----
Mk49          -----
PleosPC15     TTATATTCATGACCCAGCG/Helitron 3.9 Kb/GTACTCATATCCACATAGTATGC

PleosPC9      -----TTTGAATGAAATCCGACTCGAGCTCGCTG
Mk52          -----TTTGAATGAAATCCGACTCGAGCTCGCTG
Mk53          -----TTTGAATGAAATCCGACTCGAGCTCGCTG
Mk49          -----TTTGAATGAAATCCGACTCGAGCTCGCTG
PleosPC15     AATTCGTGACGAAGTCACGGGCTGAAACTTGTGTTGAATGAAATCCGACTCGAGCTCACTG
*****
*****

PleosPC9      AGTCTGAGCTCGCGTCGTGCACGGTCACGTTCTTGCTTTCACAGGGCTGTCGGTGTTCGG
Mk52          AGTCTGAGCTCGCGTCGTGCACGGTCACGTTCTTGCTTTCACAGGGCTGTCGGTGTTCGG
Mk53          AGTCTGAGCTCGCGTCGTGCACGGTCACGTTCTTGCTTTCACAGGGCTGTCNGTGTTCGG
Mk49          AGTCTGAGCTCGCGTCGTGCACGGTCACGTTCTTGCTTTCACAGGGCTGTCGGTGTTCGG
PleosPC15     AGTCTGAGCTCGCGTCGTGCACGGTCACGTTCTTGCTTTCACAGGGCTGTCGGTGTTCGG
** *****

PleosPC9      ACACCAAA-----CATTCAATTACGCAGTCTAACA-----CCAGAAATTCCTGATTC
Mk52          TCACCAAA-----CATTCAATTACGCAGTCTAACA-----CCAGAAATTCCTGATTC
Mk53          ACACCAAA-----CATTCAATTACGCAGTCTAACA-----CCAG--ATTCCTGATTC
Mk49          ACACCAAA-----CATTCAATTACGCAGTCTAACA-----CCAGAAATTCCTGATTC
PleosPC15     ACACCAAAGATCCACCATTCAATTACGCAGTCTAACACTCGCGCCAGAAATTCCTGATTC
*****
*****

PleosPC9      GGACAGCCTTTTAAAGTCCCTATTCCAAGCTGTGCGGCATCGATACGTCTATTGCGAAAACG
Mk52          GGACAGCCTTTTAAAGTCCCTATTCCAAGCTGTGCGGCATCGATACGTCTATTGCGAAAACG
Mk53          G--CAGCCCGTTAAGTCCCTATTCCAAGCTGTGCGGCATCGATACGTCTATTGCGAAAACG
Mk49          GGACAGCCTTTTAAAGTCCCTATTCCAAGCTGTGCGGCATCGATACGTCTATTGCGAAAACG
PleosPC15     -----CTTTTAAAGTCCCTATTCCAAGCTGTGCGGCATCGATACGTCTATTGCAAAAACG
* *****

PleosPC9      AAGTAGCCGATCATGAACCCGACGCTCCGCGGAGTATTGTCATTGGACGGATGCTGGCCT
Mk52          A-GTAGCCGATCATGAACCCGACGCTCCGCGGAGTATTGTCATTGGACGGATGCTGGCCT
Mk53          AAGTAGCCGATCATGAACCCGACGCTCCGCGGAGTATTGTCATTGGACGGATGCTGGCCT
Mk49          AAGTAGCCGATCATGAACCCGACGCTCCGCGGAGTATTGTCATTGGACGGATGCTGGCCT
PleosPC15     AAGTAGCCGATCATGAACCCGACGCTCCGCGGAGTATTGTCATTGGACGGATGCTGGCCT
* *****

```

```

PleosPC9      GAATGTATGGGCCGAGAGTGAAGGTATCCCGGCATGTTGTAATATCAATCGGTTAGAGAA
Mk52          GAATGTATGGGCCGAGAGTGAAGGTATCCCGGCATGTTGTAATATCAATCGGTTAGAGAA
Mk53          -AATGTATGGGCCGAGAGTGAAGGTATCCCGGCATGTTGTAATATCAATCGGTTAGAGAA
Mk49          GAATGTATGGGCCGAGAGTGAAGGTATCCCGGCATGTTGTAATATCAATCGGTTAGAGAA
PleosPC15     GAATGTATGGGCCGAGAGTGAAGGTATCCCGGCATGTTGTAATATCAATCGGTTAGAGAA
                *****

PleosPC9      GCGCAGCAACGTAGGACGAGAGGAGACTGAGGGGAGAGGGCAGTGAGGACTCAGC SAAGG
Mk52          GCGCAGCAACGTAGGACGAGAGGAGACTGAGGGGAGAGGGCAGTGAGGACTCAGC GAAGG
Mk53          GCGCAGCAACGTAGGACGAGAGGAGACTGAGGGGAGAGGGCAGTGAGGACTCAGC GAAGG
Mk49          GCGCAGCAACGTAGGACGAGAGGAGACTGAGGGGAGAGGGCAGTGAGGACTCAGC GAAGG
PleosPC15     GCGCAGCAACGTAGGACGAGAGGAGACTGAGGGGAGAGGGCAGTGAGGACTCAGC -----
                *****

PleosPC9      CTGGGGCTCCCCACACATCGCGCCCGATTCCCTTCACTTACGCTTTGGAATCAGCGCA
Mk52          -----CTCCCCCACACATCGCGCCCGATTCCCTTCAACTTACGCTTTGGAATCAGCGCA
Mk53          -----CTCCCCCACACATCGCGCTCNATTCCTTCAACTTACGCTTTGGAATCAGCGCA
Mk49          -----CTCCCCCACACATCGCGCCCGATTCCCTTCAACTTACGCTTTGGAATCAGCGCA
PleosPC15     -----CTCCCCCACACATCGCGCCCGATTCCCTTCAACTTACGCTTTGGAATCAGCGCA
                ** ***** * *****

PleosPC9      GACGCGGGTTAGGACGACCATCACAATGTCTCAACCTGGCTGCCAGAGGCACCTAGAGC
Mk52          GACGCGGGTTAGGACGACCATCACAATGTCAATTGACCTGGCTGCCAGAGGCACCTAGAGC
Mk53          GACGCGGGTTAGGACGACCATCACAATGTCAATTGACCTGGCTGCCAGAGGCACCTAGAGC
Mk49          GACGCGGGTTAGGACGACCATCACAATGTCAATTGACCTGGCTGCCAGAGGCACCTAGAGC
PleosPC15     GACGCGGGTTAGGACGACCATCACAATGTCAATTGACCTGGCTGCCAGAGGCACCTAGAGC
                ***** * *****

PleosPC9      -----CAACATGGAAGACTGGAGTAC
Mk52          -----CAATATGGAAGACTGGAGT--
Mk53          -----NNACATGGAAGACTGGAGTAC
Mk49          -----CAACATGGAAGACTGGAGTAC
PleosPC15     CATGTTCGGAACCTACCACCTCGAAGCCAACGACACAACAACATGGAAGACTGGAGTAC
                * *****

```

Figure S6. DNA sequence alignment of gene conversion region

Complete DNA sequence of gene conversion regions of chromosomes VI (A) and VIII (B) described respectively in 4 and 3 monokaryons. SNPs corresponding to PC15 and PC9 parental strains are indicated respectively in blue and red. HELPO2 insertions in PC15 sequence are identified by yellow colour.

Chapter III – Epigenetic and transcriptomic profiles throughout the life cycle of *Pleurotus ostreatus*

Introduction

The extraordinary increase in genomic data released during the last decade has made it possible to begin unraveling the impact of transposable elements (TEs) on a wide range of eukaryotic genomes. TEs can interrupt genes, produce rearrangements and lead to illegitimate recombination events (Doolittle and Sapienza 1980). The majority of TEs are usually present as defective copies that have accumulated mutations and deletions, which ultimately inactivate their transposition potential. Under most circumstances, TEs accumulate in centromeric and pericentromeric regions where they play a crucial role in genomic architecture and heterochromatin maintenance (Fukui et al. 2001; Slotkin and Martienssen 2007). Nevertheless, active transposons constitute a significant source of mutations that can lead to harmful effects in the host genome (McGinnis et al. 1983; Lonngren and Saedler 2002).

Thus, most eukaryotes have evolved epigenetic defense mechanisms to limit TE expansion. More specifically, transcriptional and post-transcriptional gene silencing pathways (TGS and PTGS) have been described to suppress transposon activity in plants and animals (Klahre et al. 2002; Malone and Hannon 2009). TGS operates through DNA methylation and histone modification, whereas PTGS is mediated by small RNAs and orchestrated by the RNA interference pathway (RNAi).

In the fungal kingdom, insight into epigenetic silencing mechanisms have been mainly obtained from the filamentous ascomycete *Neurospora crassa* and are beginning to be documented in other lineages. Notably, the increasing study of epigenetic phenomena in filamentous fungi have shed more light on several pathways, referred to as 'genome defense', that appear to have evolved to limit the expansion of virus and repeated elements, such as TEs. By contrast, gene regulation by epigenetic modifications observed in plants and animals (Suzuki and Bird 2008; Law and Jacobsen 2010) during development is still poorly understood in fungi. In *N. crassa*, TGS and PTGS epigenetic pathways have been identified to inactivate transposons at the vegetative and sexual stages (Dang et al. 2014). Specifically, a repeat-induced point mutation (RIP) operates during the sexual cycle to transcriptionally silence repetitive sequences through a homology-dependent manner. In *Neurospora*, it occurs during the sexual stage in haploid nuclei after fertilization but prior to meiotic DNA replication (Selker et al. 1987; Selker et al. 2003). This mechanism induces CG to AT hypermutations in these sequences during the sexual cycle, leading to their degeneration and silencing by DNA methylation. In fact, the resulting A/T-rich sequences are potent signals for subsequent *de novo* DNA methylation (Tamaru and Selker 2003).

A related mechanism, which is called methylation induced premeiotically (MIP), has been detected in *Ascobolus immersus* and displays similar hallmarks with RIP (Barry et al. 1993). Both mechanisms act to silence duplicated DNA, either by epigenetic (MIP) or a combination of epigenetic and mutagenic (RIP) marks. In filamentous fungi, DNA methylation, defined as the archetype of all epigenetics, has been described as a genome defense mechanism involved in the transcriptional repression of repeat sequences. Relevant discoveries of RIP and MIP mechanisms uncovered that in ascomycetes, transposons can be inactivated by DNA methylation linked to RIP mutations (Galagan and Selker 2004), whereas in basidiomycetes, this mechanism has been reported almost exclusively in members of the Pucciniomycotina (Horns et al. 2012) and in the ustilaginomycete *Microbotrium violaceum* (Johnson et al. 2010). Distinct patterns of DNA methylation have been observed in many fungal lineages, while in others, this mechanism seems to be absent (Malagnac and Silar 2010; Zemach and Zilberman 2010). Most filamentous fungi possess an enzymatic system that selectively target the C5 of cytosines. However, recent studies uncovered adenine methylation (6mA) associated with transcriptionally active genes in early-diverging fungal lineages (Mondo et al. 2017). Two distinct DNA methyltransferases (DMTs) have been delineated in filamentous fungi. DIM-2 in *Neurospora* (Kouzminova and Selker 2001) and its ortholog in *A. immersus* (Chernov et al. 1997) are responsible for DNA methylation and transcriptional silencing in vegetative cells. The second DMT (i.e., Masc1, which is found in *A. immersus*) is involved in development and premeiotically induced DNA methylation during the sexual stage (Malagnac et al. 1997). The ortholog RID in *Neurospora* is described for repeated-induced point (RIP) mutation during the sexual phase (Freitag et al. 2002) (Fig. 1).

Comprehensive methylome analyses carried out in five fungi that belong to the *Zygomycota*, *Ascomycota* and *Basidiomycota* groups reported a marked preference for methylation in the CG sites placed within transposons and other repeated sequences in contrast to the weak methylation levels found in gene-coding regions¹⁰. Furthermore, a similar trend was recently reported in the plant-symbiotic ascomycete *Tuber melanosporum* (Montanini et al. 2014). Regarding PTGS of transposable elements, two RNAi mechanisms have been described in *N. crassa*: quelling (Romano and Macino 1992) (related to plant co-suppression) and MSUD (meiotic silencing by unpaired DNA) (Shiu et al. 2001). Both mechanisms are based on the production of small RNAs and rely on the core components of the RNAi pathway.

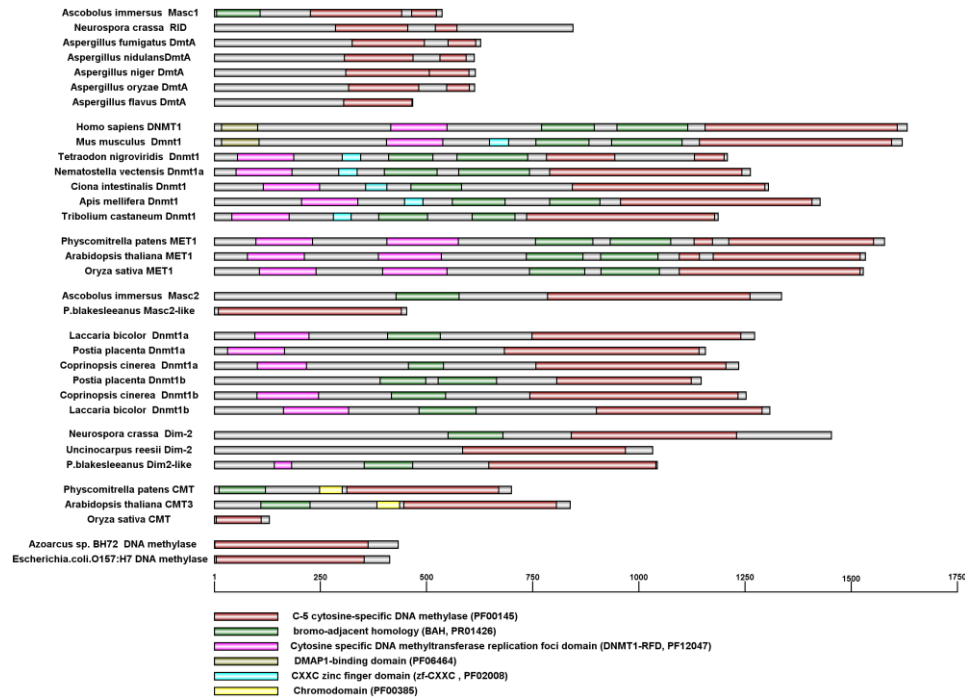


Figure 1. Conserved domains of DNA methyltransferases from 12 fungi, 2 mammals, 5 invertebrates, and 3 plants. DNA methylase domains are located on the N-terminal regions of each protein. Obtained from Liu et al. (2102).

Quelling relies on the generation of small RNA that act in *trans* to degrade targeted mRNAs molecules, leading to the silencing of repetitive DNA sequences (i.e., transposons or multi-copy genes) during vegetative growth. In this process, aberrant DNA originates from repetitive sequences, or damaged ribosomal RNA is transcribed into aberrant DNA (aDNA) by an Argonaute protein (QDE-1). Subsequently, aRNA is cleaved into 25nt small interfering RNA (siRNA) by a Dicer-like protein (Chang et al. 2012). These siRNAs bind to a second Argonaute protein (QDE-2) to degrade the homologous coding RNA (mRNA) (Figure 2). The sexual RNAi pathway triggered by unpaired DNA during meiosis (MSUD) results from the lack of sequence pairing in homologous chromosomes (Hammond et al. 2013). It has been proposed as a defense mechanism against the insertion of repeated elements that can accumulate during vegetative growth and asexual propagation, to generate unpaired regions that are silenced during meiosis (Shiu et al. 2001). In this mechanism, the unpaired DNA is transcribed into aRNA, leading to the production of double-stranded RNA (dsRNA) produced by a RNA polymerase (SAD-1). A Dicer-like 1 protein processes the dsRNA, and the resulting small noncoding RNA (sRNAs) are loaded on an Argonaute protein (SMS-2), which guides the silencing of homologous DNA sequences (Fig. 2).

Moreover, an RNAi-dependent mechanism has been recently discovered to silence transposable elements at the post-transcriptional level during sexual development (SIS, sex-induced silencing) in the basidiomycete *Cryptococcus neoformans* (Wang et al. 2010a). Similarly to quelling, SIS requires the presence of tandem arrays of the transgene sequence, and its efficiency correlates with the copy number of the transgene. Also, a significant portion of the sRNA population in *C. neoformans* originates from repetitive sequences and centromeric regions, supporting the role of the SIS mechanism in transposon control (Wang et al. 2010a).

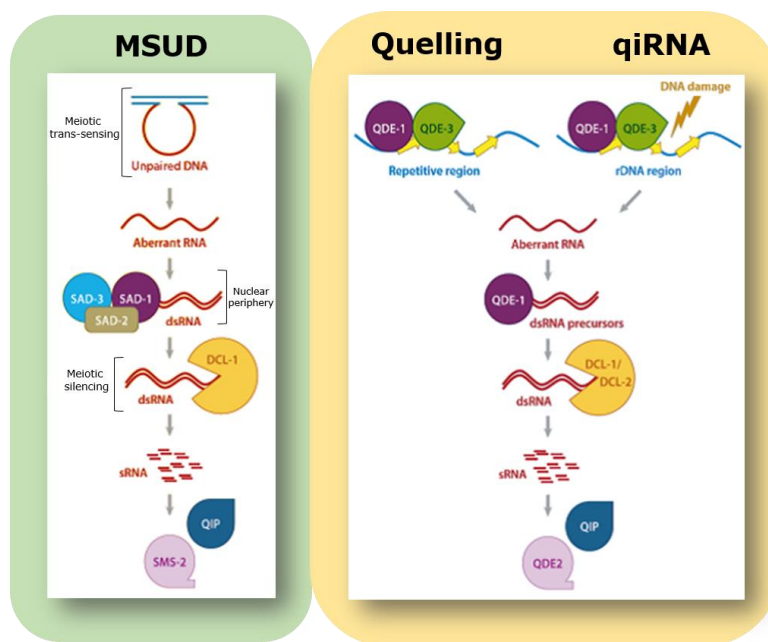


Figure 2 . RNAi pathways during sexual and vegetative stages in *N. crassa*. During meiosis (left panel) unpaired DNA is detected and induce the production of aberrant RNA (aRNA). aRNA is converted into dsRNA by RNA polymerase proteins (SAD-1, SAD-2 and SAD-3). dsRNA is cleaved by DCL-1 into sRNAs and then loaded on the Argonaute complex (SMS-2) to execute gene silencing. During the vegetative stage (right panel), both repetitive sequences (quelling) and ribosomal DNA loci (qiRNA) are involved in the production of aRNAs. aRNA is converted into dsRNA and processed by either Dcl-1 or Dcl-2. sRNAs are loaded on the Argonaute Qde-2 and transported to the the target sequence. Adapted from Chen et al. (2012).

Although epigenetic mechanisms for TE control have been extensively described in ascomycetes, particularly in the fungal model *N. crassa*, only a few studies have aimed to analyze DNA methylation (Binz et al. 1998; Foulongne-Oriol et al. 2013) and RNAi interference mechanism (Mueth et al. 2015) in basidiomycetes.

Beyond its biotechnological applications, the lignin-degrader *Pleurotus ostreatus* has gained relevance in genetic and genomics studies in recent years. The simple nature of its life cycle, the easy of cultivation under laboratory condition and the availability of an un-gapped telomere-to-telomere genome sequence makes *P. ostreatus* a good basidiomycete model. Recent comprehensive analyses carried out in our group completely characterized the landscape of transposable elements in two compatible monokaryotic strains of *Pleurotus ostreatus* (Castanera et al. 2016) (PC9 and PC15). This study uncovered the presence of 80 TE families, which encompass 2.5 and 6.2 % of the total genome sizes of PC9 and PC15, respectively. Also, most TEs were aggregated in 40 non-homologous clusters spread across the 12 chromosomes. Moreover, it was observed that genes carrying upstream or downstream TE insertions display lower transcription levels than average, especially when the genes are enclosed in TE-rich clusters. Regarding the potential RNAi activity of *P. ostreatus*, preliminary data suggests the presence of the core RNAi machinery, which has been detected *in silico* by screening for *Neurospora* MSUD and quelling orthologous proteins (Borgognone et al. 2017).

Using high-throughput sequencing, we describe the genome-wide epigenetic (DNA methylation and small RNAs) and transcriptional (mRNA) profiles of two compatible *P. ostreatus* monokaryons as well as dikaryons at different stages during fruitbody development. Our results support existing evidence of strain-specific DNA methylation and small RNAs production primarily involved in the repression of transposons activity. Also, we provide direct evidence that the TE-associated gene silencing effect previously described by Castanera *et al.* (Castanera et al. 2016) is a consequence of the extension of DNA methylation from surrounding transposons. Finally, the comparative analysis of the transcriptomes at different stages of the *P. ostreatus* life cycle provides a better understanding of the genes and functions that might be involved in the triggering of primordia formation and fruit body development.

Materials and Methods

Fungal strains and growth conditions

Here, we used four *Pleurotus ostreatus* strains: PC9 (Spanish Type Culture Collection accession CECT20312), PC15 (CECT20312), N001 (CECT20600) and N001-HyB. PC9 and PC15 are two compatible monokaryotic protoclones obtained by de-dikaryotization of the N001 commercial strain in 1999 (Larraya et al. 1999). N001-HyB is a dikaryotic sub-clone regenerated *ad hoc* by mating PC9 and PC15 in 2015, which contains the same genetic complement as N001. All strains were cultured in Erlenmeyer flasks containing 200 ml of Malt Extract (ME, 20 gr/l) in the dark, at 24°C under orbital shaking (125 rpm). After 6 days the cultures were homogenized using an Omni mixer and used as inoculum (15 ml) for Submerged Fermentation (SmF) and Solid-state Fermentation (SSF) cultures. Submerged fermentation (SmF) was carried out in Erlenmeyer flasks containing 135 ml of liquid Malt Extract medium and maintained in the dark for 7 days at 24°C. Solid-state fermentation was carried out in polycarbonate Magenta boxes containing 15 g of the total dry substrate (88% sawdust, 10% millet and 2% CaCO₃) and adjusted to an 80% water content. A total of six samples representing the main stages of the *P. ostreatus* life cycle were obtained for further analysis: i) vegetative mycelium in SmF (PC9 and PC15); ii) mycelium under fruiting induction in SSF (M_N001 and N001-HyB); iii) primordia (P_N001); and iv) and mature fruitbodies (F_N001) (Fig. 3). To induce fruiting conditions, completely colonized SSF cultures were maintained at 18 °C under a light/dark photoperiod of 12 h until fruit-body formation (~ 15 days). Three biological replicates per condition were sampled, pooled and frozen in liquid nitrogen. Pooled samples were ground in a sterile mortar in the presence of liquid nitrogen before nucleic acids extraction.

