



**HAL**  
open science

# Contrôle chromatinien et transcriptionnel de l'expression des gènes codant des effecteurs chez *Leptosphaeria* *maculans*, champignon pathogène du colza

Colin Clairet

## ► To cite this version:

Colin Clairet. Contrôle chromatinien et transcriptionnel de l'expression des gènes codant des effecteurs chez *Leptosphaeria maculans*, champignon pathogène du colza. *Phytopathologie et phytopharmacie*. Université Paris Saclay (COMUE), 2019. Français. NNT : 2019SACLS433 . tel-03024735

**HAL Id: tel-03024735**

**<https://theses.hal.science/tel-03024735>**

Submitted on 26 Nov 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Contrôle chromatinien et transcriptionnel de l'expression des gènes codant des effecteurs chez *Leptosphaeria maculans*, champignon pathogène du colza

Thèse de doctorat de l'Université Paris-Saclay  
Préparée à Université Paris-Sud

École doctorale n°567 : sciences du végétal : du gène à  
l'écosystème (SDV) Spécialité de doctorat : Biologie

Thèse présentée et soutenue à Thiverval-Grignon, le 25/11/19, par

**Colin Clairet**

Composition du Jury :

Fabienne Malagnac Professeure, Université Paris-Sud	Présidente
Philippe Silar Professeur, Université Paris-Diderot	Rapporteur
Claire Veneault-Fourrey Directrice de recherche, INRA Nancy	Rapporteuse
Nicolas Nègre Chercheur, Université Montpellier	Examineur
Isabelle Fudal Chercheuse, INRA Thiverval-Grignon	Directrice de thèse
Jessica L. Soyer Chercheuse, INRA Thiverval-Grignon	Co-Directrice de thèse

La réalisation de cette thèse au sein de l'équipe EPLM (INRA-BIOGER) dirigée par Mylène a été pour moi une aventure extrêmement formatrice. Merci à Thierry et Mylène de m'avoir accueilli à BIOGER, soutenu pendant la thèse et pour les moments passés à rigoler (ou réfléchir) au café le matin. Un remerciement spécial à Mylène qui m'a appris l'art subtil de l'entremise des champignons. Et un pour Thierry qui aura été le parfait guide de visite d'Asilomar (et de ses environs).

Merci à Isabelle et Jessica de m'avoir encadré dans la découverte du monde époustouflant de l'épigénétique fongique. Merci également de m'avoir supporté et soutenu tout au long de mon travail, d'avoir toujours été présentes et réactives, de m'avoir fait confiance et de m'avoir donnée une grande liberté dans mon travail.

Merci à Philippe Silar et Claire Veneault-Fourrey d'avoir accepté d'évaluer mon manuscrit ainsi qu'à Fabienne Malagnac et Nicolas Negre d'avoir accepté de participer à mon jury de thèse. Merci aux membres de mon comité de thèse, Muriel Viaud, Nicolas Lapalu, Thomas Guillemette, Pierre Grognet et Fredy Barnèche pour leurs conseils avisés et leur support tout au long de mon travail.

Merci à Moussa Benhamed, Cécile Raynaud, David Latrasse et Nadia Ponts pour les collaborations fructueuses ainsi que pour leur expertise en épigénétique.

Un très grand merci à la cellule bio-informatique, Nicolas Lapalu et Adeline Simon, qui m'ont beaucoup apporté, notamment les bases de la bio-informatique ainsi que du soutien sans faille lors de nos travaux inter-équipe, mon travail n'aurait pas été le même sans vous.

Un remerciement spécial au LIED (de l'époque), Philippe, Fabienne, Hervé, Sylvain, Gwenaëlle et Florence sans qui je n'aurais pas découvert le monde merveilleux des champignons (et l'art de la brasserie), et sans qui, sans doute, je n'en serai pas là. Merci pour votre encadrement depuis le début de mes études (et aussi pour les pots partagés au LIED ou ailleurs). Sylvain, ce fût un plaisir de collaborer pour notre première publi premier auteur !

Marc-Henri Lebrun et Richard Laugé pour toutes les discussions intéressantes (et parfois perchées) du matin autour du café. Merci à Marc-Henri d'avoir co-voituré avec moi pendant les grèves SNCF et pour les discussions scientifiques que nous avons pu avoir à ces occasions.

Je tiens aussi à remercier tous mes camarades thésards pour leur soutien et les échanges enrichissant que nous avons pu avoir : Maëllys Kevin, Alex Mercier, Elise Gay, Guillaume Fouché, Arthur Jallet, Morgane Grand, Yohann Petit, Jafar Kilani, Antoine Porquier, Julie Gervais, Audren Jiquel, Nacera Talbi et tous les futurs thésards du labo. Avec une mention spéciale pour Maëllys, Elise et Alex avec qui nous avons parcouru le grand Ouest Américain toujours dans la bonne humeur (et parfois un peu en mode aventure il faut bien le dire). Alex pour t'être accroché, avec moi, (depuis un moment déjà) comme deux pauvres bulots, à leur rocher qu'est la science. Et Maëllys pour avoir traversé l'océan de la thèse avec moi.

Merci à mon premier encadrant de stage, Sylvain Merlot et à toute l'équipe de Sebastien Thomine, qui m'ont fait découvrir le monde de la recherche et grâce à qui j'ai moi-même voulu en faire partie.



Merci à tous ceux qui m'ont aidé dans la réalisation des manips, Bérengère pour la microscopie, Bertrand, Martin et Yannick pour s'être occupés de mes pauvres plantes cobayes, Françoise pour les WB et autres manips au début de mon stage, Bénédicte pour son expertise inégalable en transformation et en tant d'autres choses !

Merci à Michel Meyer pour ses blagues, sa bonne humeur, les échanges de matériel ou d'astuces et les discussions parfois capilotractées mais toujours enrichissantes.

Merci à Laurent Coudard pour toutes les étiquettes que je t'ai fait imprimer, tous les primers commandés et pour ta gestion de la collection de souches.

Merci à tous les membres de l'unité BIOGER, Romain, Guillaume, Bérengère, Anaïs, Clémentine et tous les autres avec qui j'ai pu rire, partager, jouer au baby-foot ou même juste me bâfrer de fromage en gloussant.

Merci à mes vieux amis, Arthur, Manon, Alexandre, Xavier, Laura pour leur amitié et leur soutien indéfectible depuis un moment déjà (le collègue et le lycée ça date).

Merci à notre crew, les célèbres BassHeads (Alex, Vince, Juline, Cedric, Yanis, Erjan, Jeseune, Dylann, Julie, Lalita, Arthur, Nico, Keyhan, Adam, Dylan, Eloy, Marine (Qu'on croise partout.... même en hollande) et bien d'autres pour m'avoir soutenu et surtout permis de décompresser (fort) durant ma thèse. Je n'aurai sûrement pas survécu sans vous et ça a été un plaisir de tous vous rencontrer.

Et enfin merci à mes parents de m'avoir toujours poussé à apprendre et d'avoir toléré mes diverses expériences et élevages depuis longtemps déjà. Et enfin merci à mes frères et sœurs, Natacha, Hugo, Alexandre et Kevin pour tous ces moments partagés .... On a une sacrée famille quand même.



## LISTE DES ABREVIATIONS

- 5mC** : méthylation en position 5 de la cytosine (méthylation d'ADN)
- Ac** : groupement acétyle
- ADN** : acide désoxyribonucléique
- ARN** : acide ribonucléique
- ARNi** : ARN interférence
- ARNm** : ARN messenger
- ARNg** : ARN guide
- ARN Pol** : ARN Polymérase
- ChIP** : immunoprécipitation de la chromatine
- CT** : Territoire chromosomique
- DBD** : DNA binding domain (domaine de liaison à l'ADN)
- DIM-2** : defective in DNA methylation-2
- DIM-5 (KMT1)** : defective in DNA methylation-5
- ET** : Élément Transposable
- FT** : Facteur de Transcription
- FTF** : Fusarium transcription factors
- GTF** : General Transcription Factor (Facteur de transcription général)
- H3** : Histone 3
- H4** : Histone 4
- H3K4me2** : di-méthylation de l'histone H3 au niveau du résidu lysine 4
- H3K9me3** : tri-méthylation de l'histone H3 au niveau du résidu lysine 9
- H3K9ac** : acétylation de l'histone H3 au niveau du résidu lysine 9
- H3K27me3** : tri-méthylation de l'histone H3 au niveau du résidu lysine 27
- HMEs** : enzyme modificatrice d'histone
- HP1** : Heterochromatin Protein 1
- IC** : Compartiment inter-chromosomique
- Jpi (ou DPI en anglais)** : jours post infection (days post-infection)
- NRPS** : peptide synthétase non-ribosomale (Non-Ribosomal Peptide synthase)
- PKS** : polycétide synthase (PolyKetide Synthase)
- PPS (SSP en anglais)** : Petite Protéine Sécrétée (Small Secreted Protein)
- RIP** : Repeat-Induced Point mutation
- SIX** : Secreted in Xylem (Sécrétée dans le xylème)
- "Silencing"** : extinction de l'expression des gènes à l'aide d'ARNi
- TAD** : Topologically-Associated Domain (Domaine d'Association Topologique)



## LISTE DES FIGURES

### INTRODUCTION

**Figure 1** : Mode d'infection en fonction du type de nutrition des champignons phytopathogènes

**Figure 2** : Exemple de symptômes causés par différents champignons phytopathogènes

**Figure 3** : Répertoire de gènes codant des protéines sécrétées en fonction du mode de nutrition de l'espèce fongique

**Figure 4** : Schéma représentant les différents niveaux de compaction de la molécule d'ADN dans un noyau Eucaryote

**Figure 5** : Régulation allostérique des deux modifications répressives d'histone les plus étudiées (H3K9me3 et H3K27me3)

**Figure 6** : Organisation du noyau de mammifère

**Figure 7** : Interactions physiques successives de plusieurs régions chromosomiques avec la fraction active de la chromatine

**Figure 8** : Schéma explicatif de la technique de séquençage Hi-C

**Figure 9** : Schéma représentant l'action d'un facteur de transcription général (GTF)

**Figure 10** : Figure comparant des zones régulatrices chez un Eucaryotes unicellulaire par rapport à un Eucaryote complexe

**Figure 11** : Nombre de facteurs de transcription par règne et par espèce en fonction du type de FT

**Figure 12** : Cycle de vie de *Leptosphaeria maculans* en Europe

**Figure 13** : Modèle de contrôle de l'expression des gènes codant des effecteurs 'précoces' chez *Leptosphaeria maculans*

**Figure 14** : Méthode de mutagénèse dirigée CRISPR-Cas9



## CHAPITRE I: Etude genome-wide du paysage chromatinien et nucleosomal

### CHAPITRE I – ARTICLE I: Nucleosome patterns in four plant pathogenic fungi with different genome structures

**Figure 1:** Nucleosome phasing in the four fungi studied

**Figure 2:** Nucleosome density profiles in *F. graminearum*

**Figure 3:** Nucleosome density profiles in *L. maculans* 'brassicae'

**Figure 4:** Repeated di-nucleotide patterns in nucleosomal DNA

**Figure 5:** Nucleosome organization of fungal gene units

**Figure 6:** Nucleosome organization of fungal near-universal single copy orthologous gene units (BUSCO3)

**Figure 7:** Nucleosome organization at start codons/TSS vs. gene expression

**Figure S1:** Nucleosome density profiles in *B. cinerea*

**Figure S2:** Nucleosome vs. SNP density profiles in individual biological replicates of *F. graminearum*

**Figure S3:** Nucleosome density profiles vs. AT-rich regions in individual biological replicates of *L. maculans* 'brassicae'

**Figure S4:** Nucleosome density profiles vs. AT content in individual biological replicates of *L. maculans* 'lepidii'

**Figure S5:** Pearson autocorrelation coefficients of di-nucleotides frequencies (average of three biological replicates) for *L. maculans* 'brassicae'

**Figure S6:** Repeated di-nucleotide patterns in nucleosomal DNA located in AT and GC isochores of *L. maculans* 'brassicae'

**Figure S7:** Nucleosome organization at stop codons/TTS vs. gene expression



## CHAPITRE I – ARTICLE II: A comparative analysis of the genome-wide histone modification maps in two subspecies of *Leptosphaeria maculans* showing contrasted genomic organization and different host specialization

**Figure 1:** Coverage of genes, transposable elements, H3K4me2-, H3K9me3

**Figure 2:** Genome of *L. maculans* 'brassicae' harbors large TE-rich, H3K9me3-domains compared to *L. maculans* 'lepidii'

**Figure 3:** Size of the H3K4me2-, H3K9me3-, H3K27me3-domains in the genomes of *Leptosphaeria maculans* 'brassicae' (Lmb) and *L. maculans* 'lepidii' (Lml)

**Figure 4:** Number of genes associated with H3K4me2, H3K9me3 or H3K27me3 during axenic growth in A. *Leptosphaeria maculans* 'brassicae' (Lmb) and B. *L. maculans* 'lepidii'.

**Figure 5:** Genes associated with heterochromatin are less expressed than genes associated with euchromatin in the genomes of *L. maculans* 'brassicae' and *L. maculans* 'lepidii'

**Figure S1 :** Correlation analysis of the genes, transposable elements and domains enriched for H3K4me2, H3K9me3 and H3K27me3 in the genomes of A: *Leptosphaeria maculans* 'brassicae'; B: *Leptosphaeria maculans* 'lepidii'.





## CHAPITRE II: Protéines modificatrices d'histone

### CHAPITRE II – ARTICLE III: Histone methyltransferases KMT1 and KMT6 regulate effector gene expression and pathogenicity in the plant-pathogenic fungus *Leptosphaeria maculans*

**Figure 1:** Identification of *KMT6* in *Leptosphaeria maculans*

**Figure 2:** Effect of  $\Delta kmt1$  and  $\Delta kmt6$  inactivations on pathogenicity mutants of *L. maculans* inoculated on a susceptible cultivar of oilseed rape

**Figure 3:** Effect  $\Delta kmt1$  and  $\Delta kmt6$  on *L. maculans* pathogenicity on a susceptible cultivar of oilseed rape (Bristol) at 28 DPI

**Figure 4:** Effect of KMT1, KMT6 and KMT1-KMT6 deletion on *AvrLm4-7* profile of expression during early cotyledons colonization

**Figure 5:** Effect of KMT1, KMT6 and KMT1-KMT6 deletion on *AvrLm6* profile of expression during early cotyledons colonization

**Figure 6:** Effect of KMT1, KMT6 and KMT1-KMT6 deletion on *Lmb\_jn3\_11364* profile of expression during early cotyledons colonization

**Figure 7:** Effect of KMT1, KMT6 and KMT1-KMT6 deletion on *Lmb\_jn3\_00284* profile of expression during early cotyledons colonization

**Figure 8:** Effect of  $\Delta kmt1$  on 'early' effector genes expression in susceptible oilseed rape stems (Bristol) at 14DPI

**Figure S1:** Effect  $\Delta kmt6$  mutations on KMT6 protein structure and impact of *kmt1* complementation on KMT1 activity

**Figure S2:** Trimethylation of lysine 27 of histone H3 (H3K27me3) is abolished in  $\Delta kmt6$  and  $\Delta kmt1_{\Delta kmt6}$  backgrounds and recovered in complemented 2, 7, 9 and 10 transformants

### CHAPITRE II – RESULTATS COMPLEMENTAIRES

**Figure 1:** Expression profile of *Leptosphaeria maculans* avirulence genes and *Kdm8* during axenic growth and infection of oilseed rape cotyledons

**Figure 2:** Effect of the *Kdm8* mutations on *Kdm8* protein structure

**Figure 3:** Effect of *kdm8* inactivation on mycelial growth and pathogenicity of *Leptosphaeria maculans*

**Figure 4:** Influence of *Kdm8* inactivation on expression of several *L. maculans* 'early' and 'late' effector genes *in vitro*

**Figure 5:** Influence of *Kdm8* inactivation on the expression of six *L. maculans* avirulence genes during infection of oilseed rape



## CHAPITRE III: Régulation transcriptionnelle

### CHAPITRE III – ARTICLE IV: Regulation of effector gene expression as concerted waves in *Leptosphaeria maculans*: a two-players game

**Figure 1:** Identification of the Pf2 orthologue in *Leptosphaeria maculans*

**Figure 2:** Expression profile of *Leptosphaeria maculans* avirulence genes, *LmPf2* transcription factor and *Kmt1* during axenic growth and infection of oilseed rape cotyledons

**Figure 3:** Effect of the *LmPf2* mutations on *LmPf2* protein structure

**Figure 4:** Effect of the  $\Delta kmt1$  mutation on the KMT1 protein structure and function

**Figure 5:** Effect of *LmPf2* and *LmKmt1* inactivations on pathogenicity of *L. maculans*

**Figure 6:** Influence of *LmPf2* overexpression and *KMT1* inactivation on the expression of two *L. maculans* avirulence genes during axenic growth

**Figure 7:** Hierarchical clustering of the pairwise Pearson correlation between the 18 conditions

**Figure 8:** Differential control of effector genes concerted expression

**Figure 9:** Inverted regulation of *KMT1* and *LmPf2* during *Leptosphaeria maculans* life cycle

**Figure 10:** Key players of effector genes regulation as concerted waves during *Leptosphaeria maculans* colonization of oilseed rape

**Figure S1:** Morphological characterization of *LmPf2* mutants and over-expression transformants

**Figure S2:** Immunoblot analysis showing presence of H3K9me3

**Figure S3:** Level of *LmPf2* expression in the *L. maculans* transformants during axenic growth

**Figure S4:** Pathogenicity assays of WT,  $\Delta kmt1$  and  $\Delta LmPf2$  mutants after crossing with JN2-GFP

**Figure S5:** Influence of *LmPf2* overexpression and *KMT1* inactivation on the expression of two *L. maculans* avirulence genes during axenic growth

**Figure S6:** Reproducibility of RNA-Seq replicates



## CHAPITRE III – RESULTATS COMPLEMENTAIRES

**Figure 1:** Expression profile of cloned effector genes and three specific transcription factors during infection and axenic growth obtained by RNA-seq

**Figure 2:** Effect of the *Lmb\_jn3\_09123* mutations on Lmb\_jn3\_09123 protein structure

**Figure 3:** Effect of *Lmb\_jn3\_09123* inactivation on mycelial growth and pathogenicity of *Leptosphaeria maculans*

**Figure 4 :** Effect of the *Lmb\_jn3\_02170* mutations on Lmb\_jn3\_02170 protein structure

**Figure 5:** Effect of *Lmb\_jn3\_02170* inactivation on mycelial growth and pathogenicity of *Leptosphaeria maculans*

## DISCUSSION

**Figure 1:** Key players of effector genes regulation as concerted waves during *Leptosphaeria maculans* colonization of oilseed rape

## LISTE DES TABLEAUX

### INTRODUCTION

**Table 1** : Protéines modificatrices d'histone, domaines actifs, fonction et effecteur allostérique associé

## CHAPITRE I : Etude genome-wide du paysage chromatinien et nucleosomal

### CHAPITRE I – ARTICLE I : Nucleosome patterns in four plant pathogenic fungi with different genome structures

**Table 1**: Characteristics of reference genomes for the four fungal species studied

**Table S1**: Summary of sequencing and mapping metrics

**Table S2**: Nucleosome landscapes as a function of gene expression defining metrics



## CHAPITRE I – ARTICLE II: A comparative analysis of the genome-wide histone modification maps in two subspecies of *Leptosphaeria maculans* showing contrasted genomic organization and different host specialization

**Table 1:** Number of effector genes located in different compartments in *L. maculans* 'brassicae' and *L. maculans* 'lepidii'

**Table S1:** Statistics of ChIP-seq and RNA-seq datasets and alignments

**Table S2:** Correlation analysis between the different ChIP experiments generated with antibodies targeting histone modifications H3K4me2, H3K9me3 and H3K27me3 in the *L. maculans* 'lepidii', during in vitro growth

**Table S3:** Correlation analysis between the different ChIP experiments generated with antibodies targeting histone modifications H3K4me2, H3K9me3 and H3K27me3 in the *L. maculans* 'brassicae', during in vitro growth

**Table S4:** Coverage of histone modifications H3K4me2, H3K9me3 and H3K27me3 in the genome of *Leptosphaeria maculans* 'brassicae'

**Table S5:** Coverage of histone modifications H3K4me2, H3K9me3 and H3K27me3 in the genome of *L. maculans* 'lepidii'

**Table S6:** Correlation analysis between the coverage of transposable elements (TE) and coding sequences (CDS) location, H3K4me2-, H3K9me3- and H3K27me3-domains in the *L. maculans* 'brassicae' genome

**Table S7:** Correlation analysis between the coverage of transposable elements (TE) and coding sequences (CDS) location, H3K4me2-, H3K9me3- and H3K27me3-domains in the *L. maculans* 'lepidii' genome

**Table S8:** GO categories enriched in genes associated with H3K27me3 during axenic culture of *L. maculans* 'brassicae'.

**Table S9:** GO categories enriched in genes associated with H3K4me2 during axenic culture of *L. maculans* 'brassicae'

**Table S10:** GO categories enriched in genes associated with H3K4me2 during axenic culture of *L. maculans* 'lepidii'.

**Table S11:** GO categories enriched in genes associated with H3K27me3 during axenic culture of *L. maculans* 'lepidii'



## CHAPITRE II: Protéines modificatrices d'histone

### CHAPITRE II – ARTICLE III: Histone methyltransferases KMT1 and KMT6 regulate effector gene expression and pathogenicity in the plant-pathogenic fungus *Leptosphaeria maculans*

**Table 1:** Effect of *Lmkmt1* inactivation and *LmPf2* inactivation or over-expression on in vitro growth, conidia production, pathogenicity and effector gene expression

**Table S1:** Primers used for genetic constructions.

**Table S2:** Strains used for mutants purification

**Table S3:** qPCR primers used for genetic constructions



## CHAPITRE III: Régulation transcriptionnelle

### CHAPITRE III – ARTICLE IV: Regulation of effector gene expression as concerted waves in *Leptosphaeria maculans*: a two-players game

**Table 1:** Effect of *kmt1* and *LmPf2* inactivations and of *LmPf2* over-expression on mycelial growth, conidia production, pathogenicity and effector gene expression of *Leptosphaeria maculans*

**Table 2:** Effect of *LmPf2* and *KMT1* inactivation on avirulence gene expression during primary oilseed rape infection

**Table S1:** Primers used in the study

**Table S2:** Crosses performed for mutant purification

**Table S3:** qPCR primers used in the study

**Table S4:** RNA-Seq reads informations

**Table S5:** Differentially expressed genes in all transformants

**Table S6:** Differentially expressed transcription factor coding genes in *LmPf2* over-expressor harbouring different chromatinian context



LISTE DES ABREVIATIONS	4
LISTE DES FIGURES	5
LISTE DES TABLEAUX	11
INTRODUCTION	17
RESULTATS	46
CHAPITRE I	46
ETUDE GENOME-WIDE DU PAYSAGE CHROMATINIEN ET NUCLEOSOMAL	46
INTRODUCTION	47
RESULTATS	48
ARTICLE 1	48
<b>Nucleosome patterns in four plant pathogenic fungi with different genome structures</b>	<b>48</b>
ARTICLE 2	82
<b>A comparative analysis of the genome-wide histone modification maps in two subspecies of <i>Leptosphaeria maculans</i> showing contrasted genomic organization and different host specialization</b>	<b>82</b>
CHAPITRE II	118
PROTEINES MODIFICATRICES D'HISTONE	118
INTRODUCTION	119
RESULTATS	120
ARTICLE 3	120
<b>Histone methyltransferases KMT1 and KMT6 regulate effector gene expression and pathogenicity in the plant-pathogenic fungus <i>Leptosphaeria maculans</i></b>	<b>120</b>
RESULTATS COMPLEMENTAIRES	141
CHAPITRE III	146
REGULATION TRANSCRIPTIONNELLE	146
INTRODUCTION	147





RESULTATS	148
ARTICLE IV	148
Regulation of effector gene expression as concerted waves in <i>Leptosphaeria maculans</i> : a two-players game	148
RESULTATS COMPLEMENTAIRES	187
Analyse fonctionnelle de facteurs de transcription putatifs pour leur implication dans la régulation de l'expression des gènes codant des effecteurs de <i>Leptosphaeria maculans</i>	187
DISCUSSION	193
Contrôle chromatinien de l'expression des gènes codant des effecteurs	194
Contrôle transcriptionnel de l'expression des gènes codant des effecteurs	195
Double contrôle, chromatinien et transcriptionnel, de l'expression des gènes codant des effecteurs précoces	196
PERSPECTIVES	198
Conformation 3D des génomes et contrôle de l'expression des gènes codant des effecteurs chez <i>L. maculans</i>	198
Différents répertoires d'effecteurs pour différentes phases d'infection	199
Applications liées à la compréhension des mécanismes épigénétiques et transcriptionnels	199
Valorisation du matériel produit pendant la thèse	200
CONCLUSION	201
BIBLIOGRAPHIE	201



# INTRODUCTION

## Les champignons des organismes plastiques et très adaptables

Les Mycètes, organismes filamenteux (ou levures), se nourrissent par absorption de sources carbonées, possèdent une paroi chitineuse et sont présents dans la plupart des milieux terrestres. Du fait de leur hétérotrophie pour le carbone, les champignons ont dû dès leur apparition interagir avec les autres organismes terrestres (Taylor et Osborn, 1994). Ils sont dits thallophytes et cryptogames, possédant un appareil végétatif non différencié et se reproduisant par la production de spores (Silar et Malagnac, 2013). Les champignons produisent une grande variété de métabolites primaires et secondaires permettant une meilleure adaptation à leur milieu de vie et ces métabolites sont exploités par les hommes pour diverses activités (chimie, médecine, industrie ; Lindequist *et al.*, 2005 ; Gerke et Braus, 2014 ; Carhart-Harris *et al.*, 2017). Ils peuvent adopter plusieurs styles de vie afin d'assurer leur survie (**Figure 1**). Les champignons saprophytes se nourrissent de la matière organique en décomposition et participent ainsi notamment au renouvellement du sol des forêts (Kubartová *et al.*, 2009) (**Figure 1A**). Les champignons peuvent également interagir de façon symbiotique avec d'autres organismes. Ainsi on distingue deux grands types de symbioses : (i) les lichens qui sont des organismes composés par un mycobionte représentant 90 % des cellules totales et un phytobionte ou organisme photosynthétique (algue ou cyanobactérie) représentant les 10 % de cellules restantes (Muggia et Grube, 2018) et (ii) les mycorhizes qui sont formées de champignons entrant en contact avec les racines des plantes ; 80 % des plantes sont mycorhizées. La plante apporte une source de carbone au champignon tandis que ce dernier apporte une source minérale telle que le phosphore et l'azote (Harrison, 2012). Enfin les champignons peuvent être pathogènes de plantes ou d'autres types d'organismes. La plupart des champignons pathogènes identifiés se trouvent être pathogènes de plantes, résultat de leur longue histoire co-évolutive avec les végétaux (Brundrett, 2002). Depuis 10 000 ans, les hommes utilisent les plantes à des fins multiples, que ce soit pour l'alimentation humaine et animale ou bien plus récemment pour l'industrie (Volter *et al.*, 2019). L'avènement de l'ère industrielle et de la monoculture a rendu les parcelles cultivées de plus en plus vulnérables face aux menaces environnementales. Cette vulnérabilité accrue est le fruit de la faible diversité génétique de la plupart des surfaces cultivées alourdissant la pression de sélection pesant sur les agents pathogènes (Stukenbrock et McDonald, 2007). A cette faible diversité génétique des plantes cultivées s'ajoute la pression de sélection représentée par la forte dose d'intrants utilisés en agriculture (engrais et produits phytosanitaires). La plus grande des menaces rencontrées au champ est constituée des champignons phytopathogènes (Fisher *et al.*, 2012).

## 1. Différents styles de vie et déterminants de la pathogénie chez les champignons phytopathogènes

### 1.1 Stratégies infectieuses développées par les champignons phytopathogènes

Les champignons phytopathogènes établissent l'infection de leur plante hôte en suivant des étapes communes : adhésion d'une spore (sexuée ou asexuée) à la surface de la plante, germination de cette spore et pénétration dans les tissus végétaux, parfois à l'aide de structures spécialisées appelées appressoria (Deising *et al.*,



2000). Ces champignons développent ensuite différentes stratégies pour poursuivre leur cycle infectieux et sont distingués selon leur capacité à se développer sur des tissus végétaux vivants ou morts. En effet, certains champignons pénètrent la cuticule de leur plante hôte en échappant à sa reconnaissance alors que d'autres vont provoquer des nécroses directement après l'adhésion à la surface de la plante. On distingue donc plusieurs types de Mycètes en fonction de leur stratégie infectieuse.

### 1.2 Les champignons biotrophes

Les champignons biotrophes vivent en association avec leur plante hôte sans causer de symptômes jusqu'à la phase finale de leur cycle infectieux où a lieu la production de spores sexuées ou asexuées (**Figure 1B** ; **Figure 2F** ; Wipf *et al.*, 2019). On trouve des traces de leur présence sur terre au Dévonien (-400 millions d'années). Ils seraient potentiellement apparus au moment de la colonisation terrestre par les plantes (Remy *et al.*, 1994) et leur capacité à s'associer à d'autres espèces est apparue chez différents taxa : *Ascomycota*, *Basidiomycota*, *Zygomycota*, *Chytridiomycota*, *Plasmodiophoromycota* (O'Connell et Panstruga, 2006). Après adhésion à la surface de la plante, germination et pénétration dans la plante, les champignons phytopathogènes biotrophes forment généralement des structures spécialisées permettant leur nutrition (haustoria, arbuscules ; Zuccaro *et al.*, 2014). Ces structures spécialisées d'échange entre l'hôte et l'agent pathogène peuvent prendre différentes formes et localisations mais ne sont pas forcément néfastes pour l'hôte et ne mènent donc pas forcément à sa mort (cf. *Ustilago maydis* ; *Blumeria graminis f. sp. hordei* ; **Figure 2B** ; Horbach *et al.*, 2011 ; Hassing *et al.*, 2019).

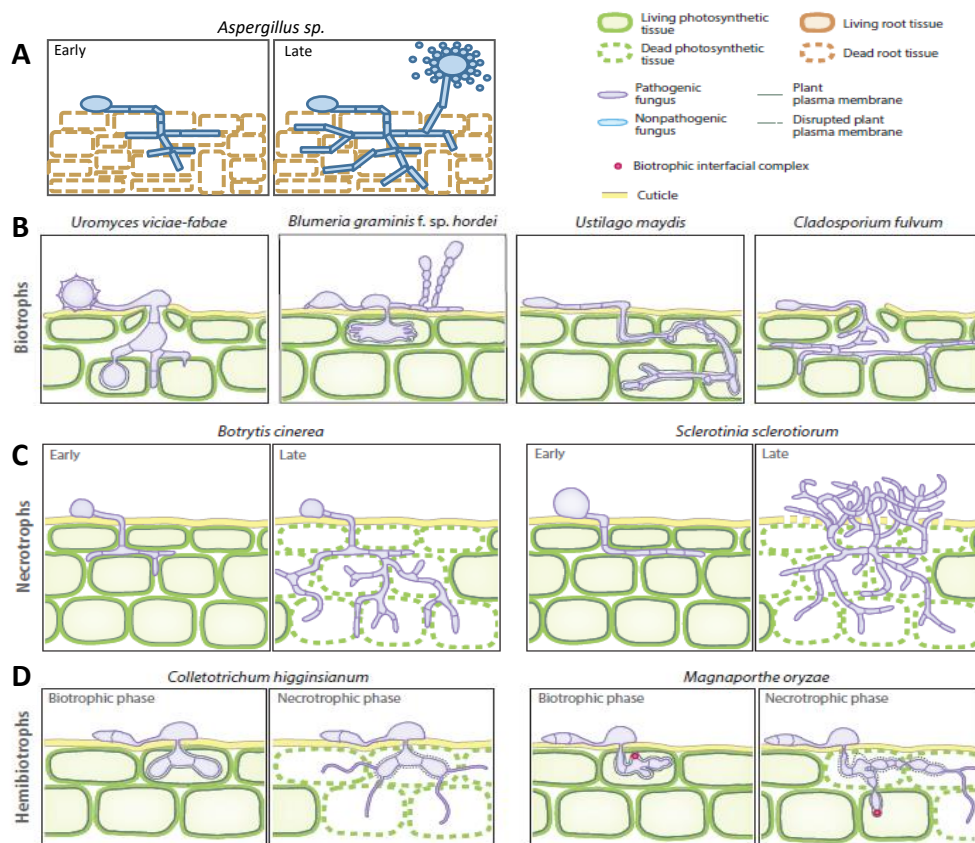
### 1.3 Les champignons nécrotrophes

Les champignons nécrotrophes, après avoir pénétré dans leur plante hôte, sécrètent un arsenal de molécules permettant la lyse cellulaire. Ils sont ensuite capables de se nourrir des cellules mortes de leur hôte (**Figure 1C** ; **Figure 2C** ; van Kan, 2006). Bien que n'ayant pas de phase asymptomatique sur plante identifiée à ce jour, les champignons nécrotrophes possèdent tout de même un arsenal de gènes codant des molécules de pathogénie très diversifié (**Figure 3** ; Lo Presti *et al.*, 2015).

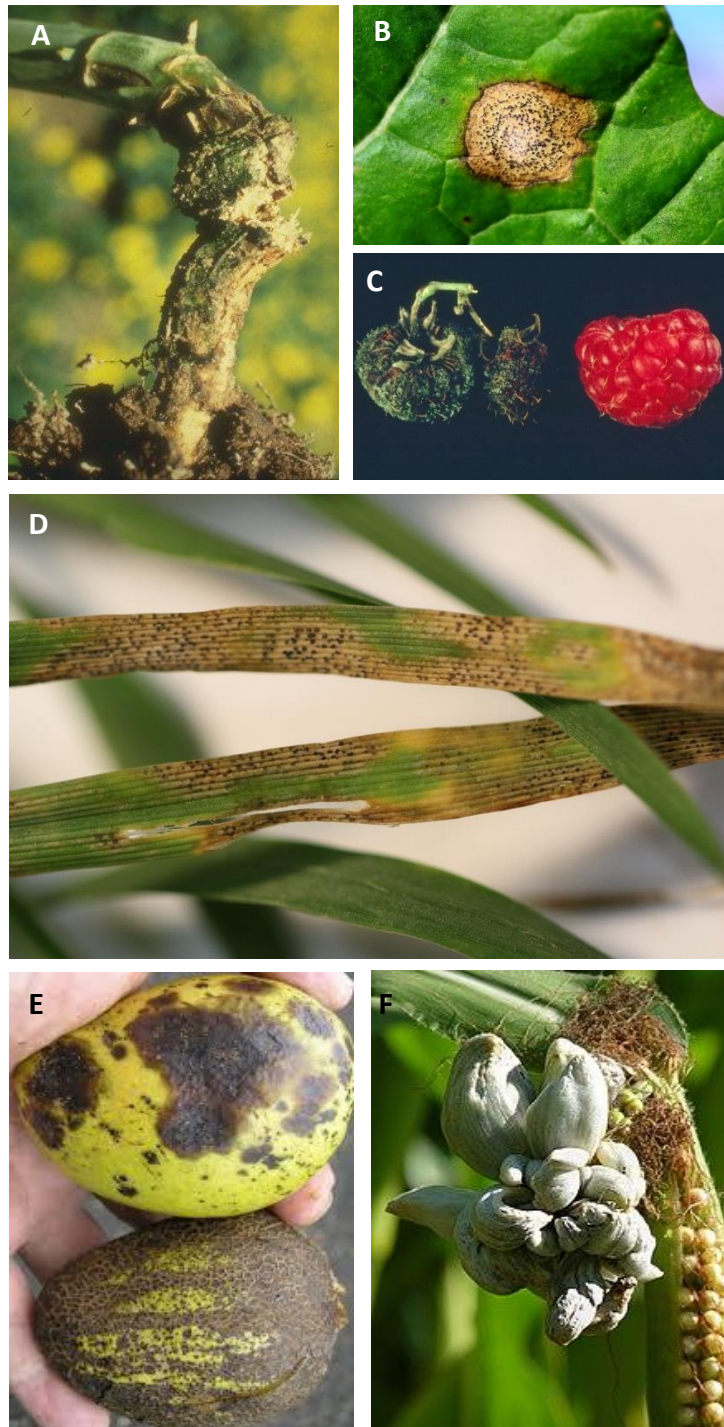
### 1.4 Les champignons hémibiotrophes

Les champignons hémibiotrophes ont un cycle de vie complexe, alternant plusieurs modes de nutrition durant ce cycle (**Figure 1D** ; **Figure 2A, B et D** ; O'Connell *et al.*, 2012). Ils initient souvent la colonisation de leur hôte de façon biotrophe puis basculent vers un mode de vie nécrotrophe, causant l'apparition de symptômes sur leur hôte, pouvant mener à la mort de ce dernier (Jupe *et al.*, 2013). Cette capacité des agents hémibiotrophes à changer drastiquement de mode de vie durant leur cycle infectieux est en partie due à une capacité de régulation génique très fine. Ces champignons expriment spécifiquement des gènes permettant d'échapper à la reconnaissance par leur hôte et de s'établir dans les tissus végétaux durant leur phase biotrophe. Ils vont par la suite exprimer un set de gènes codant des protéases, des protéines permettant la production de ROS (espèces activées d'oxygène), et d'autres protéines impliquées dans la dégradation de la biomasse ligno-cellulosique menant à la lyse cellulaire de la plante (Rohe *et al.*, 1995 ; De Wit *et al.*, 2009). Le champignon peut ensuite se nourrir de façon saprophyte sur la matière végétale morte.

Le mode de vie hémibiotrophe est sans doute le plus complexe parmi les modes d'infection possibles car il nécessite un répertoire de gènes codant des effecteurs plus large que les deux autres types (Lo Presti *et al.*, 2015). Il est également très efficace car il permet au champignon de se développer au sein de son hôte sans causer de symptôme ce qui lui permet d'atteindre de nombreuses parties de la plante avant de passer à un mode de vie nécrotrophe qui endommagera grandement voire tuera la plante hôte (Fitt *et al.*, 2006).



**Figure 1 : Mode d'infection en fonction du type de nutrition des champignons phytopathogènes. A :** Les saprophytes avec à gauche la colonisation du milieu et à droite la sporulation (ex. d'*Aspergillus sp.*). **B :** les champignons biotrophes, avec de gauche à droite *Uromyces vicia-fabae*, *Blumeria graminis f. sp. hordei*, *Ustilago maydis* et *Cladosporium fulvum*. **C :** les champignons nécrotrophes avec à gauche *Botrytis cinerea* et à droite *Sclerotinia sclerotiorum* aux stades précoces et tardifs d'infection. **D :** les champignons hémibiotrophes, à gauche *Colletotrichum higginsianum* et à droite *Magnaporthe oryzae*. D'après Lo Presti *et al.*, 2015.



**Figure 2 : Exemple de symptômes causés par différents champignons phytopathogènes.** **A** : Symptômes de *Leptosphaeria maculans* sur tige de colza. **B** : Symptômes causés par *L. maculans* sur feuille de colza (macule). **C** : Symptômes causés par *Botrytis cinerea* sur framboise, à gauche une framboise infectée et à droite une framboise saine. **D** : Symptômes causés par *Zymoseptoria tritici* sur feuille de blé tendre. **E** : Symptômes causés par *Colletotrichum gloeosporioides* sur mangue (antrachnose). **F** : Symptômes causés par *Ustilago maydis* sur épi de maïs. Photos issues de la banque d'images de l'INRA.

## 2. Déterminants de la pathogénie

Les champignons phytopathogènes ont, à l'exception des champignons biotrophes strictes, tous une partie saprophyte dans leur cycle de vie. Cette conservation du mode de vie saprophyte au sein du règne fongique indique un mode de vie ancestral de ce règne. Le dogme est qu'à l'origine les champignons étaient tous saprophytes mais certains ont acquis peu à peu un répertoire de gènes leur permettant de se nourrir de végétaux vivants (Cooke et Whipps, 1980).

### 2.1 Effecteurs fongiques et pathogénie

Pour infecter un hôte vivant, l'agent pathogène produit des effecteurs, définis comme l'ensemble des molécules sécrétées au moment de l'infection modulant l'immunité végétale et favorisant l'infection (Sánchez-Vallet *et al.*, 2018). Ces effecteurs correspondent principalement à des protéines sécrétées, mais peuvent également être des métabolites secondaires (Collemare et Seidl, 2019) ou des petits ARNs (siRNA) qui permettent à l'agent pathogène de moduler l'immunité végétale (Wang *et al.*, 2015). Chez les champignons nécrotrophes, on retrouve essentiellement des répertoires de gènes codant des cellulases et autres enzymes lytiques ainsi que des gènes impliqués dans la biosynthèse de métabolites secondaires alors que, chez les champignons biotrophes stricts, on retrouve essentiellement des gènes codant des protéines sécrétées de fonction inconnue. Enfin chez les champignons hémibiotrophes, on retrouve à la fois des gènes codant des enzymes lytiques, des gènes codant des protéines sécrétées sans fonction connue et des gènes impliqués dans la biosynthèse de métabolites secondaires (**Figure 3** ; Lo Presti *et al.*, 2015). Les effecteurs les plus étudiés chez les champignons phytopathogènes sont les petites protéines sécrétées (PPS). Celles-ci sont de petite taille (moins de 300 AA), souvent riches en cystéines permettant la formation de ponts disulfure qui améliorent la stabilité de la protéine dans l'apoplasme, et n'ont souvent pas de fonction prédite (Lo Presti et Kahmann, 2017).

### 2.2 Localisation des gènes codant des effecteurs dans les génomes fongiques

Le séquençage des génomes de champignons a révélé que les espèces associées aux plantes présentent souvent des régions génomiques contrastées : des régions évoluant rapidement comportant une grande proportion d'éléments transposables (ET) et des régions riches en gènes plus conservées. Cette observation a mené à la notion de génomes « à deux vitesses » dans lesquels la partie « plastique » des génomes contient des gènes impliqués dans l'adaptation à de nouvelles plantes hôtes ou de nouveaux génotypes. La partie moins plastique du génome étant composée des régions riches en gènes. Cette partie de génome est très conservée du fait de l'importance des gènes qui s'y trouvent. Les régions malléables permettent d'une part, une évolution plus rapide et spécifique des gènes dans ces régions et d'autre part une régulation spécifique de ces gènes (Raffaele et Kamoun, 2012). Les génomes des champignons phytopathogènes peuvent présenter une proportion globalement importante d'ET (45 à 75 % du génome, par ex. chez *Blumeria graminis* ou *Melampsora larici-populina* ; Spanu *et al.*, 2010 ; Duplessis *et al.*, 2011) et une distribution uniforme des ET dans l'ensemble du génome. Ils peuvent présenter une plus faible proportion d'ET regroupés dans des régions génomiques spécifiques telles que de longs blocs de séquences répétées (par ex. dans les génomes de *Leptosphaeria maculans* ou *Epichloe festucae* ; Rouxel *et al.*, 2011 ; Schardl *et al.*, 2013). On peut



retrouver ces régions dans des chromosomes accessoires ou dispensables comme chez *Fusarium oxysporum* ou *Nectria haematococca*, où ces chromosomes peuvent être gagnés ou perdus facilement ce qui permet à leur porteur d'acquérir un nouveau spectre d'hôte ou une capacité accrue à infecter l'hôte (Coleman *et al.*, 2009 ; Ma *et al.*, 2010 ; pour revue voir Soyer *et al.*, 2015,A) ou à des extrémités chromosomiques, appelées régions sub-téломériques. Les gènes codant des effecteurs putatifs sont sur-représentés dans ces régions dynamiques, enrichies en ET. Par exemple, chez *L. maculans*, les gènes codant des effecteurs exprimés au moment de l'infection des feuilles de colza sont

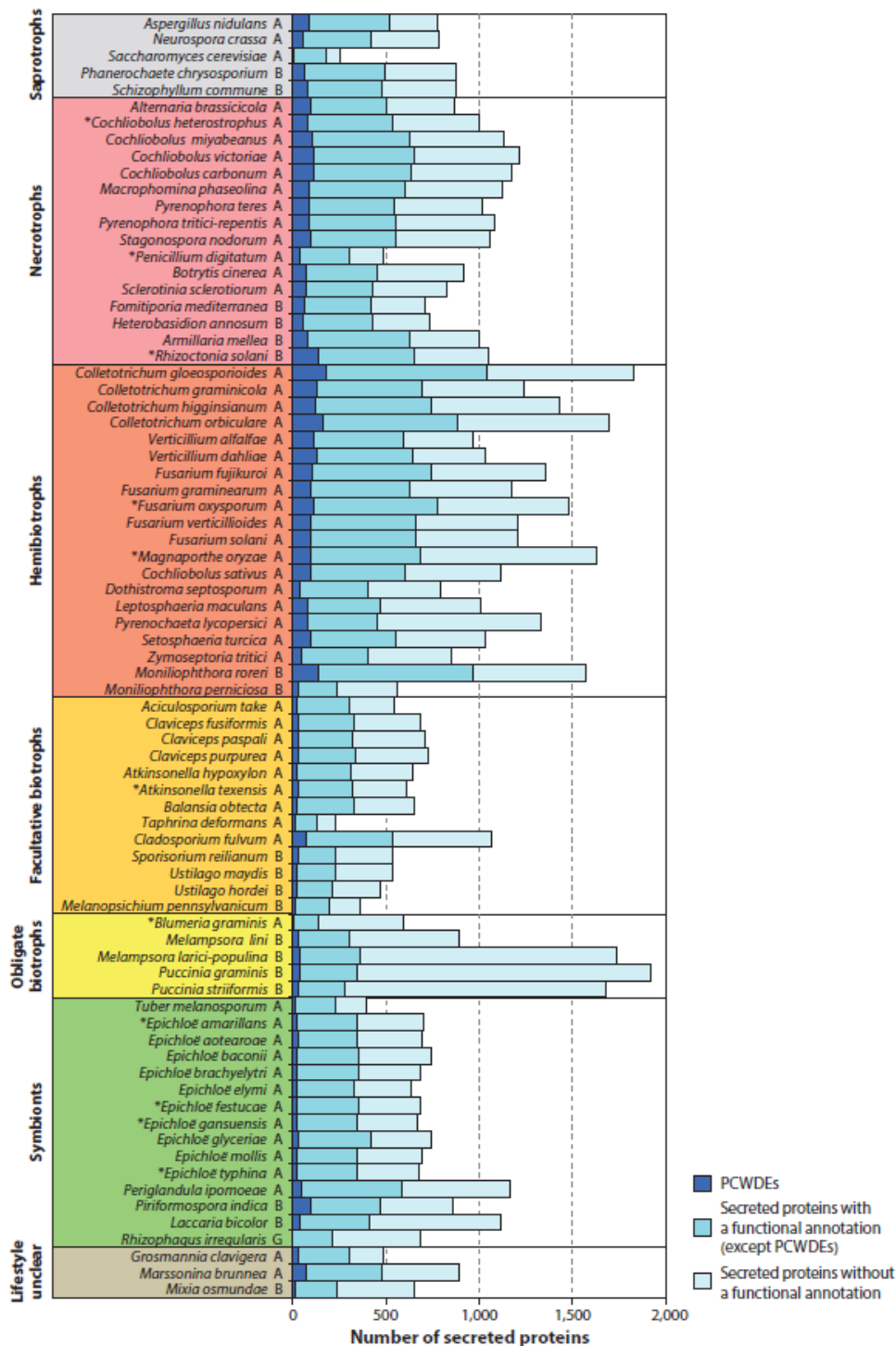


Figure 3 : Répertoire de gènes codant des protéines sécrétées en fonction du mode de nutrition de l'espèce fongique. En bleu foncé, les protéines ayant une fonction enzymatique de dégradation de la paroi végétale, en bleu plus clair, les protéines ayant une autre fonction prédite et en bleu pâle les protéines n'ayant pas de fonction prédite. D'après Lo Presti *et al.*, 2015.



majoritairement localisés dans des régions riches en ET et chez *F. oxysporum*, presque tous les gènes codant des effecteurs sont localisés dans les chromosomes accessoires (Ma *et al.*, 2010 ; Rouxel *et al.*, 2011). Les gènes impliqués dans la biosynthèse de métabolites secondaires sont, eux, souvent localisés à proximité de régions riches en ET, notamment dans les régions sub-téломériques des génomes (Ohm *et al.*, 2012). Ces régions riches en ET sont parfois inactivées via un mécanisme pré-méiotique spécifique des ascomycètes appelé RIP (Repeat-Induced Point mutation) qui induit la mutation des cytosines en thymines et entraîne l'apparition de codons Stop (Galagan et Selker, 2004).

### 2.3 Régulation de l'expression des gènes codant des effecteurs

Les gènes codant des effecteurs protéiques sont souvent peu exprimés en culture axénique et spécifiquement exprimés pendant l'infection de la plante hôte par vagues successives à des étapes stratégiques du cycle infectieux (e.g. pendant la phase de colonisation asymptomatique des feuilles, tiges ou racines ou lors du passage à un mode de vie nécrotrophe). Ceci suggère qu'une régulation fine de l'expression des effecteurs est un élément crucial du succès de l'infection / de la colonisation. Néanmoins, peu de choses sont connues concernant les bases de la régulation de l'expression des effecteurs. L'expression des gènes codant des effecteurs protéiques, tout comme celle des gènes impliqués dans la biosynthèse des métabolites secondaires, est régulée par des facteurs de transcription (TFs) spécifiques et / ou généraux (Cho *et al.*, 2013 ; Rybak *et al.*, 2017 ; Tan et Oliver, 2017). Des études récentes mettent également en évidence l'implication de facteurs physiques tels que l'encombrement stérique de la chromatine, que ce soit pour la régulation de l'expression de clusters de métabolisme secondaire ou des effecteurs protéiques (Strauss et Reyes-Dominguez, 2011 ; Connolly *et al.*, 2013 ; Soyer *et al.*, 2014 ; Chujo et Scott, 2014 ; Gacek-Matthews *et al.*, 2016 ; Soyer *et al.*, 2019). Enfin il a récemment été montré, notamment chez les mammifères et la drosophile, que l'organisation des chromosomes au sein des noyaux intervient également dans cette régulation génique et que certaines régions chromosomiques interagissent physiquement via des complexes protéiques permettant l'expression simultanée de ces régions (Cremer et Cremer, 2001 ; Hou *et al.*, 2012 ; Winter *et al.*, 2018). Il convient d'explorer l'implication de ces différents niveaux de régulation de l'expression des gènes codant des effecteurs dans la mise en place de l'infection et le déroulement du cycle infectieux.

## 3. Régulation chromatinienne et expression génique

Au sein des cellules Eucaryotes, l'ADN présente différents niveaux de compaction et l'un des premiers niveaux est la chromatine, qui est l'association entre la molécule d'ADN, des ARN et des protéines, notamment les histones (**Figure 4**). L'état de compaction de cette chromatine n'est pas figé et il peut varier pour une même région génomique pendant le cycle cellulaire. Au-delà de son importance dans la compaction de l'information génétique, c'est aussi à ce niveau que sont régulés les processus cellulaires fondamentaux tels que la réplication, la réparation de l'ADN et elle constitue le premier niveau de régulation de l'expression des gènes. La chromatine peut être dans un état dit « ouvert » qui va favoriser l'expression génique ou au contraire « fermé » qui va réprimer l'expression génique. Ces états de condensation sont contrôlés par la modification chimique de la molécule d'ADN directement ou par des protéines qui constituent la chromatine (Jenuwein et Allis, 2001 ; Bird, 2002).



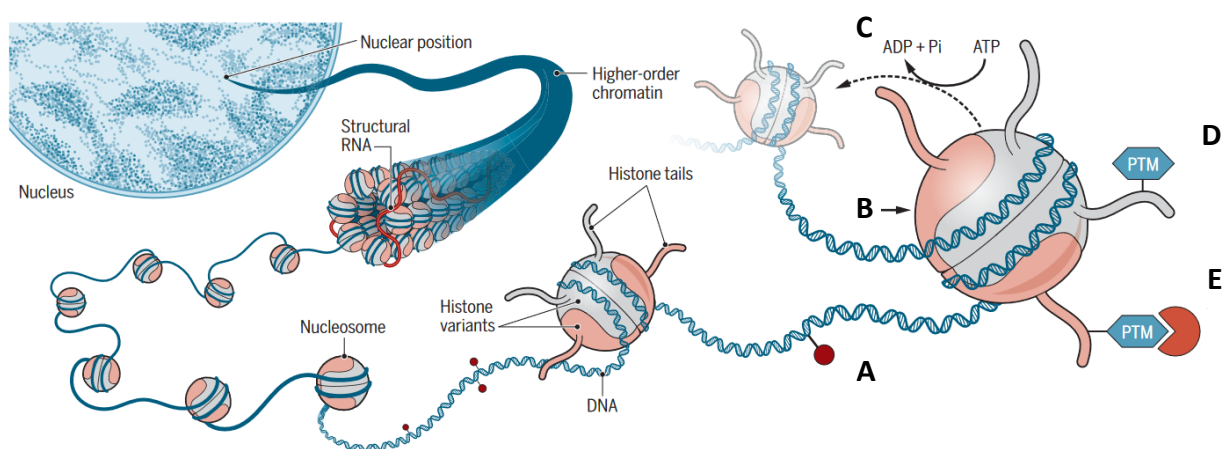


### 3.1. Composants fondamentaux de la chromatine

La chromatine se compose du double brin d'ADN associé à des protéines, les histones, constituant le composant principal du « collier de perle », le nucléosome (**Figure 4**). Environ 150 paires de bases d'ADN sont enroulées autour du nucléosome cœur constitué d'un octamère d'histones (dimère d'histones H3-H4 et H2A-H2B ; Smith, 1991 ; Luger *et al.*, 1997). Chaque nucléosome est relié par la région inter-nucléosomale formée de l'ADN de liaison qui interagit avec l'histone H1. Il existe plusieurs types de protéines histone. Ainsi, on retrouve les histones H1, H2A, H2B, H3, H4 et H5. Les gènes codant les core-histones sont co-régulés avec un pic d'expression durant la phase S du cycle cellulaire (phase durant laquelle l'organisme va doubler son matériel génétique afin de réaliser une mitose ; Marino-Ramirez *et al.*, 2011). Chaque core-histone est composé de deux structures protéiques distinctes, une partie flexible représentant 20 à 30 % de la protéine appelée « queue » et pouvant être modifiée chimiquement, et une partie globulaire autour de laquelle se structure l'ADN (**Figure 4**).

### 3.2. Etats de la chromatine : euchromatine et hétérochromatine

La chromatine peut se trouver dans deux états de condensation : on nommera euchromatine la chromatine ayant une structure décondensée et, au contraire, hétérochromatine la chromatine présentant un état de condensation élevé. Ces domaines, présentant des caractéristiques fonctionnelles et biochimiques distinctes (Grewal et Jia, 2007), avaient été identifiés initialement sur la base de critères cytologiques (Heitz, 1928). Fonctionnellement, l'hétérochromatine reste hautement condensée tout au long du cycle cellulaire, se réplique plus tardivement dans la phase S, est pauvre en gènes, comporte une grande proportion de séquences répétées et présente un faible taux de recombinaison (Richards et Elgin, 2002 ; Huisinga *et al.*, 2006, Grewal et Jia, 2007). En assurant un niveau de compaction élevé, l'hétérochromatine est garante de l'intégrité des génomes, soit structurellement, puisqu'on la retrouve au niveau des centromères et télomères, soit parce qu'elle contribue à empêcher la propagation anarchique des éléments répétés (Papamichos-Chronakis et Peterson, 2013). L'hétérochromatine, parce qu'elle est très condensée, participe également aux mécanismes d'inhibition de l'expression des gènes. Au contraire, l'euchromatine,



**Figure 4 : Schéma représentant les différents niveaux de compaction de la molécule d'ADN dans un noyau Eucaryote. A :** Modification de la molécule d'ADN (e.g méthylation des cytosines). **B :** Protéine chaperonne d'histone. **C :** Facteur de remodelage. **D :** Protéine modificatrice d'histone. **E :** Protéine 'histone reader', capable de reconnaître les modifications de queue d'histone. Issu de Yadav *et al.*, 2018.

décondensée, est plutôt « permissive » pour l'expression des gènes. Ces deux territoires sont aussi caractérisés par des modifications biochimiques de l'ADN et / ou des histones, discutées ci-après. On distingue hétérochromatine constitutive, la forme stable et très condensée de la chromatine, que l'on retrouve le plus souvent dans les régions inactives transcriptionnellement des génomes (e.g. ETs, télomères et centromères ; Saksouk *et al.*, 2015) de l'hétérochromatine facultative que l'on retrouve dans les régions plutôt actives transcriptionnellement et qui sont régulées finement (e.g. clusters de métabolisme secondaire ; Trojer et Reinberg, 2007).