Construction and sequencing of whole genome bisulfite libraries

Global DNA methylation levels were estimated by performing sodium bisulfite treatment, based on the chemical conversion of unmethylated cytosines (Frommer et al. 1992), followed by high-throughput sequencing (BS-seq). To generate BS-seq libraries, genomic DNA (gDNA) from the fungal samples was extracted using an E.Z.N.A Fungal DNA Mini Kit (Omega Bio-Tek, Norcross, GA). After additional RNase A treatment (10 mg/ml for 60 min at 37°C), gDNA was purified using phenol: chloroform solution (3:1), precipitated overnight with ethanol (2:1) and the pellet resuspended in nuclease-free water. For additional purification after extraction, gDNA was treated with the Genomic DNA Clean & Concentrator Kit (Zymo Research, Irvine, CA).

The concentrations were quantified using a Qubit® 2.0 fluorometer (Life Technology, Carlsbad, CA) and total gDNA was fragmented with a Covaris S-2 ultrasonicator to obtain fragments spanning from 150 to 300 bp size range. Library preparation was performed using the Illumina TruSeq DNA Sample Prep (Feng et al. 2011) according to the manufacturer's instruction.

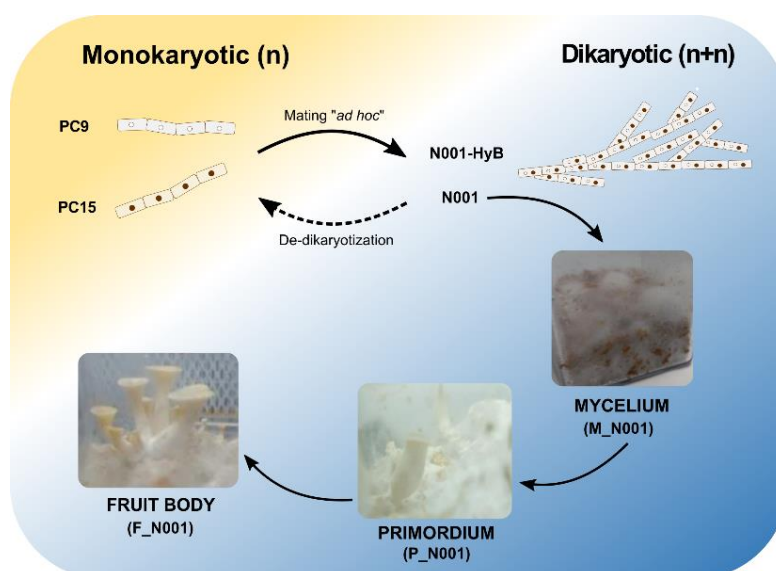


Figure 3. Summary of the samples used in this study, which represent the life cycle stages of *P. ostreatus*. PC15 and PC9 represent two monokaryotic strains that fuse to generate *ad hoc* the dikaryon N001-HyB. Given its inability to fructify, N001-HyB is examined exclusively at the mycelium stage. The N001 dikaryotic strain, bearing both PC15 and PC9 haploid nuclei and maintained under continuing subculturing during several years, is analyzed under different fruiting conditions (mycelium M_N001, primordium P_N001, and fruitbodies F_N001). N001-HyB and N001 harbor the same genetic complement although they show different fruiting ability.

Bisulfite conversion was carried out with the EpiTect Kit (QIAGEN) and two consecutive rounds of conversion for a total of 10 h were performed according to the following program: denaturation at 98°C for 2 min; 12 cycles of 98°C for 15 s, 60°C for 30 s, 72°C for 30 s; and final extension at 72°C for 5 min. All libraries obtained after bisulfite treatment were sequenced by an Illumina HiSeq2000 system (Illumina, San Diego, CA, USA) using 100 bp single-end reads.

Preparation of mRNA and small RNAs sequencing libraries

Total RNA isolation for mRNA (mRNA-seq) and small RNA (sRNA-seq) sequencing was performed using a Fungal RNA E.Z.N.A Kit (Omega Bio-Tek, Norcross, GA, USA) according to the manufacturer's guidelines. The integrity and quantity of RNA were validated by Bioanalyzer (version 2100) and Qubit® 2.0 fluorometer. mRNA-seq libraries were prepared using the TruSeq RNA Sample Prep Kit (Illumina) following the manufacturer's instruction. Total RNA was used for isolation of poly(A)-carrying mRNA molecules and synthesis of double-stranded cDNA before adapters ligations. For sRNA libraries, small RNAs molecules were resolved by electrophoresis on a 6% (w/v) polyacrylamide gel and the fraction corresponding to < 200 nt in length was eluted from the gel. Adapter-ligated molecules were reverse transcribed and enriched by PCR. The final libraries were quantified by real-time PCR in LightCycler 480 (Roche) and sequenced with an Illumina HiSeq 2000 system using 100 bp paired-end reads.

P. ostreatus reference genome

Reads from all libraries were mapped to the *P. ostreatus* reference genome (PC15 v2.0, www.genome.jgi.doe.gov/PleosPC15_2/PleosPC15_2.home.html). The PC15 v2.0 reference genome is completely assembled in twelve scaffolds (34.3 Mb genome size) corresponding to the eleven nuclear plus one mitochondrial chromosomes, as previously reported for the *P. ostreatus* karyotype(Larraya et al. 1999). In total, 12,330 genes have been annotated in this genome(Riley et al. 2014).

Bisulfite sequencing analysis

Raw reads from bisulfite treatment were checked for quality using FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc/) and aligned to the reference genome using the BS-Seeker2 pipeline(Guo et al. 2013). DNA methylation content was estimated at each cytosine level as a relative value of methylated cytosines (5mC) to unconverted (5mC) and converted (C→T) cytosines after bisulfite modification. Default parameters of the whole-genome bisulfite alignment pipeline were unchanged, except for using Bowtie2(Langmead and Salzberg 2012) as aligner and allowing unique alignments. For the following analyses, we only considered cytosines covered by at least four reads. Pairwise comparison to detect differentially methylated regions (DMR) was performed by using Metilene software(Jühling et al. 2016), setting 200 bp as the minimum window size, p -value < 0.05 and 50% minimum mean methylation difference for calling DMR.

mRNA-seq analysis

RNA-seq data were filtered for quality using FastQC and trimmed with Trimmomatic (Bolger et al. 2014) to remove sequences of adapters and low quality reads. The resulting reads were aligned to the reference genome assembly using STAR v2.3.1 (Dobin et al. 2013) restricted to single hit mapping using the following parameters: `--outReadsUnmapped Fastx --outFilterMismatchNoverLmax 0.04 --outFilterMultimapNmax 1`. Expression levels were quantified using python script `rpkmforgenes.py` (www.sandberg.cmb.ki.se/media/data/rnaseq/rpkmforgenes.py) to calculate values of reads per kilobase of transcript per million mapped reads (RPKM).

sRNA-seq analysis

Small RNA (sRNA) raw reads were first processed to trim the adapters sequence (Bolger et al. 2014), remove polyA tails and filter by size (17 to 30 nt) using Trimmomatic and Prinseq (Schmieder and Edwards 2011) software. Reads mapping to a custom database of ribosomal RNA (rRNA) and transfer RNA (tRNA) were also removed. The remaining sRNA reads were mapped to the reference genome using Butter 0.3.2, a variant of Bowtie optimized for small RNA included in the ShortStack package (Axtell 2013). This program uses iterative read assignment, and reads with multiple possible alignments were mapped to a single best location. The program was run using the parameters: `--mismatches 1 --bam2wig combined`. After mapping, sRNA reads were overlapped to features listed in TE and gene annotation datasets.

Differential genes expression analysis

Differentially expressed gene (DEG) analyses were performed using the EdgeR Bioconductor package (Robinson et al. 2010) and a dispersion parameter of 0.1. Based on gene counts, we considered a p -value < 0.01 and \log_2 (foldchange) > 3 as cutoffs for statistical significance. A heat map plot combined with hierarchical clustering was constructed with all DEG. Clusters of co-expressed genes were trimmed using the `define_clusters_by_cutting_tree.pl` tool of the TRINITY package (Grabherr et al. 2011) using 55 as the cutoff value. Gene ontology (GO) enrichment analysis of DEGs was performed using GOATOOLS software (Tang et al.). A false discovery rate (FDR) adjusted, Fisher's exact test p -value of 0.05 was used to identify significantly enriched ontologies.

Whole genome alignment

The draft genome assembly of *P. ostreatus* PC9 was aligned to the reference PC15 v2.0 genome to determine the overall nucleotide sequence identity and determine the unique regions of both genomes. PC9_V1.0 (35.6 Mb) is assembled in 572 scaffolds accounting for a total of 476 gaps covering 9.72 % of the assembly length. Unmasked assemblies were obtained from the MycoCosm browser (www.genome.jgi.doe.gov) and aligned using the NUCmer script from the MUMmer package 3.0 (Kurtz et al. 2004). NUCmer was selected to align highly conserved genome assemblies and delta-filter parameter applied to assign the optimal placement avoiding alignments to each repeat location.

General manipulation of sequencing data

Conversion between data formats and manipulation of BS-seq, RNAseq, and sRNAseq datasets was carried out with BEDTools (Quinlan and Hall 2010), SAMtools (Li et al. 2009), and custom Python scripts. All of the results generated after mapping to genomic features (methylation levels, RNAseq counts, and sRNAseq counts) have been used for subsequent analysis.

Results

DNA methylation and sRNA profiles during *P. ostreatus*' life cycle

Whole-genome bisulfite sequencing (BS-seq) was carried out in six representative samples of the *P. ostreatus* life cycle to investigate the profile of 5-cytosine DNA methylation. Sequencing of BS-seq libraries yielded an average of 60 ± 10 million total reads per sample. Reads were aligned to the PC15 v2.0 reference genome, obtaining coverages ranging from 77 to 162X (Supplementary Table S1 A). Cytosine methylation was clearly predominant in the CpG context in all samples ($5.0 \pm 1.8\%$ in CpG vs. $0.5 \pm 0.1\%$ in CHG and $0.7 \pm 0.1\%$ in CHH, Fig. 4A and Supplementary Table S1 B), and hereafter, we focus only on such context. The global methylation levels ranged from 2.4 to 6.7% of the total cytosines. *P. ostreatus* showed the lowest methylation levels in the monokaryotic stage, although the two strains tested showed differences between them (2.4% in PC15 vs. 4.4% in PC9 respectively). Interestingly, we found that the reconstructed N001-HyB dikaryotic strain (obtained by mating PC15 and PC9 protoclones) exhibited methylation levels substantially lower than the “natural” N001 dikaryotic strain (4.56 vs. 6.70%).

Also, N001-HyB displayed global methylation levels similar to the dominant PC9 profile, and this rate was even higher in N001 mycelium (M_N001), primordium (P_N001) and fruitbody (F_N001) samples, which displayed nearly identical levels (Fig. 4A). The distribution of 5-methylcytosines (5mC) was analyzed in two different genomic contexts: genes and transposable elements. The results uncovered impressive differences between these two features. Genes showed patterns of hypomethylation, with average methylation levels ranging from 2 to 5%. By contrast, TEs were heavily methylated, with levels ranging from 20 to 60% (Fig. 4B). We also observed sharp 5mC increments in the adjacent regions to both initial and terminal TE insertion sites, reaching the maximum methylation levels along the whole transposon body. By contrast, this trend was absent in protein-coding genes.

Regarding the difference between samples, the methylation levels of N001 were the highest and showed no variation during fruitbody development, neither within genes nor in TEs. Interestingly, methylation of N001-HyB across gene bodies and TEs was consistently lower than that of N001, and methylation levels of TEs showed a clear PC15-like profile (dominance of PC15 nuclei, low methylation). The genome-wide production of small RNAs was investigated in the same six samples by sRNA sequencing. We detected very important differences in the amount of sRNA reads per strain. After quality filtering, removal of rRNA and tRNA reads and size filtering, PC9 strain accounted for the highest number of sRNA reads (14.2 million) and PC15 for the lowest (1.5 million).

The dikaryotic strains N001-HyB and N001 displayed intermediate values, independently of the developmental stage (Supplementary Table S2). Most of the small RNAs originated from non-annotated genomic features, likely corresponding to heterochromatin regions (Fig. 4C). The amount of repeat-associated small interfering RNA (rasiRNA) varied between the six samples, ranging from 1.9% to 18.0 % of the total mapped reads, whereas the percentage of sRNAs mapping to genes ranged from 11.9 to 38.2%. The population of small RNAs was further characterized by analyzing the length distribution. A sharp peak was found at 21-22 nt in all strains and samples except PC15 (Fig. 4D and Supplementary Figure S1), which showed a mild descending profile from 17 to 30 nt.

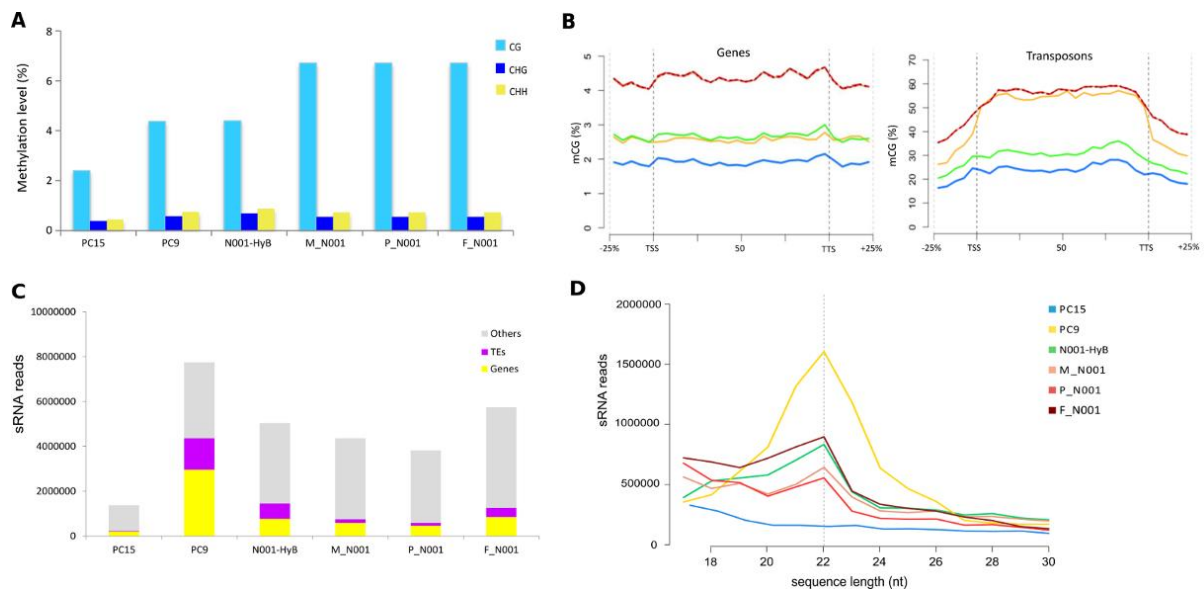


Figure 4. Global DNA methylation and small RNAs profiles within the six samples. (A) Average methylation levels of the CG, CHG and CHH contexts. **(B)** Metaplots showing DNA methylation across adjacent regions, genes and transposons body. **(C)** Distribution of sRNA total reads in transposons (TEs), genes (Genes) and heterochromatin regions (Others). **(D)** Line chart displaying the relative abundance of total sRNAs mapped reads ranging from 17 to 30-nt in length.

Genome-wide transcriptional profiles during *P. ostreatus* development

In addition to 5-cytosine DNA methylation and small RNA profiles, we performed mRNA-seq on the same six samples to analyze gene expression changes during *P. ostreatus* development. An average of 25.3 ± 2.12 million of uniquely aligned reads per sample (Supplementary Table S4) was used to calculate the transcriptional levels of all non-TE genes (genes overlapping with TEs were excluded) and perform differential expression analyses. We performed all-by-all sample comparisons and found that 3,596 out of the 11,828 genes were differentially expressed in the conditions analyzed (DGE, Table 1).

The biggest differences were found between the two monokaryotic strains PC9 and PC15 (1,393 DEGs), and the smallest between primordia (P_N001) and fruitbody (F_N001) samples of N001 (57 DEGs). To analyze the expression trends of DEGs under the six conditions, we performed a hierarchical clustering of the 3,596 genes. Two main clades were obtained in the sample dimension (horizontal axis, Fig. 5A): one formed by the PC15 and N001-HyB strains and the other by the PC9 and three N001 growing stages. In the gene dimension, we identified a total of eleven clusters of co-expression, consisting of variable numbers of genes (vertical axis, Fig. 5A). Three of the eleven clusters were further analyzed due to their relevance for the study of fruitbody development (Cluster 1 and Cluster 2) and the role of dominance in the expression profiles of the dikaryotic stage. Specifically, the expression of genes belonging to Cluster 1 (283 genes) increased in the N001 strain exclusively at primordium and fruitbody stages (Fig. 5A). Genes of Cluster 2 (113 genes) were highly expressed at all developmental stages of N001 (mycelium in fruiting induction, primordium, and fruitbody). Notably, the genes of this cluster had lower expression levels in the fruiting-impaired N001-HyB strain. As an opposite trend, the expression of genes in Cluster 11 (240 genes) was upregulated in PC15 and N001-HyB, whereas their expression decreased significantly in PC9 and N001, regardless of the growing stage.

Table 1. Differentially expressed genes in six strains at different growing stages of *P. ostreatus*.

	F_N001	N001-HyB	M_N001	PC15	PC9	P_N001
F_N001	0	1190	754	1380	1320	57
N001-HyB	1190	0	752	753	1054	915
M_N001	754	752	0	1322	992	375
PC15	1380	753	1322	0	1393	1243
PC9	1320	1054	992	1393	0	1054
P_N001	57	915	375	1243	1054	0

To better understand the link between transcription and DNA methylation in the different stages of *P. ostreatus* development, we represented the average methylation and expression levels of genes belonging to Clusters 1, 2, and 11 (Fig. 5B). Genes belonging to Cluster 1 and Cluster 2 showed a complete lack of methylation. Notably, genes included in Cluster 11 showed high methylation levels in PC9 and N001 (up to 60% average methylation level for N001 at mycelium, primordium, and fruitbody stages). Moreover, this hypermethylation pattern coincided with a low expression in the two strains. Further analysis of the genes in this cluster showed that 32% of the genes belonging to Cluster 11 were present inside the TE-rich regions defined by Castanera *et al* (Castanera *et al.* 2016). In contrast, this percentage decreased to 8 and 22% in Clusters 1 and 2, respectively (Supplementary Figure S2). Next, we retrieved the functional annotation of all DEG genes and performed Gene Ontology (GO) enrichment, focusing on genes of clusters 1, 2 and 11. This approach revealed molecular functions (MF), biological processes (BP) and cellular components (CC) significantly over-represented in the mentioned clusters (Supplementary Dataset Tables S4). Specifically, the most enriched functions of Cluster 1 (fruitbody induced genes) were monooxygenase activity (MF), lipid biosynthetic process (BP), and septin complex (CC). In Cluster 2 (genes upregulated under fruiting induction conditions and in the fruitbody), the most enriched functions were hydrolase activity (MF), carbohydrate metabolic process (BP), and lysosome (CC). Interestingly, other significantly enriched biological processes were fruiting body development and multicellular organism development (Table 2). Only one molecular function (1-aminocyclopropane-1-carboxylate oxidase activity) was found to be enriched in Cluster 11.

Table 2. Biological processes enriched in co-expression clusters upregulated in during fruitbody development.

	GO	GO description	GO / cluster	GO /genome	p-value*
Cluster 1	GO:0008610	lipid biosynthetic process	3/283	5/11828	0.001
	GO:0006334	nucleosome assembly	5/283	22/11828	0.001
	GO:0055114	oxidation-reduction process	18/283	349/11828	0.009
	GO:0007049	cell cycle	3/283	11/11828	0.009
	GO:0006108	malate metabolic process	2/283	6/11828	0.027
	GO:0006032	chitin catabolic process	2/283	8/11828	0.027
	GO:0006839	mitochondrial transport	1/283	1/11828	0.027
	GO:0000087	mitotic M phase	1/283	1/11828	0.027
	GO:0009186	deoxyribonucleoside diphosphate metabolic process	1/283	2/11828	0.045
Cluster 2	GO:0005975	carbohydrate metabolic process	8/113	199/11828	0.016
	GO:0006032	chitin catabolic process	2/113	8/11828	0.019
	GO:0030582	fruiting body development	1/113	1/11828	0.027
	GO:0019836	hemolysis by symbiont of host erythrocytes	1/113	1/11828	0.027
	GO:0006468	protein phosphorylation	7/113	287/11828	0.033
	GO:0006665	sphingolipid metabolic process	1/113	2/11828	0.033
	GO:0007275	multicellular organism development	1/113	2/11828	0.033
	GO:0007040	lysosome organization	1/113	2/11828	0.033
	GO:0009168	purine ribonucleoside monophosphate biosynthetic process	1/113	4/11828	0.040
	GO:0006508	proteolysis	6/113	272/11828	0.041

*FDR-corrected p value

These data also provided of evidence of the different transcriptional profiles between the *natural* N001 (M_N001) and the *regenerated* N001-HyB, harboring the same genetic complement (PC9 and PC15 nuclei). Looking closer, we found that a total of 752 genes were differentially transcribed between these two strains. In particular, the comparison displayed that 522 genes were upregulated in the N001-HyB and 230 in M_N001 strain (Supplementary Figure S3). To explore the functional annotation of these genes, we performed GO enrichment analysis between these two dikaryotic strains. Not surprisingly, the subset of enriched functions revealed that biological process (BP) involved in fruitbody development was exclusively represented in M_N001 strain (Supplementary Dataset Tables S5).