### 3.3 Régulation nucléosomique

La position des nucléosomes dans la séquence génique peut également influencer l'expression des gènes en modifiant l'accessibilité des promoteurs ou des séquences régulatrices aux facteurs de transcription (FTs) ou aux protéines modificatrices d'histone (HMEs ; Radman-Livaja et Rando, 2010 ; Struhl et Segal, 2013). Leur positionnement est défini par la séquence génique. Des cartes nucléosomiques ont été réalisées chez des levures et chez le champignon saprophyte *Aspergillus fumigatus* (Tsankov *et al.*, 2010 ; Nishida *et al.*, 2009). Ces résultats ont montré que les nucléosomes sont généralement déplétés dans les promoteurs, enhanceurs et terminateurs des gènes actifs transcriptionnellement ce qui indique un contrôle nucléosomique de l'expression des gènes. Cependant, aucune étude de positionnement des nucléosomes n'a été réalisée chez des champignons phytopathogènes et donc aucune relation entre positionnement des nucléosomes et régulation de l'expression des gènes et de la pathogenèse n'a pu être mise en évidence. Seules des études menées chez *Plasmodium falciparum*, agent pathogène de l'homme, ont permis d'identifier une régulation via la localisation des nucléosomes de gène de pathogénie (Ponts *et al.*, 2010 ; Bunnik *et al.*, 2014).

### 3.4 Modifications chimiques de l'ADN

Chez les champignons filamenteux, le génome est globalement peu méthylé (environ 1,5 % chez *N. crassa* ; Selker *et al.*, 1993). De plus, contrairement aux mammifères (Li *et al.*, 1993), la méthylation de l'ADN est dispensable à la vie des champignons (Foss *et al.*, 1993 ; Kouzminova et Selker, 2001). C'est l'ADN méthyltransférase DIM-2 (defective in DNA methylation-2) qui catalyse la méthylation de l'ADN (Kouzminova et Selker, 2001), presque exclusivement au niveau des régions riches en transposons affectés par le mécanisme de RIP (Selker, 2002 ; Selker *et al.*, 2003). La méthylation de l'ADN au sein du règne des Mycètes n'a été mise en évidence que chez très peu d'organismes (i.e. *N. crassa* ; Selker et Stevens, 1985 ; Foss *et al.*, 1993) ; *Mycosphaerella fijiensis* (Dhillon *et al.*, 2010) ; *Ascobolus immersus* (Goyon *et al.*, 1996) ; *Coprinus cinereus* (Zolan et Pukkila, 1985) ; *M. oryzae* (Jeon *et al.*, 2015) ; *Aspergillus flavus* (Gowher *et al.*, 2001) et il a été au contraire mis en évidence chez d'autres filamenteux que l'ADN n'était pas méthylé (i.e. *Z. tritici* (Dhillon *et al.*, 2010), *S. pombe* (Antequera *et al.*, 1984), ou encore *L. maculans* (Bewick *et al.*, 2019). Une étude récente de la méthylation de l'ADN chez de nombreux Mycètes a mis en évidence que chez ce règne le mécanisme canonique de méthylation dans le corps des gènes est absent ce qui les distingue de certains insectes et des plantes. Mais il arrive que certains champignons possèdent des clusters de méthylation d'ADN mis en place par des mécanismes indépendants de ceux précédemment décrits, mécanismes communs aux autres espèces de champignon (Bewick *et al.*, 2019).

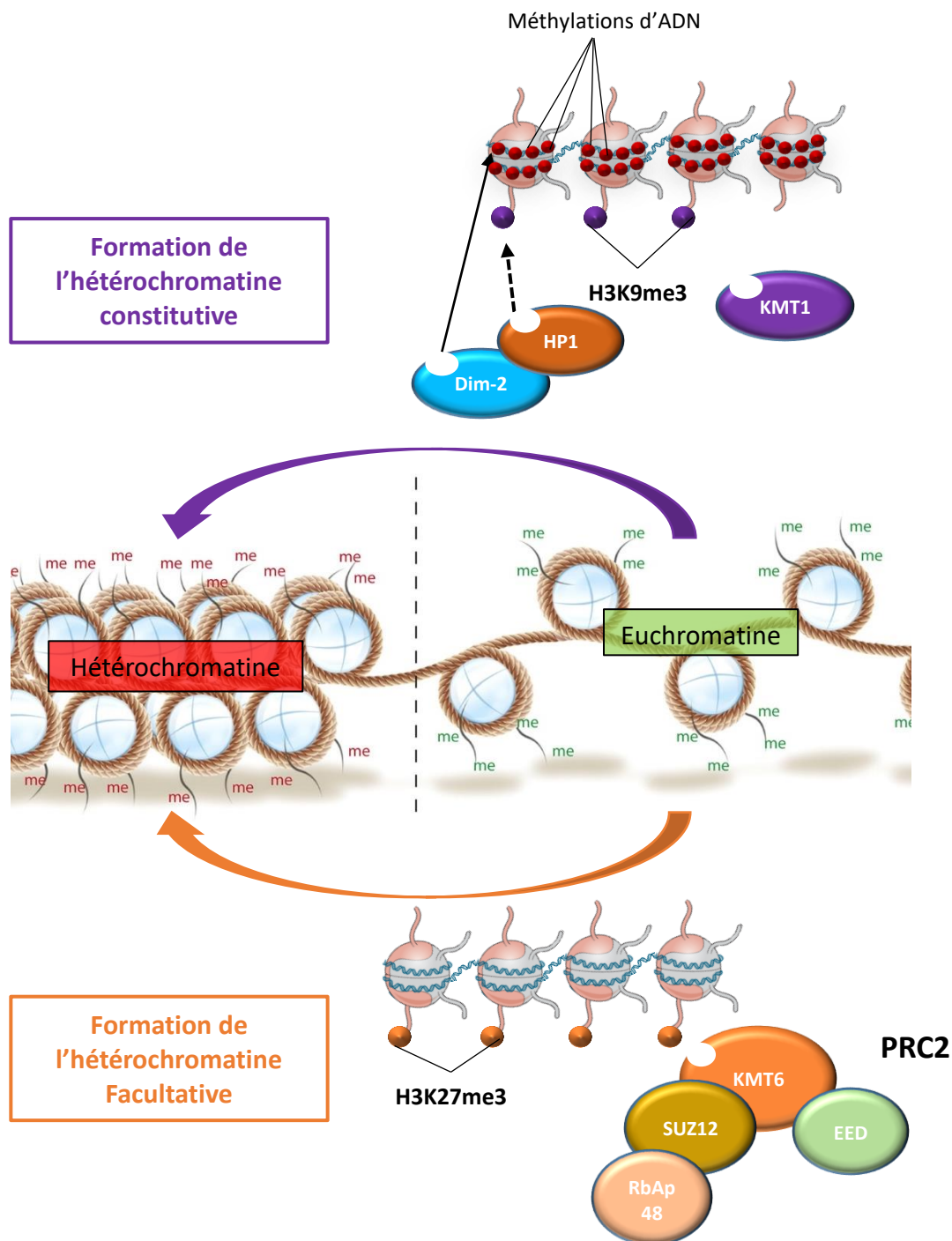


### 3.5 Modifications chimiques des queues d'histones et protéines impliquées dans ces modifications

Chaque core-histone est composée de deux structures protéiques distinctes : une partie flexible représentant 20 à 30 % de la protéine appelée « queue » et pouvant être modifiée chimiquement et une partie globulaire autour de laquelle se structure l'ADN. Les queues d'histones peuvent subir une large variété de modifications chimiques (acétylations, méthylation, sumoylation, ubiquitination etc...), qui, en fonction de l'acide aminé modifié, changera le niveau de compaction de la chromatine et ainsi l'accessibilité de l'ADN et, du même coup, son niveau de transcription (Luger et Richmond, 1998 ; Bannister et Kouzarides, 2005). Ainsi, chez de nombreux champignons, on retrouve dans les régions hétérochromatiniennes des modifications de type tri-méthylation de la lysine 9 de la queue de l'histone H3 (H3K9me3) et tri-méthylation de la lysine 27 de la queue de l'histone H3 (H3K27me3 ; Freitag, 2017). Chez les plantes et les animaux on retrouve souvent des méthylations des cytosines ou des adénines associées aux histones portant des marques répressives (Jackson *et al.*, 2002 ; Law et Jacobsen, 2010). Au contraire les régions euchromatiniennes sont plutôt marquées par des modifications de type di- ou triméthylation de la lysine 4 de l'histone H3 (H3K4me2/3) ou acétylation de la lysine 9 de l'histone H3 (H3K9ac ; **Figure 5** ; Kouzarides, 2007). De nombreuses protéines sont capables de reconnaître, d'ajouter ou de retirer des groupements chimiques aux queues d'histones, ou aux nucléotides directement, modifiant ainsi l'accessibilité de l'ADN de la région génomique concernée (Kim *et al.*, 2019). On appelle ces protéines des « histones modifying proteins (HMEs) ». Beaucoup de ces protéines possèdent des domaines de lecture des modifications de l'ADN ou des histones et permettent ainsi le recrutement d'une machinerie complexe (Kim *et al.*, 2019). Il est même souvent question d'une lecture des modifications chimiques par plusieurs protéines ce qui renforce encore la spécificité et l'efficacité de ce mécanisme (Ruthenburg *et al.*, 2007 ; Savitsky *et al.*, 2016). Des études récentes ont prouvé que les HMEs sont affectées par des composants particuliers des histones comme les modifications d'ADN/d'histone ou la longueur du linker (brin d'ADN reliant deux nucléosomes). Ces observations pourraient indiquer que ces protéines agissent comme des senseurs allostériques (**Tableau 1** ; Motlagh *et al.*, 2014). Il a par exemple été montré chez *N. crassa* que la méthylation d'ADN est induite par deux HMEs : la première, KMT1 (DIM-5), capable de déposer la marque H3K9me3 sur les queues d'histones et la deuxième, DIM-2, capable de reconnaître ces méthylations et de méthyler les cytosines associées (**Figure 5** ; Tamaru et Selker, 2001). Une autre méthyltransférase d'histone très étudiée, KMT6, responsable de la méthylation H3K27me3, est impliquée chez *F. graminearum* et *E. festucae* dans la régulation du métabolisme secondaire (**Figure 5** ; Connolly *et al.*, 2013 ; Chujo et Scott, 2014). Les méthylations répressives déposées sur les queues d'histones peuvent ensuite être retirées par des protéines appelées déméthylases. Les marques d'histones modifiant l'accessibilité de la chromatine aux facteurs de transcription et autres protéines servant à la transcription peuvent être retirées, modifiées ou remplacées ce qui permet un contrôle très fin de l'expression des gènes dans ces régions (Jamieson *et al.*, 2013). Au même titre que les méthylations déposées et retirées par les méthylases et déméthylases, on retrouve des acétylases et déacétylases, des phosphorylases et phosphatases (etc..) qui équilibrent les acétylations, les phosphorylations d'histones (etc.. ; Eberharter, 2002 ; Rossetto *et al.*, 2012). C'est cet équilibre très fin entre les différents acteurs réalisant les modifications chimiques qui permet une régulation dynamique et complexe de l'expression des gènes (Cedar et Bergman, 2009 ; Kim *et al.*, 2019). Il a également été montré chez l'homme que certaines histones peuvent porter



plusieurs types de modifications (une activatrice et une répressive) permettant de contrôler encore plus finement l'expression des gènes dans ces régions (Mauser *et al.*, 2017).



**Figure 5 : Régulation allostérique des deux modifications répressives d'histone les plus étudiées (H3K9me3 et H3K27me3).** A : Le complexe KMT1/Dim-2 responsable de la méthylation de type H3K9me3 et de la méthylation d'ADN associée. KMT1 est responsable du dépôt de H3K9me3 puis HP1 est capable de reconnaître ces modifications et de recruter Dim-2 qui va méthyler les cytosine associées (issu de Honda *et al.*, 2010). B : Le complexe PRC2 responsable de la méthylation H3K27me3, KMT6 étant responsable du dépôt des modifications chimiques (issu de Kim *et al.*, 2019). Image de nucléosome issu de Yadav *et al.*, 2018.

### 3.6 Contrôle chromatinien de l'expression génique chez les champignons

Du fait de leur possible modification chimique, les histones ainsi que la molécule d'ADN elle-même peuvent être utilisées à des fins de régulation génique (Kouzminova et Selker, 2001 ; Jamieson *et al.*, 2013). Ces modifications vont entraîner des changements d'encombrements stériques, modifiant l'accessibilité de la molécule d'ADN aux facteurs pouvant réguler l'expression des gènes. Chez certains champignons tels que *M. oryzae*, la méthylation de la molécule d'ADN est capable de réguler de nombreux processus tels que la reproduction sexuée ou la pathogénie (Jeon *et al.*, 2015). Chez d'autres champignons tels que *L. maculans*, la méthylation de la molécule d'ADN n'est pas un processus conservé et actif (Bewick *et al.*, 2019), le champignon ne peut donc utiliser que les modifications d'histone et la régulation tridimensionnelle du génome (influencée par les modifications d'histone) pour réguler l'accessibilité de ses gènes. Les marques activatrices d'histone les plus étudiées chez les champignons filamenteux sont H3K4me2, H3K36me3 et H3K9ac (di-méthylation de la lysine 4 de l'histone H3, tri-méthylation de la lysine 36 de l'histone H3 et acétylation de la lysine 9 de l'histone H3 ; Gacek et Strauss, 2012 ; Smith *et al.*, 2012 ; Janevska *et al.*, 2018). Au contraire les marques chromatiniennes répressives les plus étudiées chez les Mycètes sont H3K9me3 et H3K27me3 respectivement marqueurs de la chromatine constitutive et facultative. Chez *N. crassa*, on retrouve les marques H3K9me3 dans les régions très peu exprimées tels que les centromères ou les télomères mais ces marques ne participent pas à la conformation 3D du génome (Klocko *et al.*, 2016). Chez *Fusarium graminearum* au contraire on retrouve des modifications de type H3K27me3 dans les centromères et télomères (Connolly *et al.*, 2013). Chez *Z. tritici*, qui possède plusieurs chromosomes dispensables fortement enrichis en ETs, on retrouve la marque H3K27me3 dans ces régions fortement réprimées ainsi que dans les télomères alors que la marque H3K9me3 est spécifiquement retrouvée sur les ETs (Schotanus *et al.*, 2015). Des études récentes tendent à prouver que les modifications d'histones sont capables de réguler de nombreux processus chez les champignons ayant des styles de vie bien différents. On peut

**Tableau 1 : Protéines modificatrices d'histones, domaines actifs, fonction et effecteur allostérique associé.**

Allosteric Effectors		Chromatin Modifiers	Effector Sensing Domains/Subunits	Functions
Activators	H3K4me0	DNMT3A	ADD domain of DNMT3A	DNA methylation
		KDM5A	PHD1 domain of KDM5A	H3K4 demethylation
	H3K27me3	PRC2	EED	H3K27 methylation
	H3K36me3	Rpd3S	Eaf3	Histone deacetylation
	Length of linker DNA	Rpd3S	Not determined	Histone deacetylation
		ISWI	HSS and NegC domains of ISWI	Nucleosome spacing
		ACF	ACF1	Nucleosome spacing
	N-terminal tail of H4	ISWI	AutoN domain of ISWI	Nucleosome spacing
		ACF	ACF1	Nucleosome spacing
		Chd1	Chromodomains of Chd1	Nucleosome remodeling
Nucleosome density	PRC2	SUZ12	H3K27 methylation	
Inhibitors	H3K4me3	PRC2	RbAp48	H3K27 methylation
	H3K36me2/3	PRC2	RbAp48	H3K27 methylation

Ces protéines sont retrouvées chez la plupart des organismes. Un effecteur allostérique est une molécule se fixant à un site particulier d'une enzyme (appelé site allostérique et impactant positivement ou négativement l'activité de l'enzyme. Issu de Kim *et al.*, 2019.



ici prendre l'exemple de *N. crassa* pour lequel la modification H3K36me3 contrôle la tolérance au stress thermique et au pH (Kronholm *et al.*, 2016). Il a également été montré que les deux marques répressives H3K9me3 et H3K27me3 sont capables de contrôler la biosynthèse d'alcaloïdes et l'expression de gènes du métabolisme secondaire chez *E. festucae*, *A. nidulans* et *F. graminearum* (Gacek et Strauss, 2012 ; Connolly *et al.*, 2013 ; Chujo et Scott, 2014). De plus l'inactivation d'un des acteurs responsables du dépôt de H3K9me3, *KMT1*, peut mener à des relocalisations de H3K27me3 chez *N. crassa* ce qui induit un stress génotoxique, un défaut de croissance ainsi qu'une instabilité chromosomique (Basenko *et al.*, 2015). Chez *Z. tritici*, l'inactivation de *KMT1* induit une instabilité chromosomique résultant de la perte de H3K9me3 et de la relocalisation des marques H3K27me3 à l'endroit de ces dernières et du même coup de la déplétion de celles-ci sur les chromosomes accessoires. Au contraire l'inactivation de *KMT6* chez *Z. tritici* induit une plus grande stabilité chromosomique (Möller *et al.*, 2019). Enfin les déméthylases peuvent également agir sur les queues d'histones et ainsi jouer un rôle de répression ou d'activation de l'expression. Chez *A. nidulans*, deux déméthylases d'histones, KdmA et KdmB sont responsables de l'enlèvement respectif de la marque répressive H3K9me3 et de la marque activatrice H3K4me3 et ainsi de la régulation du métabolisme primaire et secondaire (Gacek-Matthews *et al.*, 2015 ; Gacek-Matthews *et al.*, 2016). Il a également récemment été montré que les effecteurs protéiques pouvaient être régulés par la structure de la chromatine. Ainsi, lors d'une précédente thèse dans notre équipe, un contrôle de l'expression des effecteurs de *L. maculans* via la marque répressive H3K9me3 mise en place par *KMT1* avait été mis en évidence (Soyer *et al.*, 2014) ; nous reviendrons sur ces travaux dans la partie Objectifs de la thèse.

## 4. Conformation tridimensionnelle de la chromatine et co-régulation génique

Le génome nucléaire des Eucaryotes est l'élément central de la cellule. C'est là qu'est gardée et condensée toute l'information nécessaire à la vie de l'organisme. Afin que la régulation génique soit correctement réalisée et permette à l'organisme de vivre, il faut organiser l'immense quantité d'information que représentent les chromosomes. De nombreuses protéines nucléaires sont utilisées par les organismes pour structurer les génomes au sein des noyaux (e.g. condensines, cohésines, nucleoporines etc...; Casolari *et al.*, 2004 ; Skibbens, 2019). Les modifications chimiques de l'ADN, des queues d'histones et les séquences répétées servent également de base à la structuration tridimensionnelle (3D) du génome (Bonev et Cavalli, 2016).

### 4.1 Interaction physique des régions chromosomiques

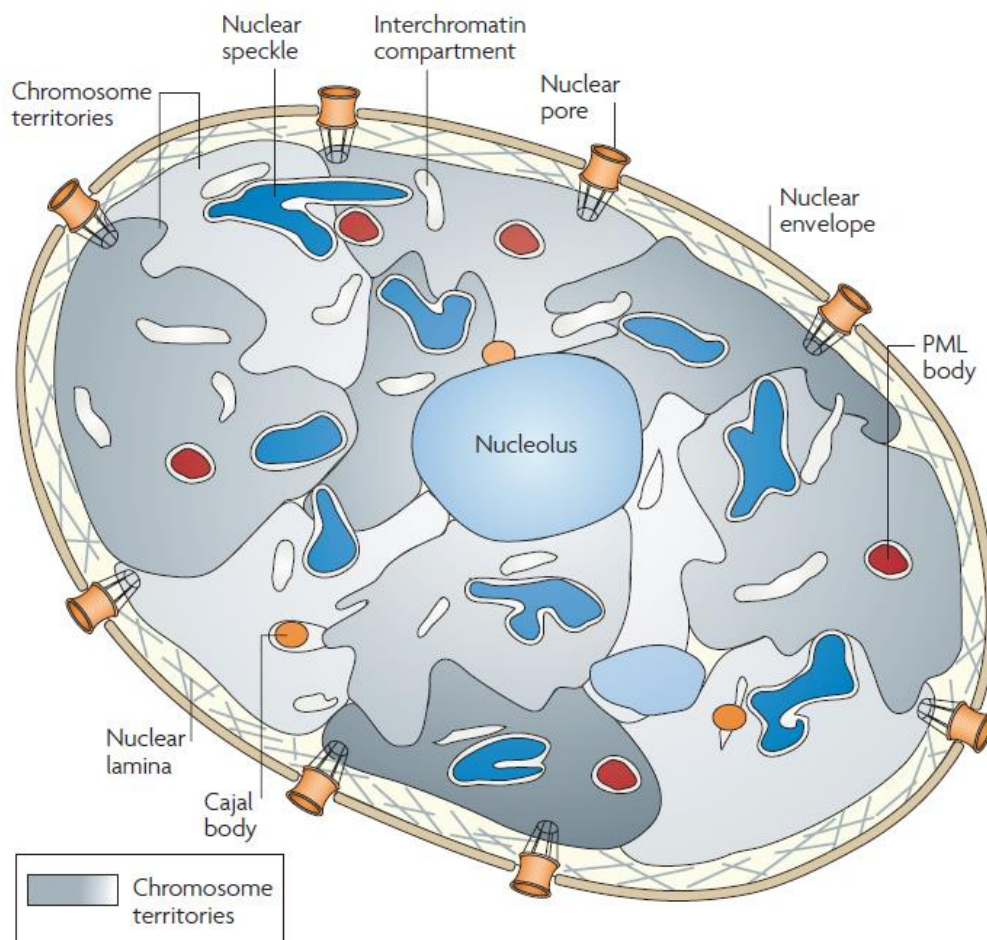
Pour permettre l'expression correcte des gènes à tous les moments de la vie de la cellule, le génome des Eucaryotes est organisé spatialement en domaines 3D appelés Territoires Chromosomiques (CTs) ainsi qu'en compartiments interchromatiniens (ICs) qui contiennent des complexes macromoléculaires nécessaires à la réplication, la transcription, l'épissage et à la réparation de l'ADN (**Figure 9** ; Cremer et Cremer, 2001). Les domaines chromosomiques interagissant physiquement sont appelés TADs pour 'Topologically Associated Domains'. Chez la levure, pendant la mitose, de larges domaines de 300 kb à 1 Mb sont formés grâce aux condensines. Ces domaines disparaissent ensuite peu à peu jusqu'à la prochaine mitose. Au contraire, de petits domaines de 30 à 40 kb formés



par les cohésines sont très stables tout au long du cycle cellulaire. Ces deux types de domaines sont régulés de façon inverse au cours du cycle cellulaire et sont mis en place indépendamment (**Figure 6** ; Tanizawa *et al.*, 2017).

## 4.2 Organisation 3D des génomes

Chez la drosophile il a été montré que les chromosomes sont organisés en domaines où la transcription est active et où on retrouve de nombreux gènes, des FTs et des protéines 'insulatrices' (protéines servant à protéger un promoteur ou une région régulatrice de l'action d'une protéine régulatrice agissant sur la région régulatrice voisine) et en domaines où la transcription est faible et où on ne retrouve pas de machinerie d'expression. Ces domaines diffèrent par leurs modifications épigénétiques et la présence de protéines spécifiques. La chromatine inactive est plutôt retrouvée dans les bordures alors que la chromatine active est plutôt retrouvée au milieu du noyau chez les mammifères à l'inverse des levures et des insectes (Lanctôt *et al.*, 2007). Des réarrangements chromosomiques peuvent avoir lieu durant le cycle cellulaire, mais ces derniers sont rares car la chromatine est très contrainte physiquement et ces réarrangements n'ont lieu que pendant de très courtes périodes de temps durant la phase G1. Une fois cette fenêtre temporelle passée, les mouvements chromatiniens sont réduits à des petits domaines subnucléaires (Lanctôt *et al.*, 2007). Chez la levure, les pores nucléaires sont d'importantes zones de transcription et

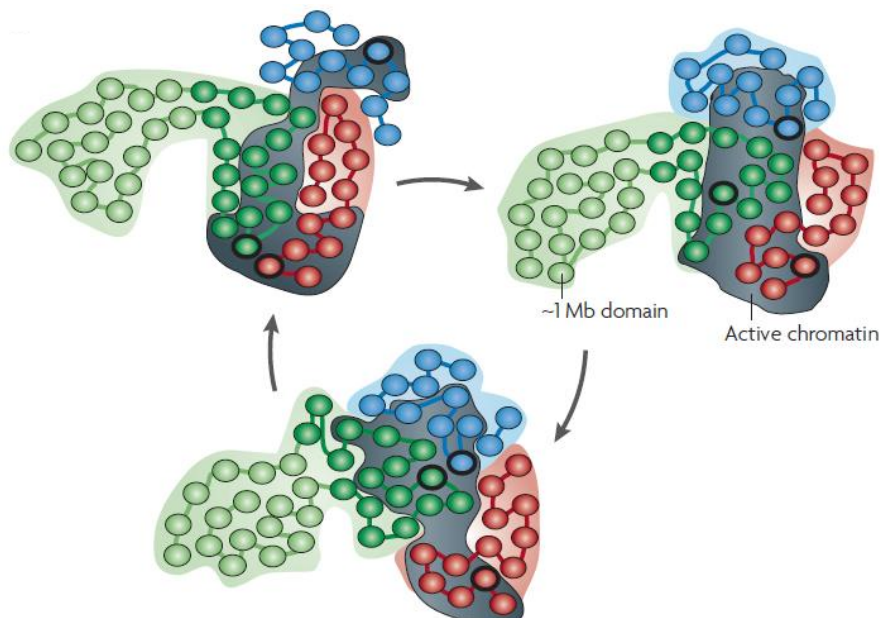


**Figure 6 : Organisation du noyau de mammifère.** Il se caractérise par une compartimentation des composants fonctionnels. L'enveloppe nucléaire qui contient des pores et repose sur un réseau de filaments intermédiaires appelés lamina nucléaire, les territoires chromosomiques et les compartiments interchromatiniens (CT-IT). La chromatine est organisée en différents CTs aussi appelés 'nuclear speckles'. (Issu de Lanctôt *et al.*, 2007).

des protéines appelées 'nuclear-pore-associated proteins' interagissent spécifiquement avec les gènes actifs. Les mammifères, au contraire, du fait de la grande taille de leur génome, empêchent la liaison des gènes actifs aux pores nucléaires par contrainte physique. Mais ces observations sont nuancées par l'existence chez la levure (*Saccharomyces cerevisiae*) de régions nucléaires périphériques transcriptionnellement inactives (Andrulis *et al.*, 1998). Les cellules de mammifères sont, elles, suspectées de subir d'autres formes de contraintes structurales du fait de la très grande quantité d'hétérochromatine qui les constituent (Casolari *et al.*, 2004; Schmid *et al.*, 2006; Lanctôt *et al.*, 2007).

### 4.3 Co-régulation des régions chromosomiques

Le cas le plus frappant de co-régulation de régions chromosomiques différentes est constitué des 'kissing-genes'. Il a été observé chez la *Drosophile* que deux versions du même gène étaient capables d'interagir physiquement notamment pour maintenir le silencing de certains gènes après la division cellulaire (Francis et Kingston, 2001). Chez *Drosophila melanogaster*, il a été montré que l'intégration ectopique de la même version d'un gène suffit à entraîner l'interaction physique entre les deux copies (Bantignies *et al.*, 2003). De plus il est connu que certains gènes interagissent avec des éléments régulateurs parfois situés très loin et même sur d'autres chromosomes (Lomvardas *et al.*, 2006). Il a été montré récemment que les gènes nécessitent d'être associés à des 'transcription factories', complexes protéiques enrichis en PolII, pour être exprimés. De plus le nombre de ces 'factories' étant limité, les gènes sont parfois contraints de partager ces complexes (Figure 7 ; Osborne *et al.*, 2004). Chez les champignons filamenteux, ces mécanismes ne sont pas encore très documentés. Le développement de nouvelles techniques permettant d'analyser les régions interagissant physiquement permette d'accroître la compréhension de l'organisation des chromosomes au sein des noyaux et de leur impact sur l'expression des gènes. On peut ici prendre l'exemple du Hi-C sequencing, technique qui se base sur le crosslinking du génome puis sur une étape de ligation, de



**Figure 7 : Interactions physiques successives de plusieurs régions chromosomiques avec la fraction active de la chromatine.** En vert, rouge et bleu, trois chromosomes différents. En gris est représentée la fraction active de la chromatine. On voit ici que la fraction active de la chromatine a une taille réduite et que pour pouvoir être exprimés correctement, les régions chromosomiques doivent interagir tour à tour avec la fraction active de la chromatine (Issu de Lanctôt *et al.*, 2007).



dé-crosslinking puis d'amplification et de séquençage. Mais alors que le 3C sequencing permet d'identifier des séquences cibles, le Hi-C permet de réaliser la carte des interactions à l'échelle du génome (**Figure 8** ; Lieberman-Aiden *et al.*, 2009). Chez *N.crassa*, il a été montré que les régions hétérochromatiniennes (surtout les régions d'hétérochromatine constitutive enrichies en H3K9me3) développent des liaisons physiques entre elles mais que l'inactivation de *KMT1* n'entraîne pas de modification de l'organisation 3D du noyau. Au contraire, les régions d'hétérochromatine facultative, enrichies en H3K27me3, sont essentielles au maintien de l'organisation 3D des génomes car elles se lient à la membrane nucléaire (Klocko *et al.*, 2016). Chez *N. crassa*, l'inactivation de *KMT1* et la perte des marques H3K9me3 entraîne une relocalisation des marques H3K27me3, qui, une fois relocalisées à l'emplacement des marques H3K9me3, ne participent pas à l'organisation 3D du génome. Ces résultats indiquent une robustesse de la structure 3D des génomes qui n'est pas mise à mal par l'inactivation de facteurs clefs impliqués dans l'établissement de la structure chromatinienne (Basenko *et al.*, 2015; Galazka *et al.*, 2016; Klocko *et al.*, 2016). L'organisation 3D du génome d'*E. festucae*, un champignon endophyte ayant une structure génomique en isochores, a été décrite récemment (Winter *et al.*, 2018). Les ETs ayant envahi le génome d'*E. festucae* ont été dégénérés par le RIP conduisant à la formation de régions riches en ETs enrichies en bases A et T (isochores AT) et de régions riches en gènes et enrichies en bases G et C (isochores GC ; Winter *et al.*, 2018). Winter *et al.* (2018) ont mis en évidence des contacts physiques entre les isochores AT et entre les isochores GC mais très peu d'interactions entre les deux types de domaines. De plus, les contacts entre régions AT sont très majoritaires (50 % de tous les contacts interchromosomiques), ce qui indique une structuration par les isochores AT dans ce type de génome. De façon intéressante, les isochores AT contiennent la plupart du temps un unique TAD alors que les isochores GC en contiennent plusieurs, ces TADs étant connus pour permettre une co-régulation génique (Ulianov *et al.*, 2016). Les isochores AT sont très enrichis en gènes fortement exprimés *in planta* et peu en condition de culture axénique, ce qui pourrait suggérer une régulation via les TADs des gènes impliqués dans la mise en place de l'infection (Winter *et al.*, 2018).

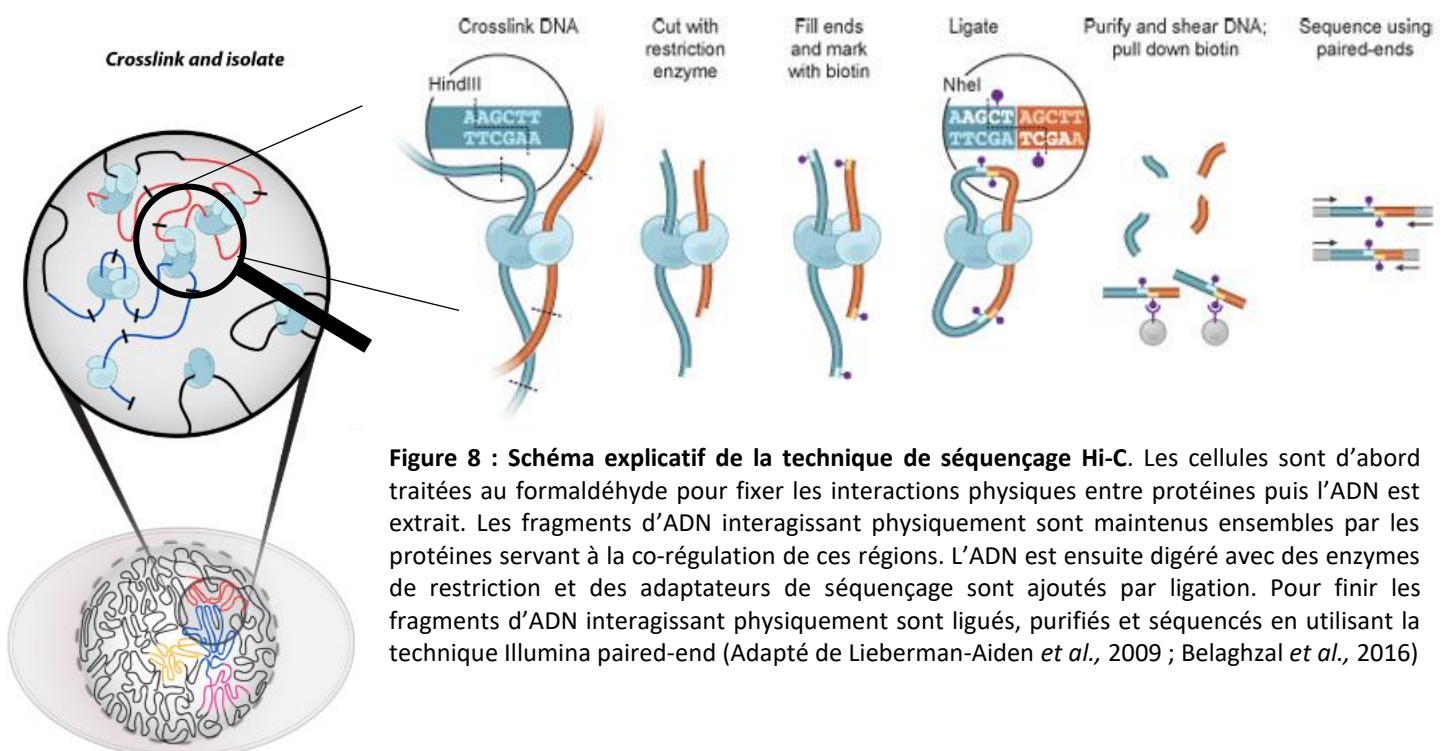
## 5. Facteurs de transcription et expression génique

Les facteurs de transcription (FTs) sont des protéines capables de se lier à l'ADN pour effectuer de nombreuses fonctions de régulation telles que le contrôle de l'expression génique, la réplication ou encore la réparation de l'ADN (Latchman, 1997). Les FTs, par ces mécanismes de liaison à l'ADN, permettent d'effectuer des fonctions essentielles à la vie de la cellule, régulant le développement, la croissance et la différenciation cellulaire (Pabo et Sauer, 1992). Les facteurs de transcription possèdent souvent plusieurs domaines dont un permettant la fixation à l'ADN et d'autres permettant de se lier à d'autres protéines ou cofacteurs. Les domaines de fixation à l'ADN (DNA-Binding Domain, DBD) sont des domaines conservés qui permettent de classer les FTs en différentes familles (Pabo et Sauer, 1992). Les facteurs de transcription agissent en amont des gènes. Ils peuvent avoir une fonction activatrice ou répressive sur le niveau de transcription du gène associé. Les FTs agissent en se fixant sur les sites de régulation de la transcription dans les promoteurs et recrutent ou empêchent la fixation des protéines de la machinerie de la transcription (**Figure 9** ; Thomas et Chiang, 2006).



## 5.1 Transcription et facteurs de transcription chez les Eucaryotes

Chez les Eucaryotes, il existe un grand nombre de polymérases capables de transcrire l'ADN en ARN. Le nombre de polymérases par génome est variable en fonction des organismes. On retrouve par exemple huit polymérases chez *Escherichia coli*, douze chez *Arabidopsis thaliana* et quinze chez l'homme (Garcia-Diaz et Bebenek, 2007). Toutes les fonctions des polymérases ne sont pas connues. Il y a trois polymérases majoritairement étudiées dont les fonctions sont bien décryptées, la Pol I, la Pol II et la Pol III. La Pol I est surtout impliquée dans la réparation de l'ADN et la production de ribosomes (Khatter *et al.*, 2017). La Pol II est responsable de la transcription des ARNm et de certains sRNA. La Pol III est responsable de la production de petits ARN comme le ribosome 5S ou les ARN de transferts (Garcia-Diaz et Bebenek, 2007). Les ARN polymérases ne sont pas capables d'initier la transcription seule. Elles ont besoin de cofacteurs, de co-activateurs et de FTs pour reconnaître le site d'initiation de la transcription, s'y fixer et initier la transcription (Patikoglou et Burley, 1997). Les Eucaryotes unicellulaires tels que les levures possèdent des régions régulatrices très courtes (quelques paires de bases à quelques centaines de paires de bases) positionnées juste à côté du promoteur (**Figure 10** ; Levine et Tjian, 2003). Chez les animaux, on retrouve des régions régulatrices plus complexes avec des éléments régulateurs en amont, en aval et dans les introns, et parfois même très éloignés de la séquence codante (jusqu'à 100 kb). Les gènes d'animaux possèdent la plupart du temps plusieurs enhanceurs avec un enhancer classique mesurant environ 500 pb. Ces régions régulatrices nécessitent l'action simultanée de complexes protéiques contenant de nombreux type de protéiques (facteurs de transcription, ARN polymérases, protéines de remodelage de la chromatine etc... ; Levine et Tjian, 2003). Chez les Eucaryotes, les FTs reconnaissent de courtes séquences dégénérées contrairement aux FTs des prokaryotes qui reconnaissent de larges motifs conservés (Wunderlich et Mirny, 2009). La reconnaissance de la région régulatrice par les FTs se fait via leur DBD. Il existe de nombreuses familles de FTs chez les Eucaryotes, une partie étant conservée entre les espèces, d'autres étant spécifiques de certains règnes (Shelest, 2008). Les gènes codant des FTs représentent 0.5 à 8 % du génome des Eucaryotes selon la complexité de

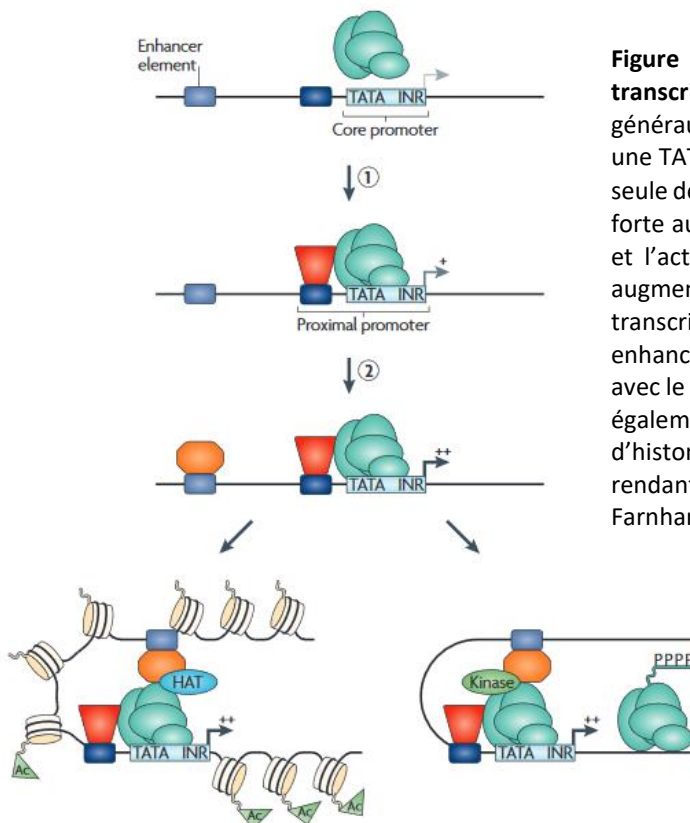


**Figure 8 : Schéma explicatif de la technique de séquençage Hi-C.** Les cellules sont d'abord traitées au formaldéhyde pour fixer les interactions physiques entre protéines puis l'ADN est extrait. Les fragments d'ADN interagissant physiquement sont maintenus ensemble par les protéines servant à la co-régulation de ces régions. L'ADN est ensuite digéré avec des enzymes de restriction et des adaptateurs de séquençage sont ajoutés par ligation. Pour finir les fragments d'ADN interagissant physiquement sont ligués, purifiés et séquençés en utilisant la technique Illumina paired-end (Adapté de Lieberman-Aiden *et al.*, 2009 ; Belaghal *et al.*, 2016)

l'organisme observé. Les familles les plus retrouvées chez les Eucaryotes sont les FTs à doigt de Zinc et les Zinc cluster (C2H2 et C4 zinc fingers ; ZnCys6/C6 zinc cluster), les helice-tour-helice (HTH), les basic leucine zipper (bZip) et les basic helice-tour-helice (bHLH ; Wunderlich et Mirny, 2009). Chez les plantes, on retrouve également les familles AP2, MADS box, WRKY et B3 (Figure 10 ; Weirauch et Hughes, 2011).

## 5.2 Les facteurs de transcription spécifiques du règne des Fungi

Les Fungi comme les autres Eucaryotes possèdent un arsenal de FTs permettant la régulation de l'expression génique. Park *et al.* (2008) ont ainsi identifié 31 832 FTs putatifs à partir de 62 espèces de champignons. Chez les champignons, on retrouve 37 superfamilles de domaines de liaison à l'ADN et parmi celles-ci 12 sont spécifiques des champignons (Figure 11 ; Shelest, 2008). On y retrouve des FTs à poing de cuivre 'copper fist domains' stabilisés par de multiples ions Cuivre et des FTs APSES qui pourraient avoir été acquis par transfert horizontal d'un KILA-N-like précurseur viral au début de l'évolution du règne des Fungi (Weirauch et Hughes, 2011). Chez *N. crassa*, un projet visant à inactiver les 10 000 gènes de son génome avait été initié : parmi les 312 gènes codant des FTs identifiés chez ce champignon, 242 ont été inactivés et 43 % des mutants obtenus avait un phénotype déficient pour au moins une fonction développementale (défauts de croissance mycélienne et/ou de conidiation), la majorité d'entre eux étant affectés dans plusieurs processus développementaux (morphologie, développement asexué, croissance, et développement sexué ; Colot *et al.*, 2006 ; Dunlap *et al.*, 2007 ; Carrillo *et al.*, 2017). On retrouve chez *N. crassa* 130 gènes codant des 'C6 binuclear cluster' qui constituent la plus grande famille de FTs chez ce champignon. La deuxième famille la plus représentée est constituée des C2H2 à doigt de Zinc avec 54 gènes. Ensuite, les familles les plus présentes chez *N. crassa* sont les bZIP (23 gènes), les MYB (16 gènes) et les bHLH (13 gènes). Une étude récente montre que chez six espèces pathogènes de l'homme seuls deux facteurs de réponse au stress oxydatif sont conservés, Yap1

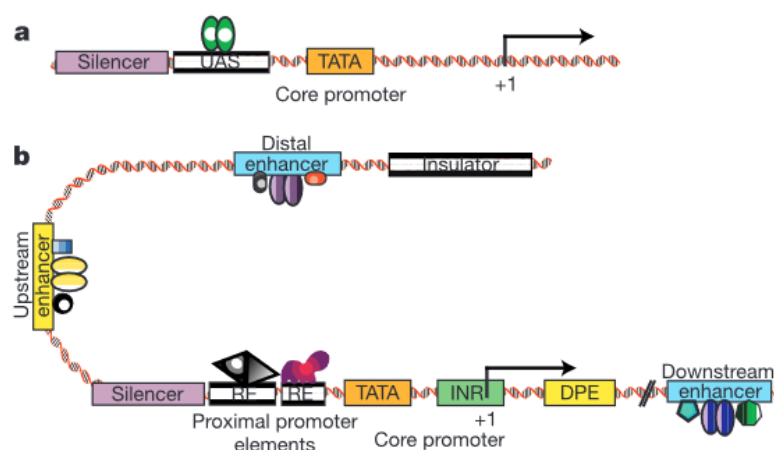


**Figure 9 : Schéma représentant l'action d'un facteur de transcription général (GTF).** Les facteurs de transcription généraux (ovales verts) se fixent au site de régulation du gène, une TATA box ou un initiateur (INR) par exemple. Mais la fixation seule de ce type de facteur de transcription n'induit pas une très forte augmentation de la transcription. Il permet le recrutement et l'action de FTs spécifiques (trapèze rouge ; étape 1) ce qui augmente fortement le niveau de transcription (+). La transcription peut être encore améliorée en recrutant un enhancer (octogone orange ; étape 2) qui stabilisera l'interaction avec le complexe d'initiation de la transcription. Enfin le GTF peut également permettre le recrutement de protéines modificatrices d'histones qui modifieront la structure de la chromatine la rendant plus accessible à la machinerie de transcription. (D'après Farnham, 2009).

et Skn7. Il semble donc que les FTs soient très espèces spécifiques en ce qui concerne le contrôle des facteurs de virulence fongiques (Pais *et al.*, 2016).

### 5.3 Contrôle transcriptionnel et pathogénie chez les champignons filamenteux

Les FTs fongiques peuvent avoir de multiples fonctions. Certains ont des rôles pléiotropes alors que d'autres ont des fonctions très ciblées et uniques. Ainsi chez *N. crassa* ou encore la levure (*S. cerevisiae*), le facteur de transcription StuA, FT général de la famille des APSES, est capable de réguler la reproduction ou encore le dimorphisme (ici, passage d'un état levure à un état filamenteux ; Gimeno et Fink, 1994 ; Aramayo *et al.*, 1996). Chez *Parastagonospora nodorum*, un champignon pathogène du blé de la famille des Dothidéomycètes, *SnStuA* a été identifié comme régulateur du métabolisme carboné et de la production de toxines. Ce facteur de transcription de la famille des APSES a été l'objet d'un transfert horizontal d'un virus vers le règne des champignons au début de leur histoire évolutive (Weirauch et Hughes, 2011 ; IpCho *et al.*, 2010). Le gène orthologue codant ce FT a été inactivé chez plusieurs autres ascomycètes (*Fusarium oxysporum*, *Glomerella cingulata*, *A. fumigatus*), entraînant une altération de la conidiation, de la croissance et même de la pathogénie (Ohara et Tsuge, 2004 ; Tong *et al.*, 2007 ; Sheppard *et al.*, 2005). Si certains facteurs de transcription comme StuA régulent de nombreuses étapes du développement fongique, d'autres tels que Ros1, un facteur de transcription de la famille des WOPR, régulent spécifiquement l'expression de gènes codant des effecteurs. Cette famille de FTs spécifique du règne des Fungi a été découverte chez *Fusarium oxysporum f. sp. lycopersici*, où ce FT avait été nommé Sge1. L'inactivation de Sge1 chez ce champignon mène à un défaut de pathogénie. Les spores du mutant Sge1 sont toujours capable de germer et de permettre la pénétration dans la tomate mais un défaut d'expression des gènes codant des effecteurs a été observé (Michielse *et al.*, 2009). Chez le basidiomycète *U. maydis*, Ros1 régule la conidiation et l'expression des effecteurs tardifs (Tollot *et al.*, 2016). Deux autres FTs regroupés au sein de la famille FTF (*Fusarium* Transcription Factors) ont été décrits chez *F. oxysporum* : FTF2 qui est présent chez tous les ascomycètes filamenteux et FTF1 qui est spécifiquement retrouvé, en plusieurs



**Figure 10 : Figure comparant des zones régulatrices chez un Eucaryote unicellulaire par rapport à un Eucaryote complexe. A :** unité transcriptionnelle chez un Eucaryote simple. Un core promoteur simple (TATA), un activateur amont (UAS), une séquence initiatrice (INR), un élément promoteur aval (DPE), des éléments promoteurs amonts (RE) et un silenseur, le tout entre 100 et 200 pb de la TATA box. **B :** Unité transcriptionnelle chez un Eucaryote complexe. Arrangement complexe de plusieurs enhancers clusterisés avec des silenceurs et insulateurs qui peuvent être situés en amont ou en aval à des distances allant de 10 kb à 50 kb de la TATA box (Issu de Levine et Tjian, 2003).

copies, chez *F. oxysporum*. L'inactivation de ces gènes induit un défaut de pathogénie ainsi qu'une diminution du niveau d'expression de gènes codant des effecteurs (effecteurs SIX, Secreted In Xylem) et de *SGE1* qui régule lui-même l'expression des gènes *SIX* (Niño-Sánchez *et al.*, 2016). Une étude récente confirme ce résultat chez *F. oxysporum*. En effet ce champignon possède un chromosome accessoire de pathogénie sur lequel sont localisés 13 FTs ainsi que de nombreux gènes codant des effecteurs. L'expression constitutive de *FTF1* ou *FTF2* induit l'expression de presque tous les gènes codant des effecteurs portés par le chromosome accessoire indiquant l'autonomie transcriptionnelle de ce chromosome qui peut être transféré horizontalement (van der Does *et al.*, 2016). Enfin un autre FT de la famille des Zn2Cys6 binuclear cluster, AbPf2, a été décrit comme impliqué dans le contrôle de l'expression de gènes codant des

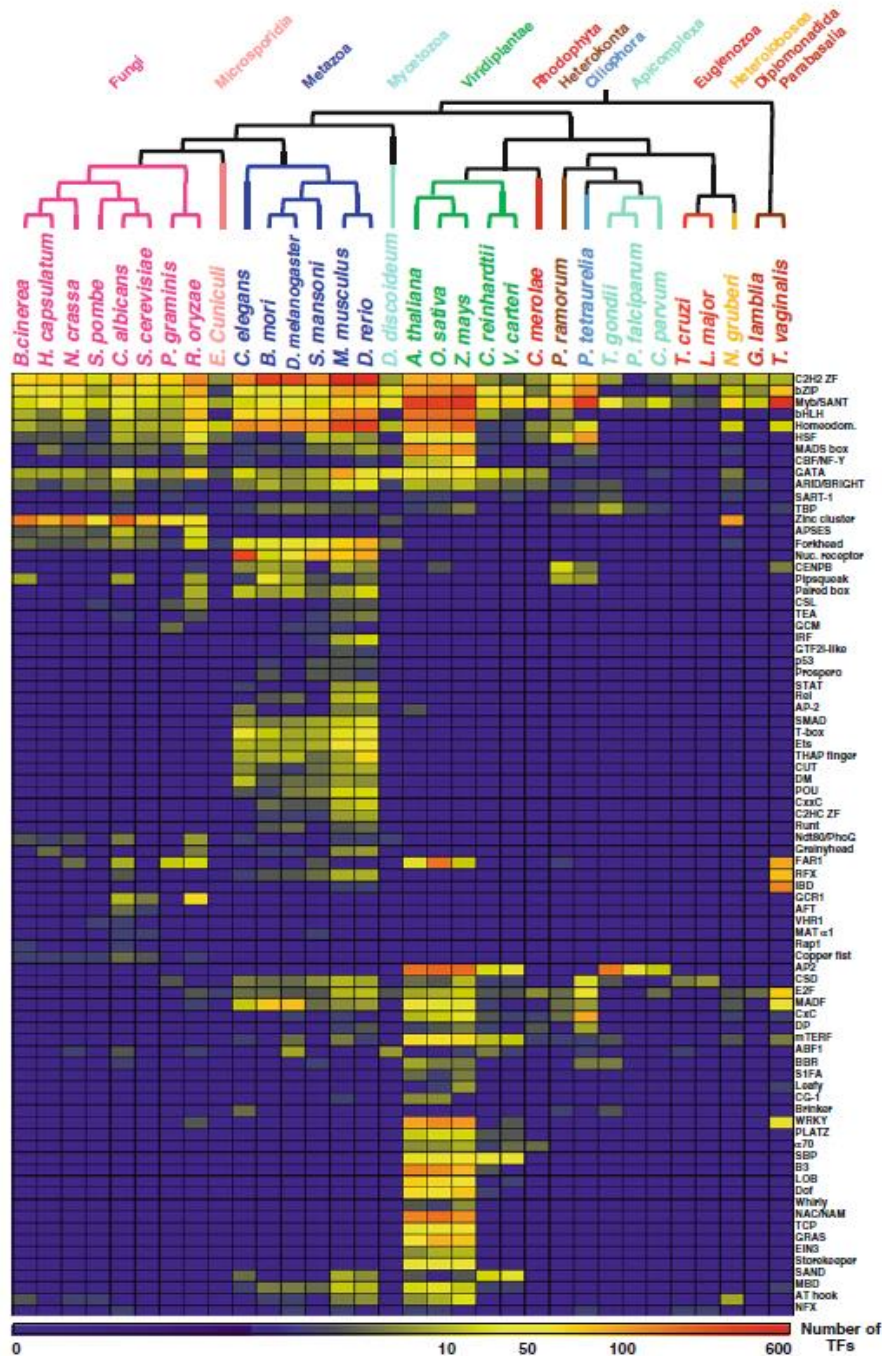


Figure 11 : Nombre de facteurs de transcription par règne et par espèce en fonction du type de FT. Heatmap réalisée à partir des domaines protéiques issus de Pfam. (D'après Weirauch et Hughes, 2011).



effecteurs chez *Alternaria brassicicola*. AbPf2 serait impliqué dans le contrôle de l'expression de 106 gènes dont 33 codent des SSPs (Cho *et al.*, 2013). Ce FT possède des homologues chez les Dothideomycetes et son inactivation chez deux espèces de Pleosporales pathogènes du blé (*P. nodorum* et *Pyrenophora tritici-repentis*) induit une diminution d'expression de gènes codant des toxines hôte-spécifique (SnToxA et PtrToxA) présentes chez les deux espèces et ayant été acquis par transfert horizontal de *P. nodorum* vers *P. tritici-repentis* (Rybak *et al.*, 2017). Cette inactivation n'entraîne ni défaut de croissance ni défaut de conidiation chez *A. brassicicola* et quelques défaut développementaux chez *P. nodorum* et *P. tritici-repentis* mais abolie totalement la capacité à infecter son hôte chez les trois espèces, ce qui indique la capacité de certains FTs à réguler uniquement la pathogénèse chez les ascomycètes phytopathogènes (Cho *et al.*, 2013 ; Rybak *et al.*, 2017).

## 6. *Leptosphaeria maculans*, champignon pathogène du colza : un bon modèle pour étudier les mécanismes contrôlant l'expression des gènes codant des effecteurs

### 6.1 Le colza, une brassicacée très cultivée mais ciblée par des agents pathogènes

Le colza, *Brassica napus*, appartient à la famille des Brassicaceae et est issu d'un croisement inter-spécifique entre deux espèces diploïdes : le chou, *Brassica oleracea* (2n = CC), et la navette, *Brassica rapa* (2n = AA), ayant eu lieu il y a 7 000 à 12 500 ans (Chalhoub *et al.*, 2014). Le colza est donc un allotétraploïde (AACC). La culture du colza remonte à l'Antiquité, notamment en Chine. Elle aurait été introduite en Europe vers le XIIIème siècle (Hebinger, 2013). Son huile servait à la fois à l'alimentation et à d'autres usages comme l'éclairage ou la savonnerie. On utilise actuellement non seulement le colza pour son huile mais également pour la production de tourteaux, galettes protéinées co-produits de l'extraction d'huile utilisés pour l'alimentation animale. Des programmes de sélection variétale ont permis au colza de perdre ses fortes teneurs en acide érucique et en glucosinolates qui provoquaient des baisses d'appétence, particulièrement chez les bovins, et induisaient des troubles thyroïdiens chez les animaux monogastriques (Bonneuil et Thomas, 2012). Suite à ce programme d'amélioration du colza, la culture du colza a connu un fort essor dès 1980. On cultive aujourd'hui en France 1,48 MHa de colza d'hiver (<http://agreste.agriculture.gouv.fr>). Le colza d'hiver est une plante annuelle. Semé de la mi-août à la fin septembre en Europe de l'Ouest, le colza connaît d'abord une première phase de levée puis reste végétatif jusqu'à fin novembre/décembre où il rentre dans une phase vernalisante. La croissance reprend à la sortie de l'hiver (février à mars), quand les températures augmentent. La tige s'allonge rapidement et de nouvelles feuilles se mettent en place ainsi que de nouvelles racines. La maturation des graines se termine en juillet (Hebinger, 2013). Les pertes de rendement dues à des bioagresseurs sont généralement causées par des champignons et des insectes ravageurs. Parmi les principaux agents pathogènes fongiques, on retrouve *Sclerotinia sclerotiorum* et *L. maculans* (agent causal de la nécrose du collet, ou phoma). D'autres champignons comme *Cylindrosporium concentricum* (agent responsable de la cylindrosporiose), *Erysiphe cruciferarum* (agent responsable de l'oïdium), *Plasmodiophora brassicae* (responsable de la hernie des crucifères) ou *Alternaria brassicae* (responsable de l'alternatiose) causent également des symptômes sur colza mais avec une incidence moindre ou une répartition locale (Source : [www.terresinovia.fr/colza/cultiver-du-colza/maladies](http://www.terresinovia.fr/colza/cultiver-du-colza/maladies)). Le sclérotinia est une maladie relativement



bien contrôlée par les fongicides à l'inverse de la nécrose du collet qui nécessite l'utilisation de variétés génétiquement résistantes pour éviter des pertes pouvant aller jusqu'à 20% de la récolte (Fitt *et al.*, 2006).

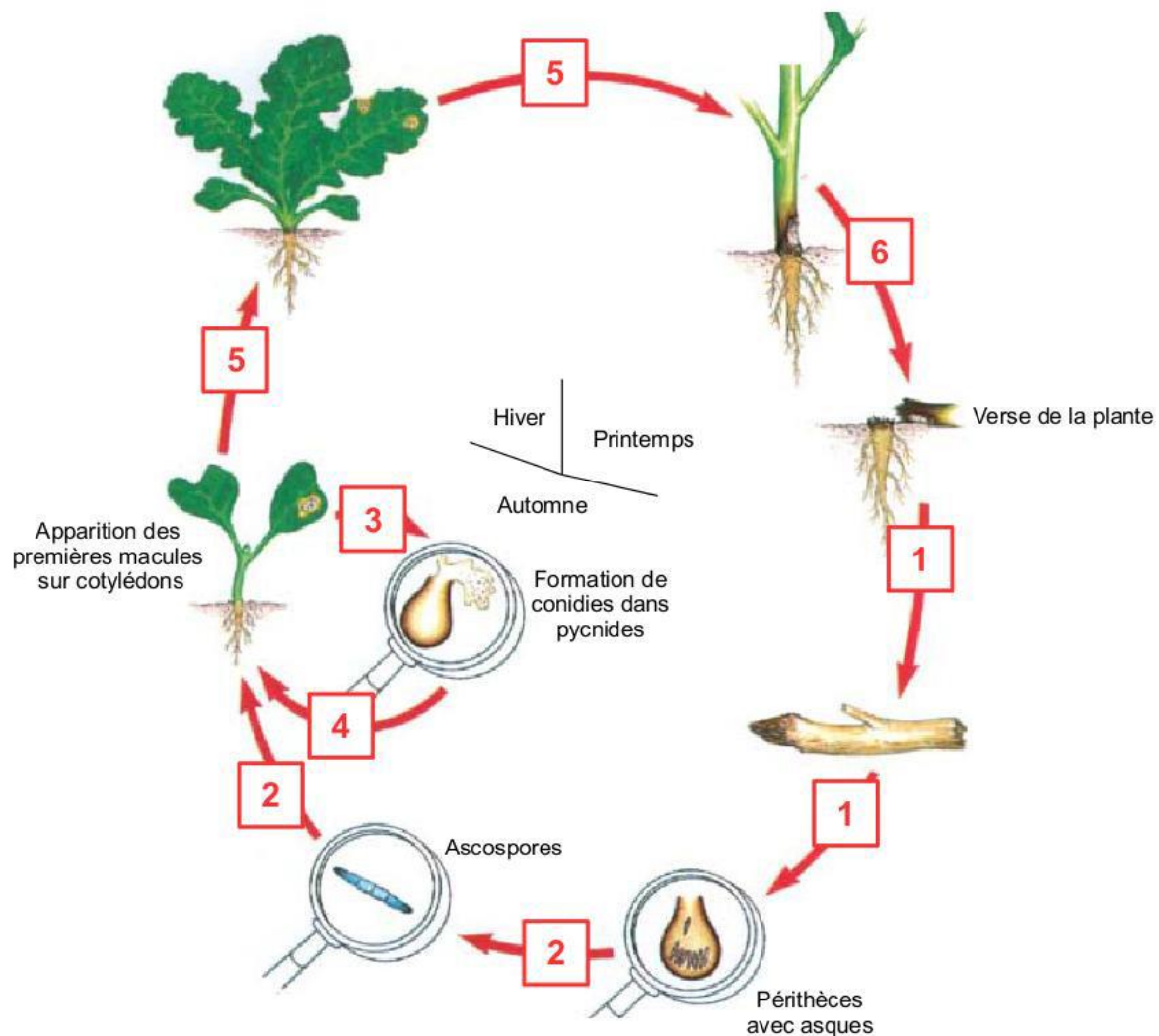
## 6.2 Le champignon pathogène responsable de la nécrose du collet (ou «blackleg disease»)

*Leptosphaeria maculans* 'brassicae' (*L. maculans* dans la suite de l'introduction) est un champignon Ascomycète, appartenant à la classe des Dothidéomycètes et à l'ordre des Pléosporales. Ce champignon est responsable de la nécrose du collet du colza et d'autres crucifères, pouvant occasionner des pertes de rendement allant jusqu'à 20 % de la récolte (Fitt *et al.*, 2006). En Europe, *L. maculans* effectue sa reproduction sexuée à l'automne sur les résidus de culture de l'année précédente puis les ascospores, dispersées par le vent, peuvent coloniser les premières feuilles au champ. Après germination des ascospores, l'hyphe mycélien pénètre dans les feuilles via les stomates ou des blessures. Après une phase de développement asymptomatique, des nécroses foliaires apparaissent. Ces macules foliaires, caractérisées par la présence de conidies (inoculum secondaire), produits de la reproduction asexuée, vont permettre de propager l'agent pathogène à courte distance. A partir de ces nécroses, le mycélium colonise de façon systémique et endophyte les espaces intercellulaires des tissus des feuilles, du pétiole et de la tige, sans provoquer de symptômes. Après cette phase endophyte de plusieurs mois, *L. maculans* devient nécrotrophe et provoque alors une nécrose de la base de la tige, qui peut conduire à la verse parasitaire des plantes (**Figure 12** ; Williams et Fitt, 1999 ; West *et al.*, 2002 ; Rouxel *et al.*, 2004). *L. maculans* possède un génome riche en ETs (33 % du génome) qui se caractérise par la présence de deux types de compartiments : des isochores GC représentant deux tiers du génome pauvres en ETs et riches en gènes (contenant 95 % des gènes) et des isochores AT, représentant un tiers du génome, pauvres en gènes et riches en ETs (Rouxel *et al.*, 2011). Une première prédiction du répertoire d'effecteurs de *L. maculans* (correspondant à des PPS de moins de 300 aa) avait permis d'identifier 651 effecteurs putatifs : 122 codés par des gènes localisés en isochores AT et 529 par des gènes localisés en isochores GC. Les gènes codant des effecteurs situés en isochores AT représentent donc 20 % des gènes prédits dans ces régions contre seulement 4 % dans le cas des isochores GC, suggérant que les isochores AT sont enrichis en gènes codant des effecteurs (Rouxel *et al.*, 2011). Plus récemment, un nouvel assemblage, une nouvelle annotation du génome de *L. maculans* et une nouvelle prédiction du répertoire de gènes codant des effecteurs ont été réalisés et utilisés pour une partie des analyses réalisées au cours de la thèse (Dutreux *et al.*, 2018 ; E.J. Gay thèse en cours).



### 6.3 Expression synchronisée des effecteurs chez *L. maculans*

Une première vague d'expression d'effecteurs (incluant les gènes d'avirulence clonés chez *L. maculans*) a été identifiée durant les étapes précoces de l'infection (effecteurs 'précoces' ; Rouxel *et al.*, 2011 ; Gervais *et al.*, 2017). Les gènes codant des effecteurs situés en isochores AT présentent une faible expression pendant la croissance mycélienne et une forte induction de leur expression pendant la phase de colonisation asymptomatique des cotylédons et des feuilles de colza, avec un pic d'expression 7 jours post infection (jpi). Les gènes codant pour des effecteurs précoces sont majoritairement localisés dans les isochores AT puisqu'à 7 jpi, 73 % des gènes codant des effecteurs en isochoire AT sont sur-exprimés contre 19 % des mêmes gènes en isochoire GC (Rouxel *et al.*, 2011). Par ailleurs, 8 % des gènes codant des effecteurs localisés en isochoire AT sont exprimés *in vitro* contre 55 % de ceux localisés en isochoire GC. Une deuxième vague d'effecteurs spécifiquement exprimés pendant les phases de colonisation systémique et de nécrose de la tige (effecteurs 'tardifs') a récemment été identifiée grâce à des analyses RNA-Seq. Les gènes codant ces effecteurs sont, eux, localisés majoritairement dans les isochores GC. 307 gènes (ont été identifiés comme spécifiquement sur-exprimés en condition tardive d'infection (Gervais *et al.*, 2017). Enfin, une



**Figure 12 : Cycle de vie de *Leptosphaeria maculans* en Europe.** 1 : phase saprophyte sur résidus de culture ; 2 : reproduction sexuée produisant des ascospores constituant l'inoculum primaire ; 3, 4 : courte phase nécrotrophe suivant l'infection primaire caractérisée par l'apparition de macules foliaires et multiplication asexuée produisant des conidies ; 5 : longue phase endophyte avec colonisation systémique de la tige et du collet ; 6 : phase nécrotrophe à l'origine de la verse de la plante. Adapté de CETIOM (1996).



analyse RNA-Seq est actuellement menée sur l'ensemble du cycle de vie de *L. maculans* (incluant infection en conditions contrôlées et infection au champ ; EJ. Gay, thèse en cours) et permet d'identifier de nouveaux profils d'expression de gènes codant des effecteurs. La localisation distincte des gènes codant des effecteurs 'précoces' ou 'tardif' au sein du génome de *L. maculans* 'brassicae' a des conséquences sur leur évolution dans les populations de *L. maculans* (les gènes codant des effecteurs 'tardifs' étant mieux conservés dans les populations de *L. maculans* que ceux codant des effecteurs précoces ; Rouxel *et al.*, 2011 ; Gervais *et al.*, 2017) et sur le contrôle de leur expression (cf. **partie 6.5**).

#### 6.4 Le complexe d'espèces *L. maculans* / *L. biglobosa*

*L. maculans* fait partie d'un complexe comprenant au moins deux groupes génétiquement distincts initialement décrits comme respectivement « virulent » et « avirulent » (McGee et Petrie, 1978). Grâce à l'avènement des outils moléculaires, ces deux groupes ont été reconnus comme étant deux espèces différentes : *L. maculans* et *Leptosphaeria biglobosa* (Shoemaker et Brun, 2001). Sur colza, ces deux espèces se développent sur les mêmes hôtes et peuvent être retrouvées simultanément sur la même plante. Elles sont toutes les deux capables de provoquer des lésions nécrotiques et de coloniser les tissus des plantes hôtes mais leur impact sur les cultures ne sont pas similaires : seule *L. maculans* est susceptible de provoquer la verse parasitaire et donc d'être dommageable pour les cultures (Williams et Fitt, 1999 ; West *et al.*, 2002 ; Rouxel *et al.*, 2004), Contrairement à *L. maculans*, *L. biglobosa* cause des lésions de la tige ne conduisant qu'à des dégâts mineurs. Chez *L. maculans*, de nombreuses interactions spécifiques (interactions gène-pour-gène) ont été montrées avec ses différents hôtes alors que pour le moment aucune interaction de ce type n'a été décrite chez *L. biglobosa* (Dilmaghani *et al.*, 2009). Des relations phylogénétiques entre les membres du complexe d'espèces ont été établies sur la base d'analyses de plusieurs séquences nucléiques (*ITS*, *actine*,  *$\beta$ -tubuline*), révélant un niveau de complexité inattendu en comparaison à la simple séparation en deux espèces (Mendes-Pereira *et al.*, 2003 ; Voigt *et al.*, 2005 ; Vincenot *et al.*, 2008). En effet, ces résultats montrent que *L. maculans* est une espèce très monomorphe avec uniquement une séparation en deux sous-clades dont l'un, *L. maculans* 'lepidii' (Lml), est très rare et représenté seulement par quelques individus isolés sur *Lepidium sp.*, une crucifère adventice. *L. biglobosa* est une espèce très polymorphe avec une séparation en six sous-clades correspondant à différentes spécificités d'hôtes ou origines géographiques : *L. biglobosa* 'brassicae' présent sur plusieurs espèces de Brassicaceae en Eurasie et aux Etats-Unis, *L. biglobosa* 'canadensis' retrouvé principalement sur les cultures de colza en Amérique et en Australie, *L. biglobosa* 'occiaustralensis' présent sur colza et chou en Australie et au Mexique, *L. biglobosa* 'australensis' présent sur colza en Australie, *L. biglobosa* 'thlaspii' isolé sur *Thlaspii arvense* au Canada et le très rare *L. biglobosa* 'erysimii' isolé sur *Erysimum spp.* au Canada (Mendes-Pereira *et al.*, 2003 ; Voigt *et al.*, 2005 ; Vincenot *et al.*, 2008 ; Van de Wouw *et al.*, 2009 ; Dilmaghani *et al.*, 2009, 2012). Étonnamment les autres espèces du complexe possèdent des répertoires d'effecteurs de tailles similaire à celui de Lmb malgré leur taille de génome plus réduite avec 875 gènes codant des effecteurs prédits chez *L. biglobosa* dont 75 d'entre eux localisés dans les isochores AT de son génome (Dutreux *et al.*, 2018). Les génomes de *L. biglobosa* 'thlaspii', *L. biglobosa* 'brassicae', *L. biglobosa* 'canadensis' et *L. maculans* 'lepidii', contiennent respectivement, 4 %, 3,2 %, 2,9 % et 2,7 % d'ETs (Grandaubert *et al.*, 2014). Mais une étude récente a permis d'identifier que le génome de *L. biglobosa* 'brassicae' pouvait contenir jusqu'à



15 % d'ETs (Dutreux *et al.*, 2018). De façon surprenante, *L. maculans* 'lepidii' possède un génome quasi identique à celui de *L. maculans* 'brassicae' avec une structuration en isochores moins marquée que celle de *L. maculans* 'brassicae' mais un répertoire d'effecteurs similaire en nombre. En revanche, malgré le nombre similaire d'effecteurs, 41 % des gènes codant des effecteurs localisés en isochores AT chez *L. maculans* 'brassicae' lui sont spécifiques, ce qui pourrait indiquer que l'invasion du génome de *L. maculans* 'brassicae' par des ETs est un mécanisme permettant de diversifier le répertoire d'effecteurs. *L. maculans* 'lepidii' possède, lui aussi, un répertoire d'effecteurs spécifique par rapport aux autres espèces du complexe (Grandaubert *et al.*, 2014).

### 6.5 Contrôle chromatinien de l'expression des gènes codant des effecteurs chez *L. maculans*

Il y a une dizaine d'années, un projet visant à déterminer l'impact de la structure de la chromatine sur la régulation de l'expression des gènes, en particulier des gènes impliqués dans la pathogénie, chez *L. maculans* a été initié. Il était alors établi que i) *L. maculans* possédait une structuration complexe de son génome, avec une alternance de régions riches en ETs et de régions riches en gènes ; ii) que les régions riches en ETs, bien que pauvres en gènes, étaient enrichies en gènes codant des protéines d'avorulence et d'autres effecteurs putatifs, et iii) que ces gènes étaient spécifiquement exprimés *in planta*. Dans ce contexte, l'hypothèse établie était que les isochores AT pouvaient correspondre à des régions hétérochromatiniennes et que la condensation de la chromatine de ces régions pouvait être à la base d'un processus efficace de régulation de l'expression de gènes codant des effecteurs permettant leur expression concertée *in planta*, et leur répression *in vitro*. Pendant la thèse de J. Soyer, une caractérisation fonctionnelle de deux protéines, HP1 et KMT1, a été menée. Ces deux protéines sont indispensables à la formation de l'hétérochromatine via la tri-méthylation H3K9me3. Le silencing par ARNi de *HP1* et *KMT1* a conduit à une augmentation d'expression *in vitro* de 41 % des gènes codant des effecteurs localisés en isochores AT, ce qui suggère que la structure de la chromatine dans ces régions contribue à une répression de l'expression de ces gènes lors des phases de croissance végétative. Cette augmentation d'expression s'accompagne pour au moins deux gènes codant des effecteurs (*AvrLm4-7* et *AvrLm1*) d'une diminution de la marque H3K9me3. Cependant, l'augmentation d'expression n'atteint pas le niveau d'expression observé lors de l'infection primaire du colza par une souche sauvage (Soyer *et al.*, 2014). L'ensemble de ces résultats suggère que la structuration de la chromatine et la régulation dynamique de cette structure ne serait pas le seul acteur impliqué dans la régulation de l'expression des gènes codant des effecteurs et qu'un, ou plusieurs, FT(s) spécifique(s) seraient également impliqués dans le contrôle de l'expression des effecteurs. Suite à cette étude, le modèle suivant a été proposé (**Figure 13**) : pendant la croissance axénique, les régions génomiques contenant les gènes codant des effecteurs situés en isochores AT sont condensées, ce qui induit une répression de l'expression génique. Au moment de l'infection, le changement de mode de vie et le shift vers la pathogénèse se traduirait par une décondensation de la chromatine (via l'élimination des marques H3K9me3), dont le déterminisme est inconnu, rendant les promoteurs géniques accessibles aux FTs. Une action de FTs spécifiques permettrait ensuite d'exprimer de façon concertée les gènes codant des effecteurs. Ce mécanisme permettrait une adaptation rapide aux nouvelles conditions d'infection en permettant une régulation similaire de l'expression de nombreux gènes, ayant pour fonction de faciliter d'infection, et localisés dans un environnement génomique similaire (Soyer *et al.*, 2015. A).



## 6.6 Contrôle transcriptionnel de l'expression des gènes codant des effecteurs chez *L. maculans*

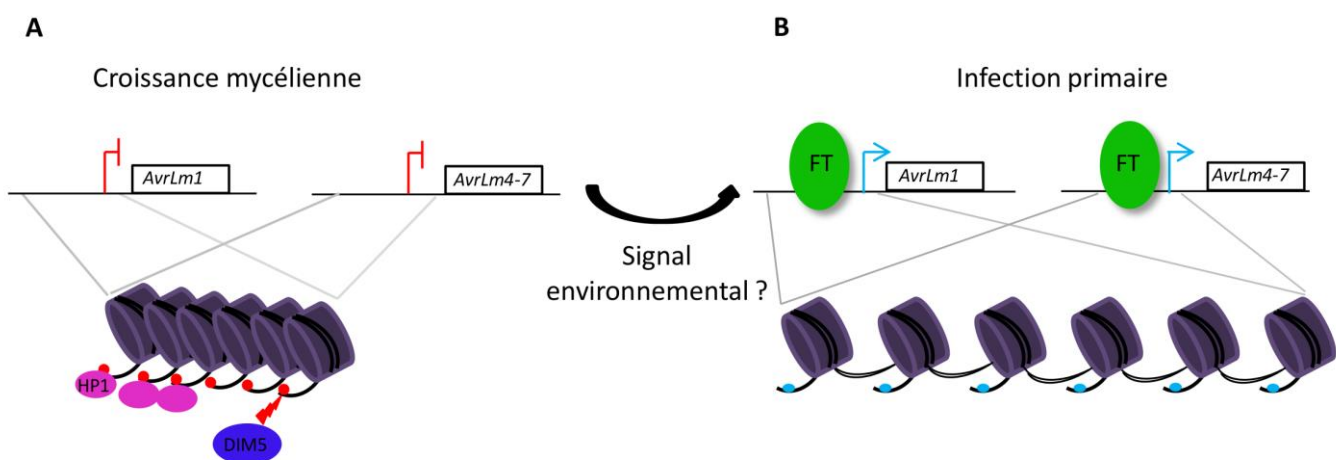
L'étude du contrôle transcriptionnel de l'expression des gènes codant des effecteurs chez *L. maculans* a également été initiée lors de la thèse de J. Soyer. Afin d'identifier un ou des FTs contrôlant l'expression des gènes codant des effecteurs de *L. maculans*, deux approches avaient été choisies : (i) une approche avec *a priori* consistant à identifier les gènes orthologues de FTs décrits chez les champignons comme impliqués dans le contrôle de l'expression des effecteurs (e.g. *Sge1*, *StuA*, cf. **partie 5.3**) et (ii) une approche sans *a priori* consistant à prédire le répertoire de FTs de *L. maculans* et à rechercher les FTs surexprimés pendant l'infection ou spécifiques de *L. maculans*. En recherchant des domaines caractéristiques des FTs comme décrit par Park *et al.* (2008), J. Soyer a pu identifier un répertoire de 682 FTs putatifs. Après avoir éliminé les FTs ayant un domaine de fixation à l'ADN mais sans prédiction d'initiation de transcription, une liste de 355 TFs putatifs a été obtenue. Après comparaison des séquences de ces FTs avec les génomes du complexe d'espèces, une liste de 44 FTs spécifiques de *L. maculans* a pu être obtenue. Par ailleurs, une analyse transcriptomique à 3, 7 et 14 jpi de cotylédons de colza a permis d'identifier 38 FTs surexprimés pendant l'infection dont deux FTs surexprimés à 3 jpi, 14 à 7 jpi et 21 à 14 jpi. Le silencing des 14 gènes codant des FTs exprimés spécifiquement à 7 jpi et de *StuA* et *Sge1* a été entrepris mais, en dehors de *StuA* et de deux FTs surexprimés pendant l'infection (gènes ID : *Lema\_P013470.1* et *Lema\_P074920.1*), le silencing n'avait pas été efficace (pas de diminution d'expression génique). L'inactivation de *LmStuA* entraînait des défauts de sporulation, de croissance et de pathogénie et une diminution d'expression d'au moins deux gènes codant des effecteurs 'précoces' (Soyer *et al.*, 2015. B).

## 6.7 Objectifs de la thèse

Actuellement, la façon dont l'expression concertée des gènes codant des effecteurs fongiques est régulée est mal comprise. *L. maculans* est un bon modèle pour l'étude du contrôle de l'expression des effecteurs : (i) il possède un cycle de vie complexe, alternant les modes de vie et les stratégies nutritionnelles au cours de son cycle. Ces capacités d'alternance entre modes de vie indiquent un système de régulation sophistiqué de ses gènes ; (ii) le génome de *L. maculans* présente une structure en isochores, alternant régions riche en gènes et régions riches en ETs enrichies en gènes codant des effecteurs 'précoces' ; (iii) au moins deux vagues de gènes codant des effecteurs sont exprimées au cours de l'infection (effecteurs 'précoces' produits pendant l'infection primaire et effecteurs 'tardifs' produits pendant la colonisation des tiges) ; (iv) *L. maculans* 'brassicae' appartient à un complexe d'espèces dont la sous-espèce la plus proche, *L. maculans* 'lepidii', ne présente pas de structure de génome en isochores, et un contenu faible en éléments répétés, et n'est pas pathogène du colza.



Le travail réalisé par J. Soyer durant sa thèse a permis d'initier une étude de l'influence de la structure de la chromatine et de FTs spécifique sur le contrôle de l'expression des effecteurs chez *L. maculans* 'brassicae'. Il a été ainsi mis en évidence un rôle de KMT1 et HP1 dans la mise en place de la marque H3K9me3 et le contrôle de l'expression d'effecteurs localisés dans les régions riches en ETs. L'inactivation par ARNi de *KMT1* et *HP1* a permis d'obtenir une diminution de la modification H3K9me3 dans le promoteur d'au moins deux gènes codant des effecteurs situés dans des régions riches en ETs, accompagnée d'une augmentation de l'expression de ces gènes *in vitro*. Néanmoins, le niveau d'expression de ces gènes *in vitro* chez les transformants n'atteignant pas le niveau d'expression observée pendant l'infection, chez la souche sauvage, ceci suggère l'intervention supplémentaire de FTs spécifiques au moment de l'infection. Ces FTs se fixeraient sur les promoteurs des gènes codant des effecteurs après décondensation de la chromatine (**Figure 13**).



**Figure 13 : Modèle de contrôle de l'expression des gènes codant des effecteurs situés en isochores AT chez *Leptosphaeria maculans*.** **A** : Pendant la croissance mycélienne, un mécanisme épigénétique reposant sur une modification des histones associée à l'hétérochromatine, H3K9me3, *via* l'action de LmHP1 et LmDIM5 (i.e. KMT1), entraîne une condensation de la chromatine dans l'environnement génomique des gènes et ainsi une répression globale de leur expression. **B** : Pendant l'infection primaire, un signal de l'environnement ou endogène, de nature inconnue, est reconnu par le champignon, permettant de lever la répression, *via* la disparition de la marque H3K9me3, et un espacement des nucléosomes, autorisant l'accès des promoteurs à un, ou plusieurs, FT(s) spécifique(s) permettant une expression concertée pendant l'infection. Boules roses, modification des histones associée à l'hétérochromatine (ici H3K9me3) ; boules bleues, modification des histones associée à l'euchromatine.

L'objectif de ma thèse est donc de comprendre le rôle de la structure de la chromatine (à plusieurs échelles : position des nucléosomes, modification des histones et organisation intra-nucléaire des chromosomes) et de FTs dans le contrôle de l'expression concertée des gènes codant des effecteurs de Lmb, et plus globalement dans le contrôle de l'expression génique et la pathogénie. Dans ce but, deux thématiques de recherche sont développées en parallèle : (i) l'étude de l'influence de la structure de la chromatine sur la régulation de l'expression génique et de la pathogénie et (ii) la recherche de facteurs de transcription impliqués dans la régulation des effecteurs au cours de l'infection. Cet objectif nécessite la mise en place d'approches complémentaires : des analyses bio-informatiques et des analyses fonctionnelles.

Afin de réaliser des analyses fonctionnelles de FTs candidats ou de protéines impliquées dans le remodelage de la chromatine, j'ai utilisé une technique de mutagenèse permettant de s'affranchir des limites de la technique d'ARNi utilisée précédemment dans l'équipe. En effet, l'analyse fonctionnelle par ARNi nécessite le criblage d'un grand nombre de transformants car le niveau de diminution d'expression est très variable. Les transformants silencés

généralisés par Soyer *et al.* (2014) permettait encore une expression résiduelle de ~20 % de *KMT1*, ce qui ne permettait pas de déterminer les effets d'une absence totale de H3K9me3 sur la structure chromatinienne et l'expression génétique. La technologie CRISPR-Cas9 a récemment été mise au point pour *L. maculans* afin de générer des mutants knock-out (**Figure 14** ; Iduurm *et al.*, 2017). J'ai pu transférer avec succès la méthode CRISPR-Cas9 au laboratoire et inactiver plusieurs gènes codant des histones méthyltransférases, des déméthylases d'histone et des FTs.

## 1. Etude de l'influence de la structure de la chromatine sur la régulation de l'expression génique et de la pathogénie

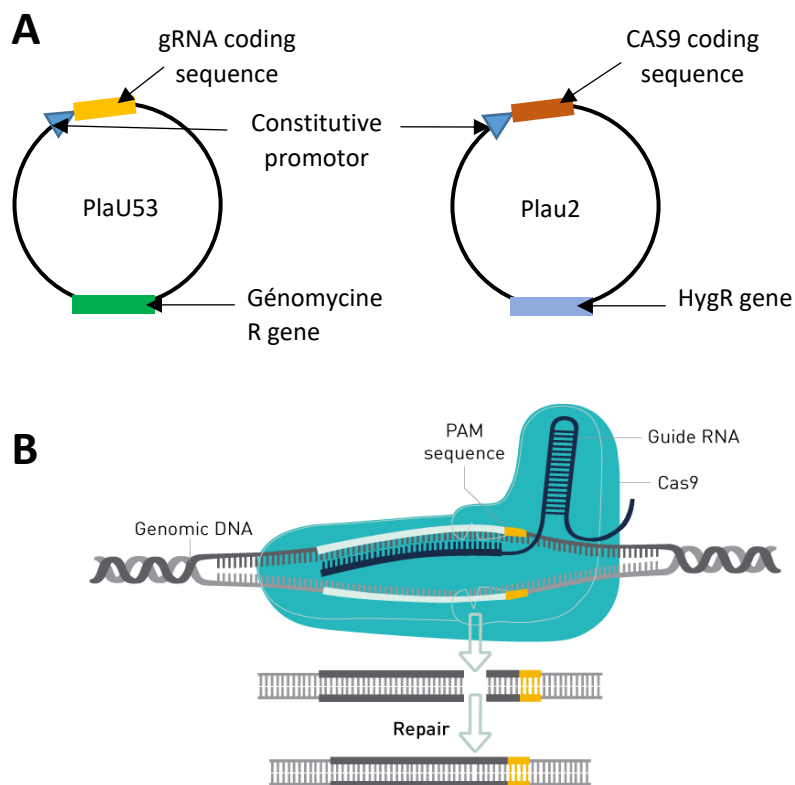
Les études réalisées précédemment dans l'équipe concernant l'organisation de la structure de la chromatine étaient menées par des approches ciblées (ChIP-qPCR) et concernaient l'effet des modifications biochimiques des histones sur l'expression des gènes. Au cours de cette thèse, grâce notamment au développement de nouvelles technologies et à la diminution des coûts de séquençage, nous avons pu aller plus loin et développer des analyses «genome-wide» et, autant que possible, comparatives, de i) la position des nucléosomes chez quatre espèces de champignons phytopathogènes présentant des modes de vie distincts (en développant des approches de MAINE-Seq et RNA-Seq), (**Chapitre 1.1**) ; ii) la localisation des modifications des histones H3K4me2, H3K9me3 et H3K27me3 chez deux espèces du complexe d'espèces ayant des structures de génome différentes, *Lmb* et *Lml* (par des approches de ChiP-Seq et RNA-Seq), (**Figure 13** ; **Chapitre 1.2**). La combinaison de ces approches permet d'obtenir une vue complète du paysage épigénomique de *L. maculans* 'brassicae' et de son influence sur l'expression des gènes, en particulier des gènes codant des effecteurs. La comparaison de l'organisation épigénomique du génome de *L. maculans* 'brassicae' avec d'autres espèces a pour objectif d'identifier des particularités d'organisation de ce génome, ou, au contraire, ce qui est conservé, et de déterminer dans quelle mesure la structure du génome et de la chromatine influent sur l'adaptation et la spécialisation à l'hôte. Par ailleurs, afin de décrypter l'influence des modifications hétérochromatiniennes H3K9me3 et H3K27me3 sur l'expression génique, des analyses fonctionnelles de trois gènes responsables de modifications chimiques de queues d'histone, *KMT1*, *KMT6* (responsable respectivement de la mise en place de la marque répressive H3K9me3 et H3K27me3) et *Kdm8* (responsable de l'enlèvement de la marque H3K36me2) ont été réalisées (**Chapitre 2.1 et 2.2**). iii) l'organisation intra-nucléaire des chromosomes (par une approche de HiC-Seq). L'étude de l'organisation intranucléaire des chromosomes, par Hi-C sequencing (**Figure 8**), a été initiée. Les données relatives à cette approche seront présentées dans le **chapitre 2.1**.

## 2. Recherche de facteurs de transcription impliqués dans la régulation des effecteurs au cours de l'infection

L'approche d'identification de FTs sans *a priori* avait permis pendant la thèse de J. Soyer l'identification et l'annotation de 355 FTs chez *Lmb*, dont 38 sont sur-exprimés pendant l'infection primaire du colza et huit sont spécifiques de *Lmb*. Parmi ces FTs, nous avons choisi de focaliser notre analyse fonctionnelle sur des FTs sur-exprimés pendant l'infection primaire, en particulier deux FTs spécifiques de *Lmb* et un portant des signatures de sélection positive au sein du complexe d'espèces *L. maculans* / *L. biglobosa*. J'ai réalisé l'analyse fonctionnelle de ces trois FTs en les inactivant grâce à la technologie CRISPR-Cas9 et en les sur-exprimant sous le contrôle d'un promoteur *EF1a* : parmi les trois FTs étudiés, l'un, *LmPf2*, s'est avéré impliqué dans la pathogénèse et le contrôle de l'expression des effecteurs précoces. En effet, l'inactivation par CRISPR-Cas9 de *LmPf2*, qui a un profil d'expression suivant celui



des effecteurs précoces, déjà décrit chez trois Dothidéomycètes (*P. tritici* 'repentis', *A. brassicicola* et *P. nodorum*) comme impliqué dans la pathogénie a été réalisée ainsi qu'une sur-expression de ce même gène dans deux fonds génétiques différents (WT et  $\Delta kmt1$ ). L'expression génique chez les transformants ainsi obtenus a été étudiée par RNA-Seq, ce qui m'a permis de déterminer que ce FT ainsi que de l'état de compaction de la chromatine étaient impliqués dans le contrôle de l'expression des gènes codant des effecteurs (**Chapitre 3.1**). L'inactivation de deux autres FTs présentant des patterns d'expression correspondant à celui des effecteurs 'précoces' n'a pas permis de mettre en évidence une implication dans la pathogénèse du colza (**Chapitre 3.2**).



**Figure 14: Méthode de mutagenèse dirigée CRISPR-Cas9. A:** Constructions utilisées pour mettre en place la mutagenèse dirigée. A gauche le plasmide portant la séquence de l'ARN guide ciblant le gène d'intérêt et à gauche le vecteur contenant la séquence de la Cas9 reconnaissant l'ARN guide. **B:** Schéma de l'action combinée de l'ARN guide reconnaissant sa séquence cible et de la Cas9 reconnaissant le gRNA et coupant l'ADN en amont de la séquence PAM. L'ADN sera ensuite réparé ce qui mènera dans certains cas à des mutations délétères pour le gène cible.

# RESULTATS

## CHAPITRE I

### Etude genome-wide du paysage chromatinien et nucleosomal



## INTRODUCTION

Dans le chapitre I nous nous intéresserons au contrôle chromatinien de l'expression des gènes, en particulier des gènes codant des effecteurs.

Dans la première partie de ce chapitre, nous avons étudié, chez différentes espèces de champignons phytopathogènes, l'organisation du paysage nucléosomique ainsi que l'impact des positions des nucléosomes sur l'expression génique. La position des nucléosomes, en condition de vie axénique, a été analysée chez quatre espèces (*L. maculans* 'brassicae', *L. maculans* 'lepidii', *Fusarium graminearum* et *Botrytis cinerea*) via une expérience de MAINE-Seq consistant à digérer la chromatine ne formant pas de nucléosomes à l'aide d'une enzyme, la micrococcale nucléase. L'ADN enroulé autour des octamères d'histones est ensuite extrait, séquencé et aligné sur le génome de chacune des espèces pour identifier les portions d'ADN au niveau desquelles sont positionnés des nucléosomes. En parallèle, une analyse transcriptomique a été réalisée dans les mêmes conditions par RNA-Seq. Nous avons pu ensuite relier paysage nucléosomique et expression chez les différents champignons étudiés et comparer ceci entre les différentes espèces. A mon arrivée au laboratoire, les expériences de MAINE avaient déjà été réalisées et les données de séquençage venaient d'être obtenues. L'analyse bio-informatique des données n'avait pas été réalisée. J'ai réalisé l'analyse des données de *L. maculans* 'brassicae' et *L. maculans* 'lepidii', participé aux réflexions sur le choix et la mise en place des outils à utiliser pour l'analyse et j'ai participé à la rédaction de l'article. Il s'agit à notre connaissance de la première analyse comparative chez des champignons filamenteux.

Dans la deuxième partie de ce chapitre nous avons mené, en phase de croissance axénique, des expériences de CHIP-Seq qui consistent, après une étape de digestion de la chromatine, à immunoprécipiter les nucléosomes avec des anticorps spécifiques des modifications des histones de type H3K9me3, H3K4me2 et H3K27me3. Ces expériences ont été réalisées chez *L. maculans* 'brassicae' et *L. maculans* 'lepidii', deux espèces ayant un contenu en ETs différent ainsi qu'une différence dans leur spectre d'hôte connu. Une analyse transcriptomique par RNA-Seq a été menée dans les mêmes conditions pour relier modifications chromatinienne et expression génique chez ces deux espèces. A mon arrivée au laboratoire, les expériences de CHIP avaient déjà été réalisées et les données de séquençage venaient d'être obtenues. J'ai participé à l'analyse des données, aux réflexions sur le choix et la mise en place des outils à utiliser pour l'analyse et j'ai participé à la rédaction de l'article.





## RESULTATS

### ARTICLE 1

#### Nucleosome patterns in four plant pathogenic fungi with different genome structures



## 1 Nucleosome patterns in four plant pathogenic fungi with different genome structures

2 C. Clairet<sup>1,2</sup>, N. Lapalu<sup>1</sup>, A. Simon<sup>1</sup>, J. L. Soyer<sup>1,3</sup>, M. Viaud<sup>1</sup>, E. Zehraoui<sup>4</sup>, B. Dalmais<sup>1</sup>, I. Fudal<sup>1,\*</sup> and N. Ponts<sup>4,\*</sup>

3 <sup>1</sup>UMR BIOGER, INRA, AgroParisTech, Université Paris-Saclay, Avenue Lucien Brétignières, BP 01, F-78850 Thiverval-Grignon, France

4 <sup>2</sup>Université Paris-Sud, 91405, Orsay, France

5 <sup>3</sup>Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, and Christian-Albrechts  
6 University of Kiel, Am Botanischen Garten 1-9, 24118 Kiel, Germany

7 <sup>4</sup>UR1264-MycSA, INRA, Villenave d'Ornon, 33140, France

8 \* To whom correspondence should be addressed.

9 Nadia Ponts Tel: +33557122482; Email: [nadia.ponts@inra.fr](mailto:nadia.ponts@inra.fr);

10 Isabelle Fudal Tel: +33130814590; Email: [Isabelle.Fudal@inra.fr](mailto:Isabelle.Fudal@inra.fr);

## 11 INTRODUCTION

12 Fungi account for a huge part of the Earth biodiversity; they rank second in terms of species number, behind  
13 the insects, with a current estimate of 2.2 to 3.8 million species (Hawksworth and Lücking, 2018). Fungi are organisms  
14 of major environmental importance as they develop beneficial symbiotic associations with plants and are able to decay  
15 dead organic matter (Taylor and Osborn, 1994). Unfortunately, fungi are also very efficient pathogens causing  
16 important damages in agriculture, human health, and environment (Fisher *et al.*, 2012). Control of fungal diseases on  
17 crops necessitates a global understanding of fungal pathogenicity determinants. Among these pathogenicity  
18 determinants, fungi secrete an arsenal of molecules known as effectors, key elements of pathogenesis which modulate  
19 innate immunity of the plant and facilitate infection, corresponding to small proteins, secondary metabolites and small  
20 RNAs (Fox and Howlett., 2008; Möbius and Hertweck, 2009; Weiberg *et al.*, 2013; Lo Presti *et al.*, 2015; Wang *et al.*,  
21 2016). Upon plant infection, fungi undergo a tightly controlled transcriptional reprogramming, and different sets of  
22 effectors are expressed at specific stages of pathogen development and host colonization. It follows that fine-tuned  
23 regulation of effector-encoding genes is important for the successful outcome of plant infection. However, so far, very  
24 little is known about the regulation of effector gene expression. Effector gene expression can be regulated by several  
25 mechanisms including specific transcription factors (reviewed in Tan and Oliver, 2017) and the structure of the  
26 genomic regions in which effector genes are located. Plant-associated fungi generally show contrasted genomic  
27 landscapes including 'plastic' regions with a high prevalence of transposable elements (TEs). These genomes either  
28 show an overall large proportion of TEs evenly distributed throughout the genome, or TEs clustered in specific regions  
29 such as long repeat-rich blocks, accessory chromosomes or subtelomeric areas (Soyer *et al.*, 2015,b). Effector-encoding  
30 genes are overrepresented in these TE-rich regions. TE-rich compartments have heterochromatin properties contrary  
31 to TE-poor regions which have euchromatin properties. The location of effector genes in regions enriched in repeats  
32 has been shown to provide a tight control of effector gene expression through chromatin remodeling. Indeed, several



33 recent studies pointed out the potential role of chromatin remodeling in the regulation of pathogenicity and the  
34 control of secondary metabolism (reviewed in Strauss and Reyes-Dominguez, 2011; and in Soyer *et al.*, 2015,b).

35 Eukaryotic chromatin is packaged into nucleosomes, each composed of DNA wrapped around a histone  
36 octamer associated with various other proteins, and separated by linker DNA (Richmond and Davey, 2003). These  
37 histone proteins are composed of histone core where the DNA is wrapped and histone tails which can be chemically  
38 modified by specific enzymes changing the chromatin 3D structure and DNA accessibility to polymerases and  
39 transcription factors (TFs). Positioning of nucleosomes throughout the genome and post-translational modifications  
40 of histones have a significant regulatory function by modifying availability of binding sites to transcription factors (TFs)  
41 and to polymerases, affecting DNA-dependent processes such as transcription, DNA repair, replication and  
42 recombination (Radman-Livaja and Rando, 2010; Struhl and Segal, 2013). Nucleosome positioning and occupancy are  
43 determined by a combination of DNA sequence features, TFs, chromatin remodelers and histones modifiers. Genome-  
44 wide maps of nucleosome occupancy and positioning are still sparse in fungi and have only been developed in  
45 Hemiascomycota (*Saccharomyces cerevisiae* (Yuan *et al.*, 2005) and other Hemiascomycota yeast species (Tsankov *et al.*  
46 *al.*, 2010), in the ascomycete *Aspergillus fumigatus* (Nishida *et al.*, 2009) and the basidiomycete *Mixia osmundae*  
47 (Nishida *et al.*, 2012). The studies revealed that nucleosomes are depleted at promoter, enhancer and terminator  
48 regions, allowing access to TFs, and that the nucleosomal DNA length distribution was similar in *M. osmundae* and *A.*  
49 *fumigatus* but differed from that of hemiascomycetous yeasts. In contrast, no comparative genome-wide analyses of  
50 nucleosome positioning have been performed in ascomycetes and notably not in plant pathogenic fungi.

51 In the present study, we investigated genome-wide nucleosome localization and its impact on gene  
52 expression, particularly on (putative) effector gene expression, in four different plant pathogenic fungi showing  
53 different host ranges, lifestyles and genomic organizations: i) *Leptosphaeria maculans* 'brassicae' (Lmb), a  
54 hemibiotrophic pathogen of oilseed rape; ii) the most closely related species of Lmb, *Leptosphaeria maculans* 'lepidii'  
55 (Lml), a pathogen of *Lepidium* spp.; iii) *Fusarium graminearum*, a hemibiotrophic pathogen of cereals and iv) *Botrytis*  
56 *cinerea*, a polyphagous necrotrophic pathogen causing grey mould on more than 1400 plant species. The genome of  
57 Lmb has been invaded by TEs (which represent more than 39% of its genome) and is composed of alternating  
58 compartments: gene-rich GC-isochores and TE-rich AT-isochores (Rouxel *et al.*, 2011; Dutreux *et al.*, 2018). On the  
59 contrary, genomes of Lml, *F. graminearum* and *B. cinerea* have a low TE-content with TE-rich regions mainly restricted  
60 to sub-telomeric domains (Grandaubert *et al.*, 2014; Amselem *et al.*, 2011; Cuomo *et al.* 2007; King *et al.* 2015). The  
61 genome of the reference strain of *B. cinerea* B05.10 contains 4% of TE which are localized essentially in the telomeric  
62 and centromeric regions of the core chromosomes, or on the two dispensable chromosomes (Porquier *et al.*, 2016).  
63 The genome of *F. graminearum* contains very little repeated sequences and TE, 1% and 0.3 % respectively (King *et al.*,  
64 2018).

65 The aim of this study is to describe nucleosome repartition patterns on whole genome depending on the  
66 lifestyle of the fungus. To gain insight into the role of nucleosome positioning and occupancy in regulating fungal  
67 pathogen transcription, we applied micrococcal nuclease digestion of mono-nucleosomes (MNase-assisted isolation



68 of nucleosomes or MAINE) with regards to mRNA abundance to concomitantly map nucleosome-rich regions and  
69 transcriptionally active regions during fungal growth in contrasted media.

## 70 MATERIAL AND METHODS

### 71 Strains and culture conditions

72 *Leptosphaeria maculans* 'brassicae' v23.1.3 and *Leptosphaeria maculans* 'lepidii' IBCN84 mycelia were  
73 inoculated into 100 mL of Fries liquid medium (1 g/L NH<sub>4</sub>NO<sub>3</sub>, 5 g/L C<sub>4</sub>H<sub>12</sub>N<sub>2</sub>O<sub>6</sub>, 1 g/L KH<sub>2</sub>PO<sub>4</sub>, 0.5 mg/L MgSO<sub>4</sub> 7H<sub>2</sub>O,  
74 130 mg/L CaCl<sub>2</sub>, 100 mg/L NaCl, 30 g/L C<sub>12</sub>H<sub>22</sub>O<sub>11</sub> and 5 g/L Yeast extract). Tissues were harvested after growing for  
75 seven days in the dark at 27°C. *Botrytis cinerea* strain B05.10 (10<sup>6</sup> spores/mL) was grown for two days on two different  
76 solid media: Malt Medium (MM, 20 g/L malt extract, 5 g/L yeast extract and 15 g/L agar), and Grape Juice (GJ, pH  
77 adjusted to 4; 15 g/L agar) media, both covered with a cellophane layer (Simon *et al.*, 2013; Kelloniemi *et al.*, 2015).  
78 The plates were incubated in a growth chamber (Sanyo MLR-350H) at 23°C with an alternation of 14 h of white light  
79 and 10 h of darkness. After two days of culture, mycelia were ground in liquid nitrogen and stored to -80°C until further  
80 processing. *Fusarium graminearum* strain CBS185.32 (Centraal Bureau voor Schimmelcultures, Utrecht, the  
81 Netherlands) was grown for three days in modified liquid MS (glucose was substituted with sucrose) as previously  
82 described (Boutigny *et al.*, 2009). All cultures were done in three biological replicates.

### 83 Preparation of nucleosomal DNA

84 Fungal material was harvested and immediately treated with micrococcal nuclease (MNase, cat. #MS0247S,  
85 New England BioLabs). For *L. maculans* 'brassicae' and 'lepidii', 100 mg of mycelium were digested with 5 µL of MNase  
86 for 10 min at 37 °C, directly followed by DNA purification (Soyer *et al.*, 2015a). For *F. graminearum*, mycelia were  
87 harvested by filtering and immediately ground for 1 min at 30 Hz using a TissueLyzer (Qiagen). Then, 100 mg of ground  
88 mycelium was digested for 10 min at 37°C with 15 µL of MNase in 600 µL of digestion buffer (0.6% v/v IGEPAL, 50 mM  
89 NaCl, 2 mM Pefabloc, 50 mM Tris-HCl pH8, 10 mM CaCl<sub>2</sub>). The reactions were stopped with EDTA and the samples  
90 treated with RNase followed by proteinase K prior DNA purification with phenol/chloroform and ethanol precipitation.  
91 For *B. cinerea*, we digested 100-200 mg of mycelium per sample with 1 µL MNase at 37°C (Soyer *et al.*, 2015,a), for 10  
92 min (MM cultures) or 30 min (GJ cultures). The reactions were stopped with EDTA and samples treated with RNase A  
93 followed by proteinase K. DNA purification was realized with the "Nucleospin Gel and PCR clean up kit" (Macherey  
94 Nagel, cat #740609.250). For all samples, nucleic acid quantification was performed by UV spectrometry using a  
95 Nanodrop-ND 1000 apparatus, and digestion profiles were checked by 2% agarose gel electrophoresis. Nucleosomal  
96 DNA for stored at -20°C waiting for DNA library preparation.

### 97 Extraction of total RNA

98 For *L. maculans* 'brassicae' and *L. maculans* 'lepidii', total RNA was extracted from mycelium grown for one  
99 week in Fries liquid medium as previously described (Fudal *et al.*, 2007). For *F. graminearum*, mycelia were harvested  
100 by filtering, rinsed twice with sterile deionized water, and flash frozen in liquid nitrogen. One milliliter of TRIzol™  
101 Reagent (Thermo Fischer Scientific) was added to 200 mg of mycelium before grinding for 1.5 min at 30 Hz using



102 a TissueLyzer (Qiagen). Total RNA was then extracted using a previously published protocol (Hallen *et al.*, 2007). For  
103 *B. cinerea*, total RNA was extracted from frozen ground mycelium using a previously published protocol (Kelloniemi *et*  
104 *al.*, 2015). All total RNA samples were stored at -80°C until sequencing RNA library preparation.

105 **Preparation of sequencing libraries, high-throughput sequencing, and read pre-processing**

106 MNase-seq libraries were prepared from purified nucleosomal DNA using the kit NEBNext Ultra DNA Library  
107 Prep Kit for Illumina (cat. # E7370L New England BioLabs) following the manufacturer's instructions. The NEBNext Ultra  
108 Directional RNA Library Prep Kit for Illumina (cat. # E7420L New England BioLabs) was used to prepare all RNA-seq  
109 libraries, following the manufacturer's instructions. Sequencing was performed by the GenomEast platform, a member  
110 of the 'France Génomique' consortium (ANR-10-INBS-0009). Samples were run in 9-plex on an Illumina HiSeq 4000 in  
111 paired mode, 2x50 bp. Initial read quality check was performed using FastQC  
112 (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Raw reads were then pre-processed with  
113 Trimmomatic v0.32 (Bolger *et al.*, 2014) to clip out any remaining sequencing adapter sequence and crop low quality  
114 nucleotides (minimum accepted Phred score of 30). Reads in pairs of 40 bp or more in length were used in the present  
115 analysis.

116 **Transcriptome analyses**

117 RNA-seq reads were mapped against their respective reference genomes (see **Table 1**) using STAR v2.5.1  
118 (Dobin *et al.*, 2013). For *F. graminearum*, genome sequence and annotation files were downloaded from FungiDB v44  
119 (Stajich *et al.*, 2012; Basenko *et al.*, 2018). TPM counts (Transcripts Per Million reads; Li *et al.*, 2010) were computed  
120 using the count TPM tool provided with the MAINE-Seq Tool Suite (MSTS) (<https://github.com/nlapalu/MSTS>). For *B.*  
121 *cinerea*, differential expression analysis was performed using SARTools commands (Varet *et al.*, 2016) with DESeq2  
122 settings (Love *et al.*, 2014).

123 **MAINE analyses**

124 MNase-seq paired-end reads were mapped using Bowtie2 software ran in very-sensitive mode (Langmead and  
125 Salzberg, 2012). MSTS (<https://github.com/nlapalu/MSTS>) was used to compute all phasograms, dinucleotide  
126 composition, as well as nucleosome density profiles of genomic compartments and/or gene lists. Lists of near-universal  
127 single copy orthologs were obtained by running BUSCO3 for Fungi (Simão *et al.*, 2015; Waterhouse *et al.*, 2017) on  
128 each reference genome studied. Graphical visualisations were computed with MSTS and Matlab R2019a (MathWorks)

129

130

131

132

133



134 **Table 1: Characteristics of reference genomes for the four fungal species studied.**

Species	<i>Leptosphaeria maculans</i> 'brassicae'	<i>Leptosphaeria maculans</i> 'lepidii'	<i>Botrytis cinerea</i>	<i>Fusarium graminearum</i>
Strain	v23.1.3	IBC84	B05.10	CBS85.32
Reference genome	v23.1.3 (Rouxel et al., 2011)	IBC84 (Grandaubert et al., 2014)	B05.10 (Amselem et al., 2011)	PH-1 vRR5 (Cuomo et al. 2007; King et al 2015 and 2018)
Availability			EnsemblFungi v1	FungiDB v44
Genome size (Mb)	45.1	31.5	42	38.1
Number of protein-coding genes	12,468	11,272	11,701	14,164
GC content (%)*	44.1	50.9	42	48.2
Transposable elements content (%)	34.2	4.0	4.0	0.29

135 \*excluding N's and mitochondrial genomes; \*\*also excluding the largely repetitive 3' end of chromosome 4

## 136 RESULTS

### 137 Establishing nucleosome landscapes of the Pezizomycotina *L. maculans* 'brassicae', *L. maculans* 'lepidii', *B. cinerea*, 138 and *F. graminearum*

139 We investigated the nucleosome landscapes of four fungal species of the Ascomycota subdivision  
140 Pezizomycotina (*L. maculans* 'brassicae', *L. maculans* 'lepidii', *B. cinerea*, and *F. graminearum*) by MAINE coupled to  
141 deep sequencing (see **Table S1** for descriptive sequencing metrics). Each sample was analysed in triplicate at more  
142 than 80-fold coverage depths and nearly 70X genome coverage by 147 bp-long nucleosome footprints (defined as the  
143 core coverage by Valouev et al. (2011), depths of sequence sufficient for in-depth characterization of nucleosome  
144 positioning (Valouev et al., 2011).

#### 145 Genome-wide nucleosome spacing

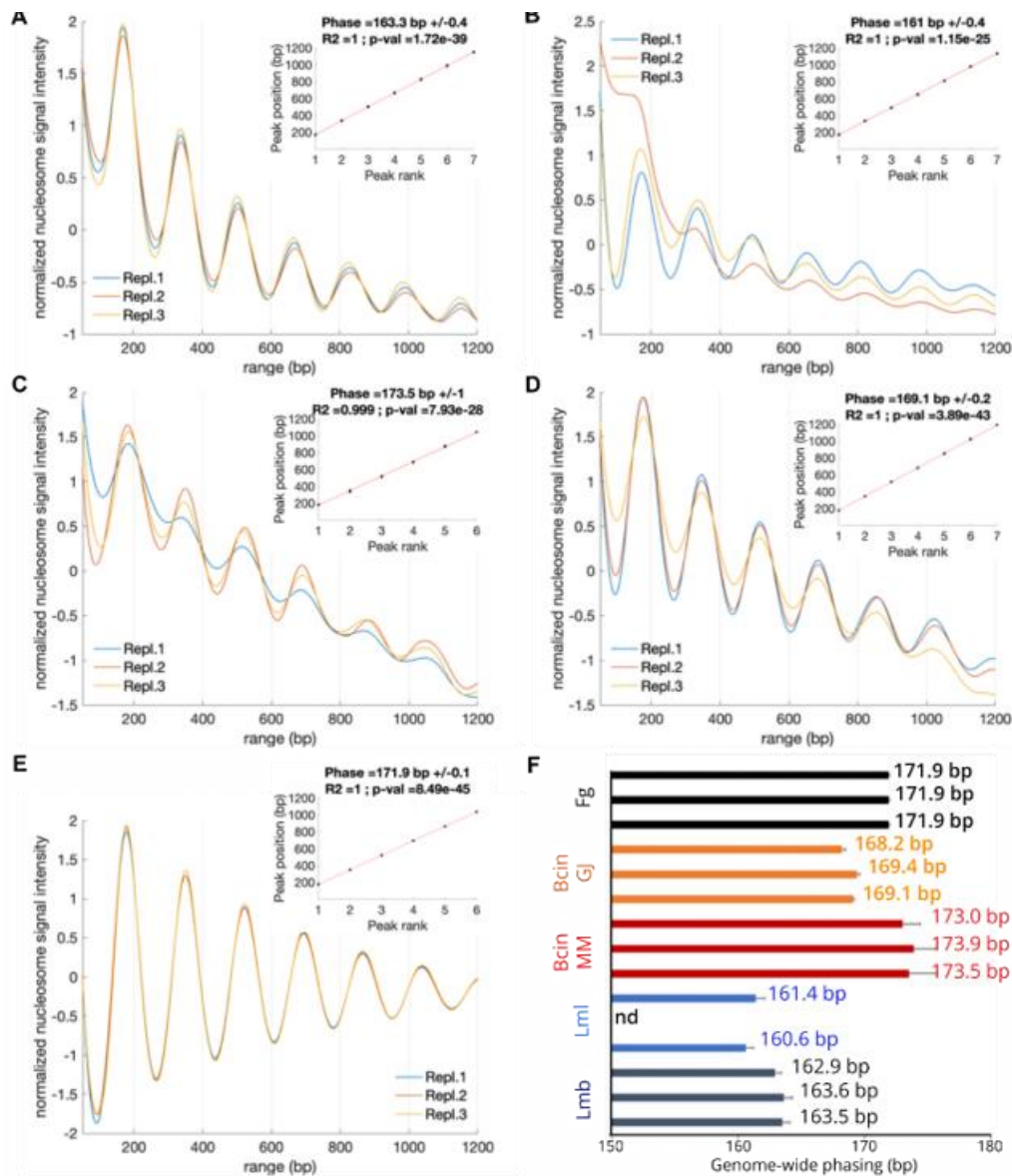
146 We first explored nucleosome landscapes in our four fungal genomes by measuring the average distance between  
147 nucleosomes genome-wide; we computed phasograms, i.e., frequency distributions of read coverage per base  
148 genome-wide for all four species (**Figure 1**). Phasograms obtained in nucleosome mapping resemble oscillating sine  
149 wave signals which period is the length of DNA bound to the histone octamer plus the length of the DNA stretch to  
150 the next nucleosome, averaged genome wide (Valouev et al., 2011). Phasing signals were observed genome-wide over



151 1,200 bp sliding windows revealing six to seven nucleosome peaks in a wave signal decaying in intensity with increasing  
 152 distance and significant linear regression on peak apex positions (**Figure 1A-E**). We found that, in the fungi studied,  
 153 nucleosomes are 161 to 173 bp distant from each other (centre to centre), also called nucleosome repeat length or  
 154 NRL (*i.e.*, the length of DNA wrapped around the histone octamer plus linker DNA), depending on the considered  
 155 species and culture condition. In *L. maculans* 'brassicae' and *L. maculans* 'lepidii' (**Figures 1A and 1B**), NRL is 163 bp  
 156 and 161 bp, respectively. In *Botrytis cinerea*, NRL is 173 bp when grown in malt-based medium (MM, **Figure 1C**), and  
 157 169 bp when a nutrient-rich grape juice-based medium (GJ) is used (**Figure 1D**). In *F. graminearum*, this distance  
 158 reaches 172 bp (**Figure 1E**). Considering these values and the canonical length of nucleosomal DNA (147 bp), linker  
 159 DNA length can be estimated to stand, in average, between 14 to 16 bp for *L. maculans* 'lepidii' and 'brassicae', 25-26  
 160 bp for *F. graminearum* and *B. cinerea* in malt medium, and 22 bp for *B. cinerea* grown in grape juice medium.  
 161 Nucleosome phasing genome-wide seems to be particularly tight in *F. graminearum*, with very little deviation in the  
 162 measured phases (**Figure 1E and 1F**). Similar tight phasing can also be observed in *B. cinerea* grown in grape juice-  
 163 based medium (**Figure 1C and 1F**) whereas higher deviations are observed in malt-based cultures (**Figure 1D and 1F**).  
 164 Similarly, phasing in grape juice-based medium is smaller (168-169 bp) than the one measured in malt medium (173-  
 165 174 bp). Examining TPM expression counts in both conditions (see Materials and Methods) for the gene coding the  
 166 unique H1 histone of the fungus (BcH1), we observed a diminution from 11,183 (sd = 202) in malt medium to 8,597  
 167 (sd = 422) in grape juice medium (data not shown), consistently with previous reports of decreased contents in linker  
 168 histones leading to shorter NRLs (Nucleosome repeat length) (Valouev *et al.*, 2011; Luque *et al.*, 2016).