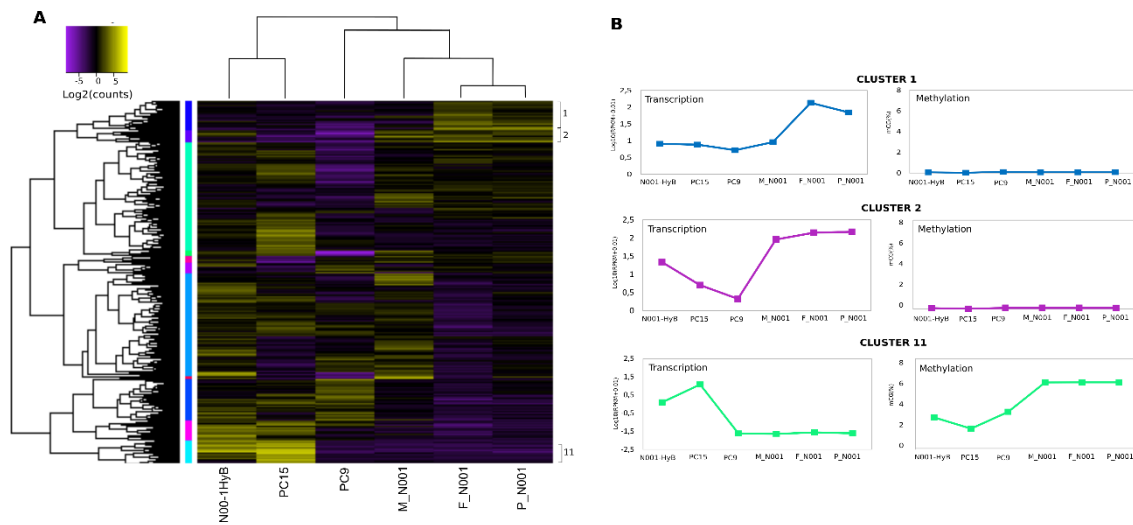


Figure 5. Hierarchical clustering of differentially expressed genes. (A) Heatmap plot illustrating differentially expressed genes (3,596 genes). Eleven main clusters are shown indicating the expression profile among samples (x-axis) and genes (y-axis). Gene expression levels from lower (violet) to higher (yellow) are indicated. Clusters 1, 2, and 11 are labeled for subsequent analysis. **(B)** Expression (left panels) and methylation (right panels) profiles of genes included in Clusters 1, 2 and 11.

Different nucleus-specific methylation profiles operate in the *P. ostreatus* dikaryotic stage

In the previous sections we have shown how the dikaryotic mycelium of the *natural* N001 strain (represented here by M_N001 sample) and the *ad hoc* generated (N001-HyB) sample displayed different methylation profiles under identical conditions, although they share the same genetic complement.

Thus, we considered the possibility of an equal contribution of each monokaryotic nucleus to the dikaryons. To test this hypothesis, we mapped BS-seq data to a new set of pseudo-genomes consisting of the nucleus-specific regions concatenated with the common regions sharing a similarity lower than 90% (one *unique* pseudo-genome for PC15 and another for PC9). To build this reference sequences set, we performed a whole genome alignment between PC15 and PC9 using the NUCmer software. The direct pairwise comparison revealed an average of 97.2 % similarity in the aligned regions, which spanned 87.3 % of PC15 assembly and 83.2% of PC9. Afterward, we determined the nucleus-specific methylation levels of M_N001 and N001-HyB using BS-Seeker2 and the *unique* reference genomes. We observed similar mappability rates between the M_N001 and N001-HyB strains when aligned to each reference genome.

Interestingly, when we looked at the global DNA methylation, we found that each nucleus contributed differentially in the two dikaryotic strains (Table 3). In fact, while genomic regions deriving from the PC9 nucleus exhibited comparable methylation levels in both the *natural* and *ad hoc* dikaryons (coefficient of variation of 13.6%), differences in regions associated with the PC15 nucleus were higher (CV of 37.7%). Specifically, methylation of the PC15 nuclei in N001-HyB was considerably lower than in M_N001 (19.22 vs. 35.12%), where the methylation levels of both nuclei in M_N001 were similar.

Table 3. Summary of BS-seq mapping to PC15 and PC9 *unique* pseudo-genomes.

Dikaryon	HyB		M_N001	
	PC15-unique	PC9-unique	PC15-unique	PC9-unique
Total raw reads	63,410,889	63,410,889	54,031,681	54,031,681
Mapped reads	4,181,804	2,108,525	3,259,529	1,937,413
Mappability (%)	6.6	3.3	6.0	3.6
Coverage (X)	83.9	72.8	74.3	72.2
mCG (%)	19.22	32.92	35.12	39.93

Given the nucleus-specific methylation profiles found in N001-HyB and M_N001, we sought to identify genomic regions showing significant differential methylation between these two strains. Using Metilene software, we detected a total of 439 differentially methylated regions (DMR), mainly extending for < 500 bp length size (Supplementary Figure S4 A). Among the resulting DMR, 426 were significantly hypermethylated in M_N001, whereas only 13 were found in N001-HyB (Supplementary Figure S4 B).

DNA methylation, small RNA production and transcriptome landscape in *P. ostreatus*

The overall distribution and levels of DNA methylation, mRNA and sRNAs production was analyzed along the twelve *P. ostreatus* chromosomes using the dikaryotic N001 strain (sample M_N001) as a reference. We noticed that DNA methylation and sRNAs production were tightly associated with TE-rich clusters spread in the genome, where the transcriptional activity displayed dramatic depletion (Fig. 6A and 6B). Moreover, we found that the levels of methylation and sRNA production varied between chromosomes. As an example, the mitochondrial genome (chromosome 12) was almost entirely targeted by methylation, whereas within the nuclear chromosomes the distribution of 5mc and sRNA varied regionally. In light of these data, we sought to analyze the correlation of methylation levels with mRNA and sRNA expression in

the whole genome. For this purpose, the genome was divided into 200 bp windows.

Windows were split into three groups according to their methylation levels (Group I: 0-20 %, Group II: 20-60% and Group III: > 60%), and plotted with their corresponding sRNA and mRNA transcriptional levels. As shown in Figure 6C, DNA methylation exhibited an inverse association with mRNA expression. Although the strongest repression was found in windows with more than 60% average methylation, the sharpest decrease in expression was found to occur when methylation exceeded 20%. Regarding DNA methylation and sRNAs expression, the two epigenetic marks were positively correlated (Fig. 6D). Moreover, we noticed that 95% of the production of sRNAs derived exclusively from 3.3% of the whole genome.

Repeat-associated DNA methylation, expression, and smallRNA profiles

A clearly opposite trend was observed in transcription and methylation levels between genes (high transcription with low methylation) and transposable elements (low transcription with high methylation) (Fig. 7A and 7B). Within the different TE orders, TIR, LTR, and DIRS were the most heavily methylated and showed the lowest expression. Helitron and LINE elements showed slightly lower methylation and higher expression values. This general trend was maintained in all samples and strains, although PC15 showed important differences in TE methylation and their associated transcriptional levels. Specifically, PC15 showed the lowest 5mC rate in all TE orders, which at the same time showed the highest expression ratios. We performed a deeper analysis accounting for the 80 TE families present in *P. ostreatus* and found that PC15 TE families were consistently hypomethylated compared to the other strains (Fig. 7C). Hierarchical clustering of TE methylation and expression grouped the six samples into the same two groups (Fig. 7C, Supplementary Figure S5) A): i) PC15 with N001-HyB and ii) PC9 with the three N001 growing stages.

The most invasive TE families (Gypsy_1, Gypsy_2, Gypsy_3, Gypsy_4 and Gypsy_5) were strongly methylated (27 to 57% on average) and transcriptionally repressed. Nevertheless, other less abundant families, such as TIR_1, showed higher methylation levels (34 to 63%) and a complete repression. Next, we analyzed the production of sRNAs by genes, TE orders, and families in the context of the previously described methylation and transcriptional levels.

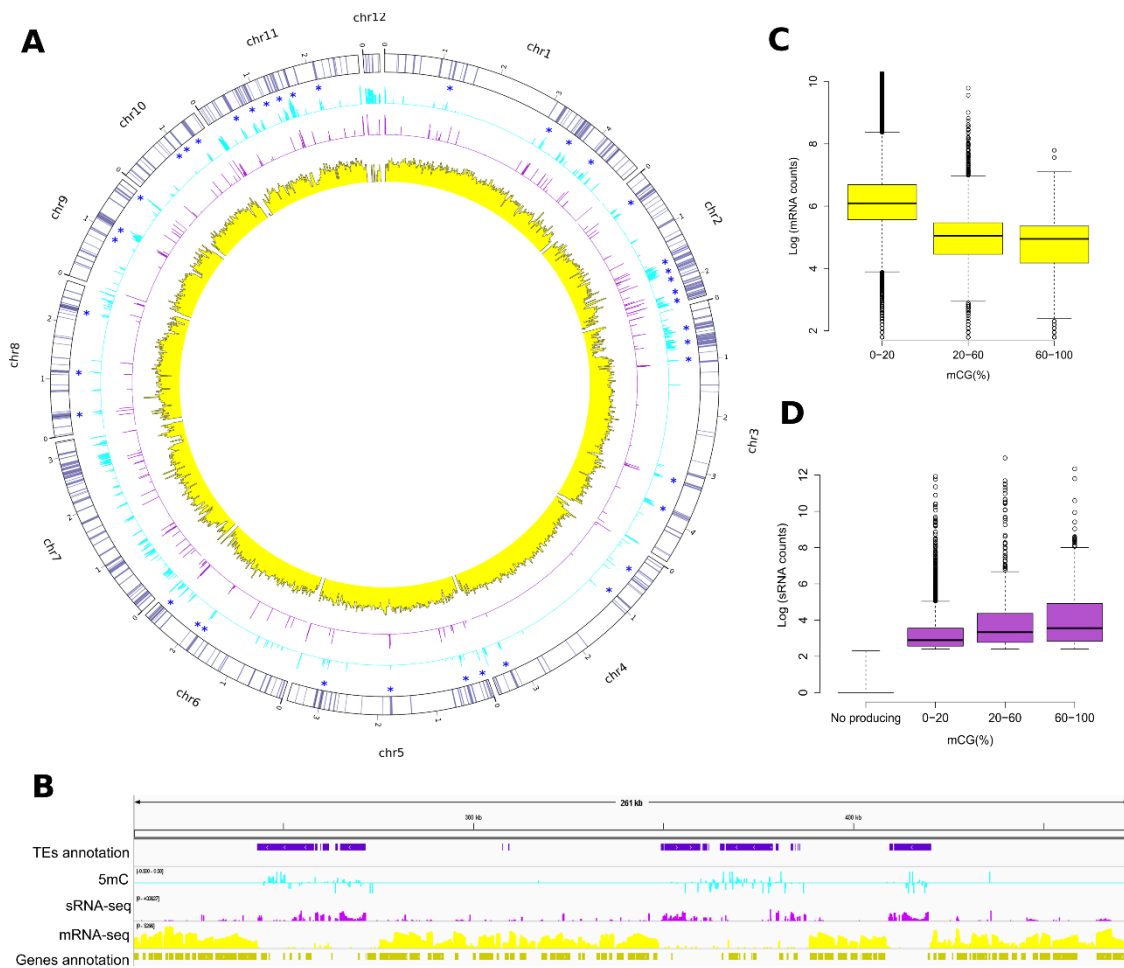


Figure 6. Global association of CpG methylation, mRNA and sRNAs expression in the M_N001 strain. Circular genome and data visualization with Circos plot were performed using a multi-mapping strategy for BS-seq, mRNA-seq and sRNA-seq reads (A). From inside to out: mRNA expression (yellow bars), small RNA production (purple bars) and DNA methylation (light blue bars). The outer track report transposable elements (dark violet bands). The presence of TE-rich clusters is indicated by blue asterisks. (B) Integrative genomics viewer (IGV) browser visualization of a 260kb region in scaffold 2 (location: 210,736–472,880 bp). Plotted from top to bottom are: transposable elements annotation, DNA methylation, sRNAs and mRNA transcription (logarithmic scale), and gene annotation. Boxplots showing the correlation of DNA methylation with mRNA (C) and sRNAs (D) transcription. Each boxplot represents the genome windows (200 bp) grouped for their methylation levels (0–20%, 20–60%, and 60–100%) and plotted for the corresponding term.

The amount of sRNAs per element (i.e., gene or TE) was higher in TEs than in genes (Fig. 7D), although it varied greatly depending on the strain and TE order. Similarly to that found in methylation levels, PC15 showed a much lower production of sRNAs than the other strains, with the only exception in LTR/Copia elements. sRNAs were abundantly produced by TIR transposons, reaching average levels up to ~7,000 RPM/copy in the PC9 strain. LTRs in the Gypsy superfamily and Helitrons also showed high sRNAs production, especially in PC9 and N001-HyB strains, and the remaining orders showed low amounts of mapped sRNAs. Considering the absolute numbers of sRNAs mapping to TEs (rasiRNAs), Gypsy and Copia retroelements and TIR elements were the main sources, although we observed great differences between families (Supplementary Figure S5 B).

We tested whether the production of rasiRNAs by the TE family was related to i) size (copy number) or ii) age (divergence between TE copies and family consensus). Despite detecting differences between strains, we found that the TE families containing more than ten copies tend to show the highest rasiRNAs expression (Supplementary Figure S6 A). This effect was especially relevant in proficient sRNAs-producing samples, such as PC9, where rasiRNA abundance (RPM/copy) was positively correlated with family copy number (Pearson correlation coefficient = 0.47, p -value = $1.853 \cdot 10^{-5}$) (Fig. 8A). To explore the relationship between rasiRNAs production and family age, we correlated rasiRNAs data to the average divergence rate, which increases linearly with family age (Supplementary Fig S6 B). Using PC9 as a reference, we found that the average family divergence and rasiRNAs expression were negatively correlated (Pearson correlation coefficient = -0.28, p -value = 0.01407) (Fig. 8B).

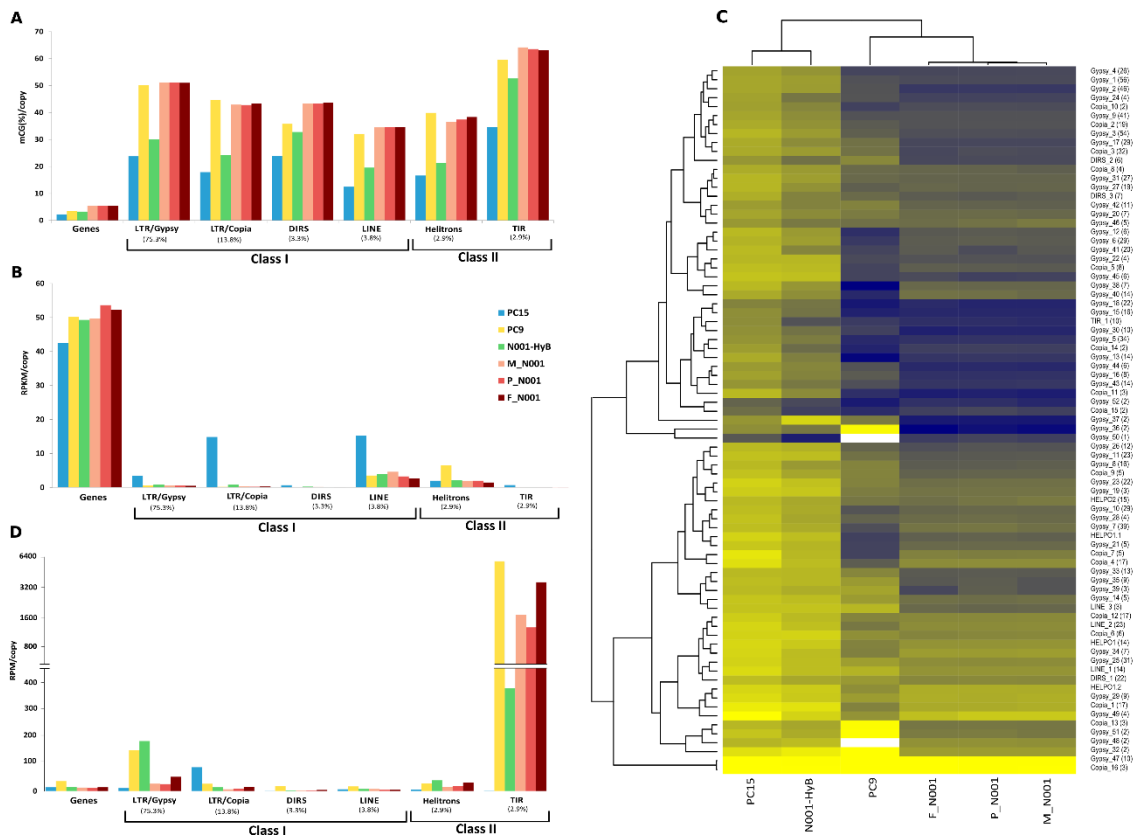


Figure 7. DNA methylation, transcriptome and sRNAs expression over annotated genomic features.

Histograms showing average methylation levels (**A**), mRNA transcription (**B**) and sRNA production (**C**) overlapping with genes and TE orders belonging to Class I and Class II. In (**C**), the y-axis is reported at different scale respectively linear in the lower part and logarithmic in the upper part. Percentage values below each order term indicate the relative occupancy in the total *P. ostreatus* TEs landscape. (**D**) Hierarchical clustering reporting DNA methylation levels within 80 TE families. Blue and yellow colors indicated high and low values respectively, expressed as mCG (%). An absence of TE families is displayed with white rectangles.

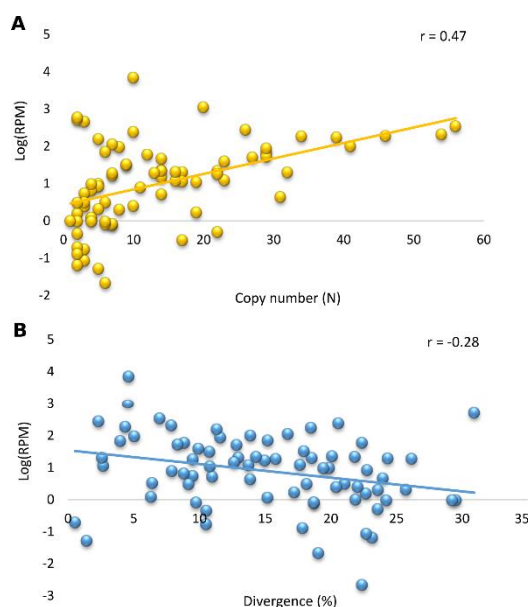


Figure 8. Association of transposons size and age with rasiRNA production.

Line charts showing the correlation of sRNAs production with copy number (A) and divergence rate (B) in 80 TE families in the PC9 strain. The coefficient of determinations is reported on the top right of each graph.

Role of epigenetic modifications on TE-mediated gene silencing

As discussed above, transposable elements represent the primary target of cytosine methylation in *P. ostreatus*. Results of our previous investigations demonstrated that TE insertions lead to a significant reduction of the expression levels of surrounding genes within 1 kb (Castanera et al. 2016). Also, we found that this TE-mediated gene silencing effect was stronger when genes were located inside any of the 40 TE clusters described in *P. ostreatus*. We explored the possibility that this phenomenon had an epigenetic explanation. In particular, we tested the hypothesis that methylation could spread outside TE boundaries reaching the surrounding genes and blocking their transcription. Therefore, we compared the methylation and expression levels of genes surrounded by transposons (Fig. 9) (within a window of 1 kb, either upstream or downstream, Group I, labeled as +TE) and genes not surrounded by TEs (Group II, labeled as Ctl). Additionally, to uncover the impact of TE clusters on this phenomenon, we split the first group of genes in two contexts: i) genes located inside a TE cluster (cluster) and ii) genes located outside TE clusters (isolated). As shown in Figures 9A and 9B using N001_HyB and M_N001 strains as a model, genes carrying a TE insertion displayed higher methylation levels than the control groups ($p < 0.05$).

In the case of genes in TE cluster (Fig. 9B), we found a bimodal distribution of methylation, with approximately half of the genes showing high 5mC levels (up to 40% in N001-HyB and 70% in M_N001). Regarding transcriptional profiles (Fig. 9C and 9D), genes surrounded by TEs showed lower expression than controls ($p < 0.05$), especially genes present inside a TE cluster (Fig. 9D). This phenomenon was present in the six strains, although with different intensities (Supplementary Figure S7 A-D and Supplementary Dataset Tables S6). Specifically, TE-mediated gene silencing was much stronger in the PC9 monokaryon than in PC15 monokaryon. Regarding the N001 dikaryotic strain, the three growing stages showed similar methylation and transcriptional profiles, being close to the PC9 profile. By contrast, N001-HyB exhibited intermediate methylation and a transcriptional state somewhere between the PC9 and PC15 parental strains. Furthermore, we analyzed the distribution of sRNA in genes grouped into the contexts described above. Notably, no significant differences, regarding sRNAs productions, were found among genes surrounded by a TE and the group control (Supplementary Figure S7 E and F).

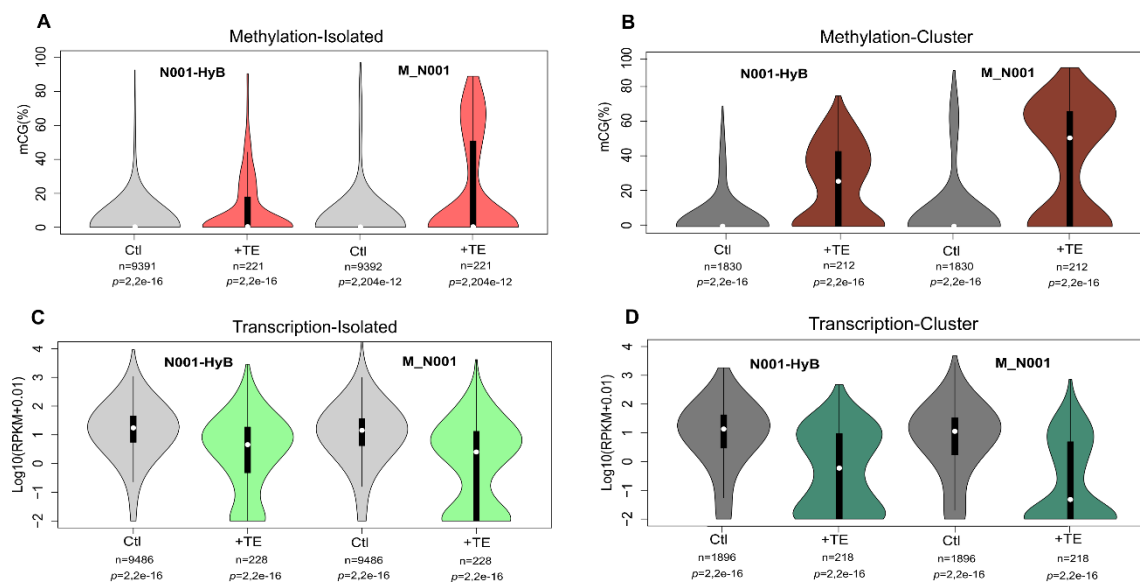


Figure 9. Influence of TE-associated methylation on nearby gene transcription. Violin plots showing methylation (A and B) and transcription (C and D) levels of genes in N001-HyB and M_N001 dikaryotic strains. Genes are classified as surrounded (+TE) or not surrounded (Ctl) by a transposon within a 1 kb-window (either upstream or downstream). Isolated (A and C) and Cluster (B and D) contexts indicate genes located outside and inside TE-rich clusters, respectively. White dot inside each plot represents the median. Below each violin plot, n represents the number of genes, and p represents the p -value of the Mann-Whitney-Wilcoxon test.