169



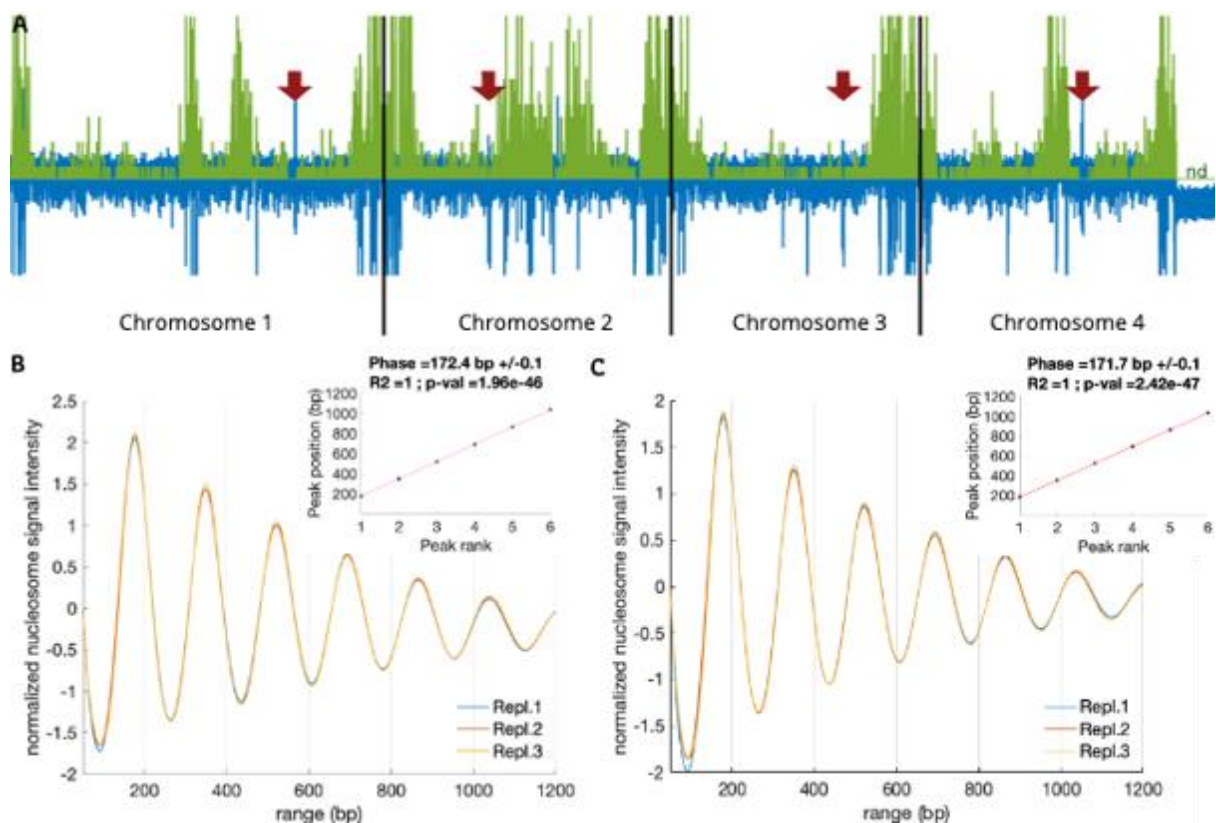


**Figure 1: Nucleosome phasing in the four fungi studied. A-E:** Main graphs display normalized (z-score) phase frequencies (y-axis) as a function of position (in base pair; x-axis). Graphs in inserts show peak positions (in base pairs; y-axis) as a function of peak order (x-axis). Phases +/- standard errors, R<sup>2</sup> (coefficient of determination), and p-values (F-test) are determined after linear regression fitting to peak positions in all replicates. Repl = replicate. **A:** *Leptosphaeria maculans* 'brassicae'; **B:** *Leptosphaeria maculans* 'lepidii'; here, phase value was measured excluding replicate #2 for which a peak apex was not clearly visible at the beginning of the profile; **C:** *Botrytis cinerea* grown in malt-based medium; **D:** *Botrytis cinerea* grown in grape juice-based medium; **E:** *Fusarium graminearum*; **F:** Genome-wide nucleosome phasing measured per individual replicate for each considered species/culture condition after fitting a linear regression model to the positions of six to seven successive peaks. Error bars are standard errors; nd = not determined. For each replicate: R<sup>2</sup> > 0.99 and p-value < 10<sup>-6</sup>.



171 Nucleosome distribution profiles

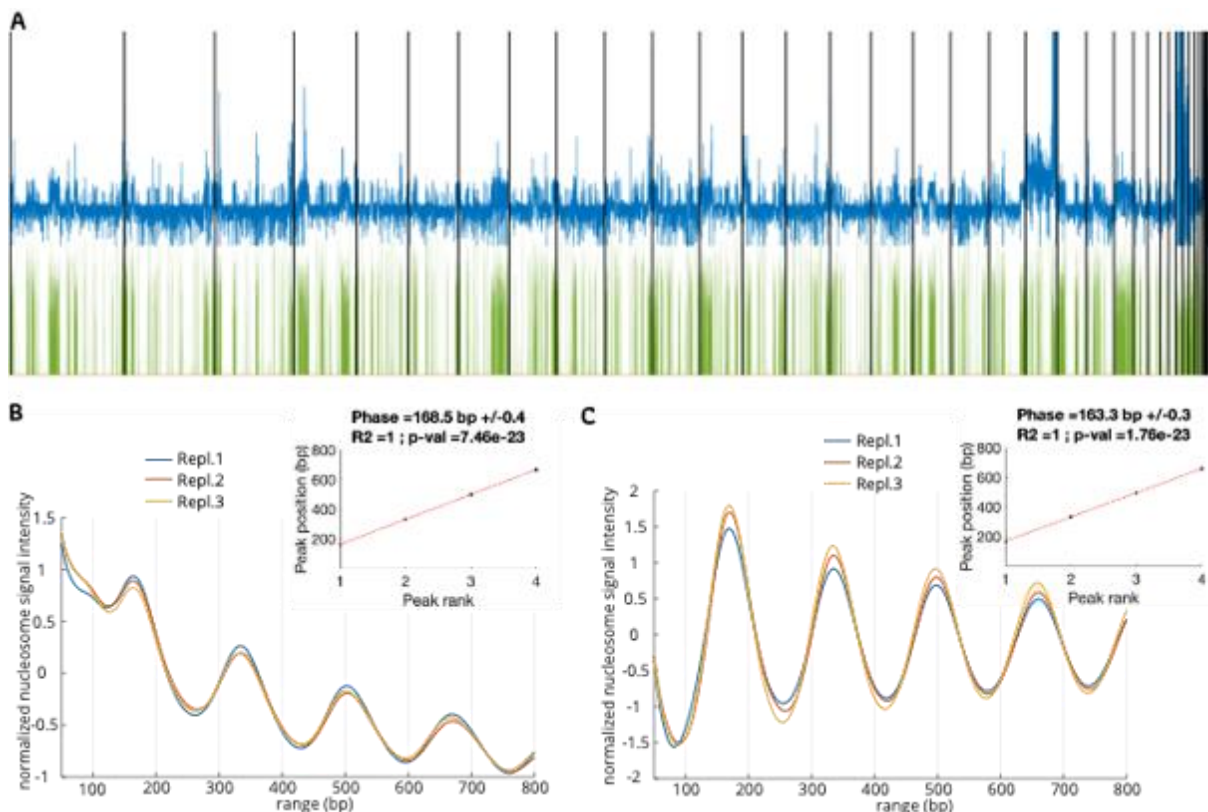
172 Read density was plotted genome-wide in one kb-long sliding non-overlapping windows along chromosomes  
173 for all four fungi (**Figure 2 and 3**, and **Figures S1 and S2**). The density profiles obtained for *B. cinerea* grown in malt or  
174 grape juice media both show remarkable regularity of nucleosome density genome-wide (**Figure S1**) whereas, in *F.*  
175 *graminearum*, regions equally packed with nucleosomes are interspaced with area with lower density (**Figure 2A** and  
176 **Figure 2**). Strikingly, this profile mirrors the previously described SNP density profiles in *F. graminearum* (Laurent *et*  
177 *al.*, 2017). We investigated whether or not this profile was the result of increased spacing between nucleosomes in  
178 regions found denser in SNPs. Phasograms were plotted as described above restricting our analysis to the SNP-  
179 enriched polymorphic islands or the rest of the genome, as defined by Laurent *et al.* (2017). Wave signals similar to  
180 the ones observed genome-wide were obtained, with phases of 172.4 bp in polymorphic islands (**Figure 2B**) and 171.7  
181 bp outside these regions (**Figure 2C**). These results indicate that nucleosomes appear well-arrayed genome-wide, with  
182 phases only slightly longer within polymorphic islands. Thus, the observed read density profile may be more the result  
183 of reduced nucleosome occupancy, *i.e.*, a measure of the stability of a nucleosome at a given position in a multiple cell  
184 sample, than a depletion in nucleosomes. This observation suggests increased frequencies of transient nucleosome  
185 positioning events in *F. graminearum* polymorphic islands and thus more relaxed chromatin structures.



**Figure 2: Nucleosome density profiles in *F. graminearum*.** **A:** Coverage density profiles were computed for non-overlapping 1 kb-long bins along the four chromosomes of *F. graminearum*. In green are plotted SNP density profiles as previously described (15). “nd” indicates the highly variable 3’ end of chromosome 4 for which SNP were not called. In blue are plotted the z-scored average nucleosome density profile (see supplementary figure 1 for individual replicate plots). The red arrows indicate centromeres (31); **B** and **C:** Nucleosome phasing in the polymorphic islands (**B**) or outside the polymorphic islands (**C**) as previously defined (15). Main graphs display phase frequencies (y-axis) as a function of position (in base pair; x-axis). Graphs in inserts show peak positions (in base pairs; y-axis) as a function of peak order (x-axis); Phases +/- standard errors, R2 (coefficient of determination), and p-values (F-test) are determined after linear regression fitting to peak positions in all replicates. Repl = replicate.



186 In *L. maculans* 'brassicae', numerous "islands" of nucleosome-dense regions can be observed at various  
187 locations of the genome (**Figure 3A** and **Figure S3**). Beside few contigs displaying higher nucleosome density, such  
188 characteristics were not observed for the closely related *L. maculans* 'lepidii' (**Figure S4**). The locations of these  
189 nucleosome-dense islands in *L. maculans* 'brassicae' parallel those of AT-rich regions of the genome (**Figure 3A**),  
190 features not visible in the genome of *L. maculans* 'lepidii' (**Figure S4**), suggesting that AT-rich regions are particularly  
191 dense in nucleosomes. Considering the remarkable GC- vs. AT-isochores organization of the genome of *L. maculans*  
192 'brassicae' (Rouxel *et al.*, 2011), absent from *L. maculans* 'lepidii', differences of nucleosome phasing and occupancy  
193 in AT-rich vs. GC-rich regions were investigated. A region was considered AT-rich if it contains less than 40 % of GC. As  
194 described earlier, AT isochores represent one third of the *L. maculans* 'brassicae' genome divided in 419 of 1 to 320  
195 kb-long regions. Examination of unprocessed mapping outputs reveals that the number of fragments (read pairs)  
196 mapped in AT- and GC- regions are very similar with 23.8 million and 24.8 million fragments, respectively, which is far  
197 from the 1/3 vs. 2/3 ratio expected. In terms of depth of coverage, mean coverage is 207 vs. 135 fragments for AT-rich  
198 and GC-rich regions, respectively, which could suggest higher nucleosome occupancy in the former. We explored this  
199 hypothesis and compared phasograms for AT-rich vs. GC-rich regions (**Figure 3B** and **3C**). NRLs were found larger in  
200 AT-rich than GC-rich compartments, measured at 168.5 bp and 163.3 bp respectively, suggesting lower nucleosome  
201 frequencies in AT-rich regions. Nonetheless, coverage density is higher in AT-rich regions, as seen on **Figure 3A** and  
202 **Figure S3**, consistently with our hypothesis of stronger nucleosome occupancy in these regions and thus less accessible



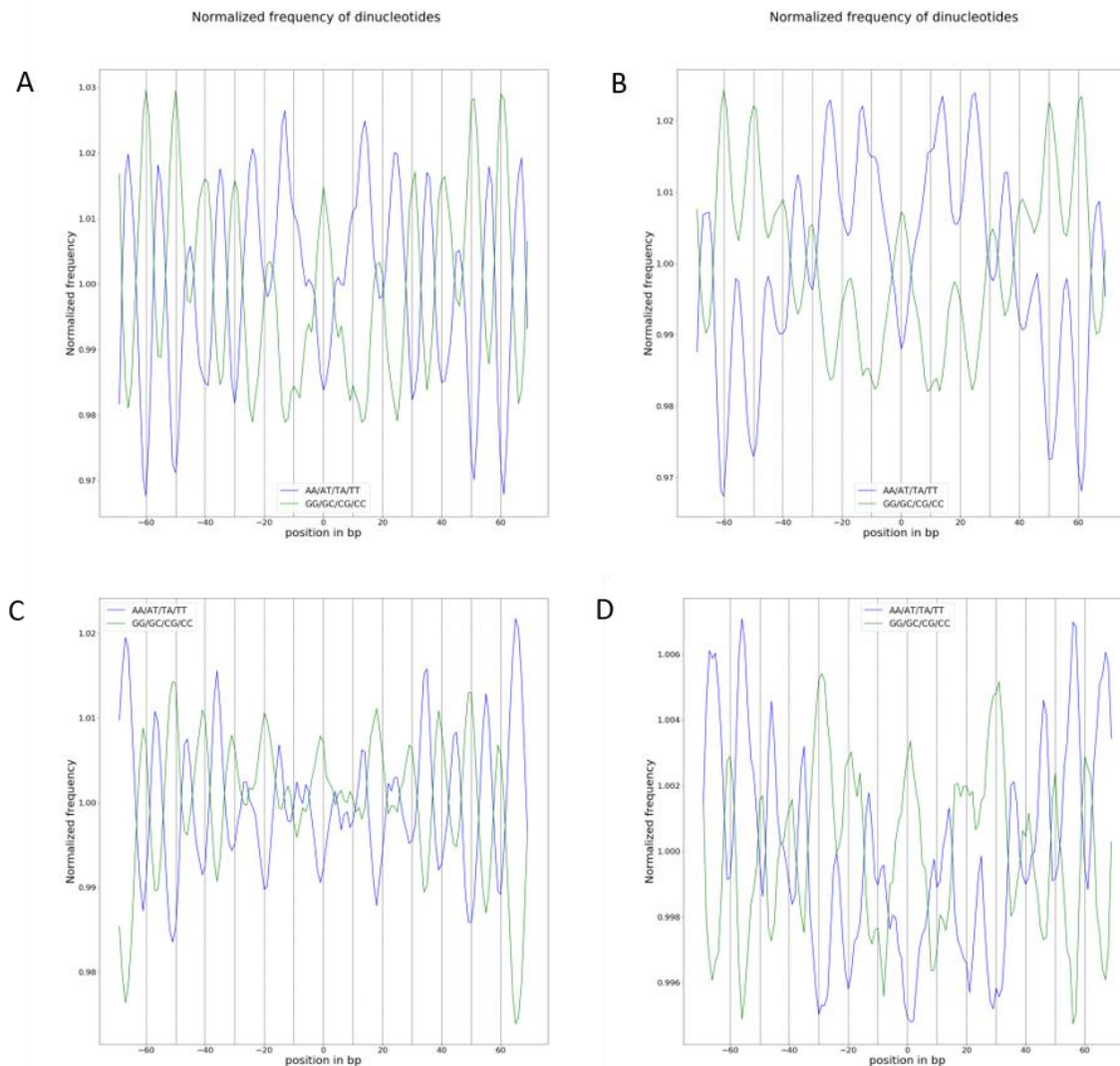
**Figure 3: Nucleosome density profiles in *L. maculans* 'brassicae'.** **A:** Coverage density profiles were computed for non-overlapping 1 kb-long bins along all supercontigs, separated by black lines. AT-rich regions are plotted in green. In blue are plotted the z-scored average nucleosome density profile (see supplementary figure 3 for individual replicate plots). **B and C:** Nucleosome phasing in AT-isochores (**B**) and GC-isochores (**C**). Main graphs display phase frequencies (y-axis) as a function of position (in base pair; x-axis). Graphs in inserts show peak positions (in base pairs; y-axis) as a function of peak order (x-axis).

203 genome compartment in heterochromatic state. Finally, signal intensity in phasograms appear more stable on the long  
204 nucleotide range in GC- than in AT-rich regions, an observation in line with the well-known destabilizing effect of AT  
205 stretches on nucleosome positioning leading to fuzzier signals (Kaplan *et al.*, 2009; (Ponts *et al.*, 2010; Tillo *et al.*, 2010;  
206 Bunnik *et al.*, 2014; Russell *et al.*, 2014; Jin *et al.*, 2018). All together, these data support a heterochromatic state of *L.*  
207 *maculans* 'brassicae' AT-rich regions mediated by strong nucleosome occupancy.

208 **Sequence composition and nucleosome positioning**

209 Literature data report abundantly that distribution of bases in nucleosome core DNA is non-random and  
210 exhibits a ~10 bp AA/TT/AT/TA offset with GG/CC/GC/CG dinucleotide frequency (Satchwell *et al.*, 1986; Segal *et al.*,  
211 2006; Segal and Widom, 2009). Here, we investigated di-nucleotide frequencies in nucleosomal DNA segments in all  
212 four fungi. Averaged di-nucleotide contents centred around all fragment (read pairs) were plotted (**Figure 4** and **Figure**  
213 **S5**). Normalized frequency graphs display the previously described ~10 bp periodicity of A+T dinucleotides oscillating  
214 out of phase with a ~10 bp G+C periodicity, the oscillation patterns being nonetheless very different between the four  
215 studied fungi. This periodicity is not constant along the entire nucleosomal DNA segment, shorter periodicities  
216 alternating with longer ones. In particular, the nucleosome dyad axis appears marked by longer periods flanked with  
217 short alternating AT and GC dinucleotide motifs. Specificities between the considered fungi can further be observed.  
218 In *L. maculans* 'lepidii', the average period along nucleosomal fragments is 10.1 bp for both AT and GC dinucleotide  
219 motifs. In *F. graminearum*, this period is shorter (8.8 bp), mostly due to very marked drops in period length around  
220 nucleosome dyads. Autocorrelation analyses of these dinucleotide frequency profiles highlight these 10 bp-  
221 periodicities in all fungi while revealing important differences in the absolute variations if the instant autocorrelation  
222 coefficients  $R_h$  (**Figure S5**). Signal is indeed very regular in *F. graminearum* and, to a lesser extent, *L. maculans*  
223 'brassicae' whereas *L. maculans* 'brassicae' and *B. cinerea* (grown in Malt medium) show more irregular  
224 autocorrelation profiles. These results are consistent with our previous observation that nucleosomes are tightly  
225 phased in *F. graminearum* whereas somehow fuzzier (higher deviation) in *B. cinerea* grown in malt medium (**Figure**  
226 **1C**). Di-nucleotide frequencies of *L. maculans* 'brassicae' AT- and GC-isochores were also inspected (**Figure S6**). As one  
227 would expect, the A+T periodicity is more pronounced in AT-isochores, but interestingly GC periodicity remains  
228 conserved.





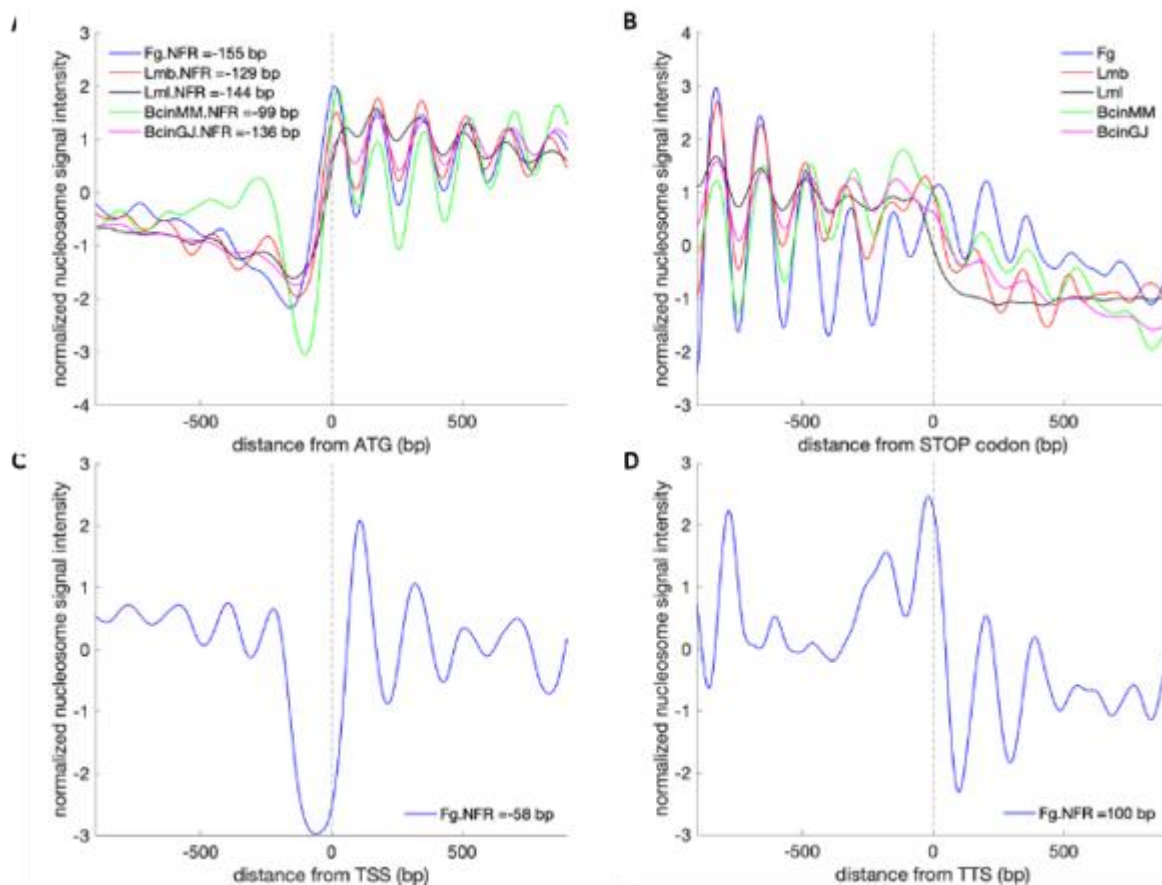
**Figure 4: Repeated di-nucleotide patterns in nucleosomal DNA.** Normalised di-nucleotides frequency plots (average of three biological replicates) for *L. maculans* 'brassicae' (A), *L. maculans* 'lepidii' (B), *F. graminearum* (C), and *B. cinerea* grown in malt medium (D).

## 229 Nucleosome landscapes of fungal gene units

230 Nucleosome occupancy profiles of gene units were investigated in all fungi (Figure 5). As previously reported  
 231 in other eukaryotes, translation start sites (the ATG codon) are preceded by a nucleosome-depleted region (NDR) and  
 232 immediately followed by a well-positioned +1 nucleosome (Figure 5A). Variations of the exact position of these  
 233 features relative to the ATG start codon, as well as the intensity of the NDR valley, are nonetheless observed when  
 234 fungal species are compared, For example, the NDR and the centre of the +1 nucleosome (or +1Nucl) are found at -  
 235 155 bp and +14 bp, respectively, in *F. graminearum* whereas they are located at -99 bp and +26 bp, respectively, in  
 236 *B. cinerea* grown in malt medium. *B. cinerea* grown in malt medium shows the deepest NDR valley among all observed  
 237 profiles. Differences can further be observed in *B. cinerea* grown in grape juice-based medium, with NDR and +1Nucl  
 238 found at -136 bp and +31 bp, respectively. In *L. maculans* 'brassicae' and 'lepidii', NDRs are found at -129 and 144 bp  
 239 from ATG, respectively, and +1Nucl at +19 bp and +55 bp. Finally, nucleosome profiles upstream of NDRs appear fuzzy  
 240 for all fungi but *L. maculans* 'brassicae', with varying degrees of fuzziness. This fuzziness is no longer visible when  
 241 nucleosome profiles are centred on TSS for *F. graminearum* ( $N_{TSS} = 6,212$ , Figure 5C and Table S2). The NDR is more

242 defined, and located immediately upstream of the TSS, with a minimum detected at -58 bp and +1Nucl found 110 bp  
243 downstream of *F. graminearum* TSS. These values are consistent with our previous measure of the average absolute  
244 distance NDR+1Nucl. Considering nucleosome environments at stop codons (**Figure 5B**), strongly arrayed  
245 nucleosomes are particularly found upstream the stop codon, fewer signal variations being observed downstream.  
246 Here, the stop codon is a clear boundary for nucleosome arraying and occupancy in all fungi and all conditions  
247 investigated. A nucleosome seems remarkably well positioned exactly on stop codons in *F. graminearum*. Repeating  
248 the analysis on TTS in *F. graminearum* ( $N_{TTS} = 5,292$ ), strong positioning of nucleosomes -2 and -1 (-178 bp and -16 bp  
249 relative to TTS, respectively) followed by a deep 3' end NDR (+100 bp downstream of TTS) can be observed (**Figure**  
250 **5D**).

251 For comparison purposes, nucleosome profiling was repeated restricting our analysis to genes identified as  
252 BUSCO lineage-specific single-copy evolutionary conserved orthologs in Fungi (Simao *et al.*, 2015; Waterhouse *et al.*,  
253 2018). Overall broad patterns remain similar to those obtained while investigating whole genomes (**Figure 6**).  
254 Nonetheless, in *F. graminearum*, whilst the distance NDR-to-TSS (-60 bp) remains very similar to the one measured  
255 earlier genome-wide (**Figure 6C**), distance NDR-to-ATG increases by 28 bp (**Figure 6A**). In *B. cinerea*, such an increase  
256 is also observed when grown in grape juice medium (+32 bp) but not in malt medium (+3 bp). Towards the 3' end of  
257 the gene unit, nucleosome signals around stop codons are similar to those obtained genome-wide for all fungi (**Figure**

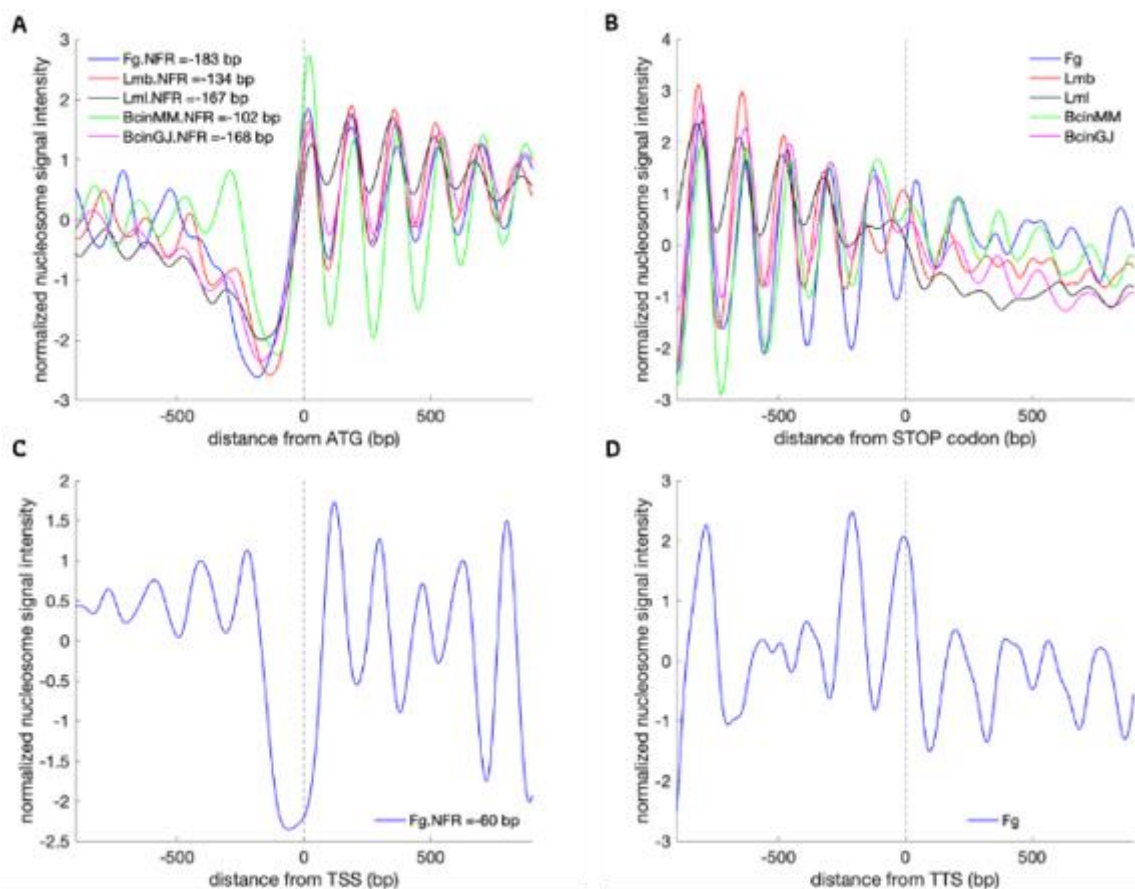


**Figure 5: Nucleosome organization of fungal gene units.** A-D: Normalised averages (three biological replicates for each fungus/condition) of nucleosome signals as a function of position (in base pairs) relative to the start codon ATG (A), the stop codon (B), TSS (C), and TTS (D). Fg = *F. graminearum*; Lmb = *L. maculans* 'brassicae'; Lml = *L. maculans* 'lepidii' = BcinMM = *B. cinerea* grown in malt medium; BcinGJ = *B. cinerea* grown in grape juice medium; NFR = nucleosome-free region.

258 **6B)**. However, the deep NDR found downstream of TTS of *F. graminearum* genome-wide can no longer be observed  
259 at “Fungi” TTS *loci* (**Figure 6D**). Similar to genome-wide profile, strongly positioned -1Nucl and -2Nucl are still visible  
260 at -5 bp and -209 bp, respectively, the latter being more intense and defined.

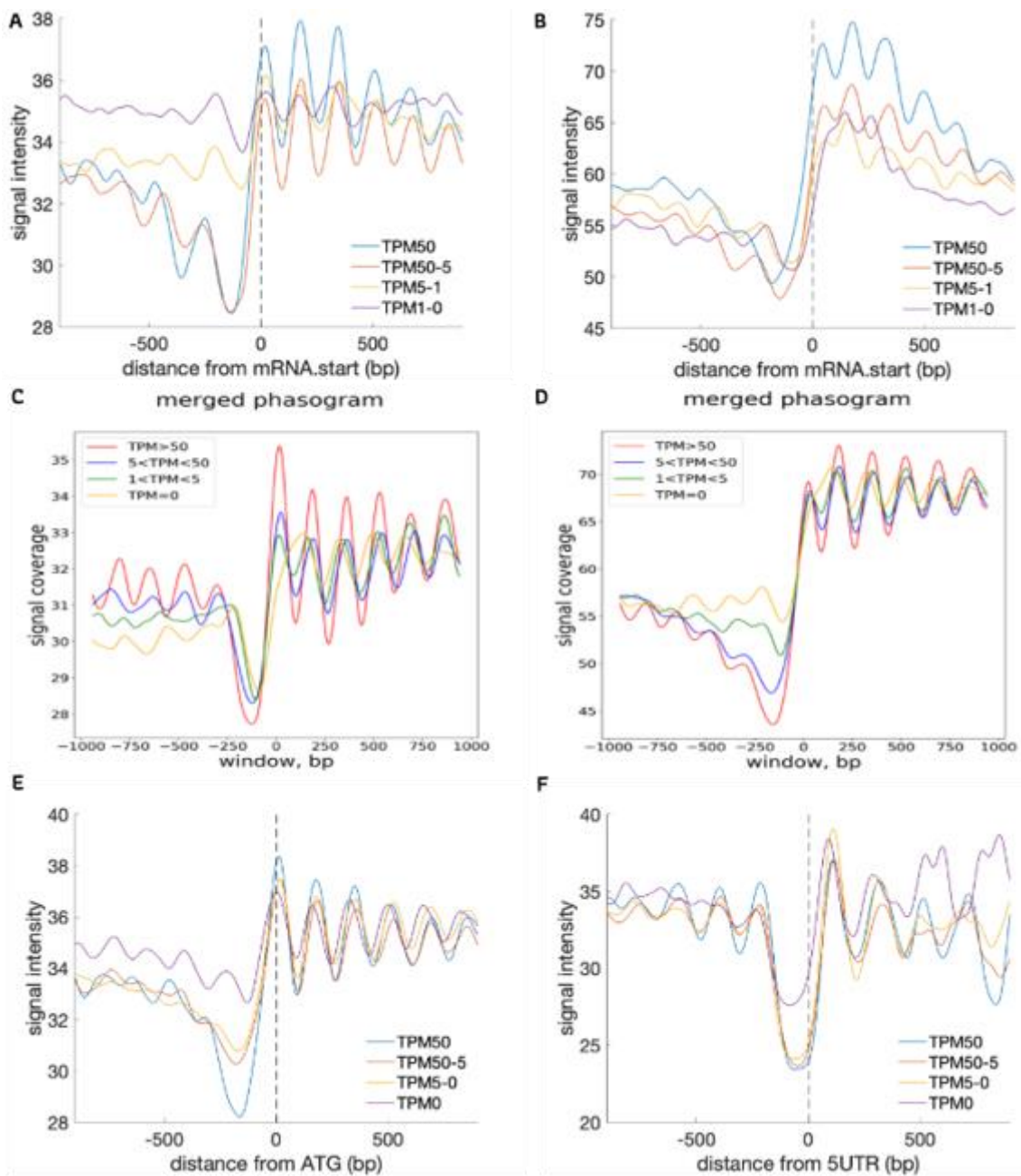
## 261 Nucleosome landscapes of gene units and gene expression

262 Same analyses were repeated for genes categorized according to their expression levels (expressed in TPM  
263 counts, see Materials and Methods). The general variations in nucleosome profiles around translation start sites are  
264 similar for all expression categories in all considered fungi and culture conditions: ATG codons are immediately  
265 followed by a well-positioned +1 nucleosome and preceded by a dip in nucleosome density (**Figure 7**). Remarkable  
266 variations are nonetheless observed with regard to positions of +1 nucleosomes and NDRs, as well as the amplitude  
267 of the nucleosome signal difference between them (here defined as  $\Delta\text{nucl} = (\text{signal}_{+1\text{nucl}} - \text{signal}_{\text{NDR}})$ , depending on gene  
268 expression. ATG-centred nucleosome profiles for genes not expressed in our conditions (TPM = 0) show remarkably  
269 reduced  $\Delta\text{nucl}$  when compared to those measured for genes more expressed (**Figure 7** and **Table S2**). Conversely,  
270 highly expressed genes display the deepest NDRs located at the furthest upstream the ATG. Similar trends are  
271 observed when profiles are centred on *F. graminearum* TSS (**Figure 7F**). Remarkably, in *B. cinerea* these differences  
272 are exacerbated when grown in grape juice medium (**Figure 7D**) compared to malt medium (**Figure 7C**). In summary,  
273 our results indicate that, in the four Ascomycota investigated here, genes highly expressed are characterised by high



**Figure 6: Nucleosome organization of fungal near-universal single copy orthologous gene units (BUSCO3).**  
A-D. Normalised averages (three biological replicates for each fungus/condition) of nucleosome signals as a function of position (in base pairs) relative to the start codon ATG (A), the stop codon (B), TSS (C), and TTS (D). Fg = *F. graminearum*; Lmb = *L. maculans* ‘brassicae’; Lml = *L. maculans* ‘lepidii’ = BcinMM = *B. cinerea* grown in malt medium; BcinGJ = *B. cinerea* grown in grape juice medium; NFR = nucleosome-free region.

274 amplitudes of nucleosome signal between the NDR valley upstream the ATG codon and the +1 nucleosome peak. In  
275 addition, these NDRs are usually found further from the ATG site than those in genes expressed at lower levels or not  
276 expressed. Finally, the nucleosome wave signal decay phenomenon is observed at distances from the ATG codons  
277 shorter in poorly expressed genes than in highly expressed genes, although the +1 nucleosome remains fairly well  
278 conserved. At the end of gene units, wave signal decay was also observed downstream of stop codons in all studied  
279 fungi (**Figure S7**). In *F. graminearum*, signal remained nonetheless stronger in genes lowly/not expressed than in those  
280 with higher expression levels (**Figure S7E**) although with species-specific features. In *L. maculans* 'brassicae' and  
281 'lepidii', strong nucleosome positioning can be found upstream the stop codon, immediately followed by a strong  
282 nucleosome valley in the case of *L. maculans* 'lepidii'. Finally, *F. graminearum* TTS are characterized by a fuzzy



**Figure 7: Nucleosome organization at start codons/TSS vs. gene expression.** Average (three biological replicates for each fungus/condition) nucleosome signal as a function of position (in base pairs) relative to the start codon ATG (and TSS for *F. graminearum*). TPM = Transcripts Per Million. **A:** *Leptosphaeria maculans* 'brassicae'; **B:** *Leptosphaeria maculans* 'lepidii'; **C:** *Botrytis cinerea* grown in malt-based medium; **D:** *Botrytis cinerea* grown in grape juice-based medium; **E:** *Fusarium graminearum*, ATG-centred; **F:** *Fusarium graminearum*, TSS-centred.

283 nucleosome wave signal expect for genes not expressed that display well-arrayed nucleosomes (**Figure S7F**).  
284 Altogether, our results show strong association of nucleosome landscapes at gene unit boundaries with expression  
285 levels in the four studied ascomycetes.

## 286 DISCUSSION

287 Nucleosome landscapes are increasingly documented in a large array of organisms, tissues, and cell types in order to  
288 understand the impact of their dynamics in gene expression. The present study explores nucleosome positioning  
289 characteristics in four different filamentous fungi, *L. maculans* 'brassicae', *L. maculans* 'lepidii', *B. cinerea*, and *F.*  
290 *graminearum*, with the initial ambition to unravel features of a fungal meta-nucleosome landscape. Here, we found  
291 features shared by all four investigated fungal genomes, as well as more specific ones.

292 Nucleosome spacing influences the formation of the higher order chromatin fibre, often referred to as the 30-nm  
293 chromatin fibre (Szerlong and Hansen, 2011). Several structural models of the chromatin fibre have been proposed,  
294 all underlining nucleosome-nucleosome interactions including the length of linker DNA fragments as major driving  
295 factors (see Zhu and Li, 2016). Notably, the chromatin fibre was found to be narrower (21-nm in diameter) for a short  
296 NRL of 167 bp (Robinson *et al.*, 2006). Similarly, increasing NRLs lead to increasingly wider fibres, reaching a highly  
297 compact 30-nm solenoid structure for an NRL of 197 bp (Robinson *et al.*, 2006; Song *et al.*, 2014). Typically, NRLs are  
298 ~175-200 bp in plants and other higher eukaryotes (Valouev *et al.*, 2008; Valouev *et al.*, 2011; Nishida *et al.*, 2012;  
299 Locke *et al.*, 2013; Zhang *et al.*, 2015), and ~165 bp and 154 bp in the yeasts *S. cerevisiae* and *S. pombe*, respectively  
300 (Yuan *et al.*, 2005; Lantermann *et al.*, 2010). In the present study, we found NRLs ranging from 161 to 173 bp in our  
301 fungal pathogens. These values are remarkably constant between biological replicates of a given fungus, a  
302 phenomenon sometimes referred to as clamping that involves ATP-dependent chromatin remodelers in purified  
303 experimental systems (Lieleg *et al.*, 2015). *F. graminearum* and *B. cinerea* show similar NRLs in the middle range (169  
304 bp to 173 bp), possibly indicating intermediate levels of compaction of the higher order chromatin fibre. These values  
305 are similar to those obtained for *Aspergillus fumigatus*, *i.e.*, linker length ranging from 21 to 27 bp, using a MNase  
306 treatment similar to the one used in the present study (Nishida *et al.*, 2012). *L. maculans* 'brassicae' and 'lepidii'  
307 distinguish themselves with shorter NRLs of only 161-163 bp, suggesting a narrower chromatin fibre structure. In *B.*  
308 *cinerea*, our data reveal a shorter NRL when grown in grape juice-based medium than when grown in malt-based  
309 medium correlating with increased histone H1 transcript levels, as previously reported in other eukaryotes (Valouev  
310 *et al.*, 2011; Luque *et al.*, 2016). Roles of histone H1 in fungi are diverse and remain largely unknown (Brosch *et al.*,  
311 2008). Here, the changes observed in *B. cinerea* grown in two different media may suggest that, beyond its structural  
312 role, H1 participates to reorganizing gene expression networks in response to environmental parameters.

313 Nucleosome positioning is believed to be particularly hard-wired to DNA sequence and especially the largely  
314 documented anti-nucleosome effect of Poly(dA:dT) tracts (Segal and Widom, 2009; Struhl and Segal, 2013). In the  
315 present study, whilst nucleosome phase was indeed found 5 bp longer in *L. maculans* 'brassicae' AT-isochores than  
316 GC-isochores, occupancy was nonetheless higher in the former leading to the formation of the previously suggested  
317 heterochromatic state of these regions (Rouxel *et al.*, 2011; Soyer *et al.*, 2014). These observations indicate the





318 mobilisation of trans-acting chromatin remodelling factors to maintain heterochromatin structures on such  
319 disfavouring sequences. Importantly, we found that GC periodicity at nucleosome dyads is preserved even within AT-  
320 isochores, suggesting such pattern in an AT-rich environment is sufficient to permit efficient wrapping of DNA around  
321 nucleosomes and strong occupancy. Conversely, our results reveal lower nucleosome occupancy in the previously  
322 described *F. graminearum* fast evolving polymorphic islands (Laurent *et al.*, 2017). Here, nucleosome dynamics may  
323 enable fast evolution of particular genome segments while regions defined by higher occupancies may secure  
324 sequence conservation.

325 The general profile of a fungal gene unit shares similarities with those previously described in various eukaryotes:  
326 the ATG codon is decorated by a well-positioned +1 nucleosome and preceded by an NDR. The NDR observed for *B.*  
327 *cinerea* grown in malt medium was remarkable in its intensity, deeper than all others. When we restricted our analysis  
328 to conserved orthologous fungal genes, the overall profiles and the intensities of NDRs were more homogenous  
329 between fungi and culture conditions. These observations suggest that NDRs positions and intensities are subjected  
330 to evolution. Similarly, whilst an NDR can be observed downstream of *F. graminearum* TTS, it is no longer visible when  
331 the analysis is restricted to fungal genes indicating again an evolutionary component.

332 From the gene expression perspective, nucleosome depletion can be the side-effect of active transcription with the  
333 binding of pre-initiation complex resulting in nucleosome eviction as previously shown in yeast (Venters and Pugh,  
334 2009). Indeed, we evidenced valleys of nucleosome signals upstream of ATG codons found deeper in highly expressed  
335 genes. On the contrary, genes not expressed show little or no NDR, depending on the considered species or culture  
336 condition. In our conditions, the amplitude between the NDR and the +1 nucleosome seems to be an informative  
337 measure of gene expression level: the more this value is high, the more genes are expressed. Strikingly, this feature is  
338 less strict when TSS are considered raising the question of different mechanisms of transcription regulation depending  
339 on gene unit structures.

## 340 ACKNOWLEDGEMENT

341 We are grateful to the Genotoul bioinformatics platform Toulouse Midi-Pyrenees for providing computing resources.  
342 J. L. Soyer was founded by a “Contrat Jeune Scientifique” grant from INRA.

## 343 FUNDING

344 This work was supported by the Plant Health and Environment division of the French National Institute for Agricultural  
345 Research [MAINE Project AAP 2014]. Funding for open access charge: Plant Health and Environment department of  
346 the French National Institute for Agricultural Research.

## 347 CONFLICT OF INTEREST

348 The authors declare no conflict of interest.

## 349 AVAILABILITY



350 The MSTs (MAINE-Seq Tool Suite) is an open source collection of tools developed by the BioinfoBIOGER platform by  
351 N. Lapalu and A. Simon, and available in the GitHub repository (<https://github.com/nlapalu/MSTS>).

## 352 ACCESSION NUMBERS

353 All sequenced reads have been deposited with the Short Read Archive under accession numbers XXXX.

## 354 REFERENCES

355 Amselem, J., Cuomo, C.A., van Kan, J.A.L., Viaud, M., Benito, E.P., Couloux, A., Coutinho, P.M., de Vries, R.P.,  
356 Dyer, P.S., Fillinger, S., Fournier, E., Gout, L., Hahn, M., Kohn, L., Lapalu, N., Plummer, K.M., Pradier, J.-M., Quévillon,  
357 E., Sharon, A., Simon, A., ten Have, A., Tudzynski, B., Tudzynski, P., Wincker, P., Andrew, M., Anthouard, V., Beever,  
358 R.E., Beffa, R., Benoit, I., Bouzid, O., Brault, B., Chen, Z., Choquer, M., Collémare, J., Cotton, P., Danchin, E.G., Da Silva,  
359 C., Gautier, A., Giraud, C., Giraud, T., Gonzalez, C., Grossetete, S., Güldener, U., Henrissat, B., Howlett, B.J., Kodira, C.,  
360 Kretschmer, M., Lappartient, A., Leroch, M., Levis, C., Mauceli, E., Neuvéglise, C., Oeser, B., Pearson, M., Poulain, J.,  
361 Poussereau, N., Quesneville, H., Rasclé, C., Schumacher, J., Ségurens, B., Sexton, A., Silva, E., Sirven, C., Soanes, D.M.,  
362 Talbot, N.J., Templeton, M., Yandava, C., Yarden, O., Zeng, Q., Rollins, J.A., Lebrun, M.-H., Dickman, M., 2011. Genomic  
363 Analysis of the Necrotrophic Fungal Pathogens *Sclerotinia sclerotiorum* and *Botrytis cinerea*. *PLoS Genet* 7.

364 Basenko, E.Y., Pulman, J.A., Shanmugasundram, A., Harb, O.S., Crouch, K., Starns, D., Warrenfeltz, S.,  
365 Aurrecochea, C., Stoeckert, C.J., Kissinger, J.C., Roos, D.S., Hertz-Fowler, C., 2018. FungiDB: An Integrated  
366 Bioinformatic Resource for Fungi and Oomycetes. *J Fungi (Basel)* 4.

367 Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data.  
368 *Bioinformatics* 30, 2114–2120.

369 Boutigny, A.-L., Barreau, C., Atanasova-Penichon, V., Verdal-Bonnin, M.-N., Pinson-Gadais, L., Richard-Forget,  
370 F., 2009. Ferulic acid, an efficient inhibitor of type B trichothecene biosynthesis and Tri gene expression in *Fusarium*  
371 liquid cultures. *Mycol. Res.* 113, 746–753.

372 Brosch, G., Loidl, P., Graessle, S., 2008. Histone modifications and chromatin dynamics: a focus on filamentous  
373 fungi. *FEMS Microbiol Rev* 32, 409–439.

374 Bunnik, E.M., Polishko, A., Prudhomme, J., Ponts, N., Gill, S.S., Lonardi, S., Le Roch, K.G., 2014. DNA-encoded  
375 nucleosome occupancy is associated with transcription levels in the human malaria parasite *Plasmodium falciparum*.  
376 *BMC Genomics* 15, 347.

377 Cuomo, C.A., Güldener, U., Xu, J.-R., Trail, F., Turgeon, B.G., Di Pietro, A., Walton, J.D., Ma, L.-J., Baker, S.E.,  
378 Rep, M., Adam, G., Antoniw, J., Baldwin, T., Calvo, S., Chang, Y.-L., Decaprio, D., Gale, L.R., Gnerre, S., Goswami, R.S.,  
379 Hammond-Kosack, K., Harris, L.J., Hilburn, K., Kennell, J.C., Kroken, S., Magnuson, J.K., Mannhaupt, G., Mauceli, E.,  
380 Mewes, H.-W., Mitterbauer, R., Muehlbauer, G., Münsterkötter, M., Nelson, D., O'donnell, K., Ouellet, T., Qi, W.,  
381 Quesneville, H., Roncero, M.I.G., Seong, K.-Y., Tetko, I.V., Urban, M., Waalwijk, C., Ward, T.J., Yao, J., Birren, B.W.,  
382 Kistler, H.C., 2007. The *Fusarium graminearum* genome reveals a link between localized polymorphism and pathogen  
383 specialization. *Science* 317, 1400–1402.



- 384 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R.,  
385 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.
- 386 Dutreux, F., Da Silva, C., d'Agata, L., Couloux, A., Gay, E.J., Istace, B., Lapalu, N., Lemainque, A., Linglin, J., Noel,  
387 B., Wincker, P., Cruaud, C., Rouxel, T., Balesdent, M.-H., Aury, J.-M., 2018. De novo assembly and annotation of three  
388 *Leptosphaeria* genomes using Oxford Nanopore MinION sequencing. *Scientific Data* 5.
- 389 Fisher, M.C., Henk, D.A., Briggs, C.J., Brownstein, J.S., Madoff, L.C., McCraw, S.L., Gurr, S.J., 2012. Emerging  
390 fungal threats to animal, plant and ecosystem health. *Nature* 484, 186–194.
- 391 Fox, E.M., Howlett, B.J., 2008. Secondary metabolism: regulation and role in fungal biology. *Current Opinion*  
392 *in Microbiology* 11, 481–487.
- 393 Fudal, I., Ross, S., Gout, L., Blaise, F., Kuhn, M.L., Eckert, M.R., Cattolico, L., Bernard-Samain, S., Balesdent,  
394 M.H., Rouxel, T., 2007. Heterochromatin-like regions as ecological niches for avirulence genes in the *Leptosphaeria*  
395 *maculans* genome: map-based cloning of AvrLm6. *Mol. Plant Microbe Interact.* 20, 459–470.
- 396 Grandaubert, J., Lowe, R.G., Soyer, J.L., Schoch, C.L., Van de Wouw, A.P., Fudal, I., Robbertse, B., Lapalu, N.,  
397 Links, M.G., Ollivier, B., 2014. Transposable element-assisted evolution and adaptation to host plant within the  
398 *Leptosphaeria maculans*-*Leptosphaeria biglobosa* species complex of fungal pathogens. *BMC genomics* 15, 891.
- 399 Hallen, H.E., Huebner, M., Shiu, S.-H., Güldener, U., Trail, F., 2007. Gene expression shifts during perithecium  
400 development in *Gibberella zeae* (anamorph *Fusarium graminearum*), with particular emphasis on ion transport  
401 proteins. *Fungal Genet. Biol.* 44, 1146–1156.
- 402 Hawksworth, D.L., Lücking, R., 2017. Fungal Diversity Revisited: 2.2 to 3.8 Million Species. *Microbiol Spectr* 5.
- 403 Jin, H., Finnegan, A.I., Song, J.S., 2018. A unified computational framework for modeling genome-wide  
404 nucleosome landscape. *Phys Biol* 15, 066011.
- 405 Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R.,  
406 Lieb, J.D., Widom, J., Segal, E., 2009. The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* 458,  
407 362–366.
- 408 Kelloniemi, J., Trouvelot, S., Héloir, M.-C., Simon, A., Dalmais, B., Frettinger, P., Cimerman, A., Fermaud, M.,  
409 Roudet, J., Baulande, S., Bruel, C., Choquer, M., Couvelard, L., Duthieuw, M., Ferrarini, A., Flors, V., Le Pêcheur, P.,  
410 Loisel, E., Morgant, G., Poussereau, N., Pradier, J.-M., Rasclé, C., Trdá, L., Poinssot, B., Viaud, M., 2015. Analysis of the  
411 Molecular Dialogue Between Gray Mold (*Botrytis cinerea*) and Grapevine (*Vitis vinifera*) Reveals a Clear Shift in Defense  
412 Mechanisms During Berry Ripening. *Mol. Plant Microbe Interact.* 28, 1167–1180.
- 413 King, R., Urban, M., Hammond-Kosack, M.C.U., Hassani-Pak, K., Hammond-Kosack, K.E., 2015. The completed  
414 genome sequence of the pathogenic ascomycete fungus *Fusarium graminearum*. *BMC Genomics* 16, 544.
- 415 King, R., Urban, M., Hammond-Kosack, K.E., 2017. Annotation of *Fusarium graminearum* (PH-1) Version 5.0.  
416 Genome Announcements 5.
- 417 Langmead, B., Salzberg, S.L., 2012. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359.
- 418 Lantermann, A.B., Straub, T., Strålfors, A., Yuan, G.-C., Ekwall, K., Korber, P., 2010. *Schizosaccharomyces*  
419 *pombe* genome-wide nucleosome mapping reveals positioning mechanisms distinct from those of *Saccharomyces*  
420 *cerevisiae*. *Nat. Struct. Mol. Biol.* 17, 251–257.