Next, we studied the distribution of methylation across the gene body and regions adjacent to the TSS (transcription start site) and TTS (transcription termination site). The aim was to understand if the methylation extension was spreading from the TE to the whole gene body or only to the promoter sequence in both contexts (TE cluster or isolated TE). We found that the methylation levels mildly decreased from the adjacent regions to the gene body. Nevertheless, gene body methylation of genes in both contexts was much higher than the control, reaching average values around 18% for genes isolated and with TE insertions and up to 30% for genes with TE insertions present in a TE cluster (Fig. 10A). Further analyses on the relationship between transcription and gene body methylation found that genes carrying TE insertions (at 1kb upstream or downstream) displayed only two activity states: i) methylated and silenced, or ii) unmethylated and transcriptionally active. As represented in Figures 10B and 10C, the vast majority of genes surrounded by a TE held the former activity state (methylated and silenced), whereas only a few copies displayed the latter (unmethylated and transcriptionally active), similarly to almost all genes included in the control group not enclosed by TE insertions (Fig. 10D).

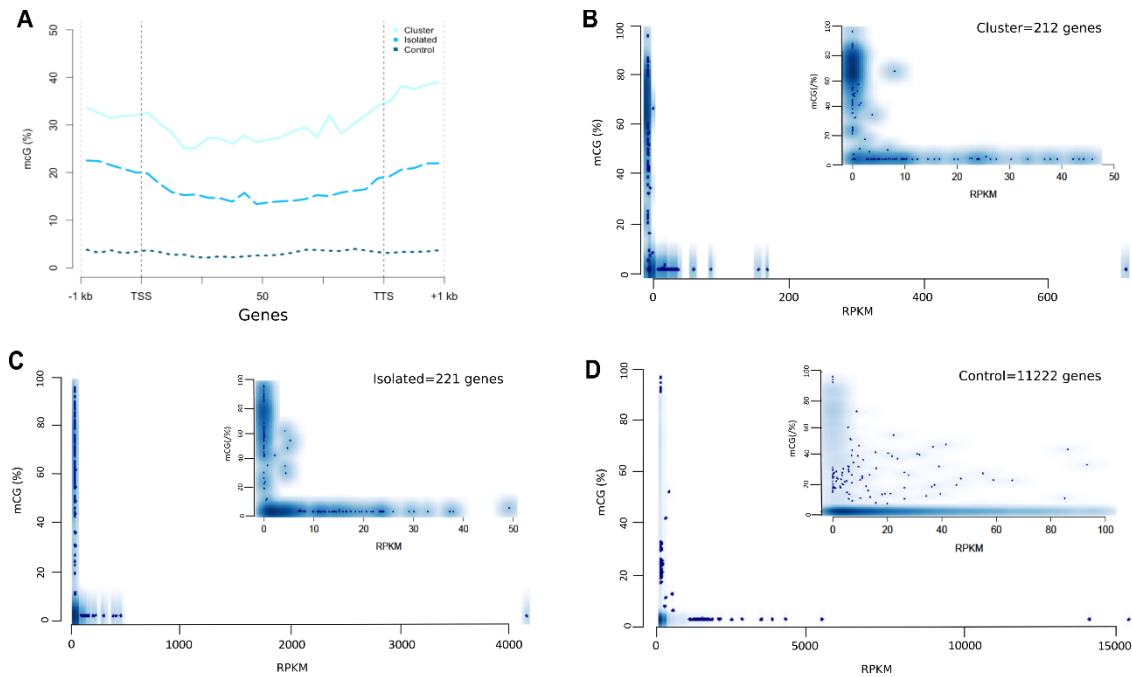


Figure 10. Association between methylation and transcription in the M_N001 strain.

(A) Metagene plot showing the average methylation levels across the genes body and adjacent region: i) surrounded by TEs at 1 kb upstream and downstream localized inside a TE-rich cluster (Cluster); ii) surrounded by TEs at 1 kb upstream and downstream localized outside a TE-rich cluster (Isolated); and iii) isolated and not surrounded by TEs (Control). Each of these groups is represented in light green, light blue and dark blue, respectively. Scatterplot distribution at different scale reporting the relationship between methylation levels and expression of genes classified in the cluster (B), isolated (C), and control (D) groups.

Discussion

Epigenetic factors contribute to *P. ostreatus* genome regulation

Previous findings have illustrated the presence and importance of epigenetic modifications in fungi (for review, see (Smith et al. 2012)). Most of the on fungal epigenetics comes from DNA methylation and RNAi studies performed in ascomycete fungi. However, few studies have report evidence of the presence of such mechanisms in basidiomycetes (Zemach et al. 2010; Foulongne-Oriol et al. 2013). Castanera *et al.* (Castanera et al. 2016) reported that Dim-2 DMTase, which is known to play an essential role in cytosine methylation in *N. crassa*, (Kouzminova and Selker 2001) was transcriptionally active in the genome of *P. ostreatus*. Our results confirm the presence of 5-cytosine methylation in *P. ostreatus*, predominantly in the CpG context and ranging from 2 to 6% in the six samples analyzed at different growing stages. These methylation levels are comparable to those described for other fungi so far (Zemach and Zilberman 2010; Jeon et al. 2015), except the highly methylated genome of *Tuber melanosporum* (Montanini et al. 2014), which displays up to 44% of cytosine methylation (Montanini et al. 2014). Previous studies in plants (Rabinowicz et al. 2003; Saze et al. 2012) and fungi (Zemach and Zilberman 2010) have demonstrated that the activity of transposons can be efficiently shut down by chromatin modifications linked to 5-cytosine methylation, as a defense mechanism aimed at controlling their expansion. In higher fungi, this trend has been reported in *T. melanosporum*, evidencing that methylation principally targets transposable elements (Montanini et al. 2014). Consistent with this observation, we found that *P. ostreatus* transposons are highly methylated whereas gene-coding regions are hypomethylated, independently of the strain and developmental stage (Fig. 4B). It is worth noting that the highest methylation levels are maintained across the whole transposon body, while they sharply decrease beyond the transposon borders.

A plausible explanation for this locally-restricted methylation might be the “mosaic methylation” described in invertebrates and plants. As reviewed by Suzuki et al (Suzuki and Bird 2008), this strategy attempts to delimit potential detrimental changes induced by methylation and prevent spurious transcriptional silencing across the genome. In addition to 5-cytosine methylation, we describe the epigenetic profiles associated with the production of small RNAs. In particular, we show that endogenous sRNAs produced by *P. ostreatus* are enriched in the 21–22 nt fraction (Fig. 4D) which could reflect the presence of transcriptional and post-transcriptional silencing mechanisms such as described in other eukaryotes (Slotkin and Martienssen 2007; Schwach et al. 2009).

In the basidiomycete *Cryptococcus neoformans*, experimental analyses showed the presence of 21–23 nt endogenous sRNAs involved in the control of mobile elements (Wang et al. 2010a).

Also, the plant-pathogen *Puccinia striiformis* produces 20–22 nt sRNA likely involved in post-transcriptional gene silencing (Mueth et al. 2015). *P. ostreatus* contains a highly conserved, transcriptionally active RNAi machinery composed of Argonaute, Dicer and RNA-dependent RNA polymerase proteins (Borgognone et al. 2017) (18 proteins sharing 93 to 100% similarity between PC9 and PC15 strains). Despite this, the PC15 strain shows a deficient production of sRNAs that is partially reverted in the dikaryotic stage. This apparent inefficacy of PC15 suggests that other non-canonical RNAi components active only in PC9 might be involved in the biogenesis of sRNAs in *P. ostreatus*.

Transposable elements are targeted by DNA methylation and RNAi machinery

TEs play a major role in genome stability and evolution (Kumekawa et al. 2001) but are also a source of mutation that can result in deleterious or lethal effects to the host. Thus, epigenetic mechanisms encoded by the host genome have been developed to silence their expression at the transcriptional and post-transcriptional levels (Slotkin and Martienssen 2007). Here, we found that DNA methylation and sRNA production were correlated with the presence of silent TEs, suggesting a role in their activity suppression. This association was even more striking in TE-rich regions. Specifically, TE clusters were shown to be transcriptionally silenced and highly methylated, similarly to that described in the basidiomycete *Laccaria bicolor* (Zemach and Zilberman 2010). In this sense, the accumulation of silent transposon knobs could correspond to heterochromatic regions matching to centromeric and pericentromeric zones, as described in plants (Gao et al. 2015). Regarding DNA methylation, both Class I and Class II transposons displayed higher methylation compared to genes. This observation supports the hypothesis that transcriptional silencing mediated by DNA methylation is primarily responsible for TE inactivation in *P. ostreatus*. Similarly to the ascomycete *T. melanosporum* (Montanini et al. 2014), we found that both LTR- retrotransposons and Class II transposons were abundantly methylated. It is interesting to note that the most methylated orders, TIR and LTR/Gypsy, also accounted for the larger portion of 22 nt rasiRNAs (Fig. 7D), especially TIR transposons belonging to the Mariner superfamily, the only cut-and-paste elements described in *P. ostreatus*. An example of this association has been reported in wheat, where most of the 21–22 nt sRNAs perfectly targeted within TIR regions of MITE elements, indicating that they are subjected to post-transcriptional control (Cantu et al. 2010).

An explanation for such an impressive amount of rasiRNA might reside in the hairpin RNA structure that elements from the TIR1 family can adopt due to the presence of Terminal Inverted Repeats (Supplementary Figure S8), a conformation that promotes the presence of dsRNA and triggers RNAi machinery. In *Caenorhabditis elegans*, it has been proposed that such conformations lead to rasiRNA production from Mariner transposons (Slotkin and Martienssen 2007). Previous studies have described that Class I transposons predominate in basidiomycetes genomes whereas Class II transposons show limited expansion. This is especially relevant in the agaricomycotina subphylum (Castanera et al. 2017), to which *P. ostreatus* belongs, and might be a consequence of the efficient replicative mechanism of Class I elements. According to our results, the under-representation of Class II transposons in the *P. ostreatus* genome might also be related to a stronger post-transcriptional inactivation in comparison to Class I elements. Thus, the lack of efficient post-transcriptional silencing of Class II TEs might account for the recently described Helitron transpositions in a subclone of PC15 strain (Borgognone et al. 2017).

In the case of LTR/Gypsy, families belonging to this superfamily account for the highest number of TE copies. Based on our results, the production of rasiRNAs positively correlates with family size and negatively correlates with mean family divergence (Fig. 8A). Thus, the production of sRNA molecules might reflect an attempt to limit the expansion of the youngest, most invasive TE families of the genome. This observation is reminiscent of co-suppression studies carried out in the basidiomycete *Cryptococcus neoformans*. In this species, mitotic-induced silencing (MIT), a quelling-like asexual mechanism operating in trans, lead to RNAi-mediated silencing of homologous sequences and repeated elements during vegetative growth (Wang et al. 2012). Also, a similar mechanism was shown to induce post-transcriptional inactivation of repetitive transgenes and transposons mediated by 21–23-nt sRNAs during the sexual reproduction (sex-induced silencing, SIS) (Wang et al. 2010a). Interestingly, the production of sRNAs in this fungal model increased according to the transgenes copy number with both mechanisms. Our finding suggests that *P. ostreatus* inactivates transposons at the transcriptional and post-transcriptional level, by epigenetic modifications associated with DNA methylation and 21–22 nt sRNAs production. The tight association between these two epigenetic marks and the presence of an active RNAi protein complex in the genome suggests that the selective suppression of TE activity could be mediated by RNA-directed DNA methylation. In this scenario, we speculate that sRNAs might be involved in the methylation of repetitive regions spread in the genome, such as transposons, guiding their silencing at the transcriptional level similarly to reported in plants (Qi et al. 2006), where non-canonical RdDM has been described to be mediated by 21–22 nt long sRNAs (Zhao et al. 2016).

Methylated-TEs induce transcriptional silencing of nearby genes

We have shown that epigenetic modifications in *P. ostreatus* act by downregulating transcriptional activity, particularly in repeated elements. Our data suggest that TE-associated transcriptional silencing of nearby genes occurs due to the extension of DNA methylation from TEs to the surrounding genes, which contain significantly higher methylation levels compared to controls. This is consistent with the results described in *Magnaporthe oryzae*, where genes with upstream, downstream or gene body methylation show lower expression than controls (Jeon et al. 2015). In our model, we observed that methylation of genes surrounded by TE insertions decreases from regions adjacent to the TSS and TTS, reaching the lowest levels in the gene body. This observation fits with the hypothesis of methylation being extended from the closest TE, repressing the transcriptional activity of neighbor genes as documented in plants (Arnaud et al. 2000) and animals (Jähner et al. 1985). Another intriguing point is to understand how much methylation is needed to impact the transcriptional activity. In light of our findings, we propose that low to intermediate methylation levels (< 20%) can prompt transcriptional silencing, although the strongest repression is found when methylation exceeds 60% (Fig. 6C and 10B). Another striking observation is that genes carrying upstream/downstream TE insertions have equal or even lower sRNA levels than control genes (Supplementary Figures 7E and F). This indicates that TE insertions do not produce post-transcriptional silencing of nearby genes, and suggests that TE-mediated gene silencing is exclusively promoted by DNA methylation.

P. ostreatus fruiting stage is associated with methylation-independent transcriptional reprogramming. The expression analyses carried out during *P. ostreatus* development yielded some insights into the transcriptional changes underlying fruitbody induction and development. The transition from vegetative mycelium to primordium stage is a complex process that requires the aggregation of cells into compact hyphal knots which later experience tissue differentiation. In *P. ostreatus*, fruiting is triggered (among other environmental conditions) by lowering temperature and introducing light-dark cycles (Arjona et al. 2009). In our study, these conditions lead to fruiting induction in N001, accompanied by the activation of a set of genes which were expressed at lower levels in monokaryons and also in N001-HyB (a strain unable to fruit), suggesting their putative role in the fruiting stage. According to the number of DEG genes, at most 10% of the *P. ostreatus* genes are potentially involved in the early induction (M_N001), while few genes are involved in changes from primordial to mature fruit bodies, similarly to that found in the basidiomycete *Coprinopsis cinerea* (Muraguchi et al. 2015).

Within this relatively small gene pool, the impressive overexpression of the pleurotolysin B (GO:0030582, > 200 fold increase in M_N001, P_N001 and F_N001 vs. any of the monokaryotic stages and N001_HyB), suggests its important role in the whole fruiting process. This observation is reinforced by the up-regulation of components involved in the interaction of pleurotolysin B with lipid rafts and cholesterol-rich domains in cell membranes, found in enriched GO terms of Cluster 2 (GO:0004348 = glucosylceramidase activity, GO:0047012 = sterol-4-alpha-carboxylate3-dehydrogenase activity and GO:0004767 = sphingomyelin phosphodiesterase activity) (Sepčić et al. 2004). In this sense, previous experimental work has shown that the expression of these hemolytic proteins is activated during the formation of primordia and young fruit bodies in *P. ostreatus* and *Agrocybe aegerita* (Berne et al. 2002; Vidic et al. 2005). Within the cluster of genes expressed during fruiting induction (Cluster 2), we also found enriched functions related to multicellular development that have been described as being involved in cell differentiation in the primitive metazoan belonging to *Giardia* genera (Wang et al. 2010b), although the most enriched biological function was related to carbohydrate metabolism. This overexpression probably reflects the fact that carbohydrates are retrieved from mycelium and abortive primordia before being degraded and provide nutrients to the growing fruiting bodies (Wösten and Wessels 2006). The expression profiles of different *P. ostreatus* growing stages suggest that glycoside hydrolase, polysaccharide lyase, and carbohydrate-binding activities are involved in this process, due to their presence in the co-expression cluster (Cluster 2), which shows a strong expression increase only in the N001 samples. Glycoside hydrolases were also found to be enriched in Cluster 1 (genes upregulated during fruit body development), along with other proteins, such as oxidative enzymes, hydrophobins involved in the aggregation of aerial hyphae, and lectins previously described to be involved in this process in basidiomycetes (Wösten and Wessels 2006; Muraguchi et al. 2015). This indicates that the core genes involved in the fruiting process are conserved across the Basidiomycota phylum. Interestingly, despite the differences in transcription of fruiting-associated genes, methylation levels were invariably low along the six samples, indicating that fruiting body formation is not subjected to epigenetic regulation.

References

- Arjona D, Aragón C, Aguilera JA, Ramírez L, Pisabarro AG (2009) Reproducible and controllable light induction of in vitro fruiting of the white-rot basidiomycete *Pleurotus ostreatus*. *Mycol Res* 113:552–558. doi: 10.1016/j.mycres.2008.12.006
- Arnaud P, Goubely C, Pélissier T, Deragon JM (2000) SINE retroposons can be used in vivo as nucleation centers for de novo methylation. *Mol Cell Biol* 20:3434–41.
- Axtell MJ (2013) ShortStack: comprehensive annotation and quantification of small RNA genes. *RNA* 19:740–51. doi: 10.1261/rna.035279.112
- Barry C, Faugeron G, Rossignol JL (1993) Methylation induced premeiotically in *Ascobolus*: coextension with DNA repeat lengths and effect on transcript elongation. *Proc Natl Acad Sci U S A* 90:4557–4561. doi: 10.1073/pnas.90.10.4557
- Berne S, Križaj I, Pohleven F, Turk T, Maček P, Sepčić K (2002) *Pleurotus* and *Agrocybe* hemolysins, new proteins hypothetically involved in fungal fruiting. *Biochim Biophys Acta - Gen Subj* 1570:153–159. doi: 10.1016/S0304-4165(02)00190-3
- Binz T, D'Mello N, Horgen PA (1998) A Comparison of DNA Methylation Levels in Selected Isolates of Higher Fungi. *Mycologia* 90:785. doi: 10.2307/3761319
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30:2114–2120. doi: 10.1093/bioinformatics/btu170
- Borgognone A, Castanera R, Muguerza E, Pisabarro AG, Ramírez L (2017) Somatic transposition and meiotically driven elimination of an active helitron family in *Pleurotus ostreatus*. *DNA Res dsw060*. doi: 10.1093/dnares/dsw060
- Cantu D, Vanzetti LS, Sumner A, Dubcovsky M, Matvienko M, Distelfeld A, Michelmore RW, Dubcovsky J (2010) Small RNAs, DNA methylation and transposable elements in wheat. *BMC Genomics* 11:408. doi: 10.1186/1471-2164-11-408
- Castanera R, Borgognone A, Pisabarro AG, Ramírez L (2017) Biology, dynamics, and applications of transposable elements in basidiomycete fungi. *Appl Microbiol Biotechnol* 101:1337–1350. doi: 10.1007/s00253-017-8097-8
- Castanera R, López-Varas L, Borgognone A, LaButti K, Lapidus A, Schmutz J, Grimwood J, Pérez G, Pisabarro AG, Grigoriev I V, Stajich JE, Ramírez L (2016) Transposable Elements versus the Fungal Genome: Impact on Whole-Genome Architecture and Transcriptional Profiles. *PLoS Genet* 12:e1006108. doi:10.1371/journal.pgen.1006108
- Chang S-S, Zhang Z, Liu Y (2012) RNA Interference Pathways in Fungi: Mechanisms and Functions. *Annu Rev Microbiol* 66:305–323. doi: 10.1146/annurev-micro-092611-150138
- Chernov A V, Vollmayr P, Walter J, Trautner TA (1997) Masc2, a C5-DNA-methyltransferase from *Ascobolus immersus* with similarity to methyltransferases of higher organisms. *Biol Chem* 378:1467–73.
- Dang Y, Zhang Z, Liu Y (2014) Small RNA-Mediated Gene Silencing in *Neurospora*. In: *Fungal RNA Biology*. Springer International Publishing, Cham, pp 269–289
- Dobin A, Davis CA, Schlesinger F, Drenkow J, Zaleski C, Jha S, Batut P, Chaisson M, Gingeras TR (2013) STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29:15–21. doi: 10.1093/bioinformatics/bts635
- Doolittle WF, Sapienza C (1980) Selfish genes, the phenotype paradigm and genome evolution. *Nature* 284:601–603. doi: 10.1038/284601a0
- Feng S, Rubbi L, Jacobsen SE, Pellegrini M (2011) Determining DNA Methylation Profiles Using Sequencing. In: *Methods in molecular biology* (Clifton, N.J.), pp 223–238
- Foulongne-Oriol M, Murat C, Castanera R, Ramírez L, Sonnenberg ASM (2013) Genome-wide survey of repetitive DNA elements in the button mushroom *Agaricus bisporus*. *Fungal Genet Biol* 55:6–21. doi: 10.1016/j.fgb.2013.04.003
- Freitag M, Williams RL, Kothe GO, Selker EU (2002) A cytosine methyltransferase homologue is essential for repeat-induced point mutation in *Neurospora crassa*. *Proc Natl Acad Sci* 99:8802–8807. doi: 10.1073/pnas.132212899
- Frommer M, McDonald LE, Millar DS, Collis CM, Watt F, Grigg GW, Molloy PL, Paul CL (1992) A genomic sequencing protocol that yields a positive display of 5-methylcytosine residues in individual DNA strands. *Proc Natl Acad Sci U S A* 89:1827–31. doi: 10.1073/PNAS.89.5.1827
- Fukui KN, Suzuki G, Lagudah ES, Rahman S, Appels R, Yamamoto M, Mukai Y (2001) Physical arrangement of retrotransposon-related repeats in centromeric regions of wheat. *Plant Cell Physiol* 42:189–96. doi: 10.1093/pccp/pce026
- Galagan JE, Selker EU (2004) RIP: The evolutionary cost of genome defense. *Trends Genet*. 20:417–423.