- 421 Laurent, B., Moinard, M., Spataro, C., Ponts, N., Barreau, C., Foulongne-Oriol, M., 2017. Landscape of genomic  
422 diversity and host adaptation in *Fusarium graminearum*. *BMC Genomics* 18, 203.
- 423 Li, B., Ruotti, V., Stewart, R.M., Thomson, J.A., Dewey, C.N., 2010. RNA-Seq gene expression estimation with  
424 read mapping uncertainty. *Bioinformatics* 26, 493–500.
- 425 Lieleg, C., Ketterer, P., Nuebler, J., Ludwigsen, J., Gerland, U., Dietz, H., Mueller-Planitz, F., Korber, P., 2015.  
426 Nucleosome spacing generated by ISWI and CHD1 remodelers is constant regardless of nucleosome density. *Mol. Cell.*  
427 *Biol.* 35, 1588–1605.
- 428 Lo Presti, L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., Zuccaro, A., Reissmann, S., Kahmann, R.,  
429 2015. Fungal Effectors and Plant Susceptibility. *Annual Review of Plant Biology* 66, 513–545.
- 430 Locke, G., Haberman, D., Johnson, S.M., Morozov, A.V., 2013. Global remodeling of nucleosome positions in  
431 *C. elegans*. *BMC Genomics* 14, 284.
- 432 Love, M.I., Huber, W., Anders, S., 2014. Moderated estimation of fold change and dispersion for RNA-seq data  
433 with DESeq2. *Genome Biol.* 15, 550.
- 434 Luque, A., Ozer, G., Schlick, T., 2016. Correlation among DNA Linker Length, Linker Histone Concentration, and  
435 Histone Tails in Chromatin. *Biophys. J.* 110, 2309–2319.
- 436 Möbius, N., Hertweck, C., 2009. Fungal phytotoxins as mediators of virulence. *Curr. Opin. Plant Biol.* 12, 390–  
437 398.
- 438 Nishida, H., Kondo, S., Matsumoto, T., Suzuki, Y., Yoshikawa, H., Taylor, T.D., Sugiyama, J., 2012. Characteristics  
439 of nucleosomes and linker DNA regions on the genome of the basidiomycete *Mixia osmundae* revealed by mono- and  
440 dinucleosome mapping. *Open Biol* 2, 120043.
- 441 Nishida, H., Motoyama, T., Yamamoto, S., Aburatani, H., Osada, H., 2009. Genome-wide maps of mono- and  
442 di-nucleosomes of *Aspergillus fumigatus*. *Bioinformatics* 25, 2295–2297.
- 443 Ponts, N., Harris, E.Y., Prudhomme, J., Wick, I., Eckhardt-Ludka, C., Hicks, G.R., Hardiman, G., Lonardi, S., Le  
444 Roch, K.G., 2010. Nucleosome landscape and control of transcription in the human malaria parasite. *Genome Res.* 20,  
445 228–238.
- 446 Porquier, A., Morgant, G., Moraga, J., Dalmais, B., Luyten, I., Simon, A., Pradier, J.-M., Amselem, J., Collado,  
447 I.G., Viaud, M., 2016. The botrydial biosynthetic gene cluster of *Botrytis cinerea* displays a bipartite genomic structure  
448 and is positively regulated by the putative Zn(II)2Cys6 transcription factor BcBot6. *Fungal Genetics and Biology* 96, 33–  
449 46.
- 450 Radman-Livaja, M., Rando, O.J., 2010. Nucleosome positioning: how is it established, and why does it matter?  
451 *Dev. Biol.* 339, 258–266.
- 452 Richmond, T.J., Davey, C.A., 2003. The structure of DNA in the nucleosome core. *Nature* 423, 145–150.
- 453 Robinson, P.J.J., Fairall, L., Huynh, V.A.T., Rhodes, D., 2006. EM measurements define the dimensions of the  
454 “30-nm” chromatin fiber: evidence for a compact, interdigitated structure. *Proc. Natl. Acad. Sci. U.S.A.* 103, 6506–  
455 6511.
- 456 Rouxel, T., Grandaubert, J., Hane, J.K., Hoede, C., van de Wouw, A.P., Couloux, A., Dominguez, V., Anthouard,  
457 V., Bally, P., Bourras, S., Cozijnsen, A.J., Ciuffetti, L.M., Degrave, A., Dilmaghani, A., Duret, L., Fudal, I., Goodwin,



- 458 S.B., Gout, L., Glaser, N., Linglin, J., Kema, G.H.J., Lapalu, N., Lawrence, C.B., May, K., Meyer, M., Ollivier, B., Poulain, J.,  
459 Schoch, C.L., Simon, A., Spatafora, J.W., Stachowiak, A., Turgeon, B.G., Tyler, B.M., Vincent, D., Weissenbach, J.,  
460 Amselem, J., Quesneville, H., Oliver, R.P., Wincker, P., Balesdent, M.-H., Howlett, B.J., 2011. Effector diversification  
461 within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations. *Nat*  
462 *Commun* 2, 202.
- 463 Russell, K., Cheng, C.-H., Bizzaro, J.W., Ponts, N., Emes, R.D., Le Roch, K., Marx, K.A., Horrocks, P., 2014.  
464 Homopolymer tract organization in the human malarial parasite *Plasmodium falciparum* and related Apicomplexan  
465 parasites. *BMC Genomics* 15.
- 466 Satchwell, S.C., Drew, H.R., Travers, A.A., 1986. Sequence periodicities in chicken nucleosome core DNA. *J.*  
467 *Mol. Biol.* 191, 659–675.
- 468 Segal, E., Fondufe-Mittendorf, Y., Chen, L., Thåström, A., Field, Y., Moore, I.K., Wang, J.-P.Z., Widom, J., 2006.  
469 A genomic code for nucleosome positioning. *Nature* 442, 772–778.
- 470 Segal, E., Widom, J., 2009. What controls nucleosome positions? *Trends in Genetics* 25, 335–343.
- 471 Simão, F.A., Waterhouse, R.M., Ioannidis, P., Kriventseva, E.V., Zdobnov, E.M., 2015. BUSCO: assessing genome  
472 assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212.
- 473 Simon, A., Dalmais, B., Morgant, G., Viaud, M., 2013. Screening of a *Botrytis cinerea* one-hybrid library reveals  
474 a Cys2His2 transcription factor involved in the regulation of secondary metabolism gene clusters. *Fungal Genet. Biol.*  
475 52, 9–19.
- 476 Song, F., Chen, P., Sun, D., Wang, M., Dong, L., Liang, D., Xu, R.-M., Zhu, P., Li, G., 2014. Cryo-EM study of the  
477 chromatin fiber reveals a double helix twisted by tetranucleosomal units. *Science* 344, 376–380.
- 478 Soyer, J.L., El Ghalid, M., Glaser, N., Ollivier, B., Linglin, J., Grandaubert, J., Balesdent, M.-H., Connolly, L.R.,  
479 Freitag, M., Rouxel, T., Fudal, I., 2014. Epigenetic Control of Effector Gene Expression in the Plant Pathogenic Fungus  
480 *Leptosphaeria maculans*. *PLoS Genetics* 10, e1004227.
- 481 Soyer, Jessica L., Möller, M., Schotanus, K., Connolly, L.R., Galazka, J.M., Freitag, M., Stukenbrock, E.H., 2015.  
482 Chromatin analyses of *Zymoseptoria tritici*: Methods for chromatin immunoprecipitation followed by high-throughput  
483 sequencing (ChIP-seq). *Fungal Genet. Biol.* 79, 63–70.
- 484 Soyer, Jessica L, Rouxel, T., Fudal, I., 2015. Chromatin-based control of effector gene expression in plant-  
485 associated fungi. *Current Opinion in Plant Biology* 26, 51–56.
- 486 Stajich, J.E., Harris, T., Brunk, B.P., Brestelli, J., Fischer, S., Harb, O.S., Kissinger, J.C., Li, W., Nayak, V., Pinney,  
487 D.F., Stoeckert, C.J., Roos, D.S., 2012. FungiDB: an integrated functional genomics database for fungi. *Nucleic Acids*  
488 *Res.* 40, D675-681.
- 489 Strauss, J., Reyes-Dominguez, Y., 2011. Regulation of secondary metabolism by chromatin structure and  
490 epigenetic codes. *Fungal Genetics and Biology* 48, 62–69.
- 491 Struhl, K., Segal, E., 2013. Determinants of nucleosome positioning. *Nat. Struct. Mol. Biol.* 20, 267–273.
- 492 Szerlong, H.J., Hansen, J.C., 2011. Nucleosome distribution and linker DNA: connecting nuclear function to  
493 dynamic chromatin structure. *Biochem. Cell Biol.* 89, 24–34.



- 494 Tan, K.-C., Oliver, R.P., 2017. Regulation of proteinaceous effector expression in phytopathogenic fungi. *PLoS*  
495 *pathogens* 13, e1006241.
- 496 Taylor, T.N., Osborn, J.M., 1996. The importance of fungi in shaping the paleoecosystem. *Review of*  
497 *Palaeobotany and Palynology* 90, 249–262.
- 498 Tillo, D., Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Field, Y., Lieb, J.D., Widom, J., Segal, E.,  
499 Hughes, T.R., 2010. High nucleosome occupancy is encoded at human regulatory sequences. *PLoS ONE* 5, e9129.
- 500 Tsankov, A.M., Thompson, D.A., Socha, A., Regev, A., Rando, O.J., 2010. The Role of Nucleosome Positioning  
501 in the Evolution of Gene Regulation. *PLoS Biology* 8, e1000414.
- 502 Valouev, A., Ichikawa, J., Tonthat, T., Stuart, J., Ranade, S., Peckham, H., Zeng, K., Malek, J.A., Costa, G.,  
503 McKernan, K., Sidow, A., Fire, A., Johnson, S.M., 2008. A high-resolution, nucleosome position map of *C. elegans*  
504 reveals a lack of universal sequence-dictated positioning. *Genome Res.* 18, 1051–1063.
- 505 Valouev, A., Johnson, S.M., Boyd, S.D., Smith, C.L., Fire, A.Z., Sidow, A., 2011. Determinants of nucleosome  
506 organization in primary human cells. *Nature* 474, 516–520.
- 507 Varet, H., Brillet-Guéguen, L., Coppée, J.-Y., Dillies, M.-A., 2016. SARTools: A DESeq2- and EdgeR-Based R  
508 Pipeline for Comprehensive Differential Analysis of RNA-Seq Data. *PLoS ONE* 11, e0157022.
- 509 Venters, B.J., Pugh, B.F., 2009. A canonical promoter organization of the transcription machinery and its  
510 regulators in the *Saccharomyces* genome. *Genome Res.* 19, 360–371.
- 511 Wang, M., Weiberg, A., Lin, F.-M., Thomma, B.P.H.J., Huang, H.-D., Jin, H., 2016. Bidirectional cross-kingdom  
512 RNAi and fungal uptake of external RNAs confer plant protection. *Nature Plants* 2, 16151.
- 513 Waterhouse, R.M., Seppey, M., Simão, F.A., Manni, M., Ioannidis, P., Klioutchnikov, G., Kriventseva, E.V.,  
514 Zdobnov, E.M., 2017. BUSCO applications from quality assessments to gene prediction and phylogenomics. *Mol. Biol.*  
515 *Evol.*
- 516 Weiberg, A., Wang, M., Lin, F.-M., Zhao, H., Zhang, Z., Kaloshian, I., Huang, H.-D., Jin, H., 2013. Fungal Small  
517 RNAs Suppress Plant Immunity by Hijacking Host RNA Interference Pathways. *Science* 342, 118–123.
- 518 Yuan, G.-C., Liu, Y.-J., Dion, M.F., Slack, M.D., Wu, L.F., Altschuler, S.J., Rando, O.J., 2005. Genome-scale  
519 identification of nucleosome positions in *S. cerevisiae*. *Science* 309, 626–630.
- 520 Zhang, T., Zhang, W., Jiang, J., 2015. Genome-Wide Nucleosome Occupancy and Positioning and Their Impact  
521 on Gene Expression and Evolution in Plants. *Plant Physiology* 168, 1406–1416.
- 522 Zhu, P., Li, G., 2016. Structural insights of nucleosome and the 30-nm chromatin fiber. *Curr. Opin. Struct. Biol.*  
523 36, 106–115.
- 524
- 525
- 526
- 527



528 **SUPPLEMENTARY DATA**

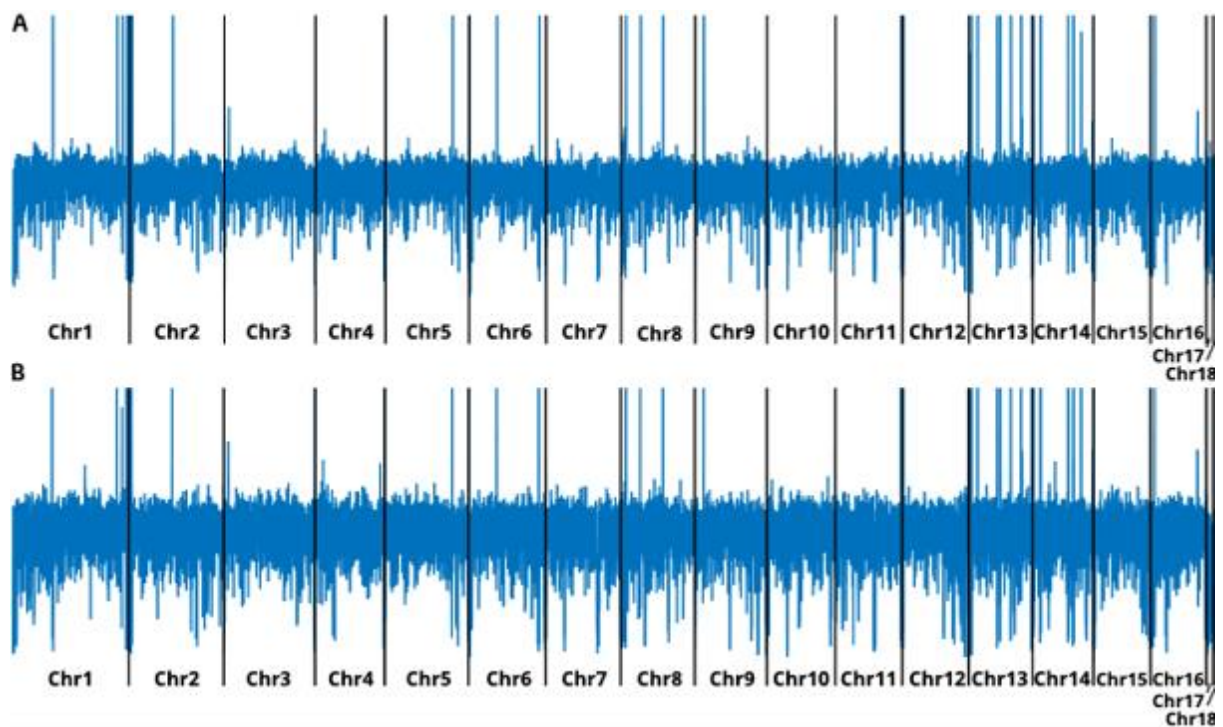
529 Supplementary Data are available at NAR online.

530

531

532

533



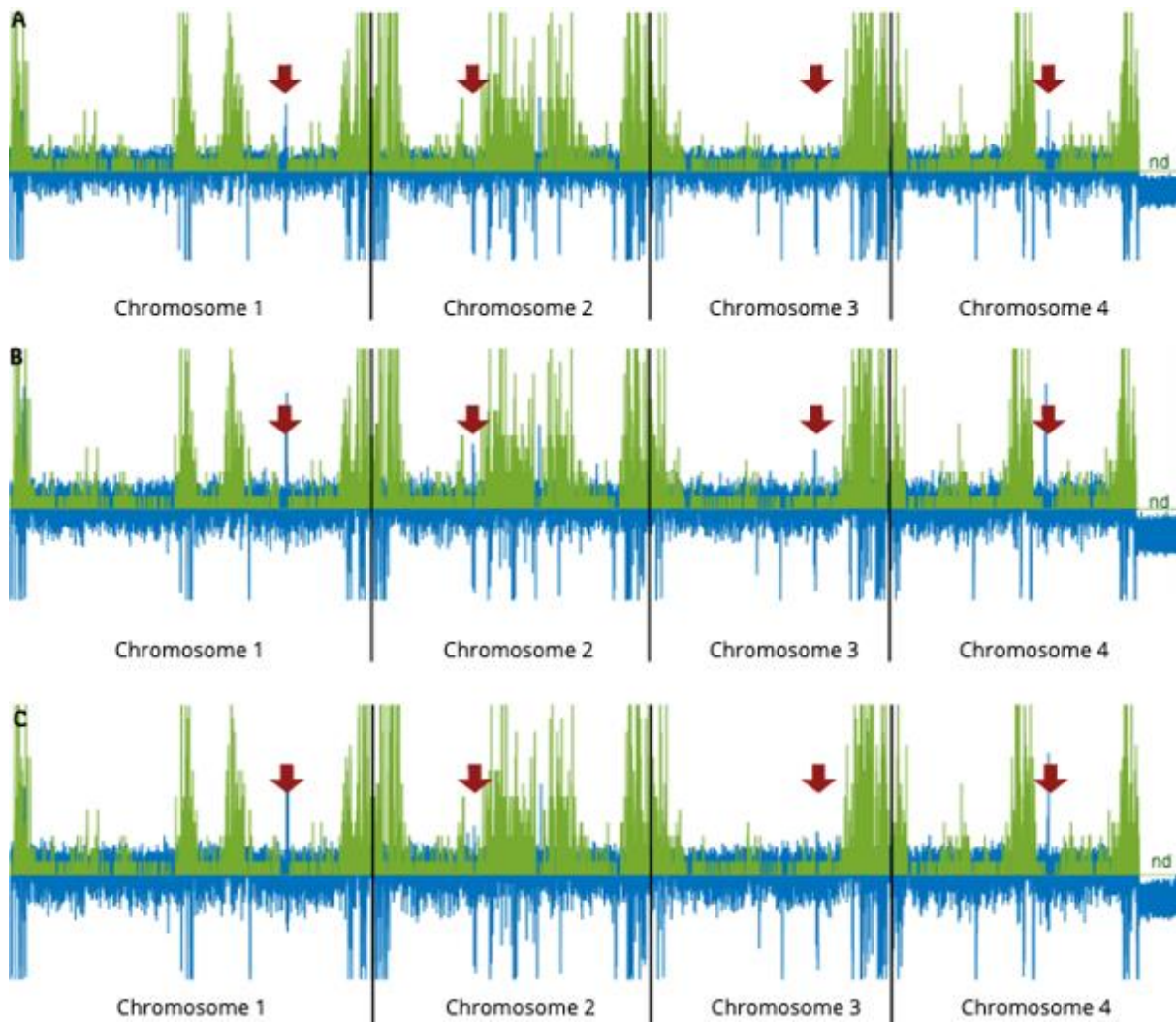
**Figure S1: Nucleosome density profiles in *B. cinerea*.** A. Coverage density profiles were computed for non-overlapping 1 kb-long bins along all chromosomes of *B. cinerea* grown in malt medium (A) or grape juice medium (B).

534

535

536

537

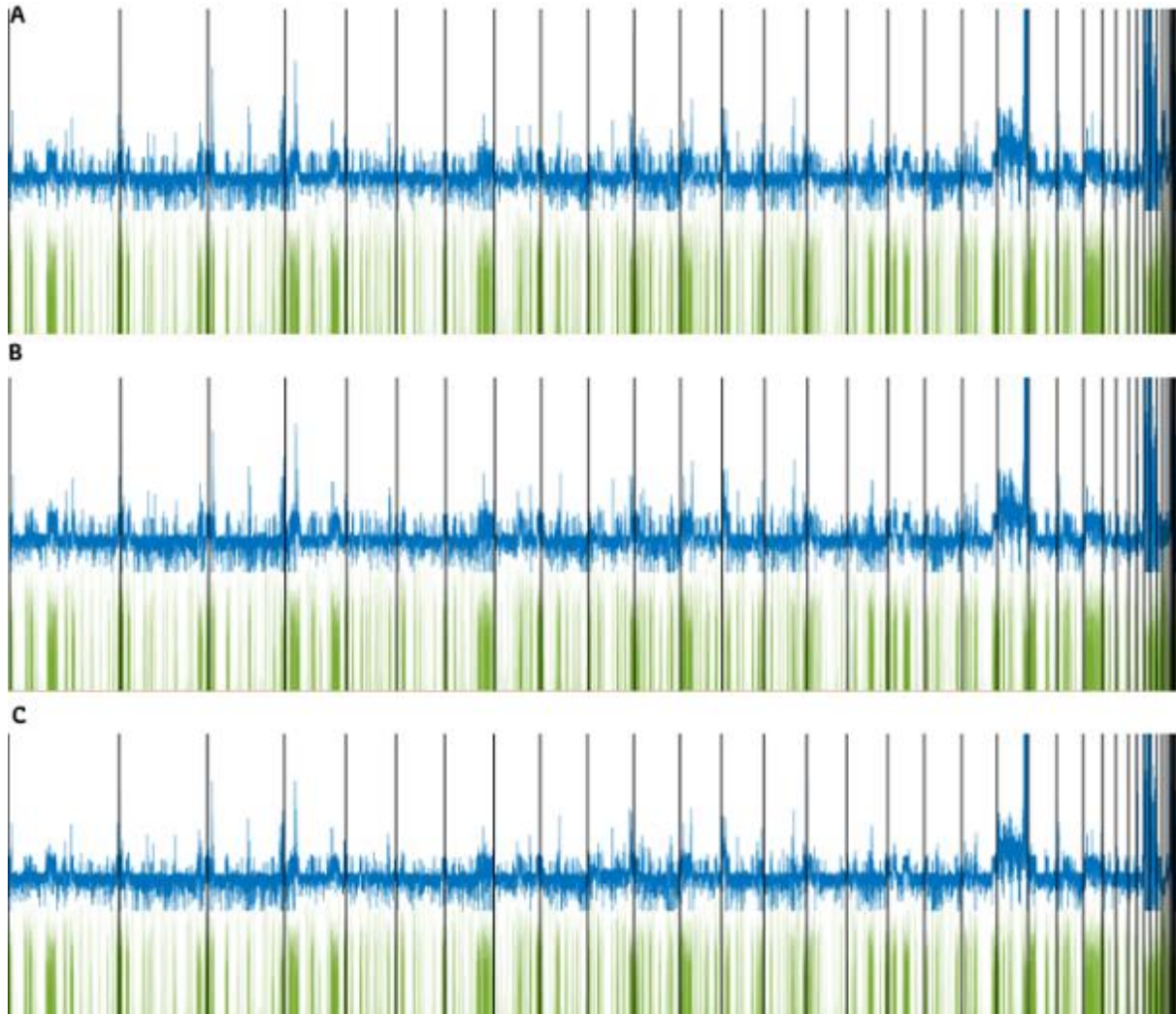


**Figure S2: Nucleosome vs. SNP density profiles in individual biological replicates of *F. graminearum*.** Coverage density profiles were computed for non-overlapping 1 kb-long bins along the four chromosomes of *F. graminearum*. In green are plotted SNP density profiles as previously described (1). “nd” indicates the highly variable 3’ end of chromosome 4 for which SNP were not called. In blue are plotted the z-scored nucleosome density profiles. The red arrows indicate centromeres (2); **A**: Biological replicate #1; **B**: Biological Replicate #2; **C**: Biological replicate #3.



538

539



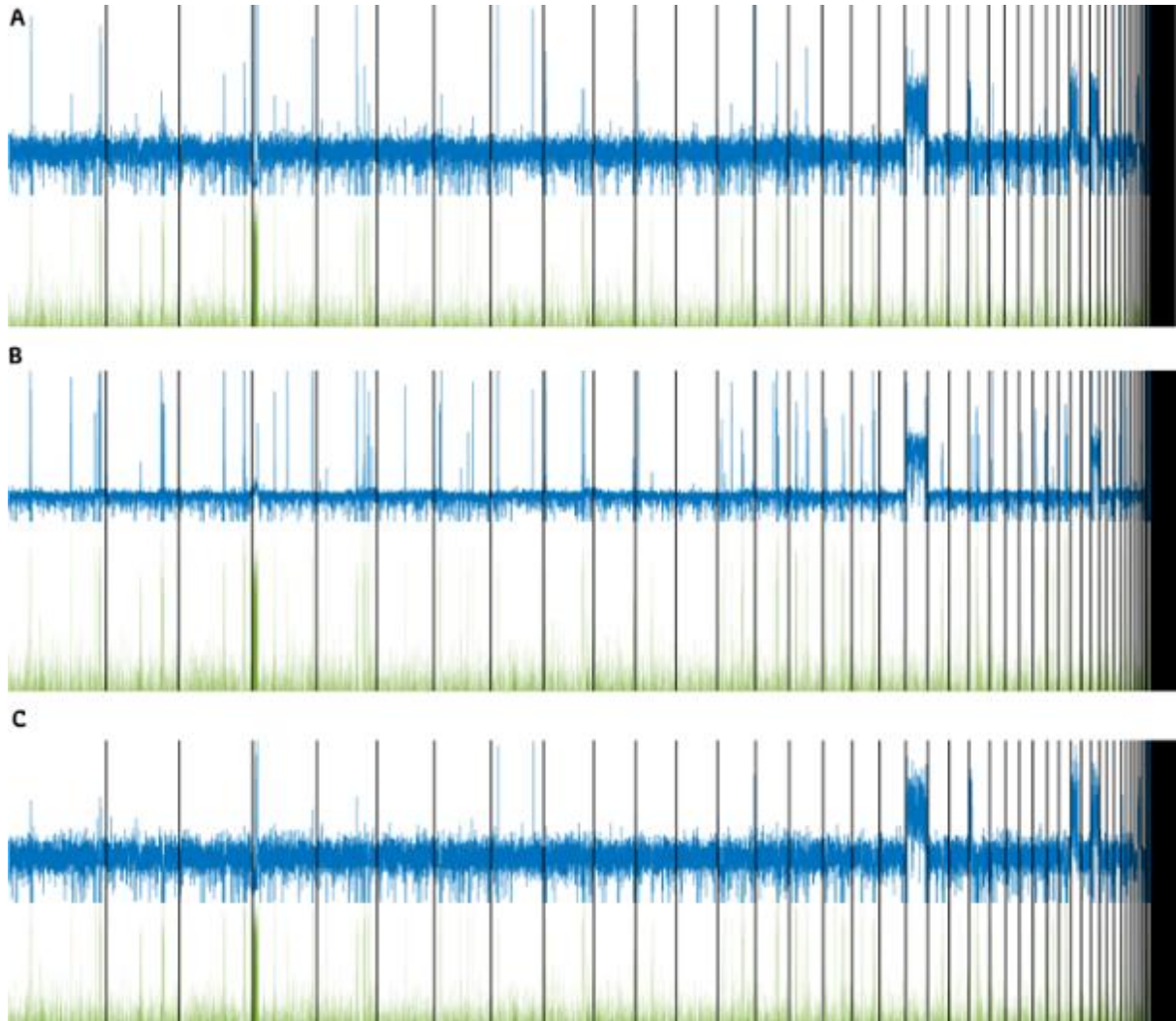
**Figure S3: Nucleosome density profiles vs. AT-rich regions in individual biological replicates of *L. maculans* 'brassicae'.** Coverage density profiles were computed for non-overlapping 1 kb-long bins along all supercontigs, separated by black lines. AT-rich 1 kb-long bins are plotted in green, isochores showing as dense regions. In blue are plotted the z-scored average nucleosome density profile **A**: Biological replicate #1; **B**: Biological Replicate #2; **C**: Biological replicate #3.

540

541

542

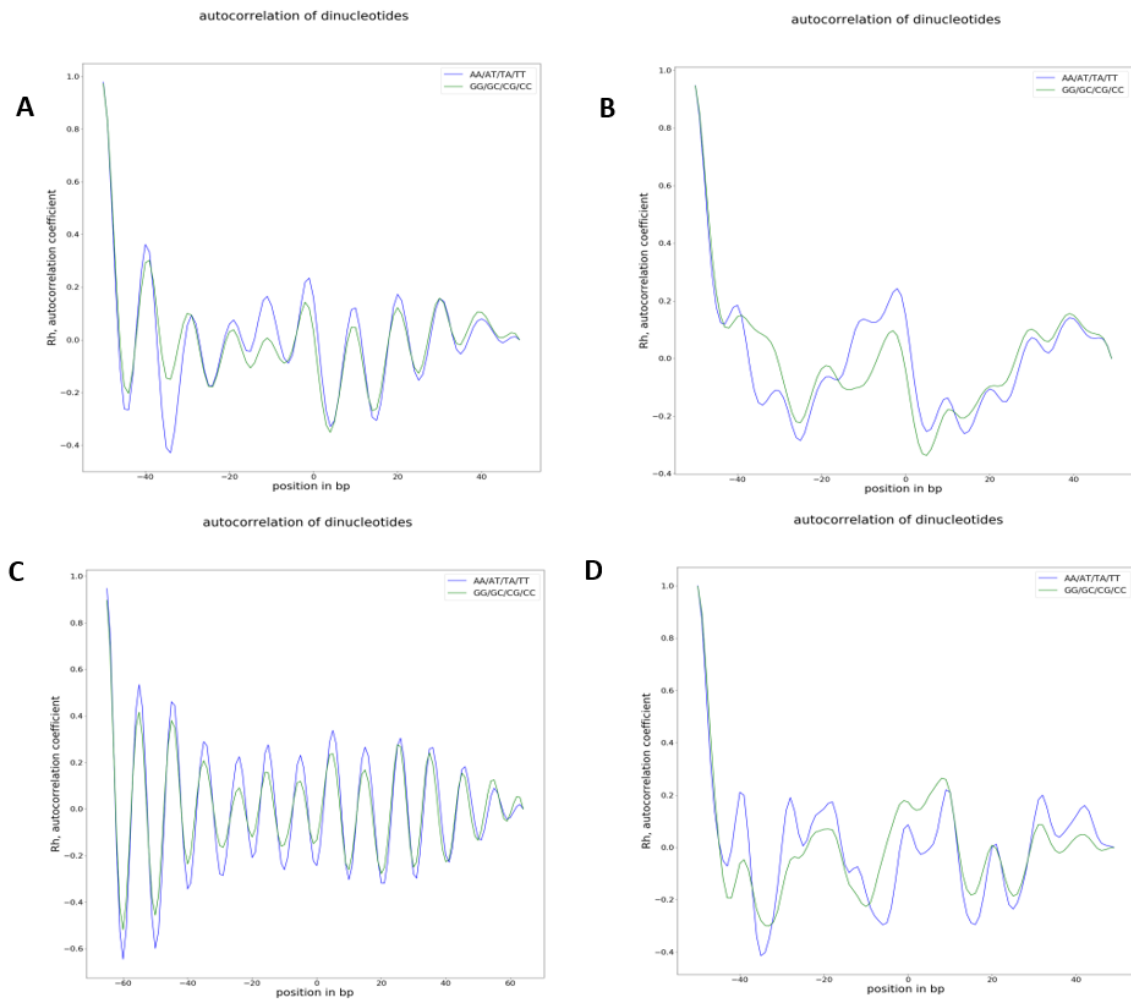
543



**Figure S4: Nucleosome density profiles vs. AT content in individual biological replicates of *L. maculans* 'lepidii'.** Coverage density profiles were computed for non-overlapping 1 kb-long bins along all supercontigs, separated by black lines. AT-rich 1 kb-long bins are plotted in green, with no isochores visible. In blue are plotted the z-scored average nucleosome density profile **A**: Biological replicate #1; **B**: Biological Replicate #2; **C**: Biological replicate #3.

544

545



**Figure S5: Pearson autocorrelation coefficients of di-nucleotides frequencies (average of three biological replicates) for *L. maculans* 'brassicae' (A), *L. maculans* 'lepidii' (B), *F. graminearum* (C), and *B. cinerea* grown in malt medium (D).**

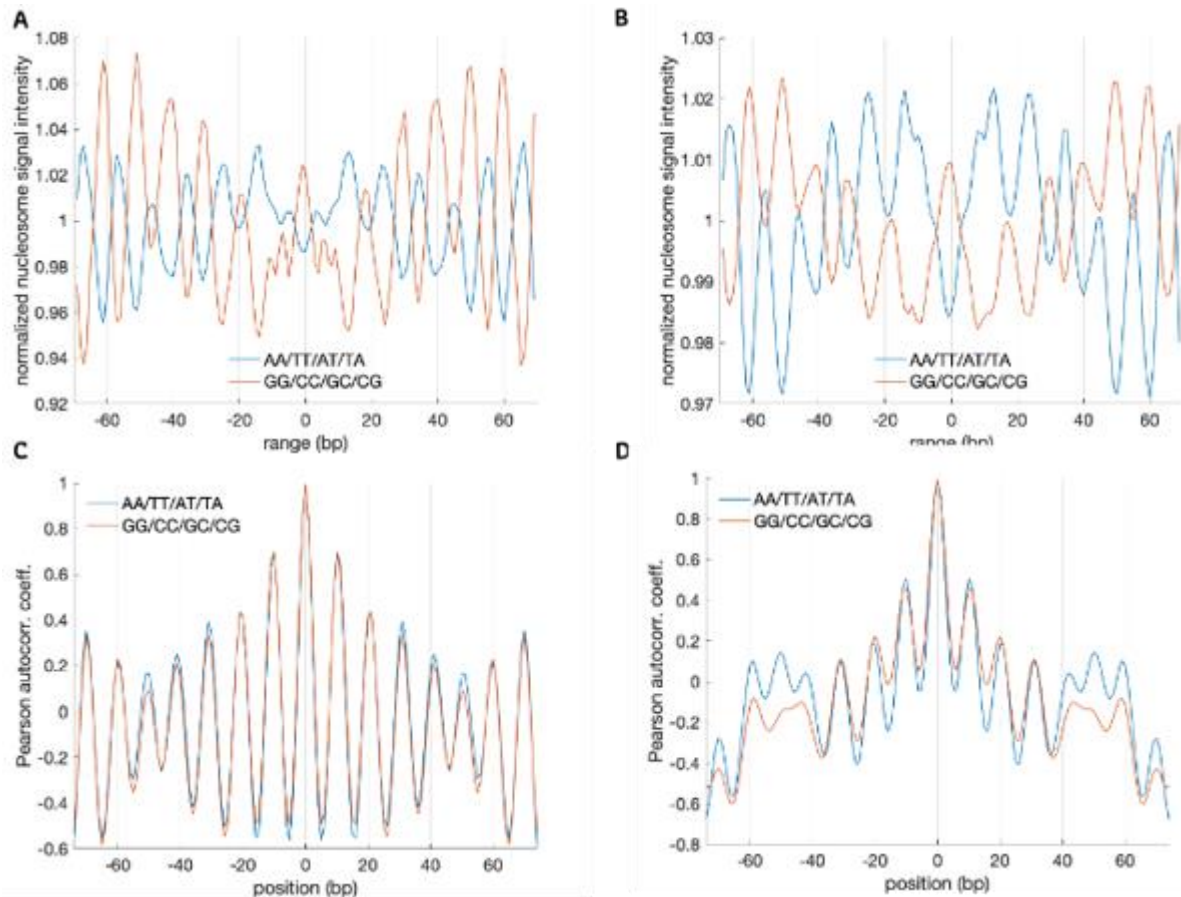
546

547

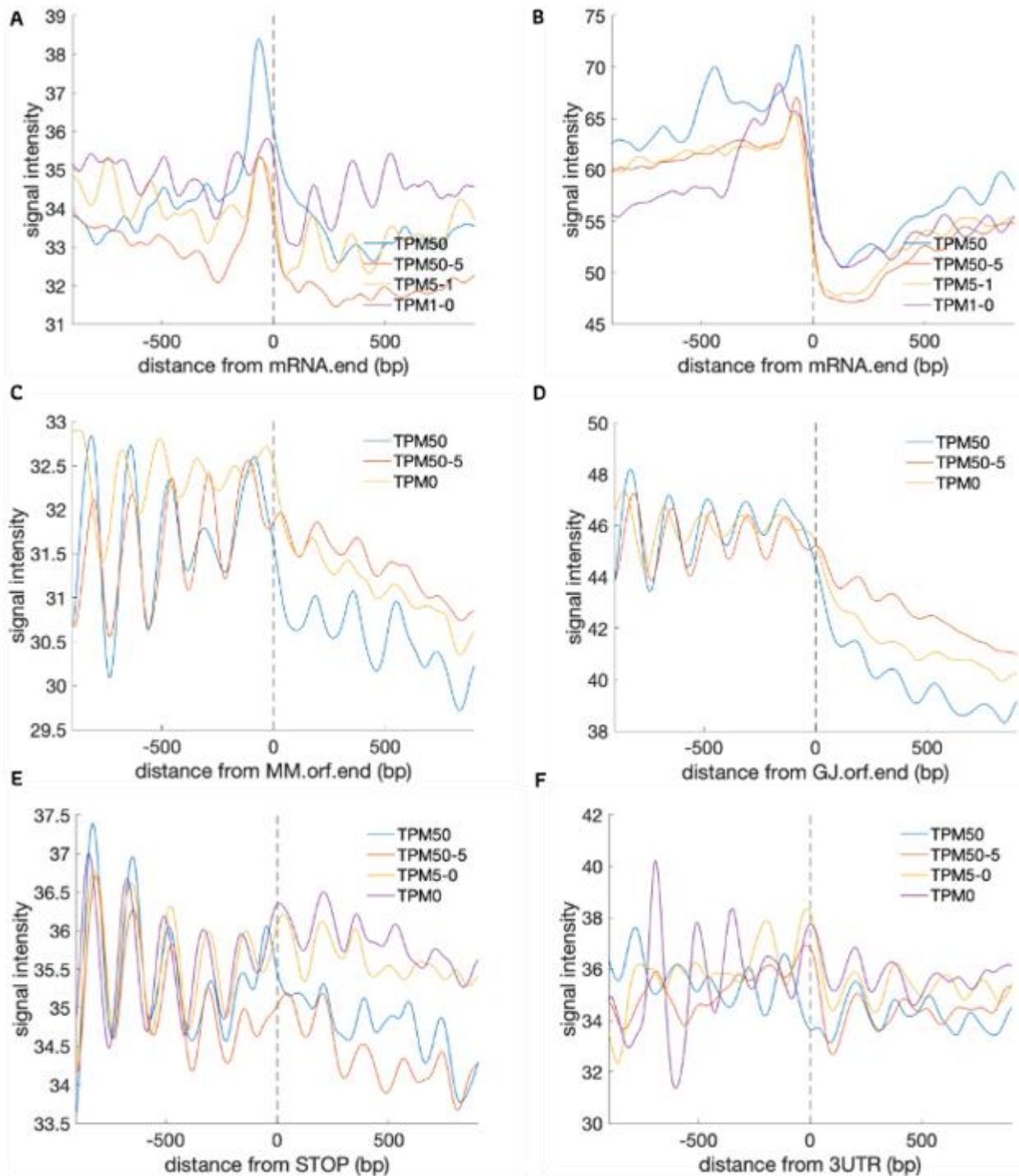
548

549

550



**Figure S6: Repeated di-nucleotide patterns in nucleosomal DNA located in AT and GC isochores of *L. maculans* 'brassicae'.** **A and B:** Normalized di-nucleotides frequency plots (average of three biological replicates) for nucleosomal DNA in AT-isochores (A) and GC-isochores (**B**); **C and D:** Pearson autocorrelation coefficients of di-nucleotides frequencies (average of three biological replicates) for nucleosomal DNA in AT-isochores (C) and GC-isochores (D).



**Figure S7: Nucleosome organization at stop codons/TTS vs. gene expression.** Average (three biological replicates for each fungus/condition) nucleosome signal as a function of position (in base pairs) relative to the stop codon (and TTS for *F. graminearum*). TPM = Transcripts Per Million. **A:** *Leptosphaeria maculans* 'brassicae'; **B:** *Leptosphaeria maculans* 'lepidii'; **C:** *Botrytis cinerea* grown in malt-based medium; **D:** *Botrytis cinerea* grown in grape juice-based medium; **E:** *Fusarium graminearum*, ATG-centred; **F:** *Fusarium graminearum*, TSS-centred.

552 Table S1: Summary of sequencing and mapping metrics

Species	Culture condition	Analysis type	Biological replicate	SRA accession number	Number of raw read pairs	Number of high-quality reads pairs*	Percentage of read pairs aligned to reference genome**
<i>Fusarium graminearum</i>	Liquid MS, 3 days	MNase-seq	Fg#1		38,25,6895	36,297,487	88.1
			Fg#2		38,122,290	36,271,479	86.6
			Fg#3		37,122,439	35,234,660	88.7
		mRNA-seq	Fg#1		34,215,941	31,728,079	98.3
			Fg#2		37,209,377	34,506,796	98.3
			Fg#3		34,108,125	31,674,167	92.7
<i>Leptosphaeria maculans</i> 'brassicae'	Fries 7 days	MNase-seq	NDPS40		54,617,195	52,061,453	83.33
			NDPS41		60,172,292	57,137,679	83.80
			NDPS42		52,577,396	49,579,029	84.65
		mRNA-seq	NDPS13		47,088,366	43,637,526	95.49
			NDPS14		46,486,817	43,749,842	95.66
			NDPS15		51,131,206	48,088,521	95.75
<i>Leptosphaeria maculans</i> 'lepidii'	Fries 7 days	MNase-seq	NDPS37		41,026,352	39,044,192	92.93
			NDPS38		54,531,542	51,203,357	90.55
			NDPS39		55,487,109	52,896,833	91.59
		mRNA-seq	NDPS10		52,095,284	43,637,526	96.58
			NDPS11		50,886,249	49,097,848	96.79
			NDPS12		50,800,169	49,014,722	96.73
		MNase-seq	Bc#1		36,861,968	34,606,209	94.8



<i>Botrytis cinerea</i>	Solid MM, 2 days		Bc#2		38,675,606	36,199,325	92.2
			Bc#3		36,684,375	34,437,003	95.0
	Solid GJ, 2 days		Bc#1		68,917,035	66,005,200	95.6
			Bc#2		38,025,909	36,154,916	96.0
	mRNA-seq		Bc#3		37,107,636	34,557,519	95.0
			Solid MM, 2 days	Bc#1		31,479,569	29,140,587
		Bc#2			38,826,778	35,931,380	97.6
		Bc#3			35,761,014	33,138,116	97.3
		Solid GJ, 2 days	Bc#1		37,895,154	35,310,616	91.0
			Bc#2		36,522,014	33,981,149	94.1
	Bc#3			37,506,128	34,587,157	94.3	

553 \*post quality trimming (see Materials and methods); \*\*concordantly exactly one time

554

555 Table S2: Nucleosome landscapes as a function of gene expression defining metrics

Species	Gene expression category <sup>1</sup>	Position of +1 nucleosome (bp from ATG)	Position of NDR (bp from ATG)	Relative nucleosome signal amplitude <sup>2</sup>
<i>Leptosphaeria maculans</i> 'brassicae'	Not expressed	25	-83	2.0
	Weakly expressed	21	-86	3.7
	Moderately expressed	18	-139	6.9
	Highly expressed	18	-131	8.7
<i>Leptosphaeria maculans</i> 'lepidii'	Not expressed	75	-88	13.4
	Weakly expressed	58	-130	11.9
	Moderately expressed	53	-146	18.8
	Highly expressed	44	-183	23.3
<i>Fusarium graminearum</i>	Not expressed	1	-130	4.3
	Weakly expressed	18	-171	6.7
	Moderately expressed	17	-179	7.2
	Highly expressed	12	-163	10.1
<i>B. cinerea</i> grown in malt medium	Not expressed	75	-90	3.7
	Weakly expressed	20	-99	4.0
	Moderately expressed	25	-104	4.4
	Highly expressed	18	-104	6.1
<i>B. cinerea</i> grown in grape juice medium	Not expressed	147 (with shoulder at 67)	-108	9.6
	Weakly expressed	34	-118	9.5
	Moderately expressed	29	-159	12.3
	<b>Highly expressed</b>	<b>21</b>	<b>-153</b>	<b>15.4</b>
Species	Gene expression category <sup>1</sup>	Position of +1 nucleosome (bp from ATG)	Position of NDR (bp from ATG)	Relative nucleosome signal amplitude <sup>2</sup>





<i>Leptosphaeria maculans</i> 'brassicae'	Not expressed	25	-83	2.0
	Weakly expressed	21	-86	3.7
	Moderately expressed	18	-139	6.9
	Highly expressed	18	-131	8.7
<i>Leptosphaeria maculans</i> 'lepidii'	Not expressed	75	-88	13.4
	Weakly expressed	58	-130	11.9
	Moderately expressed	53	-146	18.8
	Highly expressed	44	-183	23.3
<i>Fusarium graminearum</i>	Not expressed	1	-130	4.3
	Weakly expressed	18	-171	6.7
	Moderately expressed	17	-179	7.2
	Highly expressed	12	-163	10.1

556 <sup>1</sup>According to the average of TPM values for three biological replicates; not expressed: TPM = 0; weakly expressed: 0 < TPM ≤ 5; moderately  
557 expressed 5 < TPM ≤ 50; 50 < TPM

558 <sup>2</sup> $\Delta_{\text{nucl}} = |\text{signal}_{+1\text{nucl}} - \text{signal}_{\text{NDR}}|$

559

560



561

562



## ARTICLE 2

# A comparative analysis of the genome-wide histone modification maps in two subspecies of *Leptosphaeria maculans* showing contrasted genomic organization and different host specialization



1 **A comparative analysis of the genome-wide histone modification maps in two subspecies**  
2 **of *Leptosphaeria maculans* showing contrasted genomic organization and different host**  
3 **specialization**

4 J.L. Soyer<sup>1,2</sup>, C. Clairet<sup>1,3</sup>, E.J. Gay<sup>1,3</sup>, N. Lapalu<sup>1</sup>, T. Rouxel<sup>1</sup>, E.H. Stukenbrock<sup>2,‡</sup>, I. Fudal<sup>1,‡</sup>

5 <sup>1</sup>UMR BIOGER, INRA, AgroParisTech, University of Paris-Saclay, 78850 Thiverval-Grignon, France

6 <sup>2</sup>Max Planck Institute for Evolutionary Biology, August-Thienemann-Str. 2, 24306 Plön, and Christian-Albrechts University of Kiel,  
7 Am Botanischen Garten 1-9, 24118 Kiel, Germany

8 <sup>3</sup>Université Paris-Sud, 91405, Orsay, France.

9 <sup>‡</sup>These authors contributed equally to this work.

10 Corresponding author: J.L. Soyer

11 E-mail: [jessica.soyer@inra.fr](mailto:jessica.soyer@inra.fr)

12 **INTRODUCTION**

13 Each year, hundreds of millions of tons of agricultural crops are devastated by plant pathogenic fungi or made  
14 unfit for consumption due to contamination by mycotoxins. Management strategies to control fungal infection mainly  
15 involve chemical control using fungicides or breeding for naturally resistant crop cultivars. Fungal plant pathogens  
16 have however proven capable of rapidly evolving resistance against implemented fungicides (Fisher *et al.*, 2018) and  
17 to be able to overcome specific plant resistance genes within a few years (for instance in *Leptosphaeria maculans*;  
18 Rouxel & Balesdent, 2017) stimulating the need for the constant development of new control methods.  
19 Comprehending the determinants of their extreme adaptive abilities is a critical issue for the development of effective  
20 and sustainable control methods. In that respect, comparative and population genomic analyses have provided new  
21 insights into the evolutionary dynamics of fungal plant pathogens (e.g. Stukenbrock *et al.*, 2011; Grandaubert *et al.*,  
22 2014; Sánchez-Vallet *et al.*, 2018). Notably, transposable elements (TE) have been shown to play a crucial role in  
23 organizing the genome of plant-pathogenic fungi. TEs are often found to be organized in clusters, compartmentalizing  
24 the genome into gene-rich regions and TE-rich regions (e.g. in *L. maculans* or in *Mycosphaerella fijiensis*; Rouxel *et al.*,  
25 2011; Ohm *et al.*, 2012; Grandaubert *et al.*, 2014). While the gene density in TE-rich regions is low, genes located in  
26 these regions have been shown to evolve faster than genes located in TE-poor regions (Croll & McDonald, 2012; de  
27 Jonge *et al.*, 2013). Interestingly, rapidly evolving genes in TE-rich regions have in many cases been identified as  
28 effector genes. Effectors are considered to be key elements of pathogenesis, allowing pathogens to circumvent host  
29 recognition, impede defence reactions and facilitate host invasion (Sánchez-Vallet *et al.*, 2018). Plants have evolved  
30 strategies to recognize and counteract effectors, exposing them to a strong selection pressure by the host immune  
31 system (Stergiopoulos *et al.*, 2007; Chuma *et al.*, 2011; Daverdin *et al.*, 2012; Rouxel and Balesdent, 2017). Indeed, in  
32 the course of the co-evolution between a pathogen and its host, the host has set up an active defence system allowing  
33 the direct or indirect recognition of some effector molecules, which activate the host immune system triggering



34 a local cell death, called hypersensitive reaction. Effectors that can be recognized by the host are called Avrulence  
35 proteins (Lo Presti *et al.*, 2015). *Leptosphaeria maculans* 'brassicae' (hereinafter referred to as Lmb) belongs to the  
36 Dothideomycete class of Ascomycete fungi and is responsible for the stem canker of oilseed rape (*Brassica napus*).  
37 The life cycle of Lmb is complex and tightly associated to that of its host plant. Lmb displays a complex, hemibiotrophic  
38 life cycle, during which it alternates between different nutrition modes. It causes necrosis on different plant organs:  
39 leaves, more rarely on seedpods, and at the basis of the stem, causing lodging of the plant and yield losses (Rouxel  
40 and Balesdent, 2005). The most efficient method of disease control relies on the use of major resistance gene present  
41 in oilseed rape and other Brassica species. Although efficient, this control method is not sustainable, as Lmb is able to  
42 "break down" novel sources of genetic resistance rapidly (for review, Rouxel and Balesdent, 2017). So far, all described  
43 effector genes, encoding avirulence proteins, are highly expressed in the first seven days of leaf and cotyledon  
44 infection and are located in the middle of TE-rich regions (Gout *et al.*, 2006; Fudal *et al.*, 2007; Parlange *et al.*, 2009;  
45 Balesdent *et al.*, 2013; van de Wouw *et al.*, 2014; Ghanbarnia *et al.*, 2016; 2018; Plissonneau *et al.*, 2018; Petit-  
46 Houdenot *et al.*, 2019). Another set of putative effector genes are specifically expressed during stem infection and  
47 located in gene-rich regions (Gervais *et al.*, 2017). Lmb belongs to the *L. maculans* / *Leptosphaeria biglobosa* species  
48 complex, gathering species with different host specialization and genome organizations. Within the species complex,  
49 the *L. maculans* species infecting oilseed rape, Lmb, is the only one having large regions of its genome enriched in TEs  
50 while other genomes have a low TE-content (~3% while the genome of Lmb has more than 30% of TEs; Rouxel *et al.*,  
51 2011; Grandaubert *et al.*, 2014). In particular, the most closely related species to Lmb, *Leptosphaeria maculans* 'lepidii'  
52 (hereinafter referred to as Lml) is not pathogenic to oilseed rape but infects cruciferous weeds of the *Lepidium* genus  
53 (Mendes-Pereira *et al.*, 2003). Interestingly, their genomes show a high level of macrosynteny, with only a few intra-  
54 chromosomal inversions, but differ in their TE content with Lmb having undergone a massive TE expansion 5 million  
55 years ago corresponding to the speciation date (Grandaubert *et al.*, 2014). Invasion of TE in the genome of Lmb has  
56 shaped its genome, alternating TE-rich regions and gene-rich regions. On the contrary, the Lml genome has a low TE-  
57 content (~3%), with a homogeneous distribution along the chromosomes. There are increasing insights regarding how  
58 dynamic changes of DNA accessibility play fundamental role in adaptation of different individuals to abiotic and biotic  
59 stresses for instance (Gomes and Pelosi, 2013; Zogli and Libault, 2017). Importantly, evidence is accumulating that  
60 transcriptional reprogramming of effector genes or secondary metabolite gene clusters is tightly controlled by  
61 chromatin-based regulatory mechanisms as shown in Lmb, but also in distantly related fungi with different infection  
62 strategies (e.g. in *Fusarium graminearum*, *Epichloe festucae*, *L. maculans*; Connolly *et al.*, 2013; Chujo and Scott, 2014;  
63 Soyer *et al.*, 2014; Soyer *et al.*, 2015a; Soyer *et al.*, 2019). In Lmb, location of effector genes in TE-rich regions has an  
64 influence on their evolution under selection pressure and on regulation of their expression via deposition of a histone  
65 modification typical for constitutive heterochromatin, the tri-methylation of the lysine 9 of histone H3 (H3K9me3), as  
66 shown locally for two avirulence genes (*AvrLm1* and *AvrLm4-7*; Soyer *et al.*, 2014). Lmb has a complex organization of  
67 its genome and involvement of the histone modification H3K9me3 has been demonstrated in the regulation of two  
68 avirulence genes. The most closely related species of Lmb, Lml, displays a well-conserved macrosynteny at the  
69 chromosomal level but devoid of the contrasted chromosomal landscapes typical of Lmb (Grandaubert *et al.*, 2014).  
70 Comparative epigenomic analyses, at the intra- or inter-species levels, are very sparse and remain an



71 underexplored field of study, at least in fungi. In that context, the subspecies Lmb and Lml represent a promising model  
72 for the comparison of chromatin organization. We present here the first comparative epigenomic analysis in two  
73 closely related phytopathogenic fungi, in Lmb and Lml. We have performed ChIP-seq, during axenic culture, aiming at  
74 deciphering and comparing the distribution of three histone modifications, i.e. the di-methylation of the lysine 4 of  
75 histone H3, typical for euchromatin, and two histone modifications typical for heterochromatin, H3K9me3 and the tri-  
76 methylation of the lysine 27 of histone H3 (H3K27me3). We have investigated whether chromatin organization is  
77 conserved between Lmb and Lml notably by investigating whether orthologue genes are located in similar chromatin  
78 domains. The genome-wide histone modifications map generated *in vitro* also allowed investigating whether all known  
79 avirulence genes of Lmb are all located in heterochromatin domains and whether this pattern is conserved for all  
80 putative effector genes of Lmb and Lml.

## 81 MATERIAL AND METHODS

### 82 Fungal isolates

83 The isolate v23.1.3 of *L. maculans* 'brassicae' (hereinafter referred to as Lmb) and isolate IBCN84 of *L. maculans*  
84 'lepidii' (hereinafter referred to as Lml) were used throughout the analyses (Rouxel *et al.*, 2011; Grandaubert *et al.*,  
85 2014). Fungal cultures were maintained as described previously (Ansan-Melayah *et al.*, 1995). For chromatin  
86 immunoprecipitation and transcriptomic analyses performed during *in vitro* growth, mycelium of Lmb and Lml were  
87 grown on V8 agar medium at 25°C for seven days then 10 plugs of mycelium were inoculated into 100 ml of Fries liquid  
88 medium in Roux bottle. Tissue was harvested after growing for 7 days at 25°C, immediately placed in liquid nitrogen  
89 until further used.

### 90 Chromatin immunoprecipitation and high-throughput sequencing

91 ChIP was performed on mycelium grown in Fries liquid culture after seven days, as described in Soyer *et al.*  
92 (2015b), with minor modifications. ChIP was performed on native material (without crosslinking) using antibodies  
93 targeting histone modifications H3K4me2 (Merck ref. 07-030), H3K9me3 (Active Motif, Carlsbad, CA, USA; ref. 39161),  
94 H3K27me3 (Active Motif, Carlsbad, CA, USA; ref. 39155). Three different ChIP (i.e. three biological replicates) were  
95 performed for each of the histone modifications. Libraries were prepared from all biological replicates, individually,  
96 according to the Illumina TruSeq protocol "Ultra Low Input DNA library". Libraries were sequenced on an Illumina  
97 HiSeq 2000 genome analyzer at the Max Planck Genome centre Cologne (Germany). Sequencing data are available  
98 under the SRA accession number xxxx.

### 99 Read mapping and identification of significantly enriched domains

100 Analysis of ChIP-seq datasets was performed as described in Schotanus *et al.* (2015). Quality of Illumina reads  
101 was analysed using FastQC (<https://www.bioinformatics.babraham.ac.uk/projects/fastqc/>). Based on results of this  
102 analysis, ten bp were trimmed from the 5' end. Processed reads were mapped on the reference genome of Lmb  
103 (Dutreux *et al.*, 2018) or Lml (Grandaubert *et al.*, 2014) using Bowtie2 (Langmead and Salzberg, 2012) with default  
104 parameters. Peak calling analysis was performed on each ChIP sequencing dataset, to identify significantly



105 enriched domains for either H3K4me2, H3K9me3 or H3K27me3, using RSEG (Song and Smith, 2011). A significant  
106 domain for any of the histone modification was considered only when the domain was identified in at least two out of  
107 the three biological replicates. The Integrative Genome Viewer (Thorvaldsdóttir *et al.*, 2013) was used to visualize  
108 location of each domains along genomes of Lmb and Lml according to other genome features (gene annotation, TE-  
109 annotation, GC content). Coverage of any of the histone modifications was assessed in 10 kb non-overlapping sliding  
110 windows along the supercontigs and correlation analyses, using Kendall's T correlation coefficient, was performed  
111 between biological replicates to check for reproducibility of the biological replicates. In order to assess whether  
112 H3K4me2-, H3K9me3- or H3K27me3-domains are significantly enriched in certain categories of genes (such as effector  
113 genes, secondary metabolite encoding genes), a Chi<sup>2</sup> test was applied to compare the expected proportion of a given  
114 category of genes across the entire genome to the observed distribution of the gene category in the H3K4me2-,  
115 H3K9me3- or H3K27me3-domains (Soyer *et al.*, 2019). Enrichment was considered significant with a *P value* < 0.01;  
116 analyses were done in R ([www.r-project.org](http://www.r-project.org)).