- Gao D, Jiang N, Wing RA, Jiang J, Jackson SA (2015) Transposons play an important role in the evolution and diversification of centromeres among closely related species. *Front Plant Sci* 6:216. doi: 10.3389/fpls.2015.00216
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, Chen Z, Mauceli E, Hacohen N, Gnirke A, Rhind N, di Palma F, Birren BW, Nusbaum C, Lindblad-Toh K, Friedman N, Regev A (2011) Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29:644–652. doi: 10.1038/nbt.1883
- Guo W, Fizev P, Yan W, Cokus S, Sun X, Zhang MQ, Chen PY, Pellegrini M (2013) BS-Seeker2: a versatile aligning pipeline for bisulfite sequencing data. *BMC Genomics* 14:774. doi: 10.1186/1471-2164-14-774
- Hammond TM, Spollen WG, Decker LM, Blake SM, Springer GK, Shiu PKT (2013) Identification of small RNAs associated with meiotic silencing by unpaired DNA. *Genetics* 194:279–84. doi: 10.1534/genetics.112.149138
- Horns F, Petit E, Yockteng R, Hood ME (2012) Patterns of Repeat-Induced Point Mutation in Transposable Elements of Basidiomycete Fungi. *Genome Biol Evol* 4:240–247. doi: 10.1093/gbe/evs005
- Jähner D, Haase K, Mulligan R, Jaenisch R (1985) Insertion of the bacterial gpt gene into the germ line of mice by retroviral infection. *Proc Natl Acad Sci U S A* 82:6927–31.
- Jeon J, Choi J, Lee GW, Park SY, Huh A, Dean RA, Lee YH (2015) Genome-wide profiling of DNA methylation provides insights into epigenetic regulation of fungal development in a plant pathogenic fungus, *Magnaporthe oryzae*. *Sci Rep* 5:8567.
- Johnson LJ, Giraud T, Anderson R, Hood ME (2010) The impact of genome defense on mobile elements in *Microbotryum*. *Genetica* 138:313–319. doi: 10.1007/s10709-009-9419-2
- Jühling F, Kretzmer H, Bernhart SH, Otto C, Stadler PF, Hoffmann S (2016) metilene: fast and sensitive calling of differentially methylated regions from bisulfite sequencing data. *Genome Res* 26:256–62. doi: 10.1101/gr.196394.115
- Klahre U, Crété P, Leuenberger SA, Iglesias VA, Meins F, Jr. (2002) High molecular weight RNAs and small interfering RNAs induce systemic posttranscriptional gene silencing in plants. *Proc Natl Acad Sci U S A* 99:11981–6. doi: 10.1073/pnas.182204199

- Kouzminova E, Selker EU (2001) Dim-2 encodes a DNA methyltransferase responsible for all known cytosine methylation in *Neurospora*. *EMBO J* 20:4309–4323. doi: 10.1093/emboj/20.15.4309
- Kumekawa N, Ohmido N, Fukui K, Ohtsubo E, Ohtsubo H (2001) A new gypsy-type retrotransposon, RIRE7: preferential insertion into the tandem repeat sequence TrsD in pericentromeric heterochromatin regions of rice chromosomes. *Mol Genet Genomics* 265:480–8.
- Kurtz S, Phillippy A, Delcher AL, Smoot M, Shumway M, Antonescu C, Salzberg SL (2004) Versatile and open software for comparing large genomes. *Genome Biol* 5:R12. doi: 10.1186/gb-2004-5-2-r12
- Langmead B, Salzberg SL (2012) Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9:357–359. doi: 10.1038/nmeth.1923
- Larraya LM, Pérez G, Peñas MM, Baars JJ, Mikosch TS, Pisabarro AG, Ramírez L (1999) Molecular karyotype of the white rot fungus *Pleurotus ostreatus*. *Appl Environ Microbiol* 65:3413–7.
- Law JA, Jacobsen SE (2010) Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nat Rev Genet* 11:204–220. doi: 10.1038/nrg2719
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup (2009) The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25:2078–2079. doi: 10.1093/bioinformatics/btp352
- Lonnig W-E, Saedler H (2002) Chromosome rearrangements and transposable elements. *Annu Rev Genet* 36:389–410. doi: 10.1146/annurev.genet.36.040202.092802
- Malagnac F, Silar P (2010) epigenetics of eukaryotic Microbes. 185–201. doi: 10.1016/B978-0-12-375709-8.00013-7
- Malagnac F, Wendel B, Goyon C, Faugeron G, Zickler D, Rossignol JL, Noyer-Weidner M, Vollmayr P, Trautner TA, Walter J (1997) A gene essential for de novo methylation and development in *Ascombolus* reveals a novel type of eukaryotic DNA methyltransferase structure. *Cell* 91:281–290. doi: 10.1016/S0092-8674(00)80410-9
- Malone CD, Hannon GJ (2009) Small RNAs as Guardians of the Genome. *Cell* 136:656–668. doi: 10.1016/j.cell.2009.01.045
- McGinnis W, Shermoen AW, Beckendorf SK (1983) A transposable element inserted just 5' to a *Drosophila* glue protein gene alters gene expression and chromatin structure. *Cell* 34:75–84.
- Mondo SJ, Dannebaum RO, Kuo RC, Louie KB, Bewick AJ, LaButti K, Haridas S, Kuo A, Salamov A, Ahrendt SR, Lau R, Bowen BP, Lipzen A, Sullivan W, Andreopoulos BB, Clum A, Lindquist E, Daum C, Northen TR, Kunde-Ramamoorthy G, Schmitz RJ, Gryganskiy A, Culley D, Magnuson J, James TY, O'Malley MA, Stajich JE, Spatafora JW, Visel A, Grigoriev I V (2017) Widespread adenine N6-methylation of active genes in fungi. *Nat Genet*. doi: 10.1038/ng.3859
- Montanini B, Chen P-Y, Morselli M, Jaroszewicz A, Lopez D, Martin F, Ottonello S, Pellegrini M (2014) Non-exhaustive DNA methylation-mediated transposon silencing in the black truffle genome, a complex fungal genome with massive repeat element content. *Genome Biol* 15:411. doi: 10.1186/s13059-014-0411-5
- Mueth NA, Ramachandran SR, Hulbert SH (2015) Small RNAs from the wheat stripe rust fungus (*Puccinia striiformis* f.sp. *tritici*). *BMC Genomics* 16:718. doi: 10.1186/s12864-015-1895-4
- Muraguchi H, Umezawa K, Niikura M, Yoshida M, Kozaki T, Ishii K, Sakai K, Shimizu M, Nakahori K, Sakamoto Y, Choi C, Ngan CY, Lindquist E, Lipzen A, Tritt A, Haridas S, Barry K, Grigoriev I V., Pukkila PJ (2015) Strand-specific RNA-seq analyses of fruiting body development in *Coprinopsis cinerea*. *PLoS One*. doi: 10.1371/journal.pone.0141586
- Qi Y, He X, Wang X-J, Kohany O, Jurka J, Hannon GJ (2006) Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. *Nature* 443:1008–1012. doi: 10.1038/nature05198
- Quinlan AR, Hall IM (2010) BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* 26:841–2. doi: 10.1093/bioinformatics/btq033
- Rabinowicz PD, Palmer LE, May BP, Hemann MT, Lowe SW, McCombie WR, Martienssen RA (2003) Genes and transposons are differentially methylated in plants, but not in mammals. *Genome Res* 13:2658–64. doi: 10.1101/gr.1784803
- Riley R, Salamov AA, Brown DW, Nagy LG, Floudas D, Held BW, Levasseur A, Lombard V, Morin E, Otilar R, Lindquist EA, Sun H, LaButti KM, Schmutz J, Jabbour D, Luo H, Baker SE, Pisabarro AG, Walton JD, Blanchette RA, Henrissat B, Martin F, Cullen D, Hibbett DS, Grigoriev I V. (2014) Extensive sampling of basidiomycete genomes demonstrates inadequacy of the white-rot/brown-rot paradigm for wood decay fungi. *Proc Natl Acad Sci* 111:9923–9928. doi: 10.1073/pnas.1400592111
- Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26:139–140. doi: 10.1093/bioinformatics/btp616
- Romano N, Macino G (1992) Quelling: transient inactivation of gene expression in *Neurospora crassa* by transformation with homologous sequences. *Mol Microbiol* 6:3343–53.
- Saze H, Tsugane K, Kanno T, Nishimura T (2012) DNA Methylation in Plants: Relationship to Small RNAs and Histone Modifications, and Functions in Transposon Inactivation. *Plant Cell Physiol* 53:766–784. doi: 10.1093/pcp/pcs008
- Schmieder R, Edwards R (2011) Quality control and preprocessing of metagenomic datasets. *Bioinformatics* 27:863–864. doi: 10.1093/bioinformatics/btr026
- Schwach F, Moxon S, Moulton V, Dalmay T (2009) Deciphering the diversity of small RNAs in plants: The long and short of it. *Briefings Funct Genomics Proteomics* 8:472–481. doi: 10.1093/bfgp/elp024
- Selker EU, Cambareri EB, Jensen BC, Haack KR (1987) Rearrangement of duplicated DNA in specialized cells of *Neurospora*. *Cell* 51:741–752. doi: 10.1016/0092-8674(87)90097-3
- Selker EU, Tountas NA, Cross SH, Margolin BS, Murphy JG, Bird AP, Freitag M (2003) The methylated component of the *Neurospora crassa* genome. *Nature* 422:893–897. doi: 10.1038/nature01564
- Sepčić K, Berne S, Rebolj K, Batista U, Plemenitaš A, Šentjurc M, Maček P (2004) Ostreolysin, a pore-forming protein from the oyster mushroom, interacts specifically with membrane cholesterol-rich lipid domains.
- Shiu PK, Raju NB, Zickler D, Metzberg RL (2001) Meiotic silencing by unpaired DNA. *Cell* 107:905–16.
- Slotkin RK, Martienssen R (2007) Transposable elements and the epigenetic regulation of the genome. *Nat Rev Genet* 8:272–85. doi: 10.1038/nrg2072

- Smith KM, Phatale PA, Bredeweg EL, Connolly LR, Pomraning KR, Freitag M, Smith KM, Phatale PA, Bredeweg EL, Connolly LR, Pomraning KR, Freitag M (2012) Epigenetics of Filamentous Fungi. In: Encyclopedia of Molecular Cell Biology and Molecular Medicine. Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim, Germany.
- Suzuki MM, Bird A (2008) DNA methylation landscapes: provocative insights from epigenomics. *Nat Rev Genet* 9:465–76. doi: 10.1038/nrg2341
- Tamaru H, Selker EU (2003) Synthesis of signals for de novo DNA methylation in *Neurospora crassa*. *Mol Cell Biol* 23:2379–94.
- Tang H, Klopfenstein D, Pedersen B, Flick P, Sato K, Ramirez F, Yunes J, Mungall C GOATOOLS: Tools for Gene Ontology. doi: 10.5281/ZENODO.31628
- Vidic I, Berne S, Drobne D, Macek P, Frangez R, Turk T, Strus J, Sepcic K (2005) Temporal and spatial expression of ostreolysin during development of the oyster mushroom (*Pleurotus ostreatus*). *Mycol Res* 109:377–82.
- Wang X, Hsueh YP, Li W, Floyd A, Skalsky R, Heitman J (2010a) Sex-induced silencing defends the genome of *Cryptococcus neoformans* via RNAi. *Genes Dev* 24:2566–2582. doi: 10.1101/gad.1970910
- Wang X, Wang P, Sun S, Darwiche S, Idnurm A, Heitman J (2012) Transgene Induced Co-Suppression during Vegetative Growth in *Cryptococcus neoformans*. *PLoS Genet* 8:e1002885. doi: 10.1371/journal.pgen.1002885
- Wang Y-T, Pan Y-J, Cho C-C, Lin B-C, Su L-H, Huang Y-C, Sun C-H (2010b) A novel pax-like protein involved in transcriptional activation of cyst wall protein genes in *Giardia lamblia*. *J Biol Chem* 285:32213–26. doi: 10.1074/jbc.M110.156620
- Wösten HAB, Wessels JGH (2006) The Emergence of Fruiting Bodies in Basidiomycetes. In: Growth, Differentiation and Sexuality. Springer-Verlag, Berlin/Heidelberg, pp 393–414
- Zemach A, McDaniel IE, Silva P, Zilberman D (2010) Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* (80-) 328:916–919.
- Zemach A, Zilberman D (2010) Evolution of eukaryotic DNA methylation and the pursuit of safer sex. *Curr. Biol.* 20.
- Zhao J-H, Fang Y-Y, Duan C-G, Fang R-X, Ding S-W, Guo H-S (2016) Genome-wide identification of endogenous RNA-directed DNA methylation loci associated with abundant 21-nucleotide siRNAs in Arabidopsis. *Sci Rep* 6:36247. doi: 10.1038/srep36247

Supplementary information

Table S1 A. Summary of sequencing and mapping data for bisulfite-treated genomic DNA.

Strain	Number of raw reads	Number of mapped reads	# mapped in genes	# mapped in TEs	Mapability (%)	Coverage (X)
PC15	74,147,810	52,795,065	34,025,302	2,221,508	71	162
PC9	71,619,635	36,666,135	25,592,661	866,729	51	116
N001-HyB	63,410,889	26,538,399	16,816,139	1,421,868	42	78
M_N001	54,031,681	25,576,193	16,854,486	1,120,511	47	77
P_N001	61,630,234	29,172,521	19,228,014	1,277,932	47	88
F_N001	56,845,420	26,903,465	17,733,758	1,178,927	47	81

Table S1 B. Methylation levels of CG, CHG and CHH sites in the bisulfite-treated whole genome. Only cytosines with reads coverage ≥ 5 were considered for methylation levels estimation.

Strain	CG (%)	CHG (%)	CHH (%)
PC15	2.41	0.38	0.44
PC9	4.47	0.57	0.75
N001-HyB	4.56	0.68	0.88
M_N001	6.70	0.54	0.71
P_N001	6.71	0.54	0.72
F_N001	6.71	0.54	0.72

Table S2. Summary of sRNA sequencing data

Strain	Number of filtered reads (*)	Number of mapped reads	# mapped in genes	# mapped in TEs	Mapability (%)
PC15	1,525,384	1,373,886	201,833	26,492	90
PC9	14,240,465	7,744,229	2,960,910	1,397,980	54
N001-HyB	6,250,120	5,040,559	767,405	683,771	81
M_N001	5,720,475	4,359,240	583,304	161,379	76
P_N001	5,526,035	3,820,398	454,219	127,045	69
F_N001	7,583,500	5,752,373	847,192	407,569	76

(*) smallRNA reads after FastQC filtering, size selection (17-30 nt) and rRNA/tRNA removing.

Table S3. Summary of mRNA sequencing data

Strain	Number of raw reads	Number of mapped reads	# mapped in genes	# mapped in TEs	Mapability (%)
PC15	28,133,367	25,845,344	20,826,249	388,919	92
PC9	24,889,639	23,461,512	20,501,132	49,360	94
N001-HyB	28,230,991	26,934,362	22,992,547	89,643	95
M_N001	29,802,305	28,452,470	24,837,917	52,834	95
P_N001	25,863,101	24,807,071	21,603,617	34,551	96
F_N001	24,186,757	22,834,511	19,946,654	27,507	94

Dataset Tables S4. GO enrichment terms (MF: molecular function, BP: biological process, CC: cellular component) in co-expression Clusters 1, 2 and 11.

Cluster 1

GO	NS	Name	ratio_in_study	ratio_in_pop	p_value	fdr
GO:0004497	MF	monooxygenase activity	14/283	170/12330	3.7e-05	0.0013338
GO:0020037	MF	heme binding	14/283	212/12330	0.000385	0.0036575
GO:0005506	MF	iron ion binding	13/283	195/12330	0.000566	0.004608857
GO:0016615	MF	malate dehydrogenase activity	2/283	3/12330	0.00155	0.008835
GO:0008825	MF	cyclopropane-fatty-acyl-phospholipid synthase activity	2/283	3/12330	0.00155	0.008835
GO:0005199	MF	structural constituent of cell wall	3/283	26/12330	0.0211	0.0273125
GO:0016491	MF	oxidoreductase activity	18/283	438/12330	0.0209	0.0273125
GO:0004556	MF	alpha-amylase activity	2/283	6/12330	0.00741	0.0273125
GO:0004601	MF	peroxidase activity	3/283	24/12330	0.017	0.0273125
GO:0043169	MF	cation binding	2/283	9/12330	0.017	0.0273125
GO:0018560	MF	protocatechuate 3,4-dioxygenase type II activity	1/283	1/12330	0.023	0.0273125
GO:0018562	MF	3,4-dihydroxyfluorene 4,4-alpha-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018563	MF	2,3-dihydroxy-ethylbenzene 1,2-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018564	MF	carbazole 1,9a-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018565	MF	dihydroxydibenzothiophene dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018566	MF	1,2-dihydroxynaphthalene-6-sulfonate 1,8a-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018567	MF	styrene dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018568	MF	3,4-dihydroxyphenanthrene dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018569	MF	hydroquinone 1,2-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018575	MF	chlorocatechol 1,2-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018574	MF	2,6-dichloro-p-hydroquinone 1,2-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018573	MF	2-aminophenol 1,6-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018572	MF	3,5-dichlorocatechol 1,2-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018571	MF	2,3-dihydroxy-p-cumate dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018570	MF	p-cumate 2,3-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0047070	MF	3-carboxyethylcatechol 2,3-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0047074	MF	4-hydroxycatechol 1,2-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018561	MF	2'-aminobiphenyl-2,3-diol 1,2-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0004309	MF	exopolyphosphatase activity	1/283	1/12330	0.023	0.0273125
GO:0019117	MF	dihydroxyfluorene dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0019114	MF	catechol dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018542	MF	2,3-dihydroxy DDT 1,2-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018559	MF	1,1-dichloro-2-(dihydroxy-4-chlorophenyl)-(4-chlorophenyl)ethene 1,2-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018558	MF	5,6-dihydroxy-3-methyl-2-oxo-1,2-dihydroquinoline dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018555	MF	phenanthrene dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018557	MF	1,2-dihydroxyfluorene 1,1-alpha-dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0018556	MF	2,2',3-trihydroxybiphenyl dioxygenase activity	1/283	1/12330	0.023	0.0273125
GO:0008843	MF	endochitinase activity	2/283	12/12330	0.0298	0.03266538
GO:0004568	MF	chitinase activity	2/283	12/12330	0.0298	0.03266538
GO:0030060	MF	L-malate dehydrogenase activity	1/283	2/12330	0.0454	0.0454
GO:0004310	MF	farnesyl-diphosphate farnesyltransferase activity	1/283	2/12330	0.0454	0.0454
GO:0004035	MF	alkaline phosphatase activity	1/283	2/12330	0.0454	0.0454
GO:0004736	MF	pyruvate carboxylase activity	1/283	2/12330	0.0454	0.0454
GO:0008610	BP	lipid biosynthetic process	3/283	5/12330	0.000116	0.0013338
GO:0006334	BP	nucleosome assembly	5/283	22/12330	0.000117	0.0013338
GO:0006118	BP	obsolete electron transport	18/283	349/12330	0.00144	0.008835
GO:0007049	BP	cell cycle	3/283	11/12330	0.00172	0.008912727
GO:0006108	BP	malate metabolic process	2/283	6/12330	0.00741	0.0273125
GO:0006032	BP	chitin catabolic process	2/283	8/12330	0.0134	0.0273125
GO:0006839	BP	mitochondrial transport	1/283	1/12330	0.023	0.0273125
GO:0000087	BP	mitotic M phase	1/283	1/12330	0.023	0.0273125
GO:0009186	BP	deoxyribonucleoside diphosphate metabolic process	1/283	2/12330	0.0454	0.0454
GO:0000786	CC	nucleosome	5/283	21/12330	9.25e-05	0.0013338
GO:0031105	CC	septin complex	3/283	4/12330	4.7e-05	0.0013338
GO:0005618	CC	cell wall	3/283	28/12330	0.0257	0.029298

Cluster 2

GO	NS	Name	ratio_in_study	ratio_in_pop	p_value	fdr
GO:0004553	MF	hydrolase activity, hydrolyzing O-glycosyl compounds	6/113	125/12330	0,001	0.0165
GO:0008236	MF	serine-type peptidase activity	3/113	26/12330	0.00167	0.01837
GO:0008843	MF	endochitinase activity	2/113	12/12330	0.00517	0.0246675
GO:0004568	MF	chitinase activity	2/113	12/12330	0.00517	0.0246675
GO:0004674	MF	protein serine/threonine kinase activity	6/113	177/12330	0.00573	0.0246675
GO:0004672	MF	protein kinase activity	7/113	236/12330	0.00598	0.0246675
GO:0003993	MF	acid phosphatase activity	2/113	14/12330	0.00705	0.02585
GO:0046936	MF	deoxyadenosine deaminase activity	1/113	2/12330	0.0182	0.03336667
GO:0004348	MF	glucosylceramidase activity	1/113	2/12330	0.0182	0.03336667
GO:0005199	MF	structural constituent of cell wall	2/113	26/12330	0.0234	0.038016
GO:0000140	MF	acylglycerone-phosphate reductase activity	1/113	3/12330	0.0272	0.038016
GO:0004000	MF	adenosine deaminase activity	1/113	3/12330	0.0272	0.038016
GO:0047012	MF	sterol-4-alpha-carboxylate 3-dehydrogenase (decarboxylating) activity	1/113	3/12330	0.0272	0.038016
GO:0004289	MF	obsolete subtilase activity	2/113	29/12330	0.0288	0.038016
GO:0030246	MF	carbohydrate binding	2/113	29/12330	0.0288	0.038016
GO:0005507	MF	copper ion binding	2/113	32/12330	0.0345	0.03982
GO:0019239	MF	deaminase activity	1/113	4/12330	0.0362	0.03982
GO:0016811	MF	hydrolase activity, acting on carbon-nitrogen bonds, in linear amides	1/113	4/12330	0.0362	0.03982
GO:0004767	MF	sphingomyelin phosphodiesterase activity	1/113	4/12330	0.0362	0.03982
GO:0004254	MF	obsolete acylaminoacyl-peptidase activity	1/113	5/12330	0,045	0,045
GO:0008549	MF	obsolete dynamin GTPase activity	1/113	5/12330	0,045	0,045
GO:0005975	BP	carbohydrate metabolic process	8/113	199/12330	0.000481	0.015873
GO:0006032	BP	chitin catabolic process	2/113	8/12330	0.00225	0.0185625
GO:0030582	BP	fruiting body development	1/113	1/12330	0.00916	0.02748
GO:0019836	BP	hemolysis by symbiont of host erythrocytes	1/113	1/12330	0.00916	0.02748
GO:0006468	BP	protein phosphorylation	7/113	287/12330	0.0164	0.03336667
GO:0006665	BP	sphingolipid metabolic process	1/113	2/12330	0.0182	0.03336667
GO:0007275	BP	multicellular organism development	1/113	2/12330	0.0182	0.03336667
GO:0007040	BP	lysosome organization	1/113	2/12330	0.0182	0.03336667
GO:0009168	BP	purine ribonucleoside monophosphate biosynthetic process	1/113	4/12330	0.0362	0.03982
GO:0006508	BP	proteolysis	6/113	272/12330	0.0389	0.04140968
GO:0005764	CC	lysosome	1/113	2/12330	0.0182	0.03336667
GO:0005618	CC	cell wall	2/113	28/12330	0.0269	0.038016

Cluster 11

GO	NS	Name	ratio_in_study	ratio_in_pop	p_value	fdr
GO:0009815	MF	1-aminocyclopropane-1-carboxylate oxidase activity	1/240	1/12330	0.0195	0,026

Dataset Tables S5. GO enrichment terms (MF: molecular function, BP: biological process, CC: cellular component) between N001-HyB and M_N001 DEGs.

N001-HyB

GO	NS	Name	ratio_in_study	ratio_in_pop	p_value
GO:0004099	MF	chitin deacetylase activity	4/522	4/11828	3.18e-06
GO:0004553	MF	hydrolase activity, hydrolyzing O-glycosyl compounds	15/522	125/11828	0.000254
GO:0016810	MF	hydrolase activity, acting on carbon-nitrogen bonds	6/522	32/11828	0.00198
GO:0030248	MF	cellulose binding	6/522	33/11828	0.00234
GO:0004497	MF	monooxygenase activity	16/522	170/11828	0.00292
GO:0004034	MF	aldose 1-epimerase activity	2/522	3/11828	0.00522
GO:0004671	MF	protein C-terminal S-isoprenylcysteine carboxyl O-methyltransferase activity	4/522	18/11828	0.00605
GO:0020037	MF	heme binding	17/522	212/11828	0.0141
GO:0004053	MF	arginase activity	2/522	5/11828	0.0164
GO:0050236	MF	pyridoxine:NADP 4-dehydrogenase activity	2/522	5/11828	0.0164
GO:0008171	MF	O-methyltransferase activity	4/522	25/11828	0.0198
GO:0005199	MF	structural constituent of cell wall	4/522	26/11828	0.0226
GO:0016813	MF	hydrolase activity, acting on carbon-nitrogen bonds, in linear amidines	2/522	6/11828	0.0224
GO:0005506	MF	iron ion binding	15/522	195/11828	0.0284
GO:0004659	MF	prenyltransferase activity	2/522	7/11828	0.0326
GO:0016853	MF	isomerase activity	2/522	8/11828	0.0423
GO:0009815	MF	1-aminocyclopropane-1-carboxylate oxidase activity	1/522	1/11828	0.0423
GO:0004508	MF	steroid 17-alpha-monooxygenase activity	1/522	1/11828	0.0423
GO:0004132	MF	dCMP deaminase activity	1/522	1/11828	0.0423
GO:0008200	MF	ion channel inhibitor activity	1/522	1/11828	0.0423
GO:0008925	MF	maltose O-acetyltransferase activity	1/522	1/11828	0.0423
GO:0004587	MF	ornithine-oxo-acid transaminase activity	1/522	1/11828	0.0423
GO:0008119	MF	thiopurine S-methyltransferase activity	1/522	1/11828	0.0423
GO:0018820	MF	cyanamide hydratase activity	1/522	1/11828	0.0423
GO:0016564	MF	obsolete transcription repressor activity	3/522	19/11828	0.0441
GO:0005975	BP	carbohydrate metabolic process	24/522	199/11828	3.62e-06
GO:0006481	BP	C-terminal protein methylation	4/522	18/11828	0.00605
GO:0006527	BP	arginine catabolic process	2/522	5/11828	0.0164
GO:0009405	BP	pathogenesis	1/522	1/11828	0.0423
GO:0006808	BP	regulation of nitrogen utilization	3/522	19/11828	0.0441
GO:0005576	CC	extracellular region	11/522	57/11828	2.21e-05
GO:0005618	CC	cell wall	4/522	28/11828	0.00676

M_N001

GO	NS	Name	ratio_in_study	ratio_in_pop	p_value
GO:0005215	MF	transporter activity	12/230	179/11828	0.000133
GO:0050162	MF	oxalate oxidase activity	2/230	3/11828	0.00103
GO:0016787	MF	hydrolase activity	9/230	140/11828	0.00124
GO:0008236	MF	serine-type peptidase activity	4/230	26/11828	0.00128
GO:0005199	MF	structural constituent of cell wall	4/230	26/11828	0.00128
GO:0003993	MF	acid phosphatase activity	3/230	14/11828	0.00212
GO:0004767	MF	sphingomyelin phosphodiesterase activity	2/230	4/11828	0.00203
GO:0016712	MF	oxidoreductase activity	5/230	61/11828	0.00551
GO:0004601	MF	peroxidase activity	3/230	24/11828	0.00971
GO:0020037	MF	heme binding	9/230	212/11828	0.00128
GO:0004126	MF	cytidine deaminase activity	1/230	1/11828	0.0187
GO:0004185	MF	serine-type carboxypeptidase activity	2/230	13/11828	0.0236
GO:0004029	MF	aldehyde dehydrogenase (NAD) activity	2/230	13/11828	0.0236
GO:0005506	MF	iron ion binding	8/230	195/11828	0.0299
GO:0004187	MF	obsolete carboxypeptidase D activity	1/230	2/11828	0.0377
GO:0005315	MF	inorganic phosphate transmembrane transporter activity	1/230	2/11828	0.0377
GO:0046936	MF	deoxyadenosine deaminase activity	1/230	2/11828	0.0377
GO:0030599	MF	pectinesterase activity	1/230	2/11828	0.0377
GO:0018667	MF	cyclohexanone monooxygenase activity	1/230	2/11828	0.0377
GO:0030414	MF	peptidase inhibitor activity	1/230	2/11828	0.0377
GO:0008704	MF	5-carboxymethyl-2-hydroxyumuconate delta-isomerase activity	1/230	2/11828	0.0377
GO:0004497	MF	monooxygenase activity	7/230	170/11828	0.0402
GO:0016788	MF	hydrolase activity, acting on ester bonds	2/230	19/11828	0.0481
GO:0006118	BP	obsolete electron transport	14/230	349/11828	0.00716
GO:0006979	BP	response to oxidative stress	3/230	23/11828	0.00861
GO:0006810	BP	transport	12/230	302/11828	0.0144
GO:0046087	BP	cytidine metabolic process	1/230	1/11828	0.0187
GO:0030582	BP	fruiting body development	1/230	1/11828	0.0187
GO:0019836	BP	hemolysis by symbiont of host erythrocytes	1/230	1/11828	0.0187
GO:0006508	BP	proteolysis	11/230	272/11828	0.0188
GO:0042545	BP	cell wall modification	1/230	2/11828	0.0377
GO:0006685	BP	sphingomyelin catabolic process	1/230	2/11828	0.0377
GO:0006308	BP	DNA catabolic process	1/230	2/11828	0.0377
GO:0005618	CC	cell wall	5/230	28/11828	0.00015
GO:0016021	CC	integral component of membrane	12/230	343/11828	0.0384

Dataset Tables S6. Tables reporting mean, median, number of elements and p -value based on Mann-Whitney_Wolcoxon test for each sample. The condition reported are : methylation isolated (A) and in cluster (B) genes, transcription in isolated (C) and cluster (D) genes and sRNA production in isolated (E) and cluster (F) genes

A) Methylation –Isolated

Strain	Condition	Mean	Median	n	p -value
PC15	Ctl	0.01688866	0.00244186	9394	2.2E-16
	+TE	0.09262156	0.00290398	221	
PC9	Ctl	0.02499081	0.00252874	9103	4.928E-06
	+TE	0.1290807	0.00283111	184	
N001-HyB	Ctl	0.0230203	0.00278788	9391	2.2E-16
	+TE	0.1047934	0.00379121	221	
M_N001	Ctl	0.03737641	0.00246988	9392	2.904E-12
	+TE	0.2112049	0.00312883	221	
P_N001	Ctl	0.03731911	0.00246951	9390	6.6E-13
	+TE	0.2128581	0.00297297	221	
F_N001	Ctl	0.03729391	0.00246269	9392	2.2E-16
	+TE	0.2119739	0.00316832	221	

B) Methylation – Cluster

Strain	Condition	Mean	Median	n	p -value
PC15	Ctl	0.02768087	0.00252255	1830	2.2E-16
	+TE	0.1546556	0.147158	212	
PC9	Ctl	0.04871959	0.0025731	1707	2.2E-16
	+TE	0.3348772	0.153902	177	
N001-HyB	Ctl	0.04802117	0.00284803	1830	2.2E-16
	+TE	0.2487055	0.2701695	212	
M_N001	Ctl	0.08051796	0.002573315	1830	2.2E-16
	+TE	0.3904125	0.5283175	212	
P_N001	Ctl	0.08007999	0.0026087	1830	2.2E-16
	+TE	0.389367	0.526142	212	
F_N001	Ctl	0.08044475	0.00256757	1829	2.2E-16
	+TE	0.3916647	0.530193	212	

C) Transcription (mRNA) – Isolated

Strain	Condition	Mean	Median	n	p-value
PC15	Ctl	1.195509	1.330335	9486	4.417E-09
	+TE	0.6893838	0.9684209	228	
PC9	Ctl	1.004587	1.244582	9486	2.2E-16
	+TE	-0.05139848	0.1749202	228	
N001-HyB	Ctl	1.106869	1.240322	9486	2.2E-16
	+TE	0.4051247	0.6565275	228	
M_N001	Ctl	0.9833254	1.157396	9486	2.2E-16
	+TE	0.00353734	0.4018002	228	
P_N001	Ctl	1.019068	1.18716	9486	2.2E-16
	+TE	0.01889263	0.4039031	228	
F_N001	Ctl	1.001162	1.149971	9486	2.2E-16
	+TE	0.05009396	0.448767	228	

D) Transcription (mRNA) – Cluster

Strain	Condition	Mean	Median	n	p-value
PC15	Ctl	1.040034	1.27149	1896	2.2E-16
	+TE	0.1807578	0.4250232	218	
PC9	Ctl	0.7216242	1.079737	1896	2.2E-16
	+TE	-0.6113786	-1.302342	218	
N001-HyB	Ctl	0.8760176	1.132283	1896	2.2E-16
	+TE	-0.2989868	-0.2281	218	
M_N001	Ctl	0.7157091	1.051546	1896	2.2E-16
	+TE	-0.6905013	-1.319322	218	
P_N001	Ctl	0.7388166	1.045159	1896	2.2E-16
	+TE	-0.6651709	-1.466822	218	
F_N001	Ctl	0.7289426	1.024272	1896	2.2E-16
	+TE	-0.6202223	-1.274247	218	

E) sRNAs production – Isolated

Strain	Condition	Mean	Median	n	p-value
PC15	Ctl	0.1648865	0.4655984	9486	0.0002911
	+TE	-0.164538	0.1660525	228	
PC9	Ctl	0.1174258	0.2275446	9486	2.52E-02
	+TE	-0.2074446	-0.03910185	228	
N001-HyB	Ctl	0.004515428	0.1457354	9486	9.93E-02
	+TE	-0.2212737	-0.09498019	228	
M_N001	Ctl	-0.0370574	0.0633292	9486	6.67E-03
	+TE	-0.3198576	-0.1560243	228	
P_N001	Ctl	-0.08358385	0.02407964	9486	0.04312
	+TE	-0.2007154	-0.09949168	228	
F_N001	Ctl	0.01338652	0.08880531	9486	0.3078
	+TE	-0.04735829	0.02244809	228	

F) sRNAs production – Cluster

Strain	Condition	Mean	Median	n	p-value
PC15	Ctl	0.05881305	0.3411549	1896	4.873E-09
	+TE	-0.6655343	-0.1320246	218	
PC9	Ctl	0.04156783	0.1929968	1896	0.000825
	+TE	-0.09034279	-0.1052573	218	
N001-HyB	Ctl	-0.0511671	0.1457354	1896	3.11e-05
	+TE	-0.3392949	-0.2181211	218	
M_N001	Ctl	-0.1140535	0.0633292	1896	0.02023
	+TE	-0.3650636	-0.3290165	218	
P_N001	Ctl	-0.06786137	0.02407964	1896	0.02122
	+TE	-0.1955818	-0.09949168	218	
F_N001	Ctl	0.03281722	0.1463546	1896	0.6506
	+TE	0.1880618	0.08880531	218	

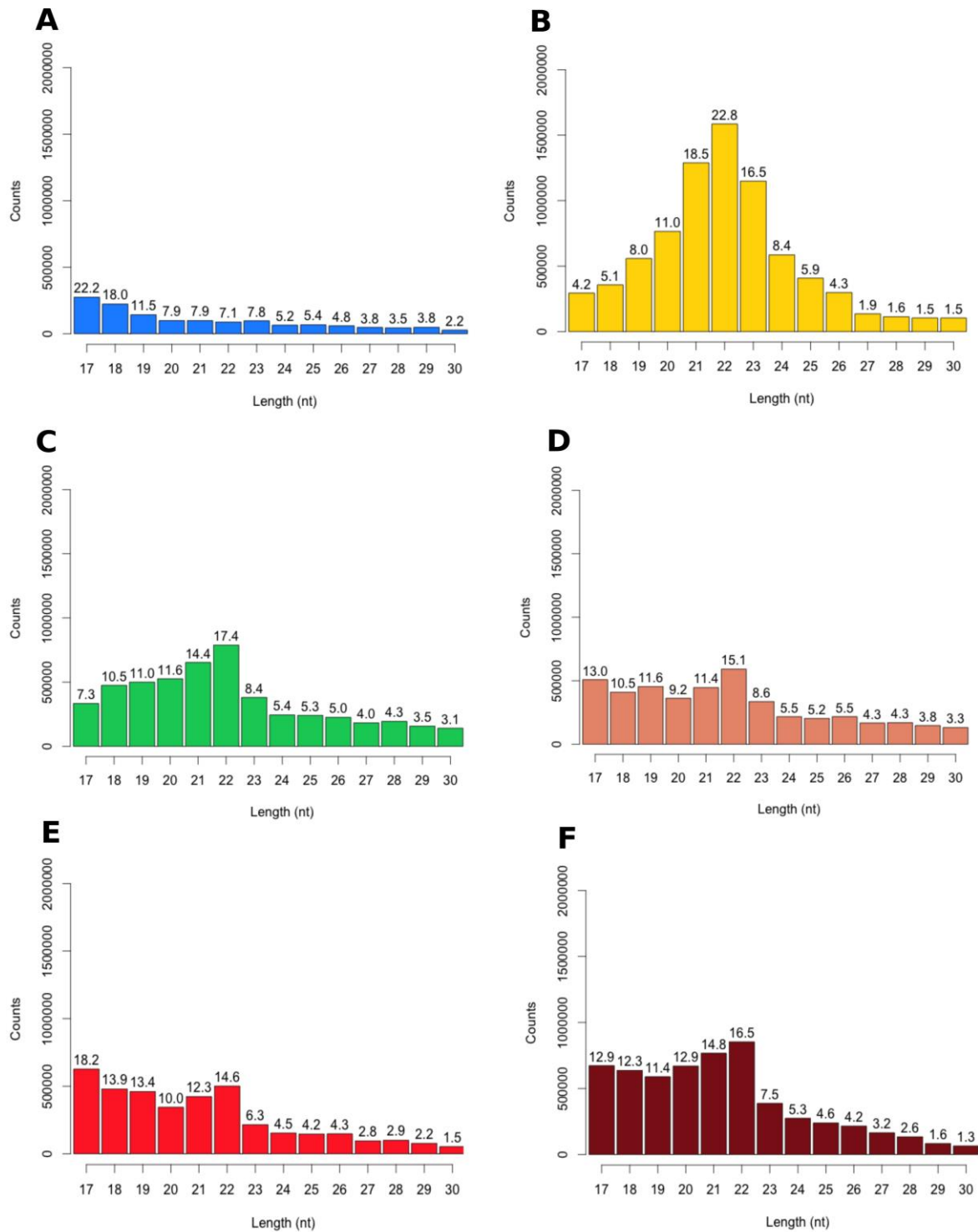


Figure S1. Histogram chart displaying the length distribution of sRNA total reads in PC15 (A), PC9 (B), N001-HyB (C), M_N001 (D), P_N001 (E) and F_N001 (F), Labels on the top of bars indicate the frequency (%) of each sRNAs class.

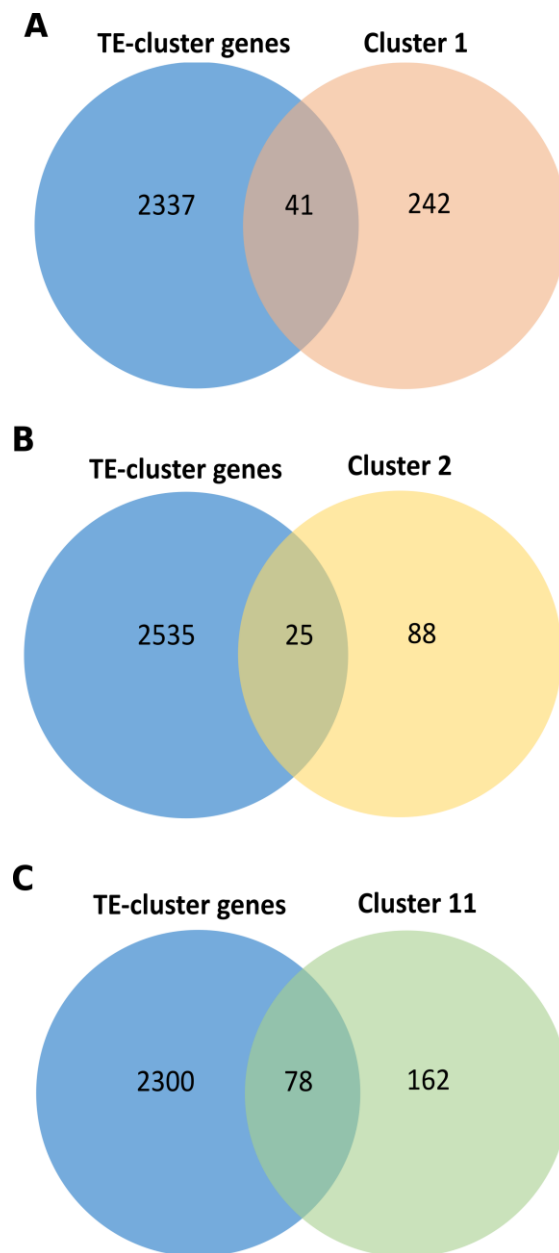


Figure S2. Venn diagrams summarizing the number of genes enclosed in TE clusters which overlap with genes in cluster 1 (A), cluster 2 (B) and cluster 11 (C).

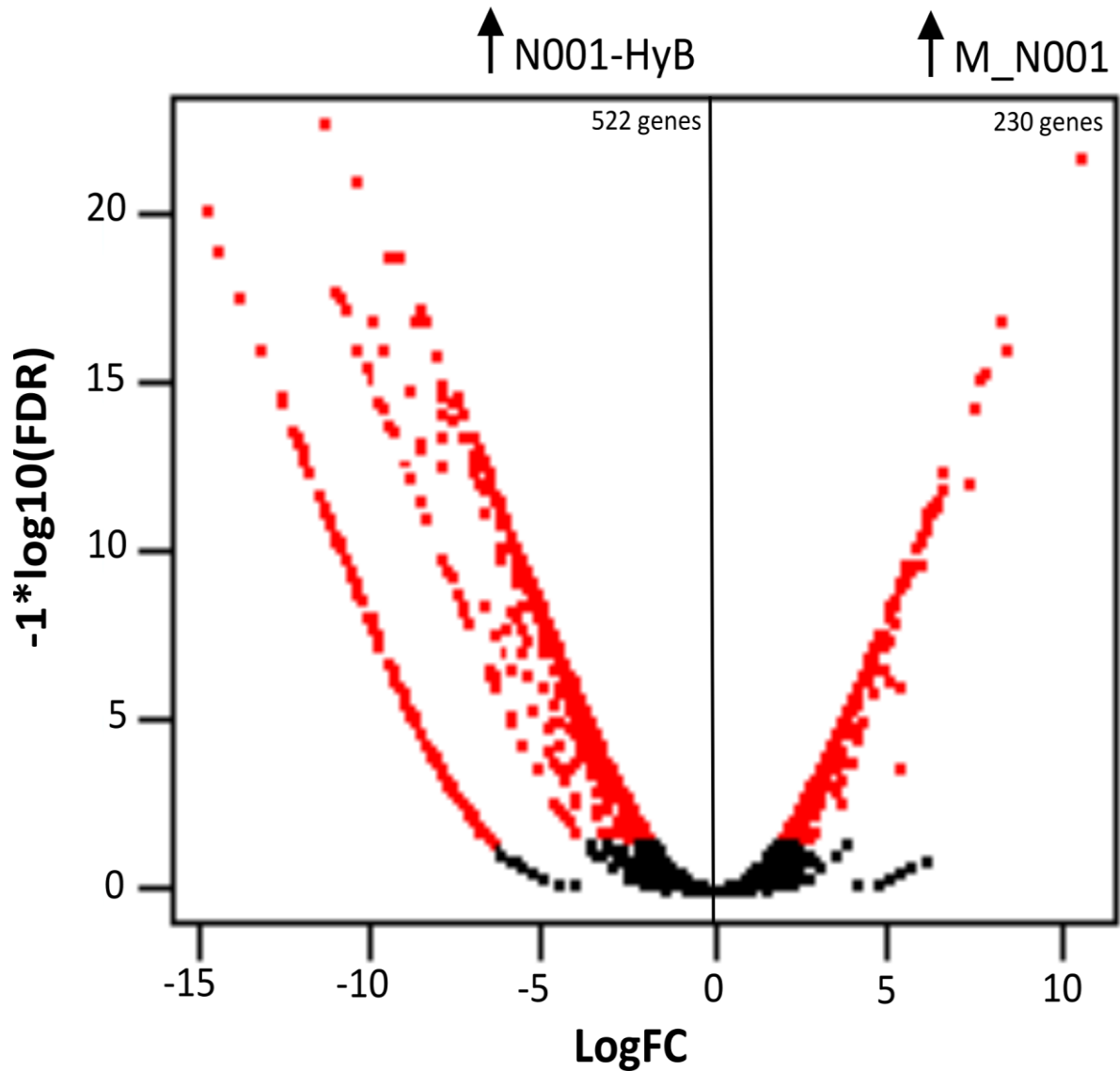


Figure S3. Volcano plot illustrating differentially expressed genes between N001-HyB and M_N001 strains. False discovery rate (\log_{10}) is reported in y-axis and the difference in expression is plotted in x-axis using ± 3 fold change as threshold of differential expression.