### 117 **RNA extraction, RNA-sequencing and expression analysis**

118 Total RNA was extracted from mycelium grown for one week in Fries liquid medium as previously described  
119 (Fudal *et al.*, 2007). For Lmb, RNA-seq libraries were prepared with the TruSeq Stranded mRNA kit (Illumina, San Diego,  
120 CA, USA) and was sequenced using 101 bp paired end reads on a HiSeq2000 Illumina sequencer. For Lml, the NEBNext  
121 Ultra Directional RNA Library Prep Kit for Illumina (cat. # E7420L New England BioLabs) was used to prepare RNA-seq  
122 libraries and sequencing was performed on an HiSeq 4000 Illumina using 50 paired end reads. Raw reads were then  
123 pre-processed with Trimmomatic (Bolger *et al.*, 2014) to remove short reads (<30 bp) and eliminate sequencing  
124 adaptors. Cleaned reads were then mapped against each genome using STAR default parameters (Dobin and Gingeras,  
125 2015). Sequencing data are available under the SRA accession number xxxx.

### 126 **Identification of orthologues and annotation of genes encoding SSP in Lmb and Lml**

127 Based on the new assembly and annotation available for the genome of Lmb (Dutreux *et al.*, 2018), an updated  
128 repertoire of putative effector genes has been predicted. Therefore, to compare effector repertoires of Lmb and Lml,  
129 the same pipeline for the prediction of putative effector-encoding genes was applied to both species. Signal peptide  
130 and subcellular localization were predicted by SignalP version 4.1 and TargetP version 1.1 respectively.  
131 Transmembrane domains were predicted by TMHMM version 2.0. The predicted secretome contains all proteins which  
132 have at most one transmembrane domain and either a predicted signal peptide or an extracellular localization. The  
133 final effector repertoires were created by applying a size cut-off of 300 amino acid on the predicted secretome. In  
134 parallel, EffectorP version 1.0 was used on the predicted secretome and, for Lml, four proteins were predicted as  
135 effector with a size superior to 300 amino-acid, these four proteins were added to the SSP set. This predicted a  
136 repertoire of 1,080 and 892 SSP-encoding genes respectively for Lmb and for Lml.

137 The genome of Lmb encodes 13,126 proteins and that of Lml encodes 11,272 proteins (Dutreux *et al.*, 2018;  
138 Grandaubert *et al.*, 2014). In order to find orthologues between Lmb and Lml, proteins of these two species were  
139 groups using OrthoFinder with default parameters (Emms and Kelly, 2015).



## 140 GO analysis and enrichment analysis

141 GO annotation of the Lmb and Lml encoding genes were retrieved from Dutreux *et al.* (2018) and Grandaubert  
142 *et al.* (2014) respectively. Gene Ontology (GO) term enrichment analysis of the H3K4me2-, H3K9me3- and H3K27me3-  
143 associated genes was performed with the plug-in Biological networks Gene ontology (BinGo; v3.0.3) of the cytoscape  
144 software (Shannon *et al.*, 2003). List of genes submitted to BINGO were considered as significantly enriched for a given  
145 GO term with an associated a False Discovery Rate  $\leq 0.01$  for the biological processes.

146 A Chi<sup>2</sup> test, or a Fisher's exact test for small sized-population, were applied to identify significant enrichment  
147 of H3K4me2-, H3K9me3- or H3K27me3-domains for certain categories of genes: the expected proportion of a given  
148 category of genes across the entire genomes of Lmb or Lml was compared to the observed distribution of the gene  
149 category in the H3K4me2-, H3K9me3- or H3K27me3-domains. Enrichment was considered significant with a *P value* <  
150 0.01. All analyses were done in R ([www.r-project.org](http://www.r-project.org)).

## 151 RESULTS

### 152 Identification of H3K4me2-, H3K9me3- and H3K27me3-enriched domains in the genomes of Lmb and Lml during 153 axenic growth

154 In order to get a genome-wide view of the epigenomic landscapes in the genomes of Lmb and Lml during  
155 axenic growth, we have performed a CHIP-seq experiment using antibodies against histone modifications H3K4me2,  
156 H3K9me3 and H3K27me3 (**Table S1**). Mapping of the sequencing data was followed by identification of the significantly  
157 enriched domains, for any of the histone modifications targeted, in the genomes of Lmb and Lml (see Materials and  
158 Methods). Three different precipitation experiments were done for each histone modification to assess the robustness  
159 of the domains identified with RSEG (Song and Smith, 2011). Domains identified for one of the H3K9me3 replicate  
160 performed for Lml were inconsistent with domains identified for this modification with the two other replicates (**Table**  
161 **S2**) and was therefore not considered for subsequent analyses. For all other biological replicates aiming at identifying  
162 H3K4me2-, H3K9me3- and H3K27me3-domains in the genomes of Lmb and Lml, location of domains was consistent  
163 (**Table S2, S3**; Kendall's *T* >0.74, *P* <2.2.10<sup>-16</sup>). Hence, for subsequent analyses, domains were considered as being  
164 significantly enriched when identified with at least two out of the three replicates, except for H3K9me3-domains of  
165 Lml for which domains were considered enriched when identified with both replicates kept for analysis. Based on the  
166 coordinates of the location of H3K4me2-, H3K9me3- or H3K27me3-domains, the number of bases associated with any  
167 of the three histone modifications was assessed, for each genome, in 1 kb sliding-windows. In the genomes of Lmb  
168 and Lml, the proportion of H3K4me2 and H3K27me3 was similar in both species while proportion of H3K9me3 in the  
169 genome of Lmb was higher compared to that of Lml (33% and 4% respectively for Lmb and Lml); (**Tables S4, S5**).

### 170 TE-rich domains are associated with H3K9me3 while gene-rich regions are enriched in H3K4me2

171 In the genomes of Lmb and Lml, H3K4me2 and H3K9me3 domains are exclusive and not associated with the  
172 same genomic regions (Kendall's *T*: -0.57 and -0.16, *P* < 2.2.10<sup>-16</sup> for respectively Lmb and Lml; **Figure S1**; **Tables S6**  
173 and **S7**). A positive correlation was measured between TEs and H3K9me3-domains (Kendall's *T*: 0.87 and 0.68, *P*



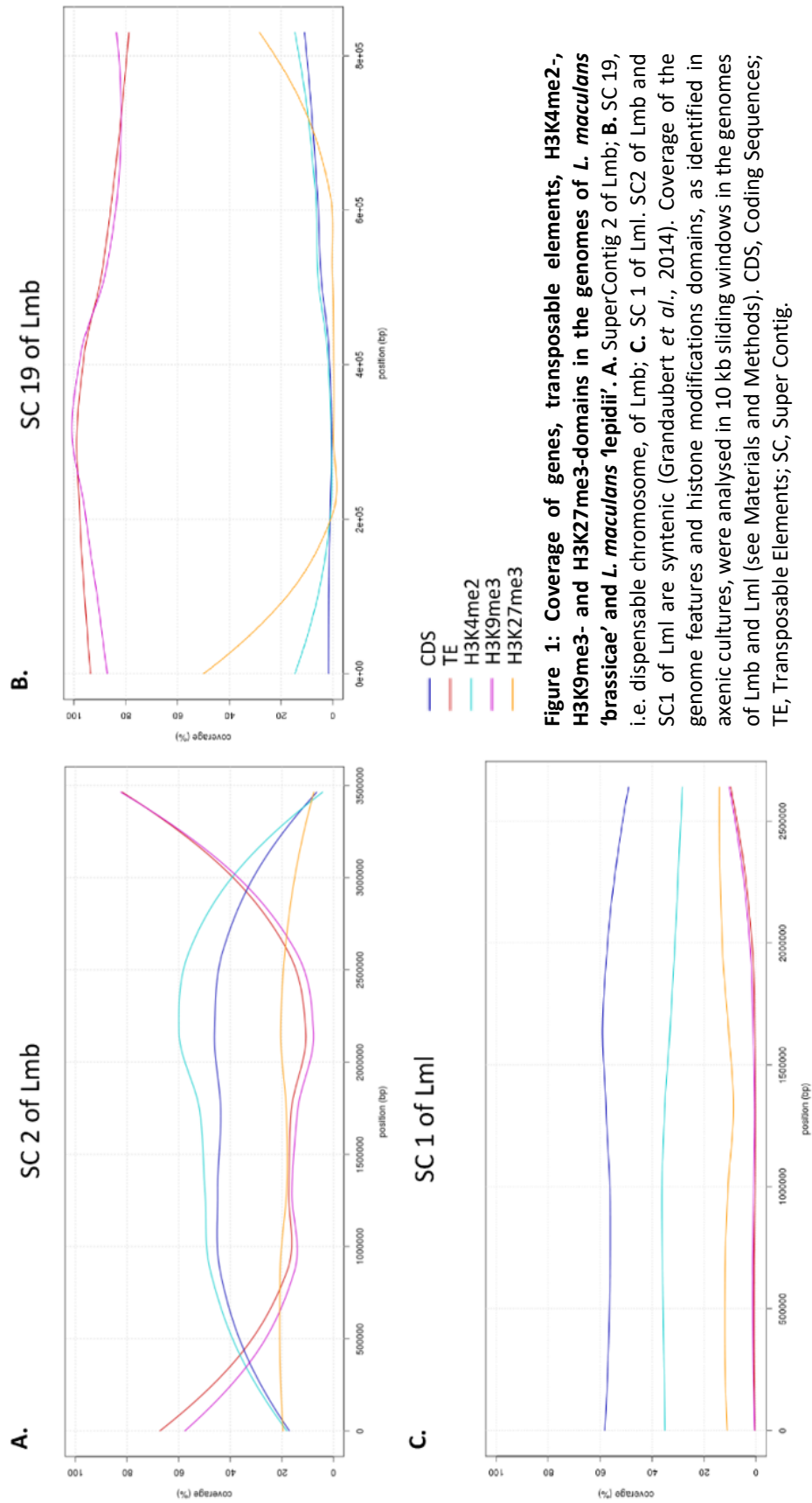


174 <  $2.2 \cdot 10^{-16}$  for respectively Lmb and Lml; **Figure 1A, 1C; Figure S1; Table S6** and S7) and between CDS and H3K4me2  
 175 (Kendall's  $T$ : 0.55 and 0.25,  $P < 2.2 \cdot 10^{-16}$  for respectively Lmb and Lml; **Figure 1A, 1C; Figure S1; Tables S6** and **S7**) in  
 176 both genomes. In the genome of Lmb, a weak correlation was measured between CDS and H3K27me3 domains  
 177 (Kendall's  $T$ : 0.12,  $P < 2.2 \cdot 10^{-16}$ ) while no clear pattern regarding location of H3K27me3 was identified, on the genome-  
 178 wide scale, in the genome of Lml (**Figure S1; Tables S6, S7**). In accordance with the association of H3K9me3 with TE,  
 179 the proportion of H3K9me3 is consistent with the TE-content of both genomes (36% and 2.8% TE in the genomes of  
 180 Lmb and Lml, respectively). The high TE-content in the Lmb genome has an impact at the structural level, as large AT-  
 181 isochores corresponding to stretches of TE-rich regions and GC-isochores alternate along the genome while it is not  
 182 the case in the Lml genome, with a homogeneous repartition of the TEs (example in **Figure 2**). Hence, while the median  
 183 size of the domains were similar in both genomes, Lmb genome displays extremely large H3K9me3-domains, with a  
 184 size up to 230 kb, while the maximum size for the H3K9me3-domains is 12 Kb in the genome of Lml (**Figure 3**).

185

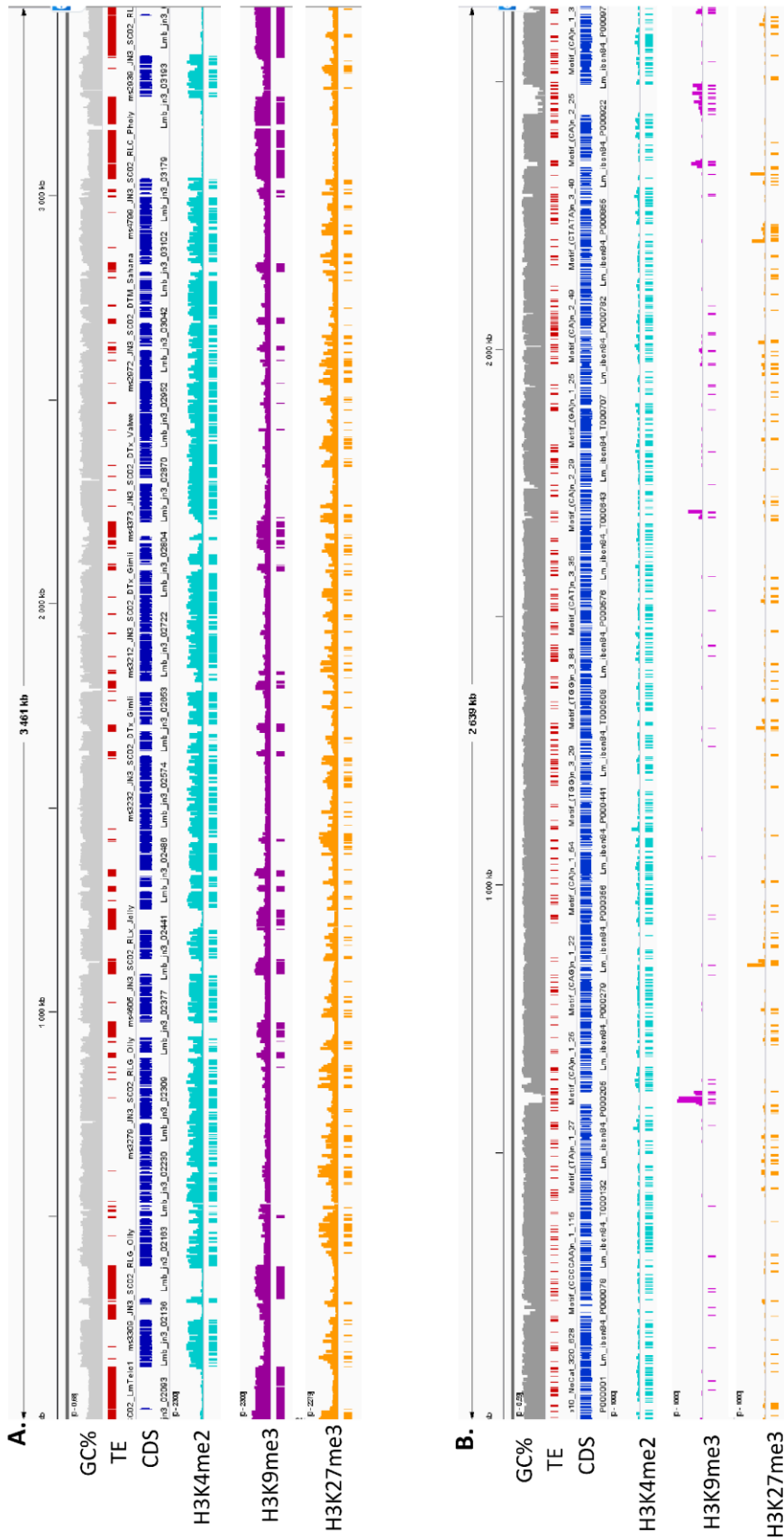
186





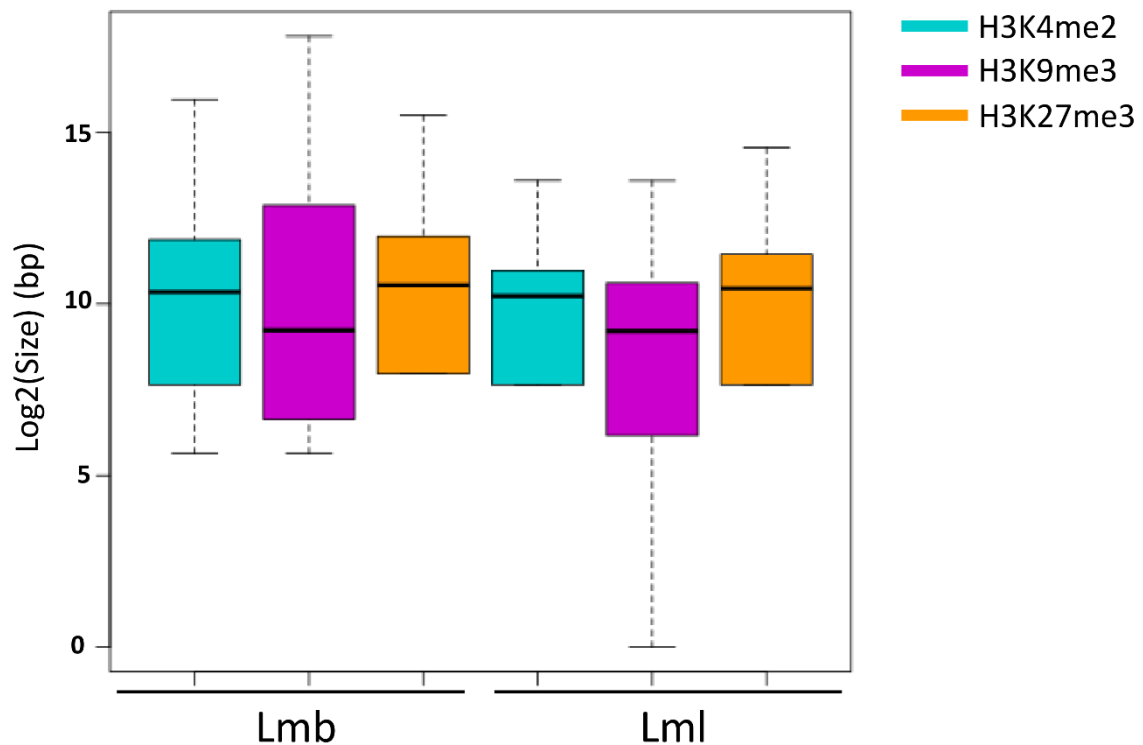
**Figure 1: Coverage of genes, transposable elements, H3K4me2-, H3K9me3- and H3K27me3-domains in the genomes of *L. maculans* 'brassicae' and *L. maculans* 'epidii'. A. SuperContig 2 of Lmb; B. SC 19, i.e. dispensable chromosome, of Lmb; C. SC 1 of Lml. SC2 of Lmb and SC1 of Lml are syntenic (Grandaubert *et al.*, 2014). Coverage of the genome features and histone modifications domains, as identified in axenic cultures, were analysed in 10 kb sliding windows in the genomes of Lmb and Lml (see Materials and Methods). CDS, Coding Sequences; TE, Transposable Elements; SC, Super Contig.**





**Figure 2. Genome of *L. maculans* 'brassicae' harbors large TE-rich, H3K9me3-domains compared to *L. maculans* 'lepidii'.** **A.** Example of SC 2 of Lmb; **B.** SC 1 of Lml. ChIP-seq was performed with antibodies targeting H3K4me2 (cyan), H3K9me3 (purple) or H3K27me3 (orange); rectangles indicate location of significantly enriched domains regions for were identified with RSEG (Song and Smith, 2014) and the number below the rectangles indicate the number of biological replicates for which the domain was identified as significantly enriched (Materials and Methods). GC%, percentage of GC along the SC; TE, Transposable element; H3K4me2, di-methylation of lysine 4 of histone H3; H3K9me3, tri-methylation of lysine 9 of the histone H3; H3K27me3, tri-methylation of the lysine 27 of histone H3.



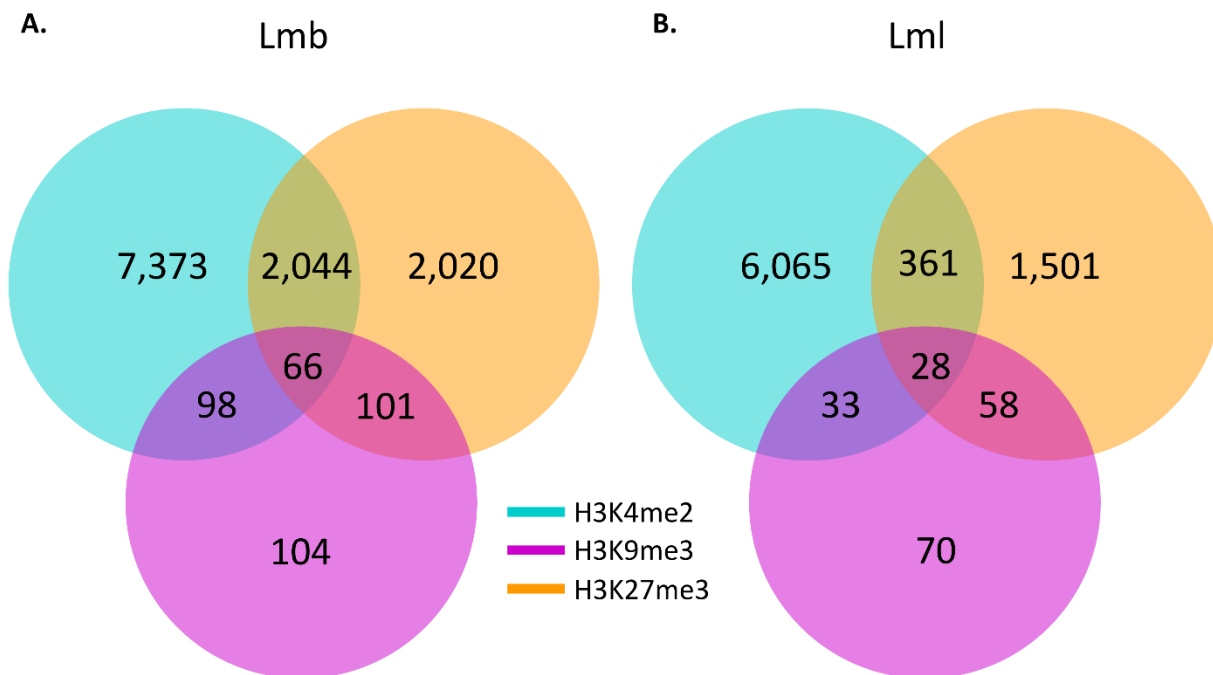


**Figure 3: Size of the H3K4me2-, H3K9me3-, H3K27me3-domains in the genomes of *Leptosphaeria maculans* 'brassicae' (Lmb) and *L. maculans* 'lepidii' (Lml).** Log2 of the Size of the domains was estimated based on the coordinates of the location of the domains, identified using RSEG (Song and Smith, 2011). Blue: H3K4me2; Purple: H3K9me3; Orange: H3K27me3. Lmb, *L. maculans* 'brassicae'; Lml, *L. maculans* 'lepidii'.

190 In the genome of Lmb, a dispensable chromosome, carrying the avirulence gene *AvrLm11*, has been identified  
 191 (Leclair *et al.*, 1996; Balesdent *et al.*, 2013; Grandaubert *et al.*, 2014), while no such accessory chromosomes was  
 192 identified in the genome of Lml. The dispensable chromosome is enriched in TEs compared to the rest of the genome  
 193 (35% of TEs in the core genome and 93% of TEs in the dispensable chromosome). Consistently, a strong enrichment in  
 194 H3K9me3 is observed on the dispensable chromosome (32% in the core genome and 90% on the dispensable  
 195 chromosome), with only a few H3K4me2- and H3K27me3-domains (**Figure 1A** and **1B**; **Figure 2**). Contrary to the  
 196 genome of Lml which is not fully assembled with large gaps remaining in its genome sequence (Grandaubert *et al.*,  
 197 2014), the complete assembly of the Lmb genome allows analysis of the sub-telomeric regions (Dutreux *et al.*, 2018).  
 198 Sub-telomeric regions of the Lmb genome, corresponding to large repeat sequences, are highly associated with  
 199 H3K9me3 and H3K27me3 overlap with H3K9me3 in the telomeric regions, while they do not on chromosome arms  
 200 (**Figure 1A** and **Figure 2**).

202 **H3K4me2- and H3K27me3-domains shelter genes involved in distinct activities**

203 Based on the location of H3K4me2-, H3K9me3- and H3K27me3-domains in the genomes of Lmb and Lml during  
204 axenic cultures, genes associated with any of these modifications were identified. Overall, the number of genes  
205 associated with any of the histone post-translational modification is similar in both species with more than 50% of the  
206 predicted genes associated with H3K4me2, ~14% of the genes associated with H3K27me3 and only a few genes  
207 associated with H3K9me3 (**Figure 4**).



**Figure 4: Number of genes associated with H3K4me2, H3K9me3 or H3K27me3 during axenic growth in A. *Leptosphaeria maculans* 'brassicae' (Lmb) and B. *L. maculans* 'lepidii'.** Location of histone modifications in the genome of Lmb and Lml were identified using RSEG (Song and Smith, 2011). Blue: H3K4me2; Purple: H3K9me3; Orange: H3K27me3. Genes were considered as associated with any of the histone modifications when at least one pb of the gene was found within the borders of the domain. Lmb, *L. maculans* 'brassicae'; Lml, *L. maculans* 'lepidii'.

208 In the genome of Lmb, a GO annotation was assigned to 5,076 genes (i.e. 39% of the predicted genes). 3,276  
209 genes located within a H3K4me2-domain had a GO annotation (i.e. 44% of the genes located in a H3K4me2-domain)  
210 while 572 genes located in a H3K27me3-domain (i.e. 28% of the genes associated with H3K27me3) and none of the  
211 genes located in a H3K9me3-domain had an annotation. This shows that genes located in heterochromatin are  
212 enriched in genes with no functional annotation ( $\chi^2$  test,  $P < 2.2 \cdot 10^{-16}$ ). H3K27me3-associated genes mainly encode  
213 proteins of unknown function and only a few GO terms were identified among genes associated with this fraction of  
214 the genome. Nevertheless, these genes are significantly enriched in GO terms involved in carbohydrate metabolic  
215 process (GO:0005975; 82 genes associated with H3K27me), oxydo-reduction process (GO:0055114; 149 genes) and  
216 transmembrane transport (GO:0055085; 104 genes), ( $P \leq 0.01$ ); (**Table S8**). This suggests that genes associated with  
217 H3K27me3 *in vitro* might encode proteins involved in pathogenesis-related function such as secretion of molecules,  
218 degradation of organic compounds, and defence against oxydo-reductant stress. On the contrary, genes associated  
219 with H3K4me2 display a wide variety of annotated function involved in primary metabolism (**Table S9**).

220 In the genome of Lml, 3,725 genes had a GO annotation (i.e. 33% of the annotated genes). As for Lmb, genes  
221 associated with heterochromatin *in vitro* were enriched in genes with no GO annotation as only 19% of the genes  
222 associated with H3K27me3 and none of the genes associated with H3K9me3 had a GO annotation ( $X^2$  test,  $P < 2.2 \cdot 10^{-16}$ ).  
223 <sup>16</sup>). On the contrary, H3K4me2-domains were enriched in genes having a GO annotation (42% of these genes). As it is  
224 the case for genes associated with H3K4me2 in Lmb, genes associated with H3K4me2 are also enriched in genes having  
225 a GO annotation involved in primary metabolism and basic cellular functions, such as translation (GO:0006412; 206  
226 genes) or cellular protein metabolic process (GO: 44267; 477 genes); (**Table S10**). As for Lmb, only a few GO annotation  
227 enrichments was detected for genes associated with H3K27me3-domains *in vitro*. Only one GO enrichment was found  
228 in common with Lmb, i.e. carbohydrate metabolic process (GO:0005975; 41 genes). Apart from that, all other  
229 enrichments were redundant in terms of definition: response to chemical (GO: 42221; 20 genes), response to nitrogen  
230 compound (GO:1901698; 10 genes), response to organophosphorus (GO: 46683, 10 genes) and others. All these GO  
231 terms indicate an enrichment in genes involved in a change of state or activity of an organism (in terms of movement,  
232 secretion, enzyme production, gene expression, etc.) as a result of different chemical stimuli. For instance,  
233 organophosphorus compounds are toxic cholinesterase inhibitors. Although most GO enrichments found among sets  
234 of genes associated with H3K27me3 in Lmb and Lml do not overlap, both types of enrichments suggest that  
235 H3K27me3-associated genes might be involved in a stress response mechanism, whether it might be an infection-  
236 related process or a defence against toxic compounds, or both. On the contrary, in both organisms, proteins encoded  
237 by the genes sheltered in H3K4me2-domains in Lmb and Lml are involved in similar biological processes involved in  
238 maintenance of basic cellular functions.

### 239 **Heterochromatin domains are enriched in effector genes and secondary metabolite-encoding genes**

240 In Lmb, so far, all cloned avirulence genes, which are highly expressed during asymptomatic infection of oilseed  
241 rape (at seven days post infection), are located within TE-rich regions of the genome. We assessed whether the  
242 particular location of known avirulence genes within TE-rich regions is also associated with a location in domains  
243 enriched for histone modifications typical for heterochromatin and whether this is a conserved feature for different  
244 sets of genes encoding pathogenicity-related proteins. In the genome of Lmb, 1,080 genes encoding effectors  
245 (including known avirulence genes; i.e. 8.2% of the genes predicted in the genome) were predicted (Dutreux *et al.*,  
246 2018; Gay *et al.*, in prep.). In the genome of Lmb, 2,478 genes (i.e. 12% of the genes) are associated with TE-rich regions  
247 (located within a 2 kb distance of a TE sequence), including 289 genes encoding effectors. Hence, while a new  
248 prediction of the effector repertoire was performed based on the new assembly of the Lmb genome, TE-rich regions  
249 are significantly enriched in effector genes, as already shown by Rouxel *et al.* (2011;  $X^2$  test,  $P = 9.6 \cdot 10^{-16}$ ). Similarly,  
250 H3K9me3- and H3K27me3-associated genes are significantly enriched in genes encoding effectors, as effector genes  
251 represent respectively 36% and 14% of the genes associated with these histone modifications *in vitro* ( $X^2$  test,  $P =$   
252  $9.6 \cdot 10^{-16}$ ; **Table 1**).



253 On the contrary, H3K4me2-associated genes contain significantly less effector genes than the rest of the  
254 genome (i.e. 6% of the genes located in H3K4me2-domains encode putative effectors; **Table 1**). In the genome of Lml,  
255 the proportion of effector genes is similar to that of Lmb (892 of the genes encode for putative effectors, i.e. 7.9% of  
256 the genome). Likewise, TE-rich regions, H3K9me3- and H3K27me3-domains are significantly enriched in putative  
257 effector genes while H3K4me2-domains are significantly deprived in such genes compared to the rest of the genome  
258 (**Table 1**).

**Table 1. Number of effector genes located in different genomic compartments in *L. maculans* 'brassicae' and *L. maculans* 'lepidii'**

		genome	TE-associated genes <sup>a</sup>	H3K4me2-domains <sup>b</sup>	H3K9me3-domains <sup>b</sup>	H3K27me3-domains <sup>b</sup>	H3K9+H3K27me3-domains <sup>b</sup>
Lmb	number of genes	13,047	2,478	7,373	104	2,020	101
	number of effectors	1,080	289	433	38	286	24
	proportion	8.2.10 <sup>-2</sup>	1.2.10 <sup>-1</sup>	5.8.10 <sup>-2</sup>	3.6.10 <sup>-1</sup>	1.4.10 <sup>-1</sup>	2.4.10 <sup>-1</sup>
	P value <sup>c</sup>	-	9.6.10 <sup>-16</sup>	6.7.10 <sup>-14</sup>	2.2.10 <sup>-16</sup>	2.2.10 <sup>-16</sup>	1.6.10 <sup>-8</sup>
Lml	number of genes	11,273	641	6,065	70	1,501	58
	number of effectors	892	79	266	14	217	18
	proportion	7.9.10 <sup>-2</sup>	1.2.10 <sup>-1</sup>	4.4.10 <sup>-2</sup>	2.10 <sup>-1</sup>	1.45.10 <sup>-1</sup>	3.1.10 <sup>-1</sup>
	P value <sup>c</sup>	-	3.5.10 <sup>-5</sup>	2.2.10 <sup>-16</sup>	1.8.10 <sup>-4</sup>	2.2.10 <sup>-16</sup>	6.9.10 <sup>-11</sup>

<sup>a</sup>Genes located up to 2 kb upstream or downstream of a transposable element sequence;

<sup>b</sup>Genes located in a H3K4me2-, H3K9me3-, H3K27me3- or H3K9me3/H3K27me3-domain *in vitro*;

<sup>c</sup>A X<sup>2</sup> test was applied to compare proportion of effector genes in the genome and in the genomic compartment analysed (Materials and Methods).

259 The genome also encodes secondary metabolites, synthesised by clusters encoding PKS/NRPS (Polyketide  
260 Synthases and Non-Ribosomal Peptide Synthases). Twenty-nine such genes have been predicted in Lmb (Rouxel *et al.*,  
261 2011; Elliott *et al.*, 2013; Grandaubert *et al.*, 2014). Among them, only one has been detected so far to be highly up-  
262 regulated at 7 days post-infection, and located within AT-isochores of the genome of Lmb, i.e. *pks5* involved in the  
263 synthesis of the abscisic-acid, along with the other genes of this cluster (Darma *et al.*, 2019). Regarding the twenty-  
264 nine genes encoding PKS/NRPS. For Lmb, H3K27me3-domains are enriched in such genes as 68% are located in such  
265 domains (20 PKS/NRPS; Fisher's exact test,  $P= 1.3.10^{-7}$ ). This included the cluster encoding abscisic acid, embedded in  
266 the middle of TEs, and for which five genes of the cluster are located in a H3K27me3-domain, including the  
267 transcription factor regulating the cluster and the PKS, while three other genes of the cluster are associated with both  
268 H3K9me3 and H3K27me3 (our results; Darma *et al.*, 2019). The gene encoding the sirodesmin, having antibacterial  
269 and antifungal properties (SirP; Gardiner *et al.*, 2004; Elliott *et al.*, 2007) is also located in a H3K27me3 domain, while  
270 not located in the middle of an AT-isochore; this gene is expressed during the infection of oilseed rape stem (Elliott *et al.*  
271 *et al.*, 2007). The PKS gene encoding phomenoic acid, *PKS7*, highly expressed during axenic culture (Elliott *et al.*, 2013),  
272 is located in a GC-isochore and associated with H3K27me3. In the genome of Lml, 25 genes are predicted as PKS/NRPS



273 (Grandaubert *et al.*, 2014). Similarly to Lmb, H3K27me3-domains of the genome of Lml are enriched in PKS/NRPS  
274 encoding genes (11 PKS/NRPS; Fisher's exact test,  $P = 1.7 \cdot 10^{-4}$ ).

275 **Heterochromatin domains of the *L. maculans* 'brassicae' and *L. maculans* 'lepidii' genomes shelter subspecies-**  
276 **specific genes**

277 Comparative genomic analyses of the Lmb and Lml genes identified that 7,393 genes were conserved between  
278 both species using Orthofinder (Emms and Kelly, 2015). We assessed location of these conserved genes and identified  
279 that H3K4me2-domains are enriched in genes conserved between Lmb and Lml (66 and 64% of the conserved genes  
280 respectively in Lmb and Lml;  $\chi^2$  test,  $P < 2.2 \cdot 10^{-16}$ ). On the contrary, H3K27me3- and H3K9me3-domains are  
281 significantly enriched in genes which are not conserved between Lmb and Lml (genes conserved between both species  
282 represent less than 1% of the genes located in these domains;  $\chi^2$  test,  $P < 2.2 \cdot 10^{-16}$ ).

283 Regarding effectors, 1,080 and 892 genes in the genome of respectively Lmb and Lml encode putative  
284 effectors, among which 274 were conserved between both sub-species (i.e. genes having an orthologue in the sister  
285 species and being predicted as a putative effector). Hence, more than two third of their effector repertoire is specific,  
286 as already described (Grandaubert *et al.*, 2014) but confirmed here with a new prediction of this repertoire. Effector  
287 genes of both species are less conserved than the rest of the predicted genes ( $\chi^2$  test,  $P < 2.2 \cdot 10^{-16}$ ) and H3K9me3- and  
288 H3K27me3-domains are enriched in sub-species-specific effector genes. Five genes encoding avirulence proteins  
289 located in the middle of a large TE-rich region of Lmb are also conserved in Lml. Four out of these five genes are located  
290 in a sub-telomeric region, including three neighbour genes (ID *Lmb\_jn3\_05822*, *Lmb\_jn3\_05823*, *Lmb\_jn3\_05824*) and  
291 consistently associated either with H3K9me3, H3K27me3 or with both histone modifications. These three putative  
292 effectors are conserved, although inverted, in Lml and are also in located a sub-telomeric region, and associated with  
293 H3K27me3. One avirulence gene of Lmb, *AvrLm10A*, is conserved in Lml. In Lmb, this gene is located in a sub-telomeric  
294 region and associated with a H3K9me3-domain. In Lml, the *AvrLm10A* orthologue is not located in a sub-telomeric  
295 region but is associated with H3K27me3.

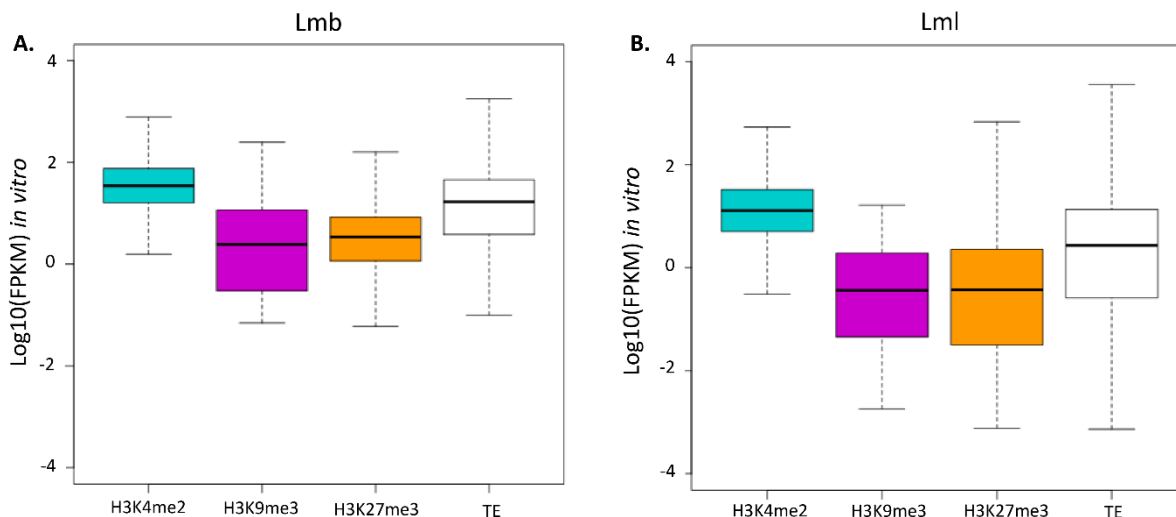
296 **All cloned Avirulence genes of *L. maculans* 'brassicae' are located in H3K9me3-domains *in vitro***

297 H3K4me2 is typically considered as permissive for gene expression while H3K9me3 and H3K27me3 are  
298 considered as inducing gene silencing. We performed a genome-wide transcriptomic analysis *in vitro*, in parallel to the  
299 analysis of histone modifications, in order to determine the role of histone modifications on the control of gene  
300 expression and whether that role was conserved between Lmb and Lml. In Lmb, 11,525 genes (i.e. 88% of the predicted  
301 genes) are expressed during *in vitro* growth (i.e. FPKM  $\geq 2$ ). Considering the 100 most expressed genes during axenic  
302 culture, 91% of them are associated with H3K4me2 while five are located within a H3K27me3-domain and none in a  
303 H3K9me3-domain. This observation was confirmed on the genome-wide scale as H3K4me2-domains are enriched in  
304 gene expressed (7,191 of the genes associated with H3K4me2 are expressed;  $P < 2.2 \cdot 10^{-16}$ ) during axenic culture while  
305 H3K9me3- and H3K27me3-domains are enriched in genes silenced *in vitro* (56 and 1290 of the genes located  
306 respectively in a H3K9me3- or a H3K27me3-domain are expressed,  $P < 2.2 \cdot 10^{-16}$ ); (**Figure 5**). The influence of the  
307 histone modifications on gene expression is similar in Lml as, among the 100 most expressed genes during axenic





308 culture, 71 are located within a H3K4me2-domain while five are located in a H3K27me3-domain. This pattern is  
309 confirmed on the genome-wide scale, as among the 7,735 genes expressed *in vitro* by Lml (i.e. 69% of the predicted  
310 genes), H3K4me2-domains are enriched in genes expressed (5,472 of the genes expressed during axenic culture are  
311 located in a H3K4me2-domain;  $P < 2.2 \cdot 10^{-16}$ ). On the contrary, H3K9me3- and H3K27me3-domains are enriched in  
312 genes which are silenced during axenic culture (17 and 403 genes expressed *in vitro* are located respectively in a  
313 H3K9me3- or a H3K27me3-domain *in vitro*;  $P < 2.2 \cdot 10^{-16}$ ). This transcriptomic analysis confirmed, in both species, that  
314 genes located in a H3K4me2-domain were more likely to be expressed while H3K9me3- and H3K27me3-domains are  
315 enriched in silent genes.



**Figure 5: Genes associated with heterochromatin are less expressed than genes associated with euchromatin in the genomes of *L. maculans* 'brassicae' and *L. maculans* 'lepidii'.** Location of histone modifications in the genome of Lmb and Lml were identified using RSEG (Song and Smith, 2011), during axenic culture. RNA-seq was performed from Lmb or Lml grown one week in FRIES media (see Materials and methods). Blue: H3K4me2; Purple: H3K9me3; Orange: H3K27me3; White: genes located within a 2 kb distance of a transposable element. Lmb, *L. maculans* 'brassicae'; Lml, *L. maculans* 'lepidii'.

316 So far, pathogenicity assays are not feasible with Lml, hence transcriptomic data *in planta* are not available.  
317 Considering Lmb, expression profile of avirulence genes is well documented during infection of oilseed rape  
318 cotyledons, showing that these genes are not, or lowly, expressed *in vitro* and highly expressed during infection of  
319 oilseed rape cotyledons, with a peak of expression seven days post-infection (Rouxel *et al.*, 2011). All the avirulence  
320 genes cloned so far in Lmb are located in a H3K9me3-domain *in vitro* consistently with the fact that these genes are  
321 located in TE-rich compartments of the genome in Lmb. Besides these avirulence genes expressed during the early  
322 stages of oilseed rape infection, 307 putative effector genes expressed during stem infection were recently described  
323 (Gervais *et al.*, 2017). A new version of the Lmb genome is now available, with a new annotation and a new effector  
324 gene prediction (Dutreux *et al.*, 2018; Gay *et al.*, in prep.), downsizing this repertoire of putative 'late' effectors to 86  
325 candidates. In contrast to avirulence genes, most of these 'late' effector genes are located in GC-isochores, as already  
326 observed by Gervais *et al.* (2017). Only 19 of these genes are located in H3K4me2-domains while 49 are located in a  
327 H3K9me3- or H3K27me3-domain ( $P < 2 \cdot 10^{-5}$ ). Altogether, these data show that, independently of the infection stage,  
328 putative effector genes are enriched in either H3K9me3 or H3K27me3 *in vitro*.

329



330 **DISCUSSION**

331 We present here the first comparative analysis of the location of three different histone modifications,  
332 typically associated with either euchromatin or heterochromatin, at the genome-wide scale, in two closely related  
333 phytopathogenic fungi displaying considerable differences in terms of genomic organisation as well as host range. Aim  
334 of this analysis was to question whether different genome organisations would shape the underlying associated  
335 epigenomic landscape and to identify conserved, or unique, features.

336 As for other fungi, domains enriched in H3K4me2, indicative of euchromatin regions, and domains enriched in  
337 H3K9me3, indicative of a heterochromatin region, are mutually exclusive in the genomes of Lmb and Lml. This was  
338 also shown for instance in *Neurospora crassa*, *Fusarium fujikuroi*, and *Zymoseptoria tritici* (Smith *et al.*, 2008; Jamieson  
339 *et al.*, 2013; Wiemann *et al.*, 2013; Schotanus *et al.*, 2015). Regarding other features, such as association of repeats  
340 with H3K9me3 or H3K27me3 or organization of the accessory chromosomes, there appears to be more and more  
341 exceptions as data are gathered. So far, H3K9me3 has always been found associated with transposable elements,  
342 whether or not sequences of transposable element span very large domains, as it is the case in Lmb (e.g. in *Z. tritici*;  
343 *Schizosaccharomyces pombe*, *F. fujikuroi*; Cam *et al.*, 2005; Schotanus *et al.*, 2015; Wiemann *et al.*, 2013). In sub-  
344 telomeric regions of Lmb, H3K9me3 and H3K27me3-domains overlap over repetitive sequences, which is a classic  
345 feature observed also in *N. crassa* or *Z. tritici* (Smith *et al.*, 2008; Connolly *et al.*, 2013; Jamieson *et al.*, 2013; Schotanus  
346 *et al.*, 2015). Only H3K27me3 has been shown to locate in sub-telomeric areas in *F. graminearum* although a genome-  
347 wide analysis of the location of H3K9me3 (although present in its genome) has not been performed yet for this fungus  
348 (Connolly *et al.*, 2013). In the genomes of Lmb and Lml, TEs consist in heterochromatin regions enriched in H3K9me3  
349 and contrary to sub-telomeric regions, no enrichment of H3K27me3 was detected associated with TEs on  
350 chromosomal arms. This is dissimilar to the organization of TE-rich regions in *Z. tritici*, enriched in both  
351 heterochromatin modifications (Schotanus *et al.*, 2015). The situation is also different for *Fusarium oxysporum*, in  
352 which most TE sequences are embedded in a H3K27me3-domain. It is worth noting that location of H3K9me3 in the  
353 genome of this fungus has not been investigated yet (Fokkens *et al.*, 2018). One can envision that, once investigated,  
354 the situation of the transposons of *F. oxysporum* might be identical to those of *Z. tritici*. In *Z. tritici*, the overlap between  
355 H3K9me3 and H3K27me3 in TE-rich regions was postulated to be a consequence of the lack of cytosine DNA  
356 methylation (Schotanus *et al.*, 2015), as the *dim-2* encoding gene is inactivated due to mutations caused by the RIP  
357 mechanisms (Dhillon *et al.*, 2010). In the genome of Lmb, almost no cytosine DNA methylation was detected as well  
358 (Bewick *et al.*, 2019), nevertheless the two heterochromatin modifications do not overlap in TE-rich regions. Therefore,  
359 overlapping of H3K9me3 and H3K27me3 observed in TE-rich regions of *Z. tritici* cannot be explained by a genome-  
360 wide lack of DNA methylation. In other fungi, such as *Aspergillus nidulans*, no H3K27me3 has been detected while  
361 H3K4me3 is associated with gene-rich regions and H3K9me3 flanks sub-telomeric regions (Gacek-Matthews *et al.*,  
362 2016; Freitag, 2017). For some other phytopathogenic fungi, such as *Botrytis cinerea*, epigenomic analysis have been  
363 set up and data from western blot analysis have helped detect H3K9me3 and H3K27me3 but a precise location is  
364 lacking so far (Schumacher *et al.*, 2019). All these data tend to indicate that epigenomic analyses are of a major interest  
365 for the scientific community interested in plant-pathogenic fungi.



366 An extreme feature of fungi, although not limited to this kingdom, is the prevalence of accessory chromosomes  
367 (also called B chromosomes or dispensable chromosomes) in their genome (Galazka and Freitag, 2014; Soyer *et al.*,  
368 2018). So far, most studies regarding accessory chromosomes have pointed them out in plant-interacting Ascomycetes  
369 although this might be a bias because their detection can be delicate. Some species harbor TE-rich accessory  
370 chromosome(s) (e.g. in *Lmb*, *Z. tritici*, *F. oxysporum*; Ma *et al.*, 2010; Rouxel *et al.*, 2011; Balesdent *et al.*, 2013; Dhillon  
371 *et al.*, 2014; Grandaubert *et al.*, 2015). Their heterochromatin nature is uncovered through the advent of ChIP-seq  
372 analyses. Accessory chromosomes of *Z. tritici*, *F. oxysporum* or *F. fujikuroi*, *F. asiaticum* are enriched in H3K27me3  
373 (Wiemann *et al.*, 2013; Galazka and Freitag, 2014; Schotanus *et al.*, 2015; Studt *et al.*, 2016; Fokkens *et al.*, 2018) as  
374 are those of *F. graminearum* while not enriched in TEs (Connolly *et al.*, 2013). Interestingly, here we shown for *Lmb*  
375 that its accessory chromosome is TE-rich and almost entirely associated with H3K9me3, which is different to other  
376 fungi.

377 Dogmas regarding conventional definitions of facultative or constitutive heterochromatin seem not to be valid  
378 in all plant pathogenic fungi. While histone modifications H3K9me3 and H3K27me3 are both a signature of  
379 heterochromatin, H3K9me3 is considered to be typical for constitutive heterochromatin, as associated with repeats  
380 and involved in genome stability (Möller *et al.*, 2019), and H3K27me3 is considered to be associated with facultative  
381 heterochromatin and considered to be easily reversed toward a euchromatin state under some stresses from the  
382 environment. As H3K9me3 is found associated with TE-rich regions, the proportion of H3K9me3 in a genome strictly  
383 reflects the TE content, as described in *Z. tritici* for instance (Schotanus *et al.*, 2015). This is also the case for *Lmb* and  
384 *Lml*. In *Lmb*, having >30% TE, 33% of the genome is associated with H3K9me3 while the genome of *Lml*, having a low  
385 TE content, presents a low enrichment in H3K9me3 (4% of H3K9me3 in the *Lml* genome). The unique dispensable  
386 chromosome of *Lmb* is mostly repetitive and 90% of this chromosome is associated with H3K9me3 (Rouxel *et al.*, 2011;  
387 Balesdent *et al.*, 2013; Our study). Co-location of H3K9me3 and TEs is in favor of a “constitutive” nature of this  
388 modification together with the fact that H3K9me3-domains encompass a small number of genes in *Lmb* and *Lml*  
389 (respectively 70 and 104 genes), as in *Z. tritici* (86 genes; Schotanus *et al.*, 2015; Soyer *et al.*, 2019). Genes associated  
390 with H3K9me3 are almost all located in the middle of repetitive elements, in sub-telomeric areas, or very close to the  
391 edge of regions enriched in repeated elements. This is in accordance with the fact that TEs are targeted by H3K9me3  
392 which might leak to neighbor genes. Nevertheless, some of these genes are heavily transcribed upon host infection  
393 (Soyer *et al.*, 2019; Chujo and Scott, 2014; Soyer *et al.*, 2014; Our study). This does not match the “constitutive” nature  
394 of this modification. Even though the largest proportion of genes is found associated with H3K4me2 from detailed  
395 ChIP-seq analyses performed so far *in vitro* in plant-interacting fungi (i.e. *F. graminearum*, *F. oxysporum*, *Z. tritici*, *Lmb*  
396 and *Lml*; Connolly *et al.*, 2013; Schotanus *et al.*, 2015; Fokkens *et al.*, 2018; Soyer *et al.*, 2019; Our study), a significant  
397 proportion of the predicted genes are found associated with H3K27me3. This is in accordance with the “facultative”  
398 nature of this histone modification. On the other side, some evidence is in favor of its important role in genome  
399 organization and stability. For instance, dispensable chromosomes of *Z. tritici* are twice as rich in TEs as core  
400 chromosomes, but although no significant enrichment of these chromosomes in H3K9me3 has been identified, they  
401 are entirely covered by H3K27me3 (Schotanus *et al.*, 2015). Besides, loss of H3K27me3 in *Z. tritici* increases stability of



402 some accessory chromosomes (Möller *et al.*, 2019). Altogether, these features seem to indicate that although  
403 H3K27me3 is an important regulator of gene expression involved in development or response to various stresses, it  
404 might also play a role in the maintenance of some parts of fungal genomes.

405 The comparative epigenomic analysis corroborates previous comparative genomic analysis and previous  
406 epigenetic studies in this species complex (Grandaubert *et al.*, 2014; Soyer *et al.*, 2014). Difference in terms of genome  
407 organization between Lmb and Lml (i.e. alternation of large stretches of TEs and large gene-rich compartments in Lmb  
408 vs. little TE content, no bipartite organization of the genome of Lml) is consistent with the organization of the  
409 epigenomic landscape between both species. In Lmb, very large H3K9me3-domains are present, spanning large TE-  
410 rich regions (the largest one being found on the dispensable chromosome of Lmb), while such extremely large  
411 H3K9me3-domains are not observed in the Lml genome. It is already known, for Lmb and other plant-interacting fungi,  
412 that TE-rich compartments of fungal genomes are enriched in genes encoding putative effectors or secondary  
413 metabolites (e.g. Ma *et al.*, 2010; Rouxel *et al.*, 2011; Grandaubert *et al.*, 2014; Faino *et al.*, 2016; Vlaardingerbroek *et*  
414 *al.*, 2016; Dallery *et al.*, 2017). Here we show that all cloned avirulence genes of Lmb are associated with H3K9me3 *in*  
415 *vitro*. Our finding regarding the genome-wide analysis of the epigenome of Lmb corroborates previous targeted  
416 epigenetic analyses performed by Soyer *et al.* (2014), in which two avirulence genes were identified associated with  
417 H3K9me3 *in vitro*, consistently with the fact that these genes are silent *in vitro*. Even outside of TE-rich regions,  
418 heterochromatin is always found associated with pathogenicity-related genes, whether they are genes encoding  
419 effectors or metabolites (Connolly *et al.*, 2013; Wiemann *et al.*, 2013; Chujo and Scott, 2014; Schotanus *et al.*, 2015;  
420 Fokkens *et al.*, 2018; Soyer *et al.*, 2019). Importantly, sets of genes up-regulated during host infection are found  
421 associated with heterochromatin, *in vitro* (Fokkens *et al.*, 2018; Soyer *et al.*, 2019). This supports our previous  
422 hypothesis that location of pathogenicity-related genes in flexible epigenomic compartments of the genome  
423 represents an efficient way to regulated expression of sets of genes scattered throughout the genomes and involved  
424 in similar biological processes (Soyer *et al.*, 2015a).

425 While importance of histone modifications in the regulation of the expression of a few effector or secondary  
426 metabolite gene clusters has already been demonstrated by ChIP-qPCR *in planta* (Chujo & Scott, 2014; Soyer *et al.*,  
427 2019) a complete view of the *in planta* chromatin dynamics of a plant-associated fungus has not been generated yet.  
428 Previous studies have assessed histone modifications at a few genomic loci of secondary metabolite gene clusters in  
429 the symbiotic fungus, *E. festucae* (Chujo and Scott, 2014). We have also performed a ChIP-qPCR analysis of targeted  
430 effector genes in *Z. tritici* during infection of wheat leaves. The experiment performed on infected wheat leaves has  
431 shown, for three putative effector genes, that their high expression at 13 dpi (corresponding to the switch from  
432 asymptomatic growth to necrotrophy for *Z. tritici*) compared to axenic culture was associated with a modification of  
433 the histone methylation patterns in their genomic environment. At 13 dpi, the level of H3K9me3 and/or H3K27me3  
434 decreased (Soyer *et al.*, 2019). Altogether, these analyses of distantly related fungi confirm that during host  
435 interaction, the specific induction of secondary metabolites / effector genes, located within heterochromatin domains  
436 *in vitro*, is associated with a dynamic re-organization of the histone modifications at their loci. So far, genome-wide  
437 scale epigenomic analyses of plant-interacting fungi have only been performed during axenic growth (Connolly *et al.*,



438 2013; Schotanus *et al.*, 2015; Fokkens *et al.*, 2018). This lack of a global comprehensive view of the chromatin dynamics  
439 of the fungal genome during plant interaction is currently due to technical limitations as discussed by Winter *et al.*  
440 (2018). Indeed, the success of performing such epigenomic analyses is jeopardized by the low amount of fungal  
441 biomass during plant infection combined with the fact that fungal genomes are small compared to massive host  
442 genomes. Our next challenge to be tackled now is to assess the chromatin dynamics *in planta* to decipher influence of  
443 histone modifications in the regulation of (effector) gene expression.

## 444 CONCLUSION

445 To the best of our knowledge, comparative analysis of histone modifications has been performed in model organisms  
446 such as mouse, in different cell types for instance (Mikkelsen *et al.*, 2010). None has been set up in phytopathogenic  
447 fungi, and nothing is known regarding how much differences in terms of location of these modifications in the genome  
448 might influence gene expression or pathogenesis. Comparative genomics has allowed analyses regarding the evolution  
449 of genes encoding effectors, their location in genomes, or secondary metabolites, of species according to their  
450 lifestyles, or the diversity of the effector repertoires according to host specialisation. The role of the epigenome is  
451 increasingly recognized in plant-pathogenic fungi, as an important regulator of the genome structure (Möller *et al.*,  
452 2019) and the expression of genes encoding effectors (Soyer *et al.*, 2014; Chujo and Scott, 2014; Soyer *et al.*, 2019).  
453 The next step forward is to set up comparative epigenomics to understand the impact of the epigenomic landscape  
454 on adaptation toward environmental rapid changes, modulation of interactions with the holobiont, host adaptation  
455 and specialization. This first comparative epigenomic analysis represent an important corpus of data to set the basis  
456 for further work aiming at investigating, on the genome wide scale, influence of histone modifications on gene  
457 evolution under selection pressure or host adaptation.