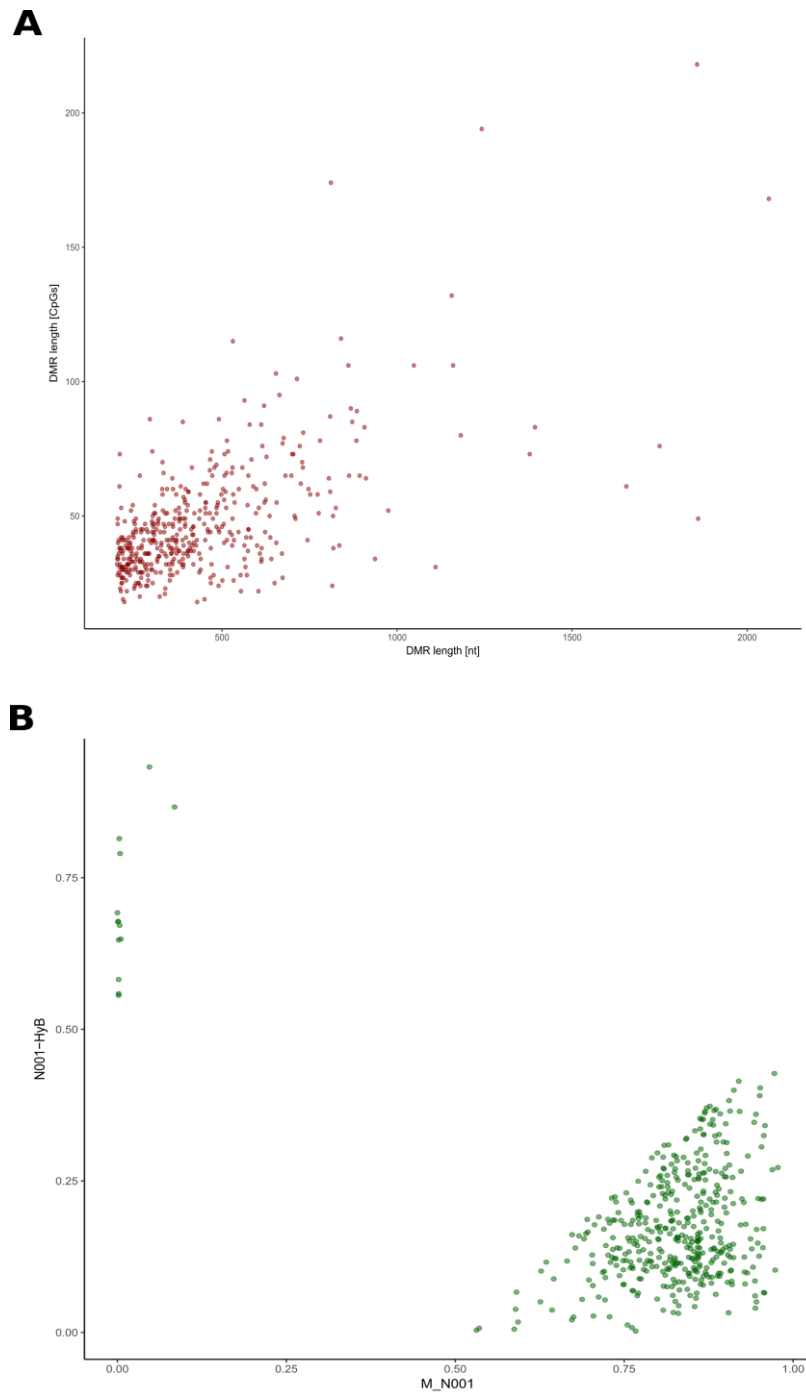


Figure S4. Distribution plot showing differentially methylated regions (DMR) between N001-HyB and M_N001 strains (**A**) and distribution of total DMR length (**B**).

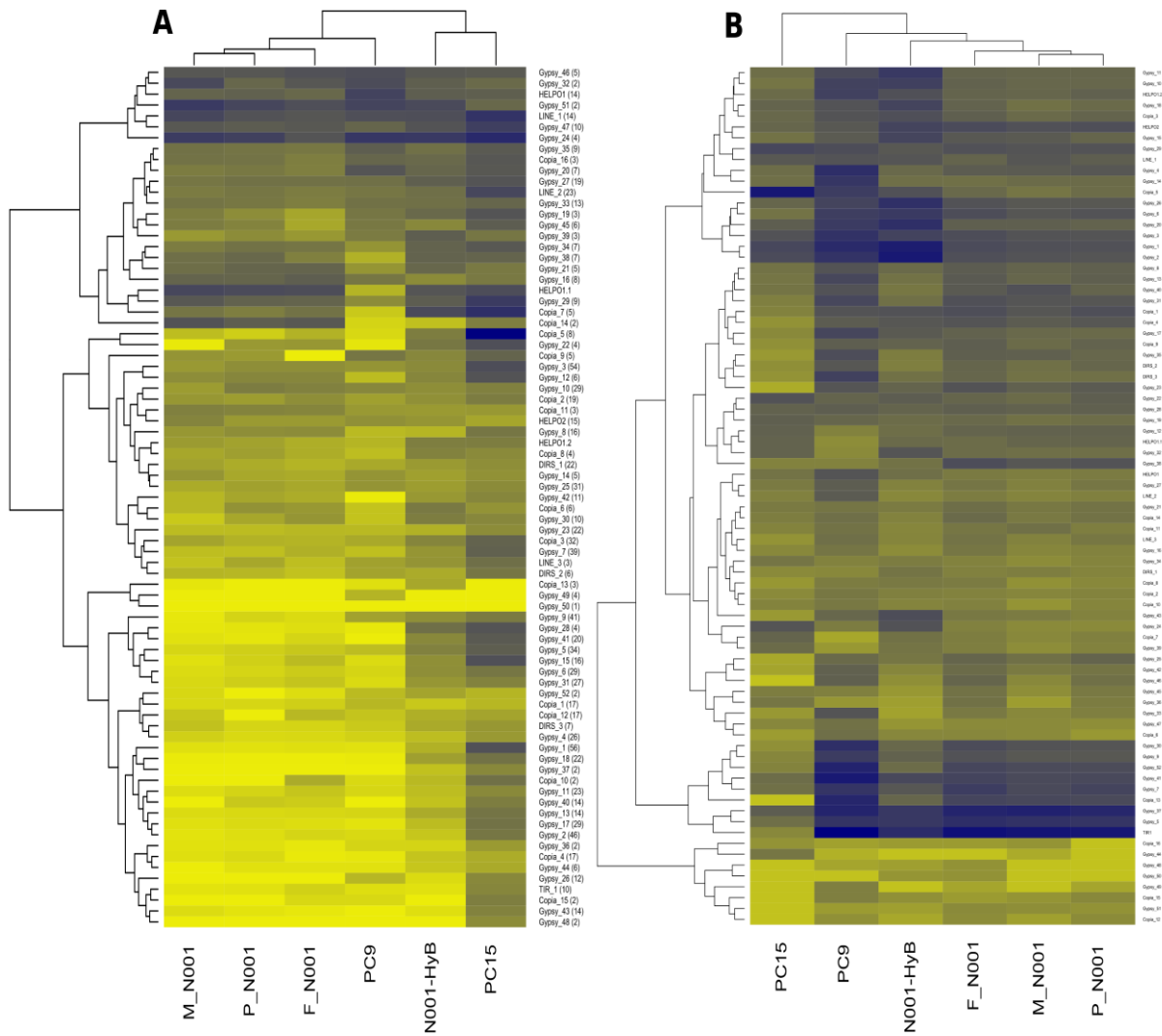


Figure S5. Heatmaps showing mRNA transcription (**A**) and sRNA production (**B**) in 80 transposon families. High (blue) and low (yellow) transcriptional levels are expressed in Log (RPKM+0.01) for mRNA and Log (RPM+0.01) for sRNA.

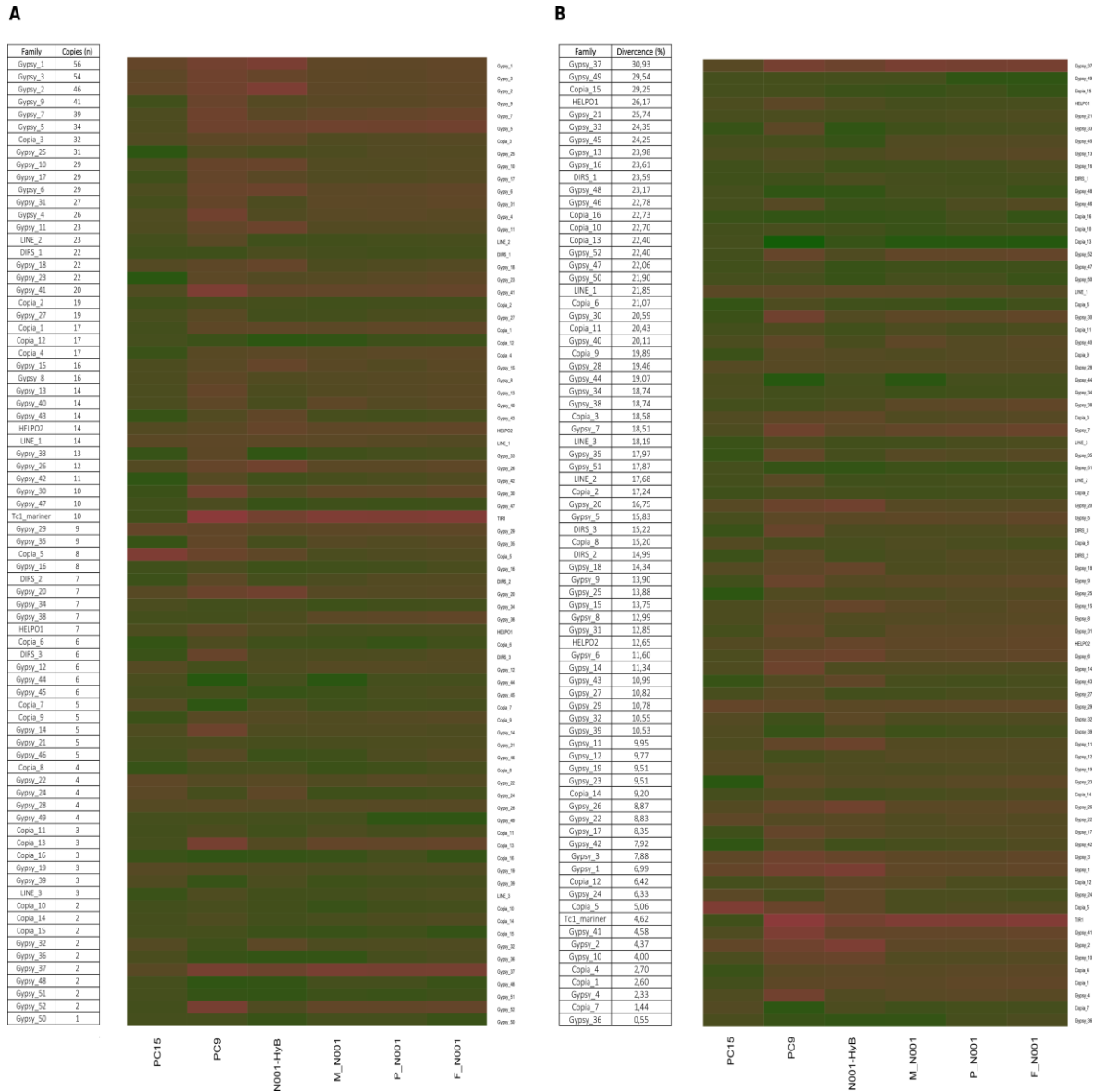


Figure S6. Non-hierarchical clustering of sRNA expression in 80 TE families classified for decreasing copy number (**A**) and divergence (**B**) values. The table next to each heatmap reports ordered TE families with respective copy number (n) and divergence (%) values. sRNA expression is reported as Log (RPM+0.01).

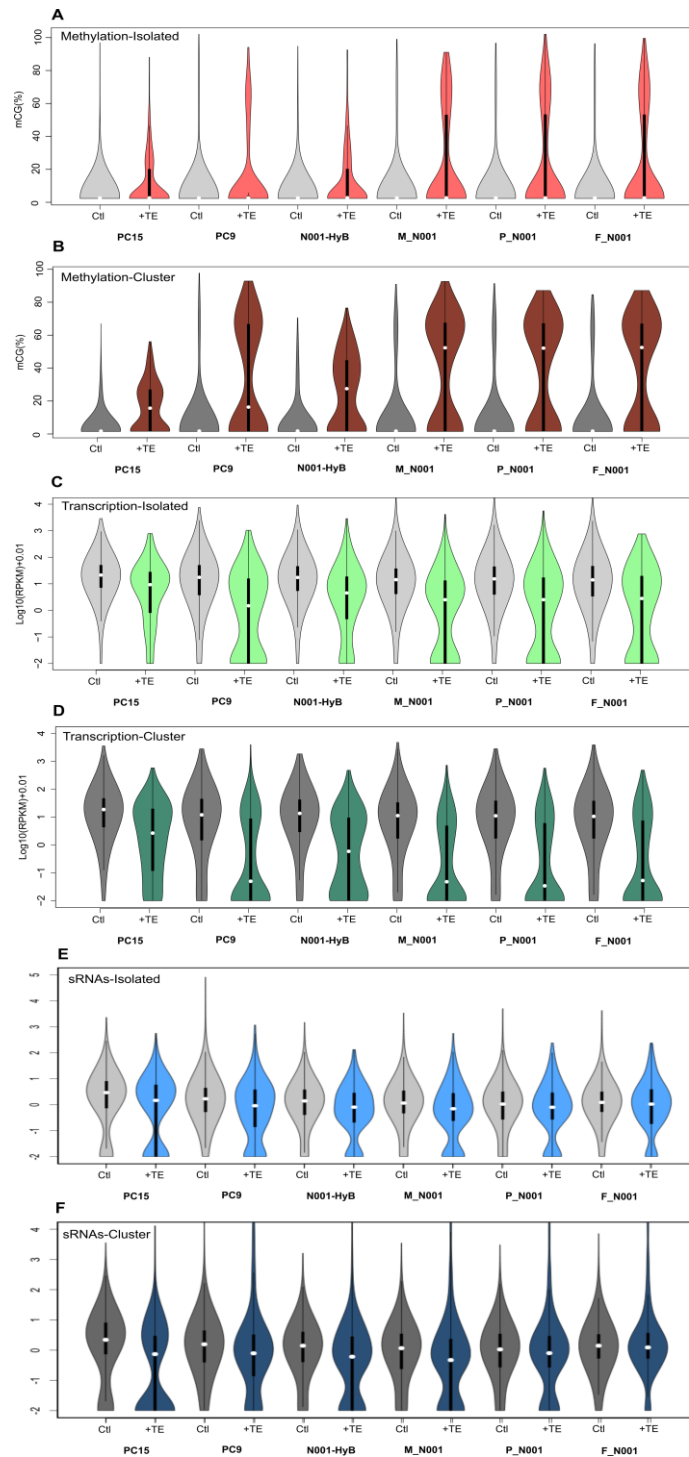


Figure S7. Violin plots representing the impact of TE-mediated gene silencing in all six samples. Genes are classified into two scenarios: localized outside (Isolated) and inside (Cluster) a TE cluster. Each term reports two sets of genes: enclosed by TEs at 1kb upstream and/or downstream (+TE) and not enclosed by a TE as control condition (Ctl). For each scenario are reported methylation (**A** and **B**), mRNA transcription (**C** and **D**) and sRNAs production (**E** and **F**).

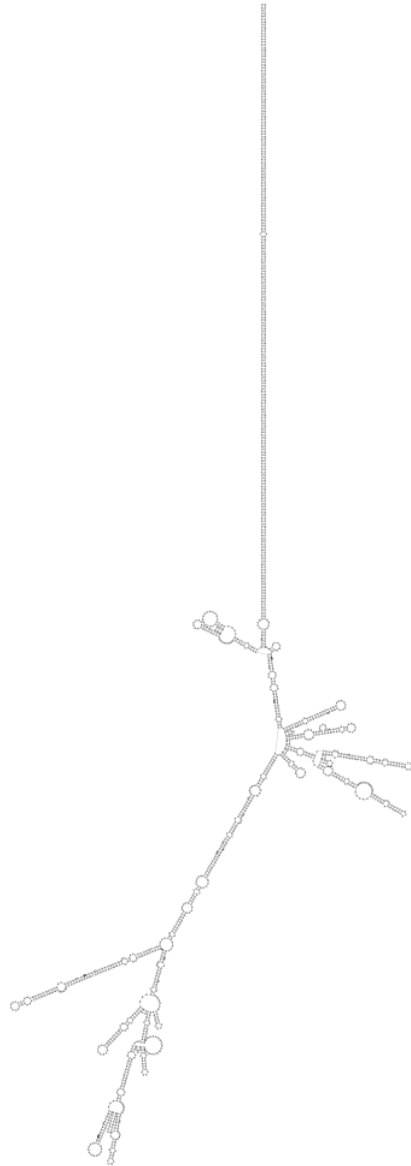


Figure S8. Putative RNA folding of a TIR1 element predicted with RNAfold software.

Final remarks and General discussion

Final Remarks and General discussion

Transposable elements play a major role in shaping eukaryotic genomes, influencing the gene regulation or even contributing to the genome stability. The dynamics of transposon evolution in the host genome depends on several factors. On one hand, the success of a TE is principally linked to its ability to increase in copy number and be inherited by the progeny, without interfering with genome viability. On the other hand, genomes have to deal with the mutagenic potential of transposons, although their presence also contribute to the stability and maintenance of certain chromosomal regions, such as centromeres and telomeres. Distinct selective forces participate in shaping the transposon landscape in genomes evolution, repressing the most invasive and mutagenic mobile elements. Hence, most eukaryotic genomes have evolved host-encoded epigenetic mechanisms to limit their proliferation. These genome defense strategies can be regulated at different levels and, in a coordinate or independent way, contribute to several cellular processes, ranging from gene regulation to transposon silencing.

In this context, several studies have been performed in filamentous fungi, providing a better understanding of the role of epigenetic modifications on transposon control. Nevertheless, several aspects have not yet been deciphered in basidiomycetes. The sequencing and genome-wide characterization of the model basidiomycete *Pleurotus ostreatus* uncovered that an important fraction of its genome is occupied by active transposable elements belonging to both Class I and Class II.

In this thesis, we describe the inheritance patterns of helitron transposons in a meiotically-obtained progeny as well as their transposition potential at both somatic and meiotic contexts.

As second objective, it was undertaken the study of *P. ostreatus* epigenetic and transcriptomic profiles at different developmental stages. Both objectives yielded new insights into the role of DNA methylation and small RNA production in genome silencing as well as the selective activation of genes involved in the fruitbody development.

Differential epigenetic and transcriptional profiles of HELPO1 and HELPO2 helitrons in *P.*

ostreatus strains

Helitrons are a group of DNA transposon discovered the past few decades “*in silico*”, in many eukaryotes by bioinformatics approaches. Several efforts have been made to describe their “rolling circle” transposition mechanism and ability to capture gene fragments. However, the experimental evidence of their mobilization was not documented until very recent in a mammal lineage (Grabundzija et al. 2016). The presence of these elements has also been deeply described in the basidiomycete *P. ostreatus* (Castanera et al. 2014), shedding light on the distribution and structural characterization of the two helitron families that populate the genome of this fungus. Both families contain putative autonomous (carry all the proteins necessary for their transposition) and non-autonomous (defective) copies. In the transposon landscape of *P. ostreatus*, helitrons (especially HELPO1 family) represent the most abundant and transcriptionally active group of DNA transposons. The high similarity between copies (100% among most HELPO2 copies) suggested that these elements could have invaded the genome in a very recent past. These facts suggest that helitrons elements have potential ability to mobilize and give rise to genome instability. In this light, we sought to analyse the inheritance pattern of both helitron families in a meiotically-obtained progeny. The distorted segregation observed in HELPO2 helitrons in comparison with the Mendelian proportion obtained for HELPO1 elements led us to examine these results under the genome profiles (presence vs. absence of helitrons) associated to each parental strain. Therefore, we retrieved and compared monokaryotic and dikaryotic subclones differing in culture time (PC9-99, PC9-14, PC15-99, PC15-14 and N001-03, N001-14) and subculture conditions (high for laboratory work vs. low while deposited in culture collection) during 10 years. We could observed that HELPO2 helitrons displayed differential distribution in PC15 and N001 subclones. Previous studies (Tittel-Elmer et al. 2010) postulated that environmental stresses (i.e., de-dikaryotization process) as well high subculturing rates could promote the increase of helitron transposition. In this sense the five-fold increment of the HELPO2 element content observed in the dikaryon N001 (highly subcultured) and its counterpart (N001 maintained at low-subculturing conditions) supports this hypothesis. However, the PC15 subclone maintained under low-subculturing condition showed higher HELPO2 content than its counterpart maintained under highly frequent subculture.

This apparently contradictory evidence suggest that, independently of the subculturing conditions, *P. ostreatus* TEs have the potential to transpose and eventually being passed to successive clonal generations by subculturing, due to some other reason as it can be the nutrient availability of the mycelium

contained in the culture stab, as described by Naas *et al.* (Naas *et al.* 1994) for IS elements present in bacteria. In contrast, the conserved pattern of HELPO1 elements found in the parental strains, along with expected inheritance pattern found in the progeny, might suggest that the transposition of helitrons belonging to this family is infrequent in *P. ostreatus* genome.

Global epigenetic and transcriptomic profiles in *P. ostreatus* genome revealed that transposon activity is tightly controlled by epigenetic modifications (5-methylcytosine methylation and small interfering RNA), which act silencing their expression. In addition, epigenetic and TE transcription variability was uncovered within monokaryotic and dikaryotic strains. Although we lack epigenetic information from all subclones, our integrated comparisons - focused on helitron elements - provide a closer picture on their epigenetic and transcriptional state in monokaryotic and dikaryotic mycelia maintained under different conditions. In all strains, HELPO2 elements displayed higher methylation rate when compared to HELPO1 helitrons. In regards to the production of small interfering RNA (rasiRNAs), a similar scenario was found, with the exception of the PC9 strain. In contrast, HELPO1 elements exhibited higher transcriptional rates compared to the nearly total silencing of HELPO2 elements (Table 1).

Table 1. Epigenetic and transcriptomic patterns in *P. ostreatus* monokaryotic and dikaryotic strains

	Helitron Family	PC15	PC9*	N001-HyB	N001
mCG (%) (BS-seq)	HELPO1	11.95	38,00	16.24	32.68
	HELPO2	21.27	41.73	26.34	40.42
RPM/copy (sRNA-seq)	HELPO1	3.78	34,52	10.89	3.49
	HELPO2	5.67	14.58	59.25	24.02
RPKM/copy (mRNA-seq)	HELPO1	4.42	12.53	4.12	3.13
	HELPO2	0.28	0.48	0.40	0.81

*degenerated non-autonomous copies

These observations provide an interesting perspective to the fact that N001 strains with the same genetic complement but differently originated (protoclones have been maintained separated in the culture collection before mating in N001-HyB or harboured in the same cytoplasm in N001 during several years) and subcultured display different epigenetic profiles related to helitron elements.

This evidence likely points to a different contribution and independent evolution of the haploid genomes (PC15 and PC9), which have cohabited for several years in the same cytoplasm (N001) and maintained separately in low subculturing condition before mating (N001-HyB).

Since HELPO2 helitrons display the ability to transpose somatically in the PC15 subclone, it was expected a higher transcriptional activity (likely derived from the expression of the helicase protein necessary for the mobilization) linked to a lower epigenetic control compared to HELPO1 helitrons. Nevertheless, our findings demonstrated the opposite trend, as HELPO1 elements showed higher transcriptional rates and lower values of DNA methylation and rasiRNA production. Although we cannot provide direct explanation, two possible scenarios can be hypothesized according to these observations:

- i. Epigenetic modifications directed to HELPO2 elements could have occurred as a consequence of their activity. It seems likely that the recent mobilization of HELPO2 helitrons at random genome position could have triggered the silencing of these elements through the activation of the epigenetic machinery in *P. ostreatus*. In fact, in Chapter III it is shown how the most invasive families (high copy number and low divergence between copies) tend to have higher methylation and rasiRNA production.
- ii. HELPO2 helitrons are constitutively silenced by epigenetic modification. Under this view, we speculate that their mobilization could have been promoted by the transcriptionally active helicase encoded by HELPO1 autonomous copies, which could identify the 5' terminus and 3' hairpin structure located at 3' end of helitron sequences for subsequent mobilization. In this sense, the conservation of the 3'-end region within helitron sequences from highly divergent species (i.e., fungi and plants) suggests that such structure plays an important role in transposition (Feschotte and Wessler 2001).

Therefore, although HELPO1 and HELPO2 elements are different families, their 5' and 3' ends show nucleotide conservation (71 and 67% of similarity in 5' and 3'- 30 bp terminal ends, respectively), which might be enough to establish the interaction with the active HELPO1 transposase required for mobilization. Although we are not able to date events of mobilization and insertion, the epigenetic profile of helitrons reflect the molecular feedback required to silence potentially deleterious mobile elements and establish a genome-specific dynamic equilibrium between TE activity and genome 'defense' strategies.

Nucleus-specific methylation is compensated in the long-term dikaryotic stage

The sample-specific profiles presented in this study provide a framework for the understanding of epigenetic and transcriptomic differentiation observed during the *P. ostreatus* life cycle. The experimental design presented in Chapter III allowed us to compare the epigenetic profiles of short-term (N001-HyB) vs. long-term (N001) cultured dikaryons. In fact, striking differences were found between these two strains, which share the same genetic complement but show clearly different methylation profiles. In particular, N001 displays heterotic TE methylation levels higher than the parental strains PC15 and PC9, whereas N001-HyB shows mid-parent level values. Our findings suggest that this difference can be explained by the different contribution of each nucleus to the overall methylation levels in the dikaryon, as in N001-HyB the PC9 nucleus is more methylated than PC15, whereas in N001 both nuclei show similar levels. The N001 strain has been sub-cultured for more than 20 years as a dikaryon, whereas N001-HyB is the result of a very recent mating (less than ten sub-cultures) between the compatible protoclones of PC15 and PC9, which were stored as isolated strains in a Culture Collection longer than 15 years. The two nuclei present in dikaryons experience co-evolution in the long-term (Anderson and Kohn 2007). In this context, the dikaryotic stage is thought to be favored over the monokaryotic, as deleterious mutations in one nucleus can be compensated with the healthy allele of the other, or even reverted by compensatory mutations such as described for *Schizophyllum commune* (Clark and Anderson 2004). Our results indicate that in addition to permanent modifications (such as mutations), epigenetic profiles are also compensated in long-lasting dikaryons. This compensation led to a higher methylation in the N001 strain, which can account for their better defense against transposable elements than their single monokaryotic counterparts. Nevertheless, this phenomenon is not observed in recently formed dikaryons such as N001-HyB, where methylation profiles of independent nuclei resemble that of their original monokaryotic parentals. To our knowledge, no similar study has reported this event in fungi. Nevertheless, in plant hybrids, a siRNA-mediated mechanism called trans-chromosomal methylation (TCM) is responsible for equilibrating the methylation levels of alleles, leading to an increase of methylation in the “low parent allele” and resulting in overall higher methylation levels in the F1 hybrids (Greaves et al. 2012). The N001 profile mimics this phenomenon, where the increase in methylation of PC15 alleles would lead to the balanced nucleus-specific methylation that we have described. The intermediate values found in N001-HyB suggest that this phenomenon could be the result of a slow process requiring progressive co-adaptation of the two nuclei in the same cytoplasm, something that seems reasonable as nuclei in dikaryons remain un-fused during most of their lives. The

case showed in Cluster 11 of Figure 3 (Chapter III) is an interesting example of how long-term compensation can impact the methylation levels of many genes. This group of genes is present in TE-rich clusters showing opposite methylation profiles in PC15 and PC9, which leads to the presence of epialleles in the early formed dikaryon N001-HyB. In this case, the hypomethylated profile of PC15 is dominant in the latter strain, but shifts to a hypermethylated and transcriptionally silent PC9-like profile in the long-term dikaryon N001. This observation suggests that the two nuclei that coexist in the recently formed dikaryon retain some degree of independence prior to establishing crosstalk interactions. Further epigenetic differences between monokaryons were found when we profiled the small RNA landscape in *P. ostreatus*. The comparison between monokaryons reveals that sRNAs were highly abundant in PC9, whereas their production seems depleted in the PC15 strain. Besides this, all dikaryotic samples showed intermediate values when compared to the parental profiles, which suggests that sRNA biogenesis in both N001-HyB and N001 is sustained by the machinery of the PC9 nuclei. This can be considered as a partial compensation of the sRNA production deficiency of PC15. In fact, it is likely that the less efficient epigenetic machinery of PC15 interferes with the control of transposon proliferation, as reflected by their overall higher TE expression compared to PC9. Based on our findings, we propose that the TE dynamics can rewire the epigenetic landscape of the fungal genome, promoting gene silencing at their surroundings and generating epialleles in the dikaryotic stage. The methylation profile of these epialleles is subjected to compensation after long-term culture, leading to balanced methylation levels in each nucleus.

RNA interference pathway in *P. ostreatus* genome

The study of epigenetic phenomena in filamentous fungi has uncovered the defensive role of RNAi mechanism pathways and its canonical pathway, in the host defense against exogenous nucleic acid and transposons (Chang et al. 2012). The action of RNAi against viruses was first documented in the fungus that causes chestnut blight *Cryphonectria parasitica*, in which the integrated multiprotein complex (dicer-like gene, *dcl2*, argonaute-like gene, *agl2*) is required for the antiviral defense response (Segers et al. 2007; Sun et al. 2009). Also, RNA-mediated silencing involved in the control of transposon activity and genome integrity has been especially studied in the ascomycete *Neurospora crassa* (Nolan et al. 2005) and in the basidiomycete *Cryptococcus neoformans* (Janbon et al. 2010), during vegetative and sexual stages. The role of RNAi pathways, as an essential defense mechanism within the genome, has been evolutionary conserved through the entire eukaryotic domain.

However, some eukaryotic organisms lack an active RNAi pathway or possess non-canonical RNAi components, whose functions has been not fully understood (Nicolás et al. 2013). Our findings show that proteins associated to the RNAi complex previously described in fungi are transcriptionally active in the genomes of both parental monokaryons (PC15 and PC9). Furthermore we found that this process is mediated by 21-22 nt sRNAs in the genome of *P. ostreatus*. These results report a strain-specific sRNA production evidencing a deficient-sRNAs profile in correspondence to PC15 strain.

Comparative transcriptomic analysis focused on the RNAi complex, reveal that, although at different level, proteins involved in this mechanism (Argonaute, Dicer and RNA-dependent RNA polymerase) are actively transcribed in all samples (Table 2). In this scenario, we propose that the apparent inefficacy of sRNAs production in PC15 strain might not rely on the transcriptional inactivity of its RNAi complex.

Table 2. Transcriptional profile of RNAi proteins in monokaryotic and dikaryotic strains of *P. ostreatus*

ID	RPKM						Description
	PC15	PC9	N001-HyB	M_N001	P_N001	F_N001	
25449	17,42	5,58	16,38	14,78	16,26	13,61	RNA-dependent RNA polymerase
1053861	10,41	24,47	17,44	23,17	15,17	11,17	RNA-dependent RNA polymerase
1076136	17,36	0,71	0,48	0,10	0,28	0,33	RNA-dependent RNA polymerase
1041769	9,83	8,57	8,54	13,25	10,59	8,18	RNA-dependent RNA polymerase
154946	27,59	6,29	32,14	40,49	19,95	10,39	RNA-dependent RNA polymerase
1033048	20,05	1,80	4,37	7,01	6,35	5,00	DICER
1064031	21,09	3,79	6,11	7,27	10,82	15,65	DICER
1093523	18,02	7,52	11,12	16,10	9,60	6,65	DICER
1060211	31,83	35,37	50,70	168,35	122,65	73,49	Argonaute
1110274	115,72	56,57	68,67	49,45	44,58	76,39	Argonaute
173501	17,98	26,35	101,45	411,22	169,42	123,65	Argonaute
44554	21,06	15,08	25,58	24,60	13,51	8,55	Argonaute
1039826	26,28	21,69	24,38	23,02	23,57	11,08	DnaQ-like exonuclease
33722	8,50	8,85	8,55	6,77	5,45	4,60	RecQ Helicase
11210	2,17	1,30	2,77	1,53	3,40	6,04	RecQ Helicase
160303	5,96	1,19	2,83	3,00	2,61	3,63	RecQ Helicase

These observations suggest that other non-canonical components required for the biogenesis of sRNAs in *P. ostreatus* might be exclusively active in the genome of PC9 (also reverted in the dikaryotic stage) but not in PC15, leading to an inefficient control of transposon activity. In this context, it can be emphasized that 21–22-nt sRNA—produced by post-transcriptional gene silencing components—are strictly associated to the transcriptional silencing of Helitron and Copia elements in plants (Garcia et al. 2012). An alternative possibility is that random mutations along the amino acid sequence of specific domains could have contribute to generate instability, resulting in a loss of function of proteins involved in the RNAi complex.

In this scenario, such changes could lead to incorrect protein folding, influencing the correct functionality. We screened the nucleotide sequence of proteins listed in Table 2, comparing in the PC15 transcriptomic profile to its reference genome. This analysis revealed the presence of SNPs located at the 3' region of the exon 6 along the sequence of a RNA-directed RNA polymerase QDE-1 required for post-transcriptional gene silencing and RNA interference (ID: 25449), which could be implied in the impaired activity of iRNA machinery in the PC15 genome. This change induced the substitution of an Asparagine (Asn) residue to Aspartic acid (Asp) (Fig. 1). Despite the conservative nature of this codon base change, such substitution can potentially induce complete loss of activity (Pakula and Sauer 1989; Castro-Chavez 2010).

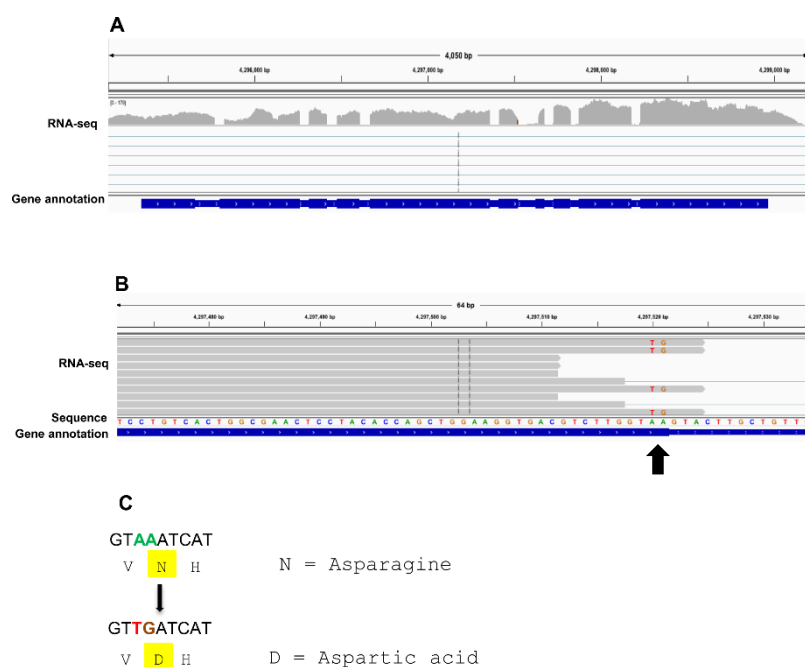


Figure 1. Genomics viewer browser visualization of a RNA-directed RNA polymerase located in scaffold 3 of PC15 genome (**A**). SNPs in exon 6 along the amino acid sequence are reported in **B**. Amino acid substitution of a asparagine residue (Asn) to Aspartic acid (Asp) induced by SNPs is shown in **C**.

This destabilizing substitution, might induce changes in the protein conformation and binding to other proteins in the specific pathway. Although further studied are needed to provide an exhaustive picture of RNAi pathways and mechanisms beyond sRNA production, these findings highlight the requirement of a functional epigenetic mechanism to fully control the mobility of transposons in the genome of *P. ostreatus*.

References

- Anderson JB, Kohn LM (2007) Dikaryons, diploids, and evolution. *Sex Fungi Mol Determ Evol Implic* ASM Press Washington, DC 333–348.
- Castanera R, Perez G, Lopez L, Sancho R, Santoyo F, Alfaro M, Gabaldon T, Pisabarro AG, Oguiza JA, Ramirez L (2014) Highly expressed captured genes and cross-kingdom domains present in Helitrons create novel diversity in *Pleurotus ostreatus* and other fungi. *BMC Genomics* 15:1071.
- Castro-Chavez F (2010) The rules of variation: amino acid exchange according to the rotating circular genetic code. *J Theor Biol* 264:711–21. doi: 10.1016/j.jtbi.2010.03.046
- Chang S-S, Zhang Z, Liu Y (2012) RNA Interference Pathways in Fungi: Mechanisms and Functions. *Annu Rev Microbiol* 66:305–323. doi: 10.1146/annurev-micro-092611-150138
- Clark TA, Anderson JB (2004) Dikaryons of the basidiomycete fungus *Schizophyllum commune*: Evolution in long-term culture. *Genetics* 167:1663–1675. doi: 10.1534/genetics.104.027235
- Feschotte C, Wessler SR (2001) Treasures in the attic: rolling circle transposons discovered in eukaryotic genomes. *Proc Natl Acad Sci U S A* 98:8923–4. doi: 10.1073/pnas.171326198
- Garcia D, Garcia S, Pontier D, Marchais A, Renou JP, Lagrange T, Voinnet O (2012) Ago Hook and RNA Helicase Motifs Underpin Dual Roles for SDE3 in Antiviral Defense and Silencing of Nonconserved Intergenic Regions. *Mol Cell* 48:109–120. doi: 10.1016/j.molcel.2012.07.028
- Grabundzija I, Messing SA, Thomas J, Cosby RL, Bilic I, Miskey C, Gogol-Döring A, Kapitonov V, Diem T, Dalda A, Jurka J, Pritham EJ, Dyda F, Izsvák Z, Ivics Z (2016) A Helitron transposon reconstructed from bats reveals a novel mechanism of genome shuffling in eukaryotes. *Nat Commun* 7:10716. doi: 10.1038/ncomms10716
- Greaves IK, Groszmann M, Dennis ES, James Peacock W (2012) Trans-chromosomal methylation. *Epigenetics* 7:800–805. doi: 10.4161/epi.20820
- Janbon G, Maeng S, Yang DH, Ko YJ, Jung KW, Moyrand F, Floyd A, Heitman J, Bahn YS (2010) Characterizing the role of RNA silencing components in *Cryptococcus neoformans*. *Fungal Genet Biol* 47:1070–1080. doi: 10.1016/j.fgb.2010.10.005
- Naas T, Blot M, Fitch WM, Arber W (1994) Insertion Sequence-Related Genetic-Variation In Resting *Escherichia-Coli* K-12. *Genetics* 136:721–730.
- Nicolás FE, Torres-Martínez S, Ruiz-Vázquez RM (2013) Loss and retention of RNA interference in fungi and parasites. *PLoS Pathog* 9:e1003089. doi: 10.1371/journal.ppat.1003089
- Nolan T, Braccini L, Azzalin G, De Toni A, Macino G, Cogoni C (2005) The post-transcriptional gene silencing machinery functions independently of DNA methylation to repress a LINE1-like retrotransposon in *Neurospora crassa*. *Nucleic Acids Res* 33:1564–73. doi: 10.1093/nar/gki300
- Pakula AA, Sauer RT (1989) Genetic Analysis of Protein Stability and Function. *Annu Rev Genet* 23:289–310. doi: 10.1146/annurev.ge.23.120189.001445
- Segers GC, Zhang X, Deng F, Sun Q, Nuss DL (2007) Evidence that RNA silencing functions as an antiviral defense mechanism in fungi. *Proc Natl Acad Sci U S A* 104:12902–6. doi: 10.1073/pnas.0702500104
- Sun Q, Choi GH, Nuss DL (2009) A single Argonaute gene is required for induction of RNA silencing antiviral defense and promotes viral RNA recombination. *Proc Natl Acad Sci U S A* 106:17927–32. doi: 10.1073/pnas.0907552106
- Tittel-Elmer M, Bucher E, Broger L, Mathieu O, Paszkowski J, Vaillant I (2010) Stress-Induced Activation of Heterochromatic Transcription. *PLoS Genet* 6:e1001175. doi: 10.1371/journal.pgen.1001175

Conclusions

Helitrons that belong to the HELPO2 family are under-represented in the meiotically-derived progeny of *P. ostreatus*; this likely reflects the activation of a specific genome defense mechanism against the recent amplification of HELPO2 elements.

The biased inheritance of HELPO2 elements with respect to the expected Mendelian segregation reflects the selective meiotic elimination of HELPO2 elements through a mechanism of allelic gene conversion.

Experimental approaches display that HELPO2 elements are active in the genome of *P. ostreatus* and mobilize in somatic cells by replicative or semi-replicative transposition.

P. ostreatus possesses highly conserved and transcriptionally active RNAi machinery, although other non-canonical RNAi components might be involved in the biogenesis of small RNAs.

The genome of *P. ostreatus* contains epigenetic hallmarks that are associated with strain-specific DNA methylation and small RNA production, which are primarily involved in the transcriptional silencing of transposable elements.

The production of repeat-associated small interfering RNAs (rasiRNAs) positively correlates with the transposon copy number and negatively correlates with the transposon divergence rate, as it is an attempt to limit the expansion of younger and more invasive elements within the genome of *P. ostreatus*.

Transposon-associated methylation but not small RNA production is directly involved in the transcriptional silencing of surrounding genes as a consequence of the extension of the methylated transposon state.

The genome of *P. ostreatus* activates a set of genes that are involved in the triggering of primordia formation and fruitbody development, similarly to other basidiomycetes. In this context, transcriptional but not methylation reprogramming triggers the fruitbody development.

The haploid nuclei that form the dikaryotic strain can undergo differential epigenetic modification when isolated before mating, and they can equilibrate their methylation levels after coexisting in the same cytoplasm for a long time.

List of publications

- **Borgognone A**, Castanera R, Morselli, Lopez-Varas L, Rubbi L, Pisabarro AG, M, Pellegrini M, Ramírez L. DNA Methylation, small RNA and transcriptomic profiles uncover epigenetic silencing throughout *Pleurotus ostreatus* life cycle. (Submitted).
- **Borgognone A**, Castanera R, Muguerza E, Pisabarro AG, Ramírez L. Somatic transposition and meiotically driven elimination of an active helitron family in *Pleurotus ostreatus*. 2017. DNA Res. doi:10.1093/dnares/dsw060.
- Castanera R, **Borgognone A**, Pisabarro AG, Ramírez L. Biology, dynamics, and applications of transposable elements in basidiomycete fungi. 2017. Appl Microbiol Biotechnol. 101: 1337. doi:10.1007/s00253-017-8097-8.
- Castanera R, López-Varas L, **Borgognone A**, LaButti K, Lapidus A, Schmutz J, et al. Transposable Elements versus the Fungal Genome: Impact on Whole-Genome Architecture and Transcriptional Profiles. 2016. PLoS Genet. 12:e1006108. doi: 10.1371/journal.pgen.1006108.

Funding

Alessandra Borgognone obtained a PhD scholarship from the Public University of Navarre (May 2013).

This work has been supported by Spanish National Research Plan :

- Effect of helitrons in genome structure and transcriptional profile of *Pleurotus ostreatus*. AGL2011-30495. Funded by the Spanish Ministry of Economy, Department of Research, Development and Innovation.
- Study of the interactions between transcriptome and methylome to explain differences in growth rate and mushroom yield in the edible fungus *Pleurotus ostreatus*. AGL2014-55971-R. Funded by the Spanish Ministry of Economy, Department of Research, Development and Innovation.