## 458 AUTHORS CONTRIBUTION

459 Conceived and designed the experiments: JLS, IF, EHS. Performed the experiments: JLS. Analysis of the data: JLS, CC,  
460 EJG, NL. Interpretation of the data: JLS, CC, IF, TR, EHS, IF. Writing of the draft: JLS. Editing of the draft: JLS, CC, TR, IF,  
461 EHS.

## 462 ACKNOWLEDGEMENT

463 Authors wish to thank all members of the “Environmental Genomics” and the “Effectors and Pathogenesis of *L.*  
464 *maculans*” groups, and M. Freitag and J. Grandaubert, for fruitful discussions. J. L. Soyer was founded by a “Contrat  
465 Jeune Scientifique” grant from INRA.

## 466 REFERENCES

467 Ansan-Melayah D., Balesdent M.-H., Buee M., Rouxel T., 1995. Genetic characterization of *AvrLm1*, the first  
468 avirulence gene of *Leptosphaeria maculans*. *Phytopathology* 85, 1525–1529.



- 469 Balesdent M.-H., Fudal I., Ollivier B., Bally P., Grandaubert J., Eber F., Chèvre A.-M., Leflon M., Rouxel T., 2013.  
470 The dispensable chromosome of *Leptosphaeria maculans* shelters an effector gene conferring avirulence towards  
471 *Brassica rapa*. *The New Phytologist*. 198, 887–898.
- 472 Bewick A.J., Hofmeister B.T., Powers R.A., Mondo S.J., Grigoriev I.V., James T.Y., Stajich J.E., Schmitz R.J., 2019.  
473 Diversity of cytosine methylation across the fungal tree of life. *Nature Ecology & Evolution*. 3, 479–490.
- 474 Bolger A.M., Lohse M., Usadel B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data.  
475 *Bioinformatics*. 30, 2114–2120.
- 476 Cam H.P., Sugiyama T., Chen E.S., Chen X., FitzGerald P.C., Grewal S.I., 2005. Comprehensive analysis of  
477 heterochromatin- and RNAi-mediated epigenetic control of the fission yeast genome. *Nature Genetics*. 37, 809–819.
- 478 Chujo T., Scott B., 2014. Histone H3K9 and H3K27 methylation regulates fungal alkaloid biosynthesis in a fungal  
479 endophyte-plant symbiosis: K9 and K27 methylation regulates symbiosis. *Molecular Microbiology* 92, 413–434.
- 480 Chuma I., Isobe C., Hotta Y., Ibaragi K., Futamata N., Kusaba M., Yoshida K., Terauchi R., Fujita Y., Nakayashiki  
481 H., Valent B., Tosa Y., 2011. Multiple translocation of the *AVR-Pita* effector gene among chromosomes of the rice blast  
482 fungus *Magnaporthe oryzae* and related species. *PLoS Pathogens*. 7, e1002147.
- 483 Connolly L.R., Smith K.M., and Freitag M., 2013. The *Fusarium graminearum* histone H3K27 methyltransferase  
484 KMT6 regulates development and expression of secondary metabolite gene clusters. *PLoS Genetics*. 9, e1003916.
- 485 Croll D., McDonald B.A., 2012. The accessory genome as a cradle for adaptive evolution in pathogens. *PLoS*  
486 *Pathogens*. 8, e1002608.
- 487 Dallery J.F., Lapalu N., Zampounis A., Pigné S., Luyten I., Amselem J., Wittenberg A.H.J., Zhou S., de Queiroz  
488 M.V., Robin G.P., Auger A., Hainaut M., Henrissat B., Kim K.T., Lee Y.J., Lespinet O., Schwartz D.C., Thon M.R., O'Connell  
489 R.J., 2017. Gapless genome assembly of *Colletotrichum higginsianum* reveals chromosome structure and association  
490 of transposable elements with secondary metabolite gene clusters. *BMC Genomics*. 18, 667.
- 491 Darma R., Lutz A., Elliott C.E., Idnurm A.2., 2019. Identification of a gene cluster for the synthesis of the plant  
492 hormone abscisic acid in the plant pathogen *Leptosphaeria maculans*. *Fungal Genetics and Biology*. 130, 62–71.
- 493 Daverdin G., Rouxel T., Gout L., Aubertot J.N., Fudal I., Meyer M., Parlange F., Carpezat J., Balesdent M.H.,  
494 2012. Genome structure and reproductive behaviour influence the evolutionary potential of a fungal phytopathogen.  
495 *PLoS Pathogens*. 8, 1003020.
- 496 de Jonge R., Bolton M.D., Kombrink A., van den Berg G.C., Yadeta K.A., Thomma B.P., 2013. Extensive  
497 chromosomal reshuffling drives evolution of virulence in an asexual pathogen. *Genome Research*. 23, 1271–1282.
- 498 Dhillon B., Cavaletto J.R., Wood K.V., Goodwin S.B., 2010. Accidental amplification and inactivation of a  
499 methyltransferase gene eliminates cytosine methylation in *Mycosphaerella graminicola*. *Genetics*. 186, 67–77.



- 500 Dhillon B., Gill N., Hamelin R.C., Goodwin S.B., 2014. The landscape of transposable elements in the finished  
501 genome of the fungal wheat pathogen *Mycosphaerella graminicola*. BMC Genomics. 15, 1132.
- 502 Dobin A., Gingeras T.R., 2015. Mapping RNA-seq Reads with STAR. Current protocols in bioinformatics. 51,  
503 11.14.1-19.
- 504 Dutreux F., Da Silva C., d'Agata L., Couloux A., Gay E.J., Istace B., Lapalu N., Lemainque A., Linglin J., Noel B.,  
505 Wincker P., Cruaud C., Rouxel T., Balesdent M.-H., Aury J.-M., 2018. De novo assembly and annotation of three  
506 *Leptosphaeria* genomes using Oxford Nanopore MinION sequencing. Scientific Data. 5, 180235.
- 507 Elliott C.E., Gardiner D.M., Thomas G., Cozijnsen A., Van De Wouw A., Howlett B.J., 2007. Production of the  
508 toxin sirodesmin PL by *Leptosphaeria maculans* during infection of *Brassica napus*. Molecular Plant Pathology. 8, 791–  
509 802.
- 510 Elliott C.E., Callahan D.L., Schwenk D., Nett M., Hoffmeister D., Howlett B.J., 2013. A gene cluster responsible  
511 for biosynthesis of phomonoic acid in the plant pathogenic fungus, *Leptosphaeria maculans*. Fungal Genetics and  
512 Biology. 53, 50–58.
- 513 Emms D.M., Kelly S., 2015. OrthoFinder: solving fundamental biases in whole genome comparisons  
514 dramatically improves orthogroup inference accuracy. Genome Biology. 16, 157.
- 515 Faino L., Seidl M.F., Shi-Kunne X., Pauper M., van den Berg G.C., Wittenberg A.H., Thomma B.P., 2016.  
516 Transposons passively and actively contribute to evolution of the two-speed genome of a fungal pathogen. Genome  
517 Research. 26, 1091–1100.
- 518 Fisher M.C., Hawkins N.J., Sanglard D., Gurr S.J., 2018. Worldwide emergence of resistance to antifungal drugs  
519 challenges human health and food security. Science. 360, 739–742.
- 520 Fokkens L., Shahi S., Connolly L.R., Stam R., Schmidt S.M., Smith K.M., Freitag M., Rep M., 2018. The multi-  
521 speed genome of *Fusarium oxysporum* reveals association of histone modifications with sequence divergence and  
522 footprints of past horizontal chromosome transfer events. BioRxiv.
- 523 Freitag M., 2017. Histone methylation by SET domain proteins in Fungi. Annual Review of Microbiology 71,  
524 413–439.
- 525 Fudal I., Ross S., Gout L., Blaise F., Kuhn M.L., Eckert M.R., Cattolico L., Bernard-Samain S., Balesdent M.H.,  
526 Rouxel T., 2007. Heterochromatin-like regions as ecological niches for avirulence genes in the *Leptosphaeria maculans*  
527 genome: map-based cloning of *AvrLm6*. Molecular Plant-Microbe Interactions. 20, 459–470.
- 528 Gacek-Matthews A., Berger H., Sasaki T., Wittstein K., Gruber C., Lewis Z.A., Strauss J., 2016. KdmB, a Jumonji  
529 Histone H3 demethylase, regulates genome-wide H3K4 trimethylation and is required for normal induction of  
530 secondary metabolism in *Aspergillus nidulans*. PLOS Genetics. 12, e1006222.



531 Galazka J.M., Freitag M., 2014. Variability of chromosome structure in pathogenic fungi--of 'ends and odds'.  
532 Curr Opin Microbiol. 20, 19–26.

533 Gardiner D.M., Cozijnsen A.J., Wilson L.M., Pedras M.S., Howlett B.J., 2004. The sirodesmin biosynthetic gene  
534 cluster of the plant pathogenic fungus *Leptosphaeria maculans*. Molecular Microbiology. 53, 1307–1318.

535 Ghanbarnia K., Fudal I., Larkan N.J., Links M.G., Balesdent M.H., Profotova B., Fernando W.G., Rouxel T., Borhan  
536 M.H. 2015, Rapid identification of the *Leptosphaeria maculans* avirulence gene *AvrLm2* using an intraspecific  
537 comparative genomics approach. Molecular Plant Pathology. 16,699–709.

538 Ghanbarnia K., Ma L., Larkan N.J., Haddadi P., Fernando W.G.D., Borhan M.-H., 2018. *Leptosphaeria maculans*  
539 *AvrLm9*: a new player in the game of hide and seek with *AvrLm4-7*. Molecular Plant Pathology. 19: 1754–1764.

540 Gervais J., Plissonneau C., Linglin J., Meyer M., Labadie K., Cruaud C., Fudal I., Rouxel T., Balesdent M.-H., 2017.  
541 Different waves of effector genes with contrasted genomic location are expressed by *Leptosphaeria maculans* during  
542 cotyledon and stem colonization of oilseed rape. Molecular Plant Pathology. 18, 1113–1126.

543 Gomes M.V., Pelosi G.G., 2013. Epigenetic vulnerability and the environmental influence on health.  
544 Experimental biology and medicine / Society for Experimental Biology and Medicine. 238, 859–865.

545 Gout L., Fudal I., Kuhn M.-L., Blaise F., Eckert M., Cattolico L., Balesdent M.-H., Rouxel T., 2006. Lost in the  
546 middle of nowhere: the *AvrLm1* avirulence gene of the Dothideomycete *Leptosphaeria maculans*. Molecular  
547 Microbiology. 60,67–80.

548 Grandaubert J., Lowe R.G., Soyer J.L., Schoch C.L., Van de Wouw A.P., Fudal I., Robbertse B., Lapalu N., Links  
549 M.G., Ollivier B., Linglin J., Barbe V., Mangenot S., Cruaud C., Borhan H., Howlett B.J., Balesdent M.H., Rouxel T., 2014.  
550 Transposable element-assisted evolution and adaptation to host plant within the *Leptosphaeria maculans*-  
551 *Leptosphaeria biglobosa* species complex of fungal pathogens. BMC Genomics. 15, 891.

552 Grandaubert J., Bhattacharyya A., Stukenbrock E.H., 2015. RNA-seq-based gene annotation and comparative  
553 genomics of four fungal grass pathogens in the genus *Zymoseptoria* identify novel orphan genes and species-specific  
554 invasions of transposable elements. G3 (Bethesda). 5, 1323–1333.

555 Jamieson K., Rountree M.R., Lewis Z.A., Stajich J.E., and Selker E.U., 2013. Regional control of histone H3 lysine  
556 27 methylation in *Neurospora*. Proceedings of the National Academy of Sciences of the United States of America. 110,  
557 6027–6032.

558 Langmead B., Salzberg S.L., 2012. Fast gapped-read alignment with Bowtie 2. Nature Methods. 9, 357–359.

559 Leclair S., Ansan-Melayah D., Rouxel T., Balesdent M., 1996. Meiotic behaviour of the minichromosome in the  
560 phytopathogenic ascomycete *Leptosphaeria maculans*. Current Genetics. 30, 541–548.

561 Lo Presti L., Lanver D., Schweizer G., Tanaka S., Liang L., Tollot M., Zuccaro A., Reissmann S., Kahmann R., 2015.  
562 Fungal effectors and plant susceptibility. Annual Review of Plant Biology. 66, 513–545.





- 563 Ma L.J., van der Does H.C., Borkovich K.A., Coleman J.J., Daboussi M.J., Di Pietro A. *et al.*, 2010. Comparative  
564 genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature*. 464,367–373.
- 565 Mendes-Pereira E., Balesdent M.-H., Brun H., Rouxel T., 2003. Molecular phylogeny of the *Leptosphaeria*  
566 *maculans*-*L. biglobosa* species complex. *Mycological Research*. 107, 1287–1304.
- 567 Mikkelsen T.S., Xu Z., Zhang X., Wang L., Gimble J.M., Lander E.S., Rosen E.D., 2010. Comparative epigenomic  
568 analysis of murine and human adipogenesis. *Cell*. 143, 156–169.
- 569 Möller M., Schotanus K., Soyer J.L., Haueisen J., Happ K., Stralucke M., Happel P., Smith K.M., Connolly L.R.,  
570 Freitag M., Stukenbrock E.H., 2019. Destabilization of chromosome structure by histone H3 lysine 27 methylation.  
571 *PLoS Genetics*. 15, e1008093.
- 572 Ohm R.A., Feau N., Henrissat B., Schoch C.L., Horwitz B.A., Barry K.W., Condon B.J., Copeland A.C., Dhillon B.,  
573 Glaser F., Hesse C.N., Kosti I., LaButti K., Lindquist E.A., Lucas S., Salamov A.A., Bradshaw R.E., Ciuffetti L., Hamelin R.C.,  
574 Kema G.H., Lawrence C., Scott J.A., Spatafora J.W., Turgeon B.G., de Wit P.J., Zhong S., Goodwin S.B., Grigoriev I.V.,  
575 2012. Diverse lifestyles and strategies of plant pathogenesis encoded in the genomes of eighteen Dothideomycetes  
576 fungi. *PLoS Pathogens*. 8, e1003037.
- 577 Parlange F., Daverdin G., Fuda I., Kuhn M.-L., Balesden M.-H., Blaise F., Grezes-Besset B., Rouxel T., 2009.  
578 *Leptosphaeria maculans* avirulence gene *AvrLm4-7* confers a dual recognition specificity by the *Rlm4* and *Rlm7*  
579 resistance genes of oilseed rape, and circumvents Rlm4-mediated recognition through a single amino acid change.  
580 *Molecular Microbiology* 71, 851–863.
- 581 Petit-Houdenot Y., Degrave A., Meye M., Blaise F., Ollivier B., Marais C.-L., Jauneau A., Audran C., Rivas S.,  
582 Veneault-Fourrey C., Brun H., Rouxel T., Fudal I., Balesdent M.-H., 2019. A two genes - for - one gene interaction  
583 between *Leptosphaeria maculans* and *Brassica napus*. *New Phytologist*. 223, 397–411.
- 584 Plissonneau C., Rouxel T., Chèvre A.-M., Van De Wouw A.P., Balesdent M.-H., 2018. One gene-one name: the  
585 *AvrLmJ1* avirulence gene of *Leptosphaeria maculans* is *AvrLm5*. *Molecular Plant Pathology*. 19,1012-1016.
- 586 Rouxel T., Balesdent M.H., 2005. The stem canker (blackleg) fungus, *Leptosphaeria maculans*, enters the  
587 genomic era. *Molecular Plant Pathology*. 6, 225–241.
- 588 Rouxel T., Grandaubert J., Hane J.K., Hoede C., van de Wouw A.P., Couloux A., Dominguez V., Anthouard V.,  
589 Bally P., Bourras S., Cozijnsen A.J., Ciuffetti L.M., Degrave A., Dilmaghani A., Duret L., Fudal I., Goodwin S.B., Gout L.,  
590 Glaser N., Linglin J., Kema G.H.J., Lapalu N., Lawrence C.B., May K., Meyer M., Ollivier B., Poulain J., Schoch C.L., Simon  
591 A., Spatafora J.W., Stachowiak A., Turgeon B.G., Tyler B.M., Vincent D., Weissenbach J., Amselem J., Quesneville H.,  
592 Oliver R.P., Wincker P., Balesdent M.-H., Howlett B.J., 2011. Effector diversification within compartments of the  
593 *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations. *Nature Communications* 2.
- 594 Rouxel T., Balesdent M.H., 2017. Life, death and rebirth of avirulence effectors in a fungal pathogen of Brassica  
595 crops, *Leptosphaeria maculans*. *New Phytologist*. 214, 526–532.



- 596 Sánchez-Vallet A., Fouché S., Fudal I., Hartmann F.E., Soyer J.L., Tellier A., Croll D., 2018. The genome biology  
597 of effector gene evolution in filamentous plant pathogens. *Annual Review of Phytopathology* 56, 21–40.
- 598 Schotanus K., Soyer J.L., Connolly L.R., Grandaubert J., Happel P., Smith K.M., Freitag M., Stukenbrock E.H.,  
599 2015. Histone modifications rather than the novel regional centromeres of *Zymoseptoria tritici* distinguish core and  
600 accessory chromosomes. *Epigenetics and Chromatin*. 8, 41.
- 601 Schumacher J., Studt L., Tudzynski P., 2019. The putative H3K36 demethylase BcKDM1 affects virulence, stress  
602 responses and photomorphogenesis in *Botrytis cinerea*. *Fungal Genetics and Biology* 123, 14–24.
- 603 Shannon P., Markiel A., Ozier O., Baliga N.S., Wang J.T., Ramage D., Amin N., Schwikowski B., Ideker T., 2003.  
604 Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Research*  
605 13: 2498-2504.
- 606 Smith K.M., Kothe G.O., Matsen C.B., Khlafallah T.K., Adhvaryu K.K., Hemphill M., Freitag M., Motamedi M.R.,  
607 Selker E.U. 2008, The fungus *Neurospora crassa* displays telomeric silencing mediated by multiple sirtuins and by  
608 methylation of histone H3 lysine 9. *Epigenetics and Chromatin*. 1, 5.
- 609 Song Q., Smith A.D., 2011. Identifying dispersed epigenomic domains from ChIP-Seq data. *Bioinformatics* 27,  
610 870–871.
- 611 Soyer J.L., El Ghalid M., Glaser N., Ollivier B., Linglin J., Grandaubert J., Balesdent M.-H., Connolly L.R., Freitag  
612 M., Rouxel T., Fudal I., 2014. Epigenetic control of effector gene expression in the plant pathogenic fungus  
613 *Leptosphaeria maculans*. *PLoS Genetics*. 10, e1004227.
- 614 Soyer J.L., Rouxel T., Fudal I., 2015a. A. Chromatin-based control of effector gene expression in plant-  
615 associated fungi. *Current Opinion in Plant Biology*. 26, 51–56.
- 616 Soyer J.L., Möller M., Schotanus K., Connolly L.R., Galazka J.M., Freitag M., Stukenbrock E.H., 2015b. Chromatin  
617 analyses of *Zymoseptoria tritici*: Methods for chromatin immunoprecipitation followed by high-throughput sequencing  
618 (ChIP-seq). *Fungal Genetics and Biology*. 79: 63–70.
- 619 Soyer J.L., Balesdent M.H., Rouxel T., Dean R.A., 2018. To B or not to B: a tale of unorthodox chromosomes.  
620 *Current Opinion in Microbiology*. 46, 50–57.
- 621 Soyer J.L., Grandaubert J., Haueisen J., Schotanus K., Stukenbrock E.H., 2019. *In planta* chromatin  
622 immunoprecipitation in *Zymoseptoria tritici* reveals chromatin-based regulation of putative effector gene expression.  
623 bioRxiv, 544627.
- 624 Stergiopoulos I., De Kock M.J., Lindhout P., De Wit P.J., 2007. Allelic variation in the effector genes of the  
625 tomato pathogen *Cladosporium fulvum* reveals different modes of adaptive evolution. *Molecular Plant-Microbe*  
626 *Interactions*. 20, 1271–1283.



627 Studt L., Rösler S.M., Burkhardt I., Arndt B., Freitag M., Humpf H.U., Dickschat J.S., Tudzynski B., 2016. Knock-  
628 down of the methyltransferase Kmt6 relieves H3K27me3 and results in induction of cryptic and otherwise silent  
629 secondary metabolite gene clusters in *Fusarium fujikuroi*. Environmental Microbiology. 18, 4037–4054.

630 Stukenbrock E.H., Bataillon T., Dutheil J.Y., Hansen T.T., Li R., Zala M., McDonald B.A., Wang J., Schierup MH,  
631 2011. The making of a new pathogen: insights from comparative population genomics of the domesticated wheat  
632 pathogen *Mycosphaerella graminicola* and its wild sister species. Genome Research. 21, 2157–2166.

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

635 Van de Wouw A.P., Lowe R.G., Elliott C.E., Dubois D.J., Howlett B.J. 2014, An avirulence gene, *AvrLmJ1*, from  
636 the blackleg fungus, *Leptosphaeria maculans*, confers avirulence to *Brassica juncea* cultivars. Molecular Plant  
637 Pathology. 15, 523–530.

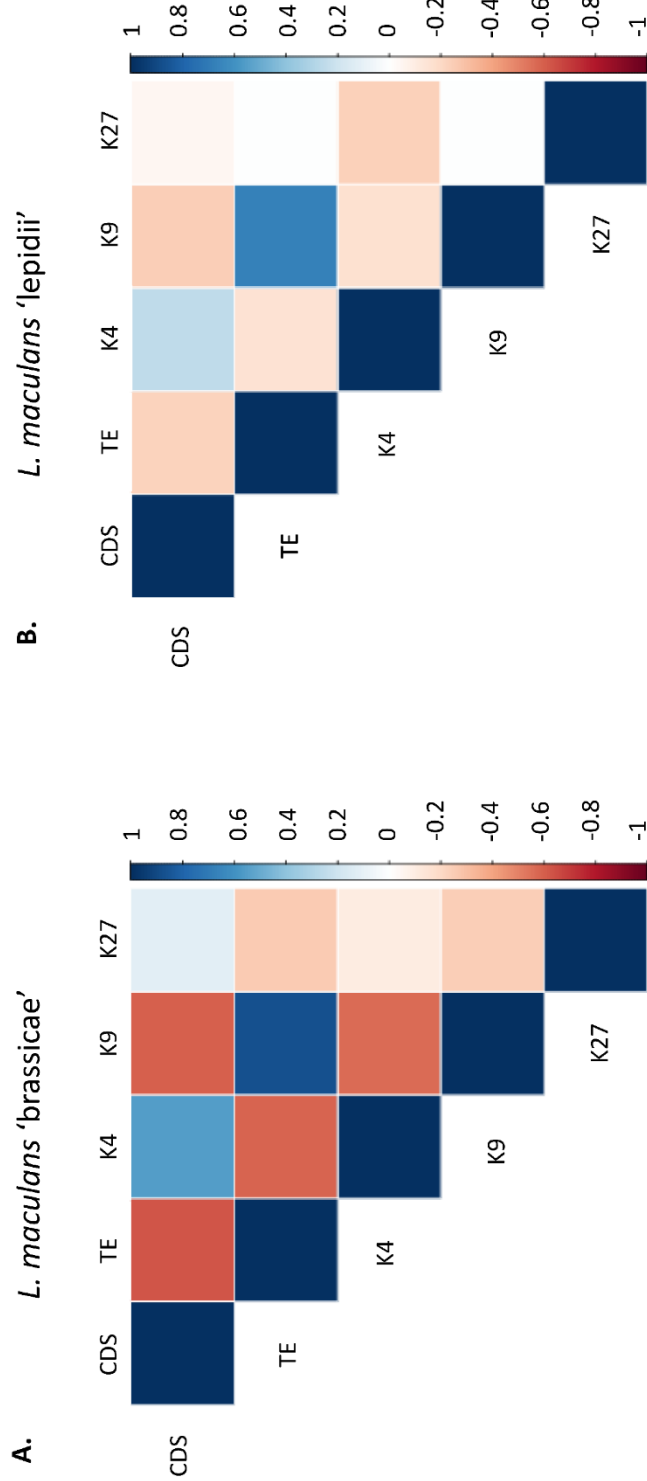
638 Vlaardingerbroek I., Beerens B., Schmidt S.M., Cornelissen B.J., Rep M., 2016. Dispensable chromosomes in  
639 *Fusarium oxysporum* f. sp. lycopersici. Molecular Plant Pathology. 17, 1455–1466.

640 Wiemann P., Sieber C.M., von Bargen K.W., Studt L., Niehaus E.M., Espino J.J., Huß K., Michielse C.B.,  
641 Albermann S., Wagner D., Bergner S.V., Connolly L.R., Fischer A., Reuter G., Kleigrew K., Bald T., Wingfield B.D., Ophir  
642 R., Freeman S., Hippler M., Smith K.M., Brown D.W., Proctor R.H., Münsterkötter M., Freitag M., Humpf H.U., Güldener  
643 U., Tudzynski B., 2013. Deciphering the cryptic genome: genome-wide analyses of the rice pathogen *Fusarium fujikuroi*  
644 reveal complex regulation of secondary metabolism and novel metabolites. PLoS Pathogens.9, e1003475.

645 Zogli P., Libault M., 2017. Plant response to biotic stress: Is there a common epigenetic response during plant-  
646 pathogenic and symbiotic interactions? Plant Sciences. 263, 89–93.

647





**Figure S1. Correlation analysis of the location of genes, transposable elements and domains enriched for H3K4me2, H3K9me3 and H3K27me3 in the genomes of A. *Leptosphaeria maculans* 'brassiccae'; B. *Leptosphaeria maculans* 'lepidii'.** Regions significantly enriched for any of the three histone modifications analysed were identified through ChIP-seq performed during axenic culture; correlation analyses were performed using a Kendall's T (see Materials and Methods). K4: H3K4me2; K9: H3K9me3; K27: H3K27me3; CDS: Coding sequences; TE: Transposable Elements.

**Table S1: Statistics of ChIP-seq and RNA-seq datasets and alignments**

Experiment	sample	ID	nb of reads	Nb not aligned	% not aligned	Nb aligned exactly one time	% aligned exactly one time	Nb aligned > 1 time	% aligned > 1 time	% aligned > 1 time	overall alignment rate
ChIP-seq_Lmb	Lmb_A_K4	1999_A	9395841	158895	1.69	9009443	95.89	227503	2.42	2.42	98.31
	Lmb_A_K9	1999_B	10919762	136322	1.25	1730751	15.85	9052689	82.9	82.9	98.75
	Lmb_A_K27	1999_C	10529291	124737	1.18	9039896	85.85	1364658	12.96	12.96	98.82
	Lmb_B_K4	1999_D	10534991	87375	0.83	10211829	96.93	235787	2.24	2.24	99.17
	Lmb_B_K9	1999_E	10654189	85037	0.80	1472180	13.82	9096972	85.38	85.38	99.2
	Lmb_B_K27	1999_F	10635233	56652	0.53	9449772	88.85	1128809	10.61	10.61	99.47
	Lmb_C_K4	1999_G	10346671	84083	0.81	10076434	97.39	186154	1.8	1.8	99.19
	Lmb_C_K9	1999_H	10459367	75082	0.72	1089336	10.41	9294949	88.87	88.87	99.28
	Lmb_C_K27	1999_I	10710240	65549	0.61	9566489	89.32	1078202	10.07	10.07	99.39
	Lmb_A_K4	1999_J	11555456	514684	4.45	11014396	95.32	26376	0.23	0.23	95.55
ChIP-seq_Lml	Lml_A_K9	1999_K	9204550	1292960	14.05	5376256	58.41	2535334	27.54	27.54	85.95
	Lml_A_K27	1999_L	10555063	993082	9.41	9370385	88.78	191596	1.82	1.82	90.59
	Lml_B_K4	1999_M	10535483	519320	4.93	9999115	94.91	17048	0.16	0.16	95.07
	Lml_B_K9	1999_N	10850527	1895009	17.46	6719265	61.93	2236253	20.61	20.61	82.54
	Lml_B_K27	1999_O	10828748	983723	9.08	9680618	89.4	164407	1.52	1.52	90.92
	Lml_C_K4	1999_P	10734950	358237	3.34	10364563	96.55	12150	0.11	0.11	96.66
	Lml_C_K9	1999_Q	10875975	1278587	11.76	4634043	42.61	4963345	45.64	45.64	88.24
	Lml_C_K27	1999_R	11402138	823586	7.22	10308659	90.41	269893	2.37	2.37	92.78
	Lml_A	NDPS10	52095284	1776449	3.41	50195802	96.35	120413	0.23	0.23	96.58
	Lml_B	NDPS11	50886249	1638537	3.22	49097848	96.49	151034	0.3	0.3	96.79
RNA-seq_Lml	Lml_C	NDPS12	50800169	1661166	3.27	49014722	96.49	121596	0.24	0.24	96.73
	Lmb_A	AMI	16800535	315850	1.88	16389267	97.55	91856	0.55	0.55	98.1
	Lmb_B	AQI	18160771	468547	2.58	17579303	96.8	110214	0.61	0.61	97.41

ChIP-seq reads were mapped against the reference genomes using Bowtie2 with default parameters; RNA-seq reads were mapped against the reference genomes using STAR (Langmead and Salzberg, 2013; Dobin and Gingeras, 2015; see Materials and Methods). Table shows the number of reads used for the alignment, the number of reads mapping only at one location, number of reads aligned more than once and the unmapped reads against the genome of *L. maculans* 'brassicae' (Dutreux *et al.*, 2018) or *L. maculans* 'lepidii' (Grandaubert *et al.*, 2014).



652

653

654

**Table S2: Correlation analysis between the different ChIP experiments generated with antibodies targeting histone modifications H3K4me2, H3K9me3 and H3K27me3 in the *L. maculans* 'lepidii', during *in vitro* growth.**

		H3K4me2			H3K9me3			H3K27me3		
		1999_J	1999_M	1999_P	1999_K	1999_N	1999_Q	1999_L	1999_O	1999_R
H3K4me2	1999_J	1.00								
	1999_M	0.94	1.00							
	1999_P	0.89	0.90	1.00						
H3K9me3	1999_K	-0.21	-0.21	-0.23	1.00					
	1999_N	0.22	0.23	0.24	-0.88	1.00				
	1999_Q	-0.26	-0.26	-0.28	0.78	-0.77	1.00			
H3K27me3	1999_L	-0.21	-0.21	-0.22	-0.05	0.08	0.07	1.00		
	1999_O	-0.22	-0.22	-0.23	-0.06	0.09	0.05	0.89	1.00	
	1999_R	-0.30	-0.30	-0.31	-0.05	0.07	0.05	0.74	0.80	1.00

Correlation analyses were performed to analyze location of the significantly enriched domains, identified using RSEG (Song and Smith, 2011), from the three different biological replicates generated and analyzed separately. A Kendall's T correlation test was performed using R.

655

**Table S3: Correlation analysis between the different ChIP experiments generated with antibodies targeting histone modifications H3K4me2, H3K9me3 and H3K27me3 in the *L. maculans* 'brassicae', during *in vitro* growth.**

		H3K4me2			H3K9me3			H3K27me3		
		1999_A	1999_D	1999_G	1999_B	1999_E	1999_H	1999_C	1999_F	1999_I
H3K4me2	1999_A	1								
	1999_D	0.90	1.00							
	1999_G	0.87	0.91	1.00						
H3K9me3	1999_B	-0.62	-0.58	-0.55	1.00					
	1999_E	-0.61	-0.58	-0.55	0.97	1.00				
	1999_H	-0.62	-0.58	-0.56	0.98	0.97	1.00			
H3K27me3	1999_C	-0.11	-0.13	-0.14	-0.23	-0.24	-0.24	1.00		
	1999_F	-0.11	-0.12	-0.14	-0.25	-0.25	-0.25	0.94	1.00	
	1999_I	-0.10	-0.11	-0.13	-0.25	-0.26	-0.25	0.95	0.94	1.00

Correlation analyses were performed to analyze location of the significantly enriched domains, identified using RSEG (Song and Smith, 2011), from the three different biological replicates generated and analyzed separately. A Kendall's T correlation test was performed using R.

656

657

658

659



**Table S4: Coverage of histone modifications H3K4me2, H3K9me3 and H3K27me3 in the genome of *Leptosphaeria maculans* 'brassicae'.**

ID_SC <sup>a</sup>	SC size (bp) <sup>a</sup>	H3K4me2 <sup>b</sup>		H3K9me3 <sup>b</sup>		H3K27me3 <sup>b</sup>	
		bp	%	bp	%	bp	%
JN3_BAC01	21008	0	0.00	21000	99.96	1251	5.95
JN3_SC00	3635781	1307171	35.95	1102622	30.33	762986	20.99
JN3_SC01	3650661	1483193	40.63	1047969	28.71	687488	18.83
JN3_SC02	3465289	1568755	45.27	821441	23.70	640962	18.50
JN3_SC03	3116548	1193866	38.31	999254	32.06	574441	18.43
JN3_SC04	3080038	1425610	46.29	665755	21.62	540171	17.54
JN3_SC05	2723307	963377	35.38	963038	35.36	579405	21.28
JN3_SC06	2669349	1086759	40.71	921773	34.53	488919	18.32
JN3_SC07	2437616	1197276	49.12	522765	21.45	587443	24.10
JN3_SC08	2329530	727752	31.24	1123307	48.22	306359	13.15
JN3_SC09	2121901	748451	35.27	886766	41.79	379620	17.89
JN3_SC10	1953324	918611	47.03	482165	24.68	439910	22.52
JN3_SC11	1933728	1000894	51.76	400992	20.74	273848	14.16
JN3_SC12	1809099	748253	41.36	486740	26.91	486379	26.89
JN3_SC13	1787499	765814	42.84	565221	31.62	372365	20.83
JN3_SC14	1545248	808577	52.33	318036	20.58	342119	22.14
JN3_SC15	1485921	587315	39.53	451809	30.41	387118	26.05
JN3_SC16	1416674	556218	39.26	545236	38.49	243093	17.16
JN3_SC17	1391278	480336	34.52	588903	42.33	415859	29.89
JN3_SC18	1351186	562433	41.63	490141	36.27	258335	19.12
JN3_SC19	832888	44716	5.37	748487	89.87	56517	6.79
JN3_SC20	252801	0	0.00	197121	77.97	32524	12.87
JN3_SC21	237616	0	0.00	197118	82.96	18507	7.79
JN3_SC22	206368	0	0.00	178511	86.50	4003	1.94
JN3_SC23	157065	0	0.00	0	0.00	0	0.00
JN3_SC24	81359	0	0.00	54108	66.51	751	0.92
JN3_SC25	68927	0	0.00	65503	95.03	5502	7.98
JN3_SC26	58575	0	0.00	57354	97.92	3254	5.56
JN3_SC27	50646	0	0.00	26005	51.35	1502	2.97
JN3_SC28	36829	0	0.00	34500	93.68	6753	18.34
JN3_SC29	33591	0	0.00	22753	67.74	15503	46.15
JN3_SC30	23239	0	0.00	20901	89.94	0	0.00
JN3_SC31	21588	0	0.00	21001	97.28	0	0.00
genome	45986477	18175377	39.52	15028295	32.68	8912887	19.38

<sup>a</sup>Genome as published in Dutreux *et al.* (2018). <sup>b</sup>Location of H3K4me2, H3K9me3 and H3K27me3 was determined through ChIP-seq analysis, in vitro, and regions significantly enriched for any of the modifications was identified using RSEG (Song and Smith, 2011).

660

661

662



**Table S5: Coverage of histone modifications H3K4me2, H3K9me3 and H3K27me3 in the genome of *L. maculans* 'lepidii'.**

ID_SC <sup>a</sup>	SC size <sup>a</sup>	H3K4me2 <sup>b</sup>		H3K9me3 <sup>b</sup>		H3K27me3 <sup>b</sup>	
		bp	%	bp	%	bp	%
scaffold00001	2643178	888592	33.62	59684	2.26	313773	11.87
scaffold00002	1959821	699265	35.68	58666	2.99	233310	11.90
scaffold00003	1983546	608230	30.66	41111	2.07	242118	12.21
scaffold00004	1736283	575205	33.13	62432	3.60	172672	9.94
scaffold00005	1623222	501923	30.92	65093	4.01	189501	11.67
scaffold00006	1544579	536762	34.75	12113	0.78	153263	9.92
scaffold00007	1533342	565981	36.91	36346	2.37	120062	7.83
scaffold00008	1418419	388288	27.37	18392	1.30	190887	13.46
scaffold00009	1356343	424898	31.33	45357	3.34	231893	17.10
scaffold00010	1127952	315229	27.95	32413	2.87	165274	14.65
scaffold00011	1096957	373438	34.04	17581	1.60	137278	12.51
scaffold00012	1109454	444873	40.10	5018	0.45	87649	7.90
scaffold00013	1030750	285020	27.65	66076	6.41	138272	13.41
scaffold00014	907115	319426	35.21	33535	3.70	97454	10.74
scaffold00015	913738	323415	35.39	29781	3.26	74040	8.10
scaffold00016	782495	246356	31.48	46437	5.93	105244	13.45
scaffold00017	748245	222170	29.69	28898	3.86	92057	12.30
scaffold00018	696173	219971	31.60	5438	0.78	92845	13.34
scaffold00019	605627	248734	41.07	35469	5.86	219071	36.17
scaffold00020	573543	186512	32.52	28680	5.00	67433	11.76
...	...	...	...	...	...	...	...
scaffold00111	2197	0	0	2000	91.03	0	0
scaffold00112	2174	0	0	2000	92.00	0	0
scaffold00113	2174	0	0	2000	92.00	0	0
scaffold00114	2171	1001	46.11	0	0	0	0
scaffold00115	2150	0	0	2000	93.02	0	0
scaffold00116	2137	0	0	2000	93.59	800	37.44
scaffold00117	2134	0	0	1001	46.91	0	0
scaffold00118	2111	0	0	2000	94.74	0	0
scaffold00119	2093	0	0	2000	95.56	0	0
scaffold00120	2077	0	0	2000	96.29	0	0
scaffold00121	2060	0	0	2000	97.09	1000	48.54
scaffold00122	2041	801	39.25	951	46.59	0	0
scaffold00123	2008	0	0	951	47.36	0	0
genome	31530973	10147334	32.18	1232561	3.91	3989471	12.65

<sup>a</sup>Genome as published in Grandaubert *et al.* (2014). <sup>b</sup>Location of H3K4me2, H3K9me3 and H3K27me3 was determined through ChIP-seq analysis, in vitro, and regions significantly enriched for any of the modifications was identified using RSEG (Song and Smith, 2011).





664

665

666

667

668

**Table S6: Correlation analysis between the coverage of transposable elements (TE) and coding sequences (CDS) location, H3K4me2-, H3K9me3- and H3K27me3-domains in the *L. maculans* 'brassicae' genome.**

	CDS	TE	H3K4me2	H3K9me3	H3K27me3
CDS	1	-0.63	0.55	-0.60	0.12
TE	-0.63	1	-0.58	0.87	-0.25
H3K4me2	0.55	-0.58	1	-0.57	-0.11
H3K9me3	-0.60	0.87	-0.57	1	-0.24
H3K27me3	0.12	-0.25	-0.11	-0.24	1

The coverage of transposable elements (TE), coding sequences (CDS) and histone modifications (H3K9me3, H3K27me3, H3K4me2) along the genome of *L. maculans* 'brassicae' was analyzed in 1000 pb sliding windows. A Kendall's T correlation analysis was done using R (Materials and Methods).

**Table S7: Correlation analysis between the coverage of transposable elements (TE) and coding sequences (CDS) location, H3K4me2-, H3K9me3- and H3K27me3-domains in the *L. maculans* 'lepidii' genome.**

	CDS	TE	H3K4me2	H3K9me3	H3K27me3
CDS	1.00	-0.22	0.25	-0.25	-0.05
TE	-0.22	1.00	-0.16	0.68	0.00
H3K4me2	0.25	-0.16	1.00	-0.16	-0.23
H3K9me3	-0.25	0.68	-0.16	1.00	0.01
H3K27me3	-0.05	0.00	-0.23	0.01	1.00

The coverage of transposable elements (TE), coding sequences (CDS) and histone modifications (H3K9me3, H3K27me3, H3K4me2) along the genome of *L. maculans* 'lepidii' was analyzed in 1000 pb sliding windows. A Kendall's T correlation analysis was done using R (Materials and Methods).



**Table S8: GO categories enriched in genes associated with H3K27me3 during axenic culture of *L. maculans* 'brassiccae'.**

GO-ID	FDR	nb. of genes with the GO term within H3K27me3	nb. of genes with a GO annotation within H3K27me3	proportion in the H3K27me3 -genes	nb. of genes with the given GO term in the genome	nb. of genes with a GO annotation in the genome	proportion in the genome	description
005975	1.54E-06	82	572	1.43E-01	362	5076	7.13E-02	carbohydrate metabolic process
0055114	4.08E-04	149	572	2.60E-01	911	5076	1.79E-01	oxidation reduction process
0055085	3.10E-03	104	572	1.82E-01	608	5076	1.20E-01	transmembrane transport

GO annotation of the Lmb genes were retrieved from Dutreux *et al.* (2018). Analysis of GO enrichment among the genes associated with H3K27me3 during axenic culture of Lmb was performed using Cytoscape (Shannon *et al.*, 2003).

671

672

**Table S9: GO categories enriched in genes associated with H3K4me2 during axenic culture of *L. maculans* 'brassicae'.**

GO-ID	FDR	nb. of genes with the GO term within H3K4me2	nb. of genes with a GO annotation within H3K4me2	proportion in the H3K4me2-genes	nb. of genes with the given GO in the genome	nb. of genes with a GO annotation in the genome	proportion in the genome	description
9987	3.60E-13	1976	3276	6.03E-01	2700	5076	5.32E-01	cellular_process
44237	6.85E-06	1585	3276	4.84E-01	2203	5076	4.34E-01	cellular_metabolic_process
71840	6.85E-06	540	3276	1.65E-01	661	5076	1.30E-01	cellular_component_organization_or_biogenesis
10467	1.10E-05	498	3276	1.52E-01	606	5076	1.19E-01	gene_expression
1901576	4.13E-05	698	3276	2.13E-01	899	5076	1.77E-01	organic_substance_biosynthetic
44249	4.13E-05	690	3276	2.11E-01	888	5076	1.75E-01	cellular_biosynthetic_process
44085	4.76E-05	336	3276	1.03E-01	392	5076	7.72E-02	cellular_component_biogenesis
6807	1.06E-04	1498	3276	4.57E-01	2102	5076	4.14E-01	nitrogen_compound_metabolic
9058	1.06E-04	727	3276	2.22E-01	950	5076	1.87E-01	biosynthetic_process
34645	2.43E-04	398	3276	1.21E-01	487	5076	9.59E-02	cellular_macromolecule_biosynthetic
9059	4.92E-04	402	3276	1.23E-01	497	5076	9.79E-02	macromolecule_biosynthetic_process
44271	4.92E-04	475	3276	1.45E-01	600	5076	1.18E-01	cellular_nitrogen_compound_biosynthetic_process
44267	6.15E-04	562	3276	1.72E-01	726	5076	1.43E-01	cellular_protein_metabolic_process
44260	1.37E-03	834	3276	2.55E-01	1129	5076	2.22E-01	cellular_macromolecule_metabolic
1901566	1.44E-03	445	3276	1.36E-01	565	5076	1.11E-01	nitrogen_compound_biosynthetic

GO annotation of the Lmb genes were retrieved from Dutreux *et al.* (2018). Analysis of GO enrichment among the genes associated with H3K27me3 during axenic culture of Lmb was performed using Cytoscape (Shannon *et al.*, 2003).



**Table S10.: GO categories enriched in genes associated with H3K4me2 during axenic culture of *L. maculans* 'lepidii'.**

GO-ID	FDR	nb. of genes with the GO term within H3K4me2	nb. of genes with a GO annotation within H3K4me2	proportion in the H3K4me2-genes	nb. of genes with the given GO in the genome	nb. of genes with a GO annotation in the genome	proportion in the genome	description
9987	3.69E-06	1564	2581	6.06E-01	2044	3725	5.49E-01	cellular_process
44267	4.48E-04	477	2581	1.85E-01	557	3725	1.50E-01	cellular_protein_metabolic_process
6807	1.25E-03	1107	2581	4.29E-01	1433	3725	3.85E-01	nitrogen_compound_metabolic_process
190156 4	1.83E-03	829	2581	3.21E-01	1048	3725	2.81E-01	organonitrogen_compound_metabolic_process
190156 6	1.87E-03	414	2581	1.60E-01	486	3725	1.30E-01	organonitrogen_compound_biosynthetic_process
44237	1.87E-03	1291	2581	5.00E-01	1704	3725	4.57E-01	cellular_metabolic_process
19538	2.97E-03	580	2581	2.25E-01	713	3725	1.91E-01	protein_metabolic_process
190157 6	3.65E-03	585	2581	2.27E-01	722	3725	1.94E-01	organic_substance_biosynthetic_process
44260	3.68E-03	647	2581	2.51E-01	808	3725	2.17E-01	cellular_macromolecule_metabolic_process
43170	3.68E-03	863	2581	3.34E-01	1108	3725	2.97E-01	macromolecule_metabolic_process

GO annotation of the Lmb genes were retrieved from Grandaubert *et al.* (2014). Analysis of GO enrichment among the genes associated with H3K4me2 during axenic culture of Lml was performed using Cytoscape (Shannon *et al.*, 2003).

1

2

**Table S11: GO categories enriched in genes associated with H3K27me3 during axenic culture of *L. maculans* 'lepidii'.**

GO-ID	FDR	nb. of genes with the GO term among the genes associated with H3K27me3	nb. of genes with a GO annotation among the genes associated with H3K27me3	proportion in the H3K27me3-genes	nb. of genes with the given GO in the genome	nb. of genes with a GO annotation in the genome	proportion in the genome	description
5975	2.44E-03	41	288	1.42E-01	250	3725	6.71E-02	carbohydrate metabolic process
42221	2.44E-03	20	288	6.94E-02	88	3725	2.36E-02	response to chemical
51591	2.44E-03	10	288	3.47E-02	25	3725	6.71E-03	response to cAMP
46683	2.44E-03	10	288	3.47E-02	25	3725	6.71E-03	response to organophosphorus
10033	2.44E-03	11	288	3.82E-02	31	3725	8.32E-03	response to organic substance
10243	2.44E-03	10	288	3.47E-02	26	3725	6.98E-03	response to organonitrogen compound
1901698	2.44E-03	10	288	3.47E-02	26	3725	6.98E-03	response to nitrogen compound
14070	2.44E-03	10	288	3.47E-02	26	3725	6.98E-03	response to organic cyclic compound
14074	2.44E-03	10	288	3.47E-02	26	3725	6.98E-03	response to purine-containing compound

GO annotation of the Lmb genes were retrieved from Grandaubert *et al.* (2014). Analysis of GO enrichment among the genes associated with H3K27me3 during axenic culture of Lml was performed using Cytoscape (Shannon *et al.*, 2003).





## CHAPITRE II

### Protéines modificatrices d'histone



## INTRODUCTION

Dans le chapitre II, nous aborderons l'implication d'acteurs clés de la régulation chromatinienne de l'expression des gènes via une approche d'analyse fonctionnelle. Pour cela j'ai transféré la technique de mutagenèse CRISPR-Cas9, mise au point pour *L. maculans* par A. Idnurm (Idnurm *et al.*, 2017), avec succès au laboratoire. Nous avons choisi d'inactiver plusieurs HMEs et déméthylases d'histones et d'étudier l'impact de leur inactivation sur la régulation génique et la structure du génome. Les protéines modificatrices d'histones sont de plus en plus étudiées et il s'avère que leur capacité à réguler des mécanismes complexes ainsi que des pans entiers du métabolisme chez les champignons filamenteux est de plus en plus documenté. L'objectif des analyses fonctionnelles menées chez *L. maculans* est de valider de façon expérimentale les conclusions que nous pouvons faire basées sur les analyses « genome-wide » quant à l'importance de l'organisation du génome sur la régulation de la pathogénie. A mon arrivée au laboratoire, les plasmides CRISPR-Cas9 n'étaient pas encore disponibles. J'ai donc initié des expériences de RNAi, puis ai réalisé en deuxième année de thèse l'inactivation des gènes quand les vecteurs ont été disponibles. J'ai ainsi mené les expériences d'inactivation génique, la caractérisation des transformants obtenus, et généré des échantillons afin d'obtenir des données de ChIP-Seq, RNA-Seq et Hi-C-Seq pour aller plus loin dans la caractérisation des transformants obtenus. Le séquençage des échantillons pour les analyses ChIP-Seq, RNA-Seq et Hi-C-Seq est actuellement en cours.

Dans une première partie, que j'ai rédigée sous forme d'article, je présenterai l'effet de l'inactivation de *KMT1* et *KMT6* responsables respectivement de la mise en place des marques répressives H3K9me3 et H3K27me3, sur la croissance mycélienne, la sporulation, la capacité à infecter le colza et la régulation de l'expression des gènes codant des effecteurs. Ces résultats sont préliminaires et les expériences de séquençage en cours apporteront beaucoup à l'avancée de ce projet.

Dans une deuxième partie, nous aborderons l'influence de Kdm8, une histone déméthylase dont le profil d'expression correspond à celui des gènes *AVR* clonés et dont la fonction est encore inconnue chez *L. maculans*. Le gène codant Kdm8 a été inactivé et les mutants ainsi obtenus ont été testés pour leur capacité à infecter le colza, leur croissance et la régulation de gènes codant des effecteurs.





## RESULTATS

### ARTICLE 3

**Histone methyltransferases KMT1 and KMT6 regulate effector gene expression and pathogenicity in the plant-pathogenic fungus *Leptosphaeria maculans***



# 1 Histone methyltransferases KMT1 and KMT6 regulate effector gene expression and 2 pathogenicity in the plant-pathogenic fungus *Leptosphaeria maculans*

3 C. Clairet<sup>1,2</sup>, F. Blaise<sup>1</sup>, J.L. Soyer<sup>1\*</sup> and I. Fudal<sup>1\*</sup>

4 <sup>1</sup>UMR BIOGER, INRA, AgroParisTech, Université Paris-Saclay, Avenue Lucien Brétignières, BP 01, F-78850 Thiverval-Grignon, France

5 <sup>2</sup>Université Paris-Sud, Orsay, France

6 \* To whom correspondence should be addressed.

## 7 INTRODUCTION

8 During infection, plant pathogenic fungi secrete a set of molecules, collectively known as effectors, targeting  
9 the apoplast or the cytoplasm of the plant where they are mainly involved in overcoming the host immune defense  
10 system, nutrient uptake and, eventually, symptom development (Lo Presti *et al.*, 2015; Lo Presti and Kahmann, 2017).  
11 In some cases, effectors are recognized by specific resistance genes from the plant, triggering plant defense reactions,  
12 and are then called avirulence proteins. Proteinaceous effectors are referred to as Small Secreted Proteins (SSPs) as  
13 they mainly correspond to small proteins, potentially secreted, often cysteine-rich, with no homology to other known  
14 proteins in the databases (Stergiopoulos and de Wit, 2009). Effector genes of filamentous plant pathogens are often  
15 located close to dispersed transposable elements (TEs) or in TE-rich regions of the genome, such as dispensable  
16 chromosomes, subtelomeric regions or AT-isochores (Rouxel *et al.*, 2011; Sánchez-Vallet *et al.*, 2018). Analyses of  
17 transcriptomic data, generated at different stages of infection, over the course of infection in different pathosystems  
18 highlighted concerted waves of effector gene expression (Sánchez-Vallet *et al.*, 2018). The regulation of such sets of  
19 genes expressed concertedly during infection remains poorly understood. In fungi, a few transcription factors (TFs)  
20 have been identified as regulating expression of effector genes or secondary metabolites (for review, see Tan and  
21 Oliver, 2017). As effector or secondary metabolite encoding genes are often found in TE-rich compartments of the  
22 fungal genomes, histone modifying enzymes have recently been investigated for their role in effector gene regulation.  
23 Two different histone methylations typical for heterochromatin formation have attracted attention in fungi, the tri-  
24 methylation of the lysine 9 of histone H3 (H3K9me3) and the tri-methylation of the lysine 27 of histone H3  
25 (H3K27me3), respectively deposited by KMT1 and KMT6. Recent studies tend to prove that these two histone  
26 methyltransferases are able to control effector gene and secondary metabolite gene cluster expression in several plant-  
27 associated fungi (Connolly *et al.*, 2013; Chujo and Scott, 2014; Soyer *et al.*, 2014). *Leptosphaeria maculans* is an ideal  
28 model to study involvement of these histone modifications in the control of effector gene expression: (i) *L. maculans*  
29 is a hemibiotrophic fungus causing stem canker (blackleg disease) on oilseed rape. Its complex life cycle presents  
30 alternating phases of saprophytism, biotrophy and necrotrophy. It first infects oilseed rape leaves and undergo an  
31 asexual sporulation leading to leaf spots. Then, the fungus grows into the stem without visible symptom and finally  
32 provokes stem canker at the end of the oilseed rape growing season (Rouxel and Balesdent, 2005). (ii) The *L. maculans*  
33 genome displays a bipartite structure with gene-rich regions (GC-isochores) and transposable element (TE)-rich  
34 regions (AT-isochores). One hundred and twenty-two putative effector-encoding genes were predicted in its



35 genome, and regions enriched in histone modifications typical for heterochromatin are significantly found in such  
36 genes (Dutreux *et al.*, 2018; Gay *et al.*, in prep.; Soyer *et al.*, in prep., **Article 2, Chapter 1**). *L. maculans* expresses  
37 several waves of effector genes during infection (Rouxel *et al.*, 2011; Gervais *et al.*, 2017; Gay *et al.*, in prep.). Of  
38 particular interest, all known avirulence genes are expressed concertedly during asymptomatic stage occurring on  
39 both leaves and stems. Avirulence genes of *L. maculans* are specifically located in TE-rich regions, associated with  
40 H3K9me3 (Rouxel and Balesdent, 2017; Soyer *et al.*, in prep., **Chapter 1; Article 2**). In contrast, gene-rich regions  
41 contain three hundred and seven 'late' effector candidates, enriched in H3K27me3-domains and under-expressed in  
42 the early colonization stages and over-expressed in the infected stems (Gervais *et al.*, 2017; Soyer *et al.*, unpublished).  
43 In *L. maculans*, RNAi-silencing of *KMT1* led to an over-expression of more than 30% of the genes located in TE-rich  
44 environments, normally silenced *in vitro* in the WT strain, notably effector genes. CHIP-qPCR analyses showed that  
45 over-expression of at least two avirulence genes was associated with a decrease of the repressive histone modification  
46 H3K9me3 in the genomic environment of these genes (Soyer *et al.*, 2014). The silenced transformants used by Soyer  
47 *et al.* (2014) still expressed ~20% of *KMT1* which did not allow an in-depth investigation of a complete lack of H3K9me3  
48 on the chromatin structure and gene expression. The CRISPR-Cas9 technology has recently been established for *L.*  
49 *maculans* to generate knock-out mutants (Idnurm *et al.*, 2017). We used that technology to inactivate *KMT1*, and also  
50 *KMT6*, responsible for H3K27me3 deposition, since H3K27me3 domains were enriched in 'late' effector genes *in vitro*.  
51 We also generated *KMT1/KMT6* double mutant. Mutants were characterized for their ability to grow in axenic culture  
52 and *in planta*. Pathogenicity and effector gene expression was also assessed in the mutants and compared to the WT.  
53 In this study we present preliminary results on chromatin-based regulation of effector genes by two histone methylase,  
54 *KMT1* and *KMT6*. Part of the results about *KMT1* regulation are presented in the **chapter 3 (Article 4)**.

## 55 MATERIAL AND METHODS

### 56 Fungal culture

57 The reference isolate JN3 was used as host for genetic transformations (Balesdent *et al.*, 2001). Fungal cultures  
58 and conidia production were performed as previously described (Ansan-Melayah *et al.*, 1995). For DNA/RNA  
59 extractions, mycelium was grown on V8-juice agar medium at 25°C in the dark for seven days and then plugs were  
60 transferred into 150 ml of Fries liquid medium in 500 ml Roux flasks. Tissues were harvested after growing for seven  
61 days at 25°C.

### 62 Pathogenicity and growth assays

63 Pathogenicity assays were performed on cotyledons of 15-day-old plantlets of a susceptible cultivar of *B.*  
64 *napus*, Es-Astrid. Plants were incubated in a growth chamber at 16/24°C (night/day) with a 16h photoperiod.  
65 Symptoms were scored on 10-12 plants, with two biological replicates, 13 days post inoculation (dpi) using the  
66 IMAScore rating scale comprising six infection classes (IC), where IC1 to IC3 correspond to various levels of resistance  
67 of the plant and IC4 to IC6 to susceptibility (Balesdent *et al.*, 2001). Pathogenicity on adult plants was tested by  
68 infecting one-month old susceptible oilseed rape (Bristol) plants as described in Gervais *et al.* (2017). Growth assays  
69 were performed by deposition of a 5 mm plug at the center of 90 mm Petri dish (containing 20 ml of V8-juice



70 agar medium or MMII medium) and radial growth was measured nine days after incubation in a growth chamber  
71 (25°C) on four biological replicates. Statistical analyses were performed using Kruskal-Wallis test (Guo *et al.*, 2013)

## 72 DNA and RNA manipulation

73 For PCR, genomic DNA was extracted from conidia or from mycelium grown in Fries liquid culture with the  
74 DNAeasy 96 plant Kit (Qiagen S.A., Courtaboeuf, France). PCR amplifications were performed as previously described  
75 (Fudal *et al.*, 2007). To identify mutation arising in the sequence of the gene targeted by the CRISPR-Cas9 strategy,  
76 sequencing was performed by Eurofins Genomics (Anzinger, Ebersberg, Germany). Total RNA was extracted from  
77 mycelium grown for one week in Fries liquid medium, from cotyledons of Es-Astrid infected by *L. maculans* at seven  
78 dpi and from stems infected by *L. maculans* at 14 dpi as previously described (Fudal *et al.*, 2007). Total gDNA was  
79 extracted for biomass quantification from 28 days infected oilseed rape stems as described in Fudal *et al.* (2007) using  
80 alkaline phenol instead of acid phenol and treating with RNase instead of DNase.

## 81 Gene annotation and domain prediction

82 The KMT6 orthologue of *L. maculans* was identified by using the KMT6 protein from *N. crassa* (Jamieson *et al.*,  
83 2013) and the NCBI BLAST program (Altschul *et al.*, 1990). Functional domains were identified using Pfam as described  
84 in Finn *et al.* (2014) (<https://pfam.xfam.org/>). Alignments were performed with COBALT (Papadopoulos and Agarwala,  
85 2007).

## 86 Vector construction and fungal transformation

87 Vectors pLAU2 and pLAU53 conferring respectively hygromycin and geneticin resistance were used to perform  
88 CRISPR-Cas9 gene inactivation, as described by Idnurm *et al.* (2017). DNA fragments coding for guide RNA which target  
89 gene of interest were designed using the CRISPOR prediction tool and *L. maculans* as reference genome  
90 (<http://crispor.tefor.net/>; **Table S1**; Dutreux *et al.*, 2018). The guide RNA were chosen not to match on any other genes.  
91 The DNA fragment coding for guide RNA was amplified using primers MAI0309 and MAI0310 and then inserted into  
92 the *Xho*I site of plasmid pLAU53 using Gibson assembly (**Table S1**; Silayeva and Barnes, 2018). Hence, plasmid pLAU53-  
93 KMT6 was generated to inactivate *KMT6*. pLAU2 containing *Cas9* gene was used without modification.

94 Complementation plasmids were obtained using pBht2 vector conferring resistance to hygromycin (Mullins *et al.*  
95 *et al.*, 2001). Native sequences of *KMT1* and *KMT6* were amplified with ComplKMT1\_F, ComplKMT1\_R, ComplKMT6\_F,  
96 ComplKMT6\_R primers using genomic DNA of JN3 isolate and then inserted into *Hind*III digested pBht2 using Gibson  
97 assembly (New England Biolabs; **Table S1**) to obtain pBht2-complKMT1 or pBht2-complKMT6 vectors.

98 The constructs were introduced into the *Agrobacterium tumefaciens* strain C58-pGV2260 by electroporation  
99 (1.5 kV, 200 ohms and 25 mF). *A. tumefaciens* mediated transformation (ATMT) of *L. maculans* was performed as  
100 previously described (Gout *et al.*, 2006). Transformants were plated on minimal medium complemented with geneticin  
101 (50 mg/l) for pLAU53-gRNA or hygromycin (50 mg/l) for pLAU2-Cas9 and cefotaxime (250 mg/l). For the CRISPR-Cas9  
102 gene inactivation, construct containing Cas9 (pLAU2-Cas9) was first introduced into a WT strain JN3 and transformants  
103 were selected for hygromycin resistance. Transformants transformed with pLAU2-Cas9 construct were then



104 transformed with pLAU53-gRNA directed against *KMT6*. Transformants were selected for geneticin, hygromycin and  
105 cefotaxime resistance. Mutations in the targeted gene were checked in the transformations by PCR amplifying with  
106 specific primers (**Table S1**) and sequencing.

### 107 **Fungal crosses**

108 Purifications in order to eliminate *Cas9* gene and guide RNA coding gene from CRISPR-Cas9 mutants were  
109 performed by crossing mutants with a WT isolate from the opposite mating type (**Table S2**). Crosses were performed  
110 as described in Balesdent *et al.* (2002). Progeny was harvested and plated on V8-juice agar medium. Mycelium was  
111 collected and DNA extracted. PCR and sequencing were performed to select progeny without *Cas9* and guide RNA  
112 which had integrated the CRISPR-Cas9 mutation.

### 113 **Quantitative RT-PCR**

114 Quantitative RT-PCR (qRT-PCR) were performed using a model CFX96 Real Time System (BIORAD; Hercules,  
115 CA, USA) and Absolute SYBR Green ROX dUTP Mix (ABgene, Courtaboeuf, France) as previously described (Fudal *et al.*,  
116 2007). For each condition tested, two different RNA extractions from two different biological samples and two reverse  
117 transcriptions for each biological replicate were performed. Primers used for qRT-PCR are described in **Table S1**. Ct  
118 values were analyzed as described by Muller *et al.* (2002) or using the  $2^{-\Delta\Delta Ct}$  method (Livak and Schmittgen, 2001).  *$\beta$ -*  
119 *tubulin* was used as a constitutively expressed reference gene. Biomass of *L. maculans* in infected plants was analyzed  
120 by generating a standard curve with known concentrations of *L. maculans* and *B. napus* gDNA. Total gDNA was  
121 extracted from adult plants 28 dpi and a  $10^{-1}$  dilution was used for qPCR. The Ct values obtained were reported on the  
122 standard curve in order to determine fungal and plant gDNA proportion.

### 123 **Native Chromatin Immunoprecipitation (N-ChIP)**

124 We subjected *L. maculans* to “native” ChIP (no crosslinking) and isolated mono- and dinucleosomes after  
125 digesting 150 mg of mycelium per sample with micrococcal nuclease (MNase) for 13 min at 37°C (Soyer *et al.*, 2015).  
126 For immunoprecipitation, three antibodies were used, against H3K9me3 and H3K27me3 (respectively Active Motif  
127 #39161 and Active Motif #39155). The Input material, i.e. the DNA extracted directly after digestion, was used to  
128 normalize data for qPCR, as described in Soyer *et al.* (2014).

129



130 **Western Blot**

131 Total proteins were extracted from 10-100 mg lyophilized mycelium. Mycelium was ground using tungsten  
132 beads and Mixer Mill MM 400 (Retsch, Éragny, France). Total proteins were treated as described in Petit-Houdenot *et*  
133 *al.* (2019). Proteins were transferred on a PVDF membrane according to the manufacturer's protocol (Trans-Blot®  
134 Turbo™ Rapid Transfer System, BIORAD, Les-Ulis, France) using a small protein transfer program. The PVDF membrane  
135 was incubated in TBS 1X containing 5% milk, 0.05% tween 20 for one hour to saturate membrane. The membrane was  
136 then incubated at 4°C over-night in TBS 1% containing 0.05% Tween 20, milk 1% and H3K9me3 antibody (1:5000;  
137 39062 ActiveMotif, La Hulpe, Belgium Germany). PVDF membrane was washed as described in Petit-Houdenot *et al.*  
138 (2019) and was then incubated in TBS 1X + 0.05% Tween 20 + milk 1% + antiRabbit IgG II<sup>R</sup> (goat anti-rabbit, Santa Cruz  
139 Biotechnology, Heidelberg, Allemagne) and washed as previously. Finally, membrane was incubated 1min in 1 mL  
140 enzyme solution (Clarity™ Western ECL, BIORAD, Les-Ulis, France) and 1 mL Luminol/enhancer solution (Clarity™  
141 Western ECL, BIORAD, Les-Ulis, France) and observed using ChemiDoc (ChemiDoc™ Imaging Systems, BIORAD, Les-  
142 Ulis, France).

143

**Table 1: Effect of *Lmkmt1* inactivation and *Lmpf2* inactivation or over-expression on in vitro growth, conidia production, pathogenicity and effector gene expression.**

Isolates/ transformants <sup>a</sup>	radial growth 9 Days V8 (mm)	conidiation	Effector gene expression <sup>d</sup>			
			AvrLm6	AvrLm4-7	<i>Lmb_jn3_11364</i>	<i>Lmb_jn3_00284</i>
JN3	46 (±0.9)	normal	1	1	1	1
<i>Δkmt1_A</i>	50.25 (±1.9)	normal	9.70E-02 (± 0.22)	2.37E+02 (± 27)	4.78E-02 (±0.03)	9.16E-01(±0.06)
<i>Δkmt6_A</i>	49.5 (±1.8)	normal	1.01E+00 (±0.28)	6.87E-01 (±0.13)	5.25E-01 (± 0.01)	1.10E+00 (±0.08)
<i>Δkmt6_B</i>	54.8 (±0.6)	normal	4.78E-02 (± 0.01)	1.79E-02 (±0.01)	1.15E+00 (±	2.07E+00 (±0.018)
<i>Δkmt1_Δkmt6</i>	65 (±4.2)	normal	1.60E-01 (±0.12)	2.55E+02 (±6.91)	1.23E-01 (±0.01)	6.30E+00 (±0.05)
complKMT1 3	nd*	normal	nd*	nd*	nd*	nd*
complKMT6 9	nd*	normal	1.01E+00 (±0.32)	1.07E+00 (±0.25)	3.56E-01(± 0.02)	9.07E-01 (±0.05)

<sup>a</sup>JN3 corresponding to WT; *Δkmt1\_A* is *KMT1* mutant; *Δkmt6\_A* and *Δkmt6\_B* are *KMT6* mutants; *Δkmt1/Δkmt6* is *KMT1 KMT6* double mutant obtained by crossing; ComplKMT1 3 and ComplKMT6 9 are complemented mutants of respectively *KMT1* and *KMT6*.

<sup>b</sup>Measured *in vitro* by RNA-Seq using TMM normalisation method and represented as RPKM. nd\*: not-tested.

146 **RESULTS**

147 **Identification of the *KMT6* orthologue in *Leptosphaeria maculans***

148 In order to analyze involvement of *KMT6* in the regulation of gene expression, we first identified *KMT6* in *L.*  
149 *maculans*. The gene encoding *KMT6* in *N. crassa* was identified by Jamieson *et al.* (2013) and phylogenetic analysis  
150 showed that an orthologue of *NcKMT6* (Genbank Accession XP\_965043.2) was found in *L. maculans* (Genbank  
151 Accession XP\_003838658.1). This Genbank Accession corresponds to gene ID *Lmb\_jn3\_06106* located on SuperContig  
152 6 (Dutreux *et al.*, 2018). A bidirectional Best Hit with BLASTp was performed to confirm orthology between *NcKMT6*  
153 and *LmKMT6* (**Figure 1A**). The *L. maculans LmKMT6* gene (hereafter referred to as *KMT6*) encodes a 1,342 amino acids

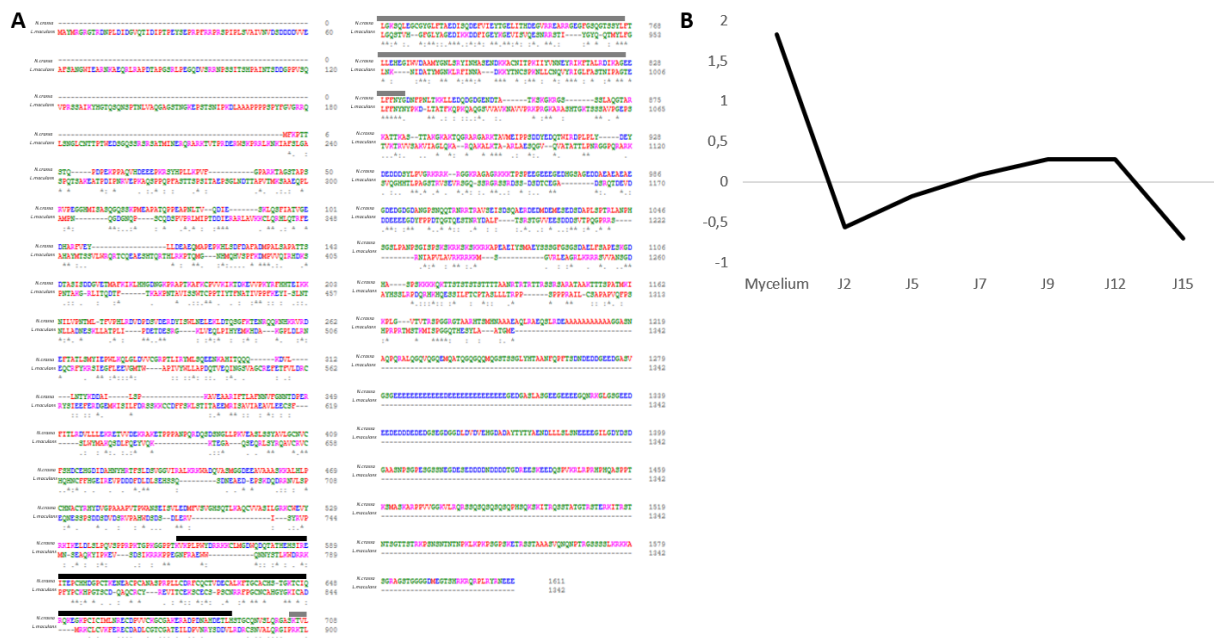


154 protein sharing 26.3% identity with *NcKMT6*. The two domain CXC and SET are conserved in both proteins (IP551633;  
155 IP550280; **Figure 1A**).

156 We investigated the expression profile of *KMT6* during axenic culture and at different stages of primary  
157 infection of oilseed rape (**Figure 1B**). We observed that this gene is strongly expressed *in vitro* during mycelium growth  
158 and poorly expressed during *in planta* infection. In our previous study, we identified that *KMT1* is inversely regulated  
159 compared to avirulence genes (Clairet *et al.*, unpublished; **Chapter III, Article 4**).

160 ***KMT6*, *KMT1* and *KMT1/KMT6* inactivations do not induce morphology or conidiation defect in *L. maculans***

161 In a previous study, we determined that *KMT1* inactivation did not lead to growth or conidiation defect in *L.*  
162 *maculans* (Clairet *et al.*, in prep.; **Chapter III; Article 4**). Inactivation of *KMT6* was performed using the CRISPR-Cas9  
163 strategy as detailed in the Materials and Methods. Twenty transformants resistant to hygromycin and geneticin were  
164 obtained and sequenced for the *KMT6* gene. Out of the 20 transformants, 17 had no mutations in the coding sequence  
165 of *KMT6* compared to the WT, two had a 10-bp and one had a 2-bp deletion near the cleavage site (**Figure S1**). These  
166 mutations resulted, at the protein level, in frame-shifts leading to two different truncated proteins: 372 amino-acids  
167 and 401 amino-acids respectively for the 10-bp and the 2-bp deletions compared to the 1,342 amino-acids length of  
168 the WT protein (**Figure S1**). The two mutants were crossed with JN2 in order to obtain purified  $\Delta kmt6$  mutants without  
169 *Cas9* and CRISPR guide RNA (**Table S2**). The two corresponding mutants are hereinafter referred to as  $\Delta kmt6\_A$  and  
170  $\Delta kmt6\_B$  (A corresponding to the 10-bp and B to the 2-bp deletion). The two  $\Delta kmt6$  mutants were able to produce as  
171 conidia as the WT after 14 days of growth on V8-juice agar medium (**Table 1**).  $\Delta kmt6\_A$  was then crossed with JN2



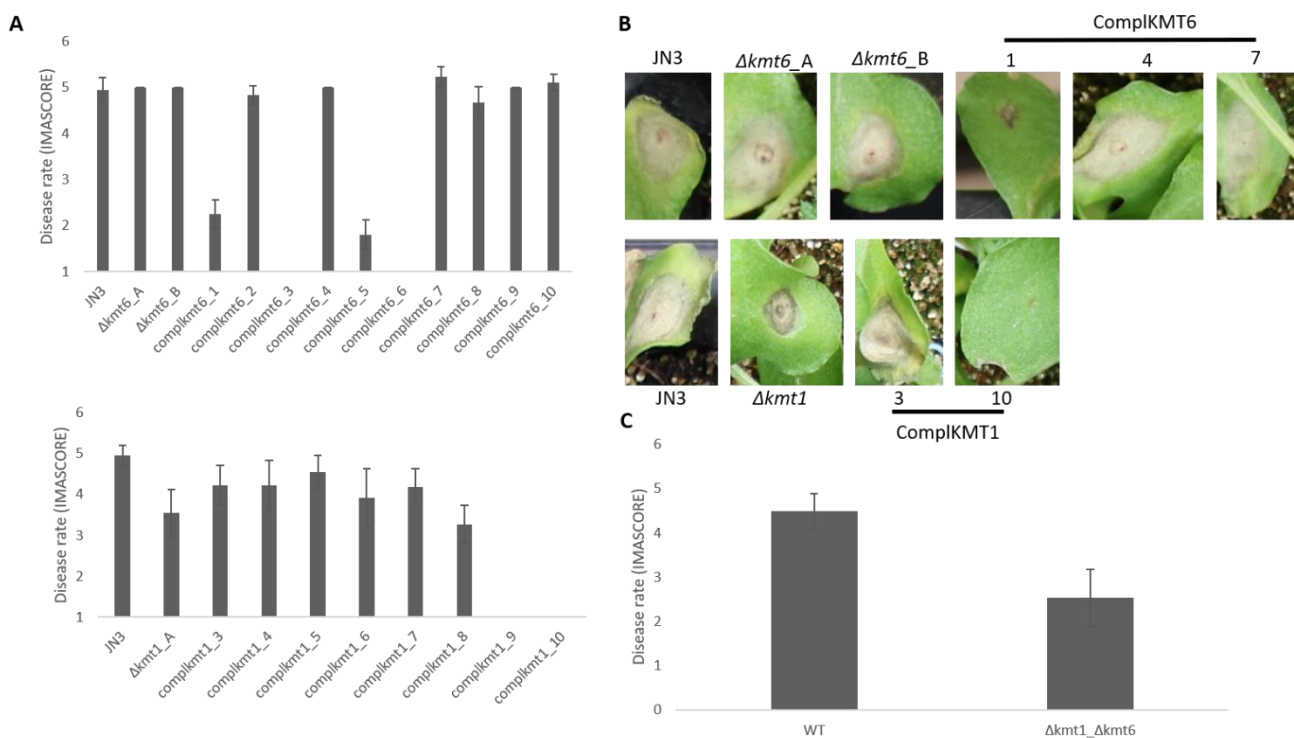
**Figure 1: Identification of *KMT6* in *Leptosphaeria maculans*.** **A:** Alignment between *KMT6* of *L. maculans* and *Neurospora crassa* was performed using COBALT (Papadopoulos and Agarwala, 2007). The black bar corresponds to the CXC domain and the grey bar corresponds to the SET domain, both identified using Pfam as described in Finn *et al.*, 2014 (<https://pfam.xfam.org/>). **B:** Mycelium was obtained by growing JN2 strain on V8 agar medium for 7 days. Cotyledons of the susceptible cultivar Darmor were inoculated by JN2 and sampled at 7 different time points (2, 5, 7, 9, 12 and 15 days post infection). Two biological replicates per condition were generated. Total RNA was extracted and sequenced. The expression level of *KMT6* in RNA-seq data is represented by the log<sub>2</sub> of RPKM (Reads Per Kilobase Per Million mapped reads) (Gay *et al.*, in prep.).



172 transformed with pLAU2-Cas9 and pLAUS3-KMT1. Progeny harboring the two mutations were used for subsequent  
173 analyses. We also verified that our  $\Delta kmt1_{\Delta kmt6}$  mutants were defective in H3K9me3 and H3K27me3 deposition  
174 (**Figure S1; Figure S2**). After nine days of growth on V8-agar plate, no mutant showed growth defect indicating that  
175 KMT6 and KMT1/KMT6 deletions do not lead to growth defect. To summarize, no major defect in conidia production  
176 or growth rate was associated with the inactivation of *KMT6* or *KMT1/KMT6*.

177 **H3K27me3 is not required for establishment of pathogenesis whereas H3K9me3 is crucial for infection**  
178 **establishment**

179 After studying impact of H3K27me3 on *L. maculans* growth we wanted to investigate the impact both  
180 repressive mark H3K9me3 and H3K27me3 on pathogenicity. We previously identified that KMT1 via the deposition of  
181 H3K9me3 is crucial for pathogenesis establishment (Soyer *et al.*, 2014; Clairet *et al.*, in prep.; **Chapter III; Article 4**). On  
182 the contrary, the two  $\Delta kmt6\_A$  and  $\Delta kmt6\_B$  mutants do not harbor any pathogenicity defect on susceptible cultivar  
183 Es-Astrid at 13 dpi (**Figure 2A, B**). The  $\Delta kmt1_{\Delta kmt6}$  double mutant produces even less symptoms on oilseed rape  
184 than  $\Delta kmt1$  (**Figure 2C**).



**Figure 2: Effect of  $\Delta kmt1$  and  $\Delta kmt6$  inactivations on pathogenicity of *L. maculans*.** **A:** pathogenicity assays of the selected mutants and complemented strains 14 days after inoculation on the susceptible cultivar Es-Astrid of *B. napus*. Results are expressed as a mean disease scoring using the IMAScore rating scale comprising six infection classes (IC), where IC1 to IC3 correspond to various levels of resistance of the plant and IC4 to IC6 to susceptibility (Balesdent *et al.*, 2001). Error bars indicate the standard deviation of two biological and technical replicates. **B:** Photos of symptoms at 13 DPI on cotyledons infected by mutants, complemented mutants and WT strains. Each photo is representative of at least 24 inoculated plants. **C:** pathogenicity assays of  $\Delta kmt1_{\Delta kmt6}$  14 days after inoculation on the susceptible cultivar Es-Astrid of *B. napus*. Results are expressed as the mean scoring using the IMAScore rating scale as described in Balesdent *et al.* (2001). Error bars indicate the standard deviation of two biological and technical replicates.

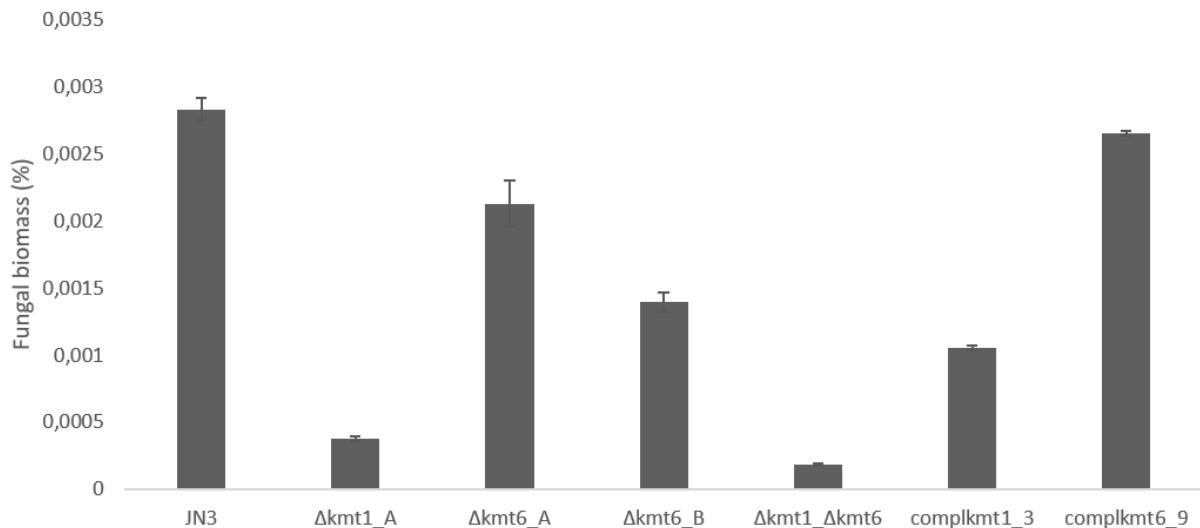
185

186





187 We also infected adult plants with the mutants to study the relative impact of *KMT1* and *KMT6* on stem  
188 colonization. After 28 days post inoculation, fungal biomass was assessed to evaluate whether the mutants were able  
189 to colonize the stem. Plants infected with the  $\Delta kmt1$  strains contained from 5 to 10 times less fungal DNA than plants  
190 infected with the WT. Interestingly, adults plants infected with the  $\Delta kmt6$  strains contained between 1/3 and 1/2 less  
191 fungal gDNA than the plants infected with the WT. Plants infected with the  $\Delta kmt1_{\Delta kmt6}$  mutants contained even  
192 less *L. maculans* gDNA than plants infected with  $\Delta kmt1$  (**Figure 3**).

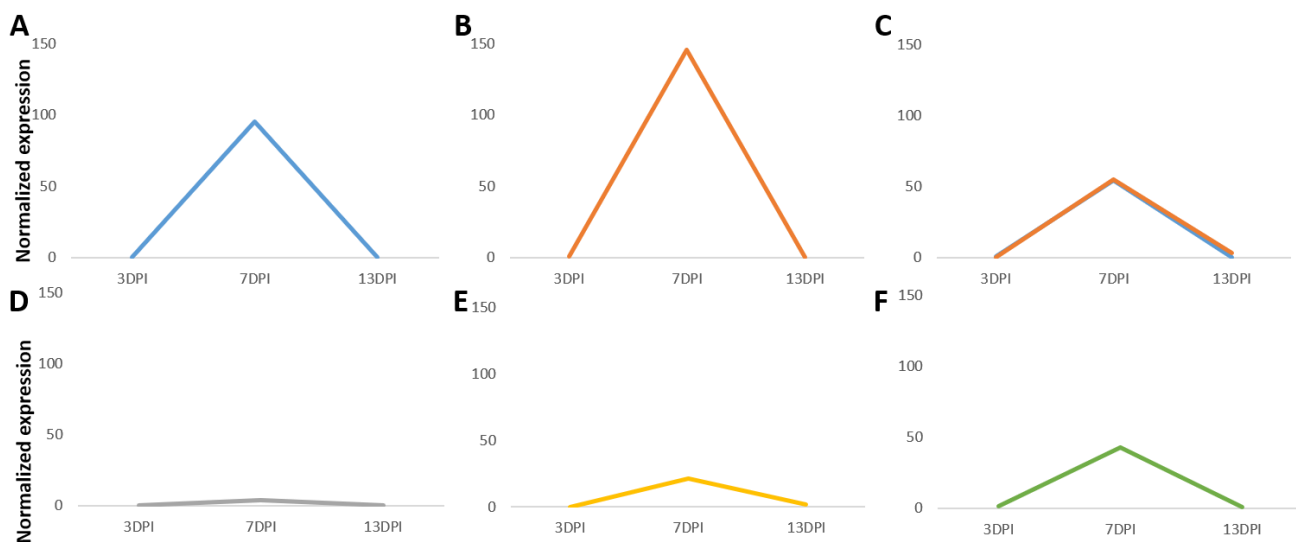


**Figure 3: Effect of *KMT1* and *KMT6* inactivation on *L. maculans in planta* growth.** 28 days old adult plants were inoculated with 10 $\mu$ L of a 10<sup>7</sup>/mL conidia suspension as described in Gervais *et al.* (2017). Stems were then harvested after 28 days of infection and total DNA was extracted. Fungal biomass was measured by qPCR.

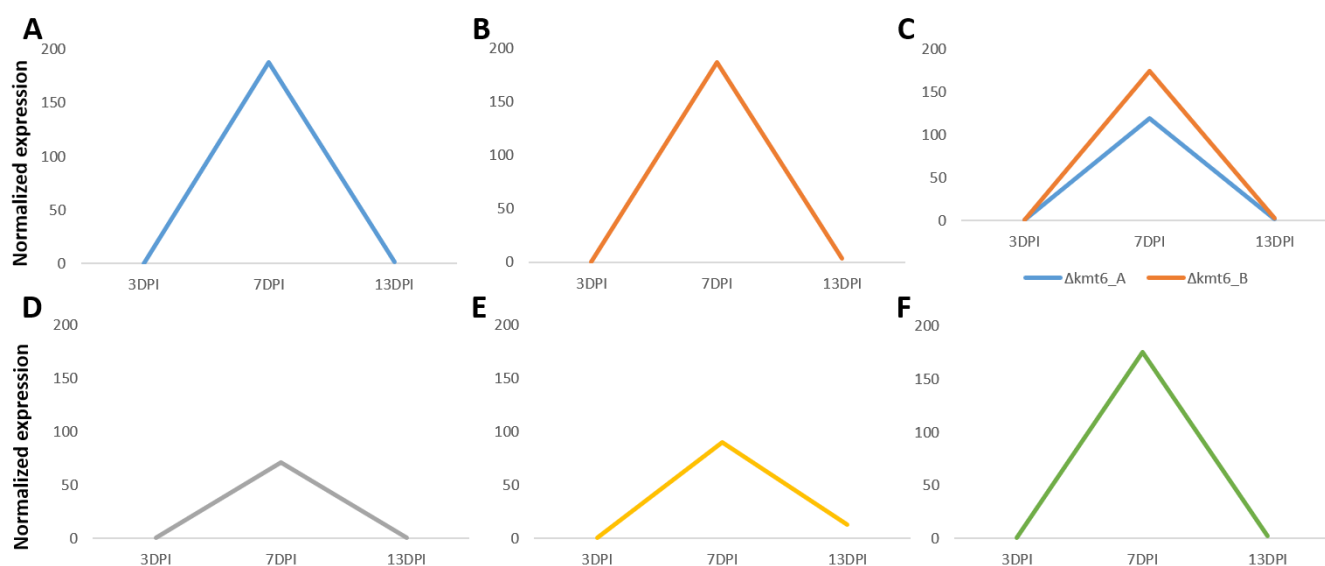
### 193 Effector gene expression is modified in $\Delta kmt1$ , $\Delta kmt6$ and $\Delta kmt1_{\Delta kmt6}$ double mutants

194 Avirulence genes are usually not expressed during axenic growth and highly expressed during infection of  
195 cotyledons (Rouxel *et al.*, 2011). Previous analyses showed these avirulence genes, located in the middle of large TE-  
196 rich regions, are associated with H3K9me3 *in vitro*, consistently with their low expression under this growth condition  
197 (Soyer *et al.*, 2014; Soyer *et al.*, in prep.). We measured expression of avirulence genes *AvrLm6* and *AvrLm4-7* and  
198 tardive genes *Lmb\_jn3\_11364* and *Lmb\_jn3\_00284 in vitro* in our different mutants during axenic culture. *AvrLm6*  
199 expression is not deregulated in any of the mutants (**Table 1**). Inactivation of *KMT1* increased expression of *AvrLm4-7*  
200 up to 200 times during *in vitro* growth as previously described in Clairet *et al.*, in prep. The  $\Delta kmt1_{\Delta kmt6}$  double  
201 mutant harbored the same *AvrLm4-7* up-regulation as in  $\Delta kmt1$  (**Table 1**). *Lmb\_jn3\_11364* and *Lmb\_jn3\_00284* was  
202 reduced up to 20 time in  $\Delta kmt1$  and up-regulated up to 6 time in  $\Delta kmt1_{\Delta kmt6}$  while not altered in  $\Delta kmt6$  mutants  
203 (**Table 1**).

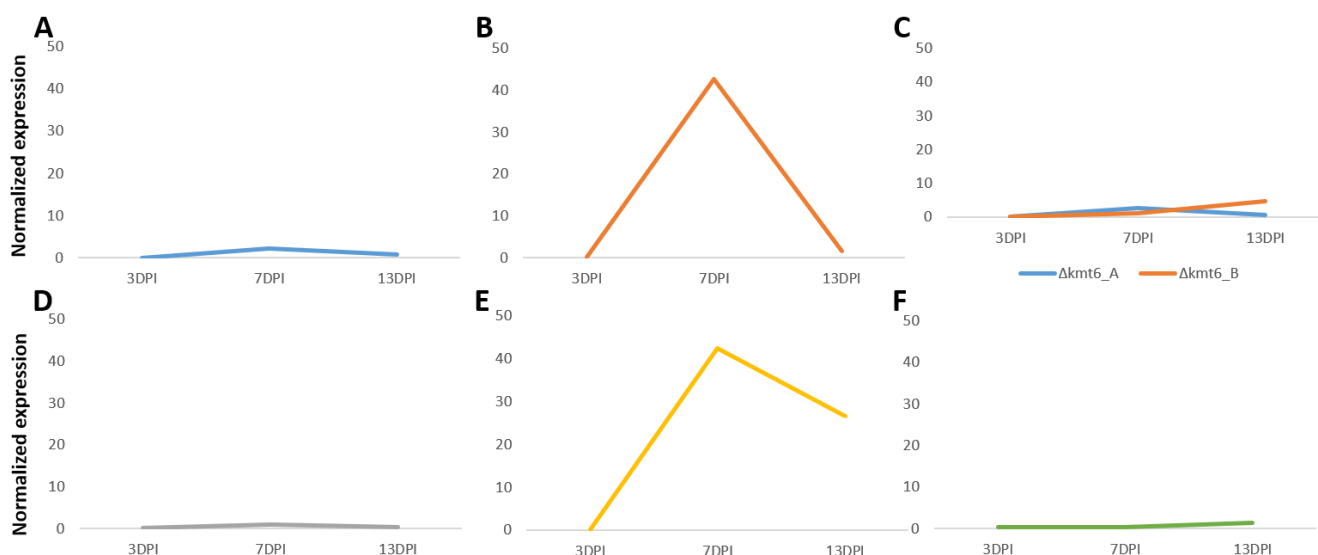
204 We measured expression of selected effector genes during early infection of cotyledons, in our different  
205 mutants, at 3, 7 and 13 dpi. We assessed expression of *AvrLm6* and *AvrLm4-7*, located in a H3K9me3-domain, and  
206 expression of two putative effector genes located in a H3K27me3-domains (*Lmb\_jn3\_00284* and *Lmb\_jn3\_11364*) and  
207 expressed respectively at 3 dpi of oilseed rape cotyledons and during stem infection. Inactivation of *KMT1* led to a 1.5-  
208 fold over-expression of *AvrLm4-7* and inactivation of *KMT6* increased twice its expression indicating that H3K9me3 is  
209 repressing *AvrLm4-7* expression whereas H3K27me3 is involved in correct expression of *AvrLm4-7*, but only *in planta*.  
210 The  $\Delta kmt1_{\Delta kmt6}$  double mutant did not express *AvrLm4-7*, indicating involvement of these two histone  
211 modifications in *AvrLm4-7* correct regulation (**Figure 4**). Interestingly *AvrLm6* expression was not altered by *KMT1* and  
212 *KMT6* inactivation whereas *KMT1/KMT6* double inactivation lead to four-time less expression of *AvrLm6* at 7 dpi  
213 compared to the WT indicating the involvement of these two histone modifications in *AvrLm6* correct regulation  
214 (**Figure 5**). The putative effector *Lmb\_jn3\_11364* is usually expressed in stems, during late stages of infection (Gervais  
215 *et al.*, 2017). Although located in a H3K27me3-domain *in vitro* (Soyer *et al.*, unpublished), inactivation of *KMT1*  
216 increased expression of this gene 40 times at 7 dpi. On the contrary, inactivation of *KMT6* did not modify expression  
217 of *Lmb\_jn3\_11364*. Finally,  $\Delta kmt1_{\Delta kmt6}$  double mutant was not able to express *Lmb\_jn3\_11364* (**Figure 6**). The  
218 *Lmb\_jn3\_00284* gene is expressed at early stage of infection (3 dpi). Inactivation of *KMT6* increased expression of  
219 *Lmb\_jn3\_00284* twice indicating that



**Figure 4: Effect of *KMT1*, *KMT6* and *KMT1-KMT6* deletions on *AvrLm4-7* expression pattern during early cotyledon colonization.** Gene expression levels are relative to *Lmbtubulin* and calculated as described by Muller *et al.* (2002). Each value is the average of two biological replicates (two extractions from different biological replicates) and two technical replicates (two RT-PCR). Expression was measured at 3, 7 and 13 DPI. DPI= Days Post Inoculation. Expression profile (**A**) in the WT, (**B**) in the  $\Delta kmt1$  mutant, (**C**) in the  $\Delta kmt6_{A}$  and B, (**D**) in the  $\Delta kmt1_{\Delta kmt6}$ , (**E**) in the ComplKMT1\_3 and (**F**) in the ComplKMT6\_9.



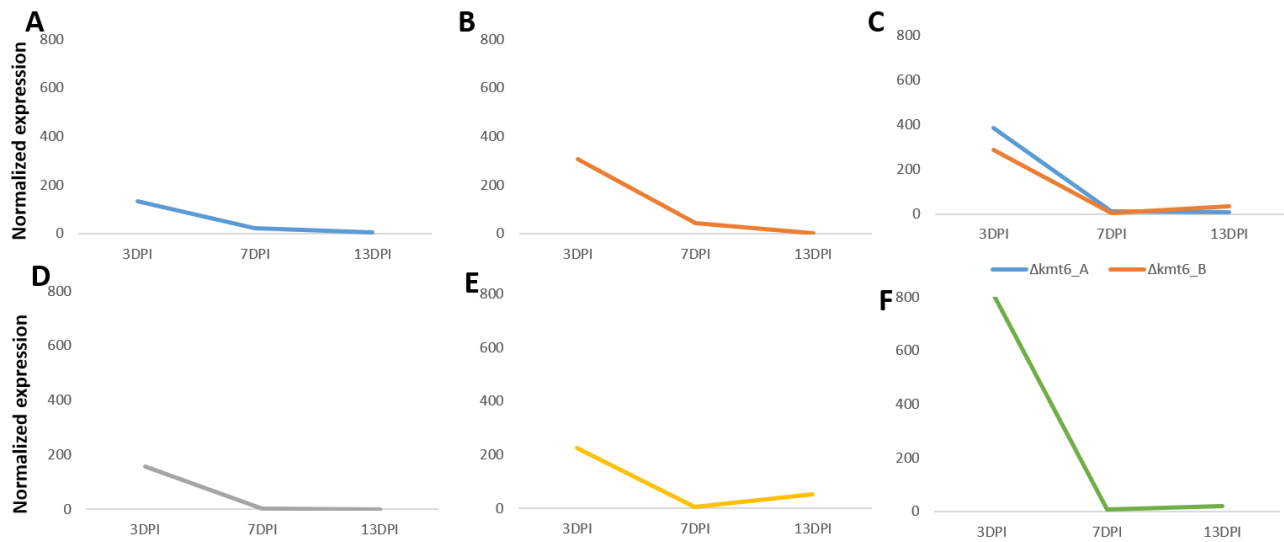
**Figure 5: Effect of *KMT1*, *KMT6* and *KMT1-KMT6* deletion on *AvrLm6* expression pattern during early cotyledon colonization.** Gene expression levels are relative to *Lm $\beta$ tubulin* and calculated as described by Muller *et al.* (2002). Each value is the average of two biological replicates (two extractions from different biological replicates) and two technical replicates (two RT-PCR). Expression was measured at 3, 7 and 13 DPI. DPI= Days Post Inoculation. Expression profile (A) in the WT, (B) in the  $\Delta kmt1$  mutant, (C) in the  $\Delta kmt6\_A$  and B, (D) in the  $\Delta kmt1\_ \Delta kmt6$ , (E) in the ComplKMT1\_3 and (F) in the ComplKMT6\_9.



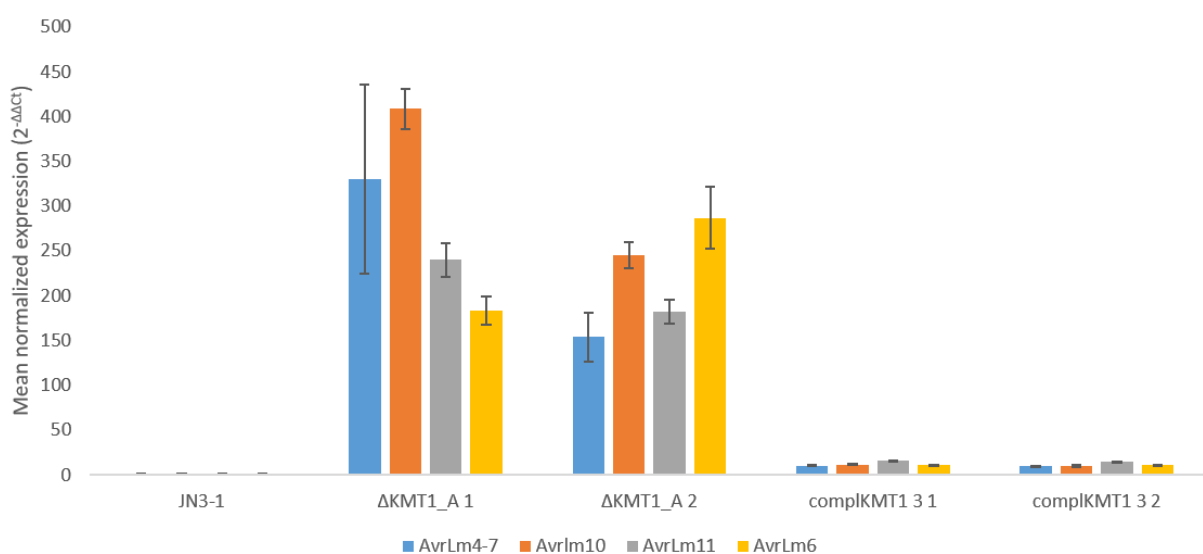
**Figure 6: Effect of *KMT1*, *KMT6* and *KMT1-KMT6* deletion on *Lmb\_jn3\_11364* expression pattern during early cotyledon colonization.** Gene expression levels are relative to *Lm $\beta$ tubulin* and calculated as described by Muller *et al.* (2002). Each value is the average of two biological replicates (two extractions from different biological replicates) and two technical replicates (two RT-PCR). Expression was measured at 3, 7 and 13 DPI. DPI= Days Post Inoculation. Expression profile (A) in the WT, (B) in the  $\Delta kmt1$  mutant, (C) in the  $\Delta kmt6\_A$  and B, (D) in the  $\Delta kmt1\_ \Delta kmt6$ , (E) in the ComplKMT1\_3 and (F) in the ComplKMT6\_9.



221 H3K27me3 alone is able to repress *Lmb\_jn3\_00284*. Inactivation of *KMT1* had no effect on *Lmb\_jn3\_00284* expression  
222 (Figure 7). We previously determined that H3K9me3 is repressing avirulence gene expression during *in vitro* growth  
223 and that *KMT1* is able to inhibit expression between 7 and 12 dpi of cotyledons (Clairet *et al.*, in prep.). To get further  
224 insight about role of *KMT1* in the regulation of avirulence gene expression, we measured expression of four *AvrLm*  
225 genes in adult plants at 14 dpi of stems. As already reported, the four genes were not expressed during infection of  
226 the stem by the WT strain. Inactivation of *KMT1* induced expression of the four avirulence genes with an increase  
227 expression of up to 400 times in the mutants suggesting involvement of *KMT1* in the repression of avirulence gene  
228 expression during the late stages of infection (Figure 8).



**Figure 7: Effect of *KMT1*, *KMT6* and *KMT1-KMT6* deletion on *Lmb\_jn3\_00284* expression pattern during early cotyledon colonization.** Gene expression levels are relative to *Lmbtubulin* and calculated as described by Muller *et al.* (2002). Each value is the average of two biological replicates (two extractions from different biological replicates) and two technical replicates (two RT-PCR). Expression was measured at 3, 7 and 13 DPI. DPI= Days Post Inoculation. Expression profile (A) in the WT, (B) in the  $\Delta kmt1$  mutant, (C) in the  $\Delta kmt6_A$  and B, (D) in the  $\Delta kmt1_\Delta kmt6$ , (E) in the ComplKMT1\_3 and (F) in the ComplKMT6\_9.



**Figure 8: Effect of *KMT1* inactivation on avirulence gene expression during oilseed rape stem infection (cultivar Bristol) 14DPI.** Expression of four avirulence genes was measured in the WT,  $\Delta kmt1_A$  and Complkmt1\_3 (complementation of  $\Delta kmt1_A$ ). Expression levels in each transformants is compared to the expression in the WT strain measured by RT-qPCR using *LmbTubulin* as a constitutive gene and using the  $2^{-\Delta\Delta Ct}$  method (Livak and Schmittgen, 2001). Error bars indicate the standard deviation of technical replicates.

## 229 **Complementation of *KMT1* and *KMT6* restored the WT phenotype**

230 Inactivated mutants were complemented using the WT genes inserted randomly in the purified mutants (**Table**  
231 **S1; Table S2**). This restored the ability of each protein to deposit H3K9me3 or H3K27me3 (**Figure S1; Figure S2**). The  
232 insertion of the WT version of the *KMT1* gene led to a partial complementation of the pathogenicity abilities. We  
233 selected two transformants, ComplKMT1\_3 and ComplKMT6\_9, which presented the closest pathogenicity and  
234 histone modifications phenotype to the WT (**Figure 2; Figure S1; Figure S2**). The two complemented transformants  
235 presented partially restored fungal biomasses during late infection on stem at 28 dpi (**Figure 3**). Interestingly, *KMT1*  
236 complementation led to *AvrLm4-7* and *AvrLm6* repression compared to the WT during cotyledon infection (**Figure 4,**  
237 **5**). ComplKMT1\_3 transformant also harbored altered *Lmb\_jn3\_11364* expression indicating that *KMT1* re-  
238 introduction do not restore correct expression (**Figure 6**). Intriguingly, Compl\_KMT6\_9 transformant presented  
239 *Lmb\_jn3\_00284* four times up-regulation indicating that *KMT6* re-introduction alter normal *Lmb\_jn3\_00284*  
240 expression which could be explained by the higher H3K27me3 deposition on *Lmb\_jn3\_00284* in ComplKMT6\_9 than  
241 in the WT (**Figure 7; Figure S2**). Finally, ComplKMT1\_3 transformants did not present avirulence gene up-regulation in  
242 stems at 14 dpi indicating that *KMT1* re-introduction lead to correct repression of avirulence gene between 7 dpi and  
243 14 dpi (**Figure 8**).

## 244 **DISCUSSION**

245 Here we demonstrated that *KMT1*, via the deposition of H3K9me3, and *KMT6*, via the deposition of  
246 H3K27me3, are involved in the regulation of the expression of at least some effector genes and are key regulators of  
247 pathogenicity. On the contrary, no growth or conidiation defect were associated to loss of H3K9me3, H3K27me3 or  
248 both histone modifications.

### 249 ***KMT1* and *KMT6* are not required for correct growth of *L. maculans***

250 *KMT1* and *KMT6* are crucial proteins in many organisms. They guarantee genome stability and integrity in *N.*  
251 *crassa*, *Z. tritici* and *E. festucea* (Connolly *et al.*, 2013; Jamieson *et al.*, 2013; Schotanus *et al.*, 2015). Inactivation of  
252 *KMT1*, responsible for H3K9me3 deposition, resulted in relocation of H3K27me3 in sites normally enriched with  
253 H3K27me3 in *Z. tritici* and *N. crassa* (Basenko *et al.*, 2015; Möller *et al.*, 2019). In contrast, lack of H3K27me3 increased  
254 stability of accessory chromosomes of *Z. tritici* (Möller *et al.*, 2019). In this study, we inactivated *KMT1*, *KMT6* and  
255 constructed the double  $\Delta kmt1_{\Delta kmt6}$  mutant. We did not identify any growth or conidiation defects in our mutants  
256 indicating that *KMT1* and *KMT6* are not crucial for correct growth of *L. maculans* during axenic culture.

### 257 ***KMT1* is essential for pathogenesis establishment whereas *KMT6* inactivation does not alter *L. maculans* ability to** 258 **produce symptoms on *B. napus***

259 We determined in this study that *KMT1* inactivation led to pathogenicity defects and deregulation of  
260 avirulence genes whereas *KMT6* inactivation had limited impact on pathogenicity and effector gene expression.  
261 *KMT1/KMT6* inactivation led to critical pathogenicity defects and effector gene deregulation indicating that H3K9me3  
262 at least is critical for correct effector gene regulation. *KMT1* and *KMT6* are known to be able to regulate effector



263 genes and secondary metabolites gene clusters in several fungi (Gacek and Strauss, 2012; Connolly *et al.*, 2013; Chujo  
264 and Scott, 2014). Furthermore, previous studies on *L. maculans* determined that KMT1 was involved in the regulation  
265 of at least two avirulence genes (Soyer *et al.*, 2014; Clairet *et al.*, in prep). In this study, we confirmed the involvement  
266 of KMT1 in the control of *L. maculans* avirulence gene expression and determined that KMT6 is also involved in effector  
267 gene regulation but no clear involvement in the repression of avirulence gene expression. In addition, we previously  
268 determined that KMT1 is involved in the regulation of the switch between the asymptomatic colonization and the  
269 necrotrophic growth on oilseed rape cotyledons (Clairet *et al.*, in prep.; **Chapter III, Article 4**). In this study, we  
270 determined that in absence of H3K9me3, avirulence genes are not actively repressed during stem infection, and that  
271 *KMT1* complementation restored repression of these genes during stem infection.

## 272 **Are H3K9me3 and H3K27me3 involved in the maintenance of genome stability in *L. maculans*?**

273 H3K9me3- and H3K27me3-domains were recently reported for their ability to structure genome organization  
274 (Klocko *et al.*, 2016; Winter *et al.*, 2018). Furthermore, *KMT1* inactivation was associated to H3K27me3 re-localization  
275 and to genome organization destabilization in *N.crassa* and *Z. tritici* (Basenko *et al.*, 2015; Möller *et al.*, 2019). But  
276 impact of KMT1 and KMT6 on genome organization of filamentous fungi and especially on phytopathogenic fungi are  
277 poorly documented. In this study we will get cutting edge results about ability of two histone modifications to structure  
278 *L. maculans* genome and their impact on TADs (Transcriptionally Active Domains) regulation (Lanctôt *et al.*, 2007). In  
279 *N. crassa*, constitutive heterochromatin domains were identified for their ability to interact together without the  
280 involvement of H3K9me3 modifications. On the contrary H3K27me3 is essential for genome stability in *N. crassa*  
281 (Klocko *et al.*, 2016). In *Z. tritici*, H3K27me3 have been demonstrated to be enriched in sub-telomeric compartment  
282 and to be essential to maintain interaction between telomeres and nuclear membrane. *KMT1* inactivation in this  
283 species lead to massive H3K27me3 relocation at H3K9me3 loci leading to massive genome reorganization. At the  
284 contrary *KMT6* inactivation induces limited chromosome rearrangement (Möller *et al.*, 2019). However, H3K27me3 in  
285 *Fusarium oxysporum* have been recently demonstrated to be chromatin niche for effector genes and to allow rapid  
286 evolution of this set of genes. Chromosome rearrangements can occur between core chromosomes and accessory  
287 chromosomes marked by H3K27me3 allowing rising and diversification of effector gene repertoire in two-speed  
288 genomes (Fokkens *et al.*, 2018).

## 289 **ACCESSION NUMBERS**

290 All sequenced reads have been deposited with the Short Read Archive under accession numbers XXXX.

## 291 **REFERENCES**

- 292 Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. Journal  
293 of Molecular Biology 215, 403–410.
- 294 Ansan-Melayah, D. (inra, Balesdent, M.H., Buee, M., Rouxel, T., 1995. Genetic characterization of AvrLm1, the  
295 first avirulence gene of *Leptosphaeria maculans*. Phytopathology (USA).



- 296 Balesdent, M.H., Attard, A., Ansan-Melayah, D., Delourme, R., Renard, M., Rouxel, T., 2001. Genetic Control  
297 and Host Range of Avirulence Toward Brassica napus Cultivars Quinta and Jet Neuf in *Leptosphaeria maculans*.  
298 *Phytopathology* 91, 70–76.
- 299 Balesdent, M.H., Attard, A., Kühn, M.L., Rouxel, T., 2002. New Avirulence Genes in the Phytopathogenic Fungus  
300 *Leptosphaeria maculans*. *Phytopathology* 92, 1122–1133.
- 301 Basenko, E.Y., Sasaki, T., Ji, L., Prybol, C.J., Burckhardt, R.M., Schmitz, R.J., Lewis, Z.A., 2015. Genome-wide  
302 redistribution of H3K27me3 is linked to genotoxic stress and defective growth. *Proceedings of the National Academy*  
303 *of Sciences* 112, E6339–E6348.
- 304 Chujo, T., Scott, B., 2014. Histone H3K9 and H3K27 methylation regulates fungal alkaloid biosynthesis in a  
305 fungal endophyte-plant symbiosis: K9 and K27 methylation regulates symbiosis. *Molecular Microbiology* 92, 413–434.
- 306 Connolly, L.R., Smith, K.M., Freitag, M., 2013. The *Fusarium graminearum* Histone H3 K27 Methyltransferase  
307 KMT6 Regulates Development and Expression of Secondary Metabolite Gene Clusters. *PLoS Genetics* 9, e1003916.
- 308 Dutreux, F., Da Silva, C., d'Agata, L., Couloux, A., Gay, E.J., Istace, B., Lapalu, N., Lemainque, A., Linglin, J., Noel,  
309 B., Wincker, P., Cruaud, C., Rouxel, T., Balesdent, M.-H., Aury, J.-M., 2018. De novo assembly and annotation of three  
310 *Leptosphaeria* genomes using Oxford Nanopore MinION sequencing. *Scientific Data* 5.
- 311 Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm,  
312 L., Mistry, J., Sonnhammer, E.L.L., Tate, J., Punta, M., 2014. Pfam: the protein families database. *Nucleic Acids Res.* 42,  
313 D222-230.
- 314 Fokkens, L., Shahi, S., Connolly, L.R., Stam, R., Schmidt, S.M., Smith, K.M., Freitag, M., Rep, M., 2018. The multi-  
315 speed genome of *Fusarium oxysporum* reveals association of histone modifications with sequence divergence and  
316 footprints of past horizontal chromosome transfer events. *bioRxiv*.
- 317 Fudal, I., Ross, S., Gout, L., Blaise, F., Kuhn, M.L., Eckert, M.R., Cattolico, L., Bernard-Samain, S., Balesdent,  
318 M.H., Rouxel, T., 2007. Heterochromatin-Like Regions as Ecological Niches for Avirulence Genes in the *Leptosphaeria*  
319 *maculans* Genome: Map-Based Cloning of *AvrLm6*. *Molecular Plant-Microbe Interactions* 20, 459–470.
- 320 Gacek, A., Strauss, J., 2012. The chromatin code of fungal secondary metabolite gene clusters. *Appl Microbiol*  
321 *Biotechnol* 95, 1389–1404.
- 322 Gervais, J., Plissonneau, C., Linglin, J., Meyer, M., Labadie, K., Cruaud, C., Fudal, I., Rouxel, T., Balesdent, M.-  
323 H., 2017. Different waves of effector genes with contrasted genomic location are expressed by *Leptosphaeria*  
324 *maculans* during cotyledon and stem colonization of oilseed rape. *Molecular plant pathology* 18, 1113–1126.
- 325 Gout, L., Fudal, I., Kuhn, M.-L., Blaise, F., Eckert, M., Cattolico, L., Balesdent, M.-H., Rouxel, T., 2006. Lost in the  
326 middle of nowhere: the *AvrLm1* avirulence gene of the Dothideomycete *Leptosphaeria maculans*. *Molecular*  
327 *Microbiology* 60, 67–80.
- 328 Guo, S., Zhong, S., Zhang, A., 2013. Privacy-preserving Kruskal-Wallis test. *Comput Methods Programs Biomed*  
329 112, 135–145.
- 330 Idnurm, A., Urquhart, A.S., Vummadi, D.R., Chang, S., Van de Wouw, A.P., López-Ruiz, F.J., 2017. Spontaneous  
331 and CRISPR/Cas9-induced mutation of the osmosensor histidine kinase of the canola pathogen *Leptosphaeria*  
332 *maculans*. *Fungal Biology and Biotechnology* 4.



- 333 Jamieson, K., Rountree, M.R., Lewis, Z.A., Stajich, J.E., Selker, E.U., 2013. Regional control of histone H3 lysine  
334 27 methylation in *Neurospora*. *Proceedings of the National Academy of Sciences* 110, 6027–6032.
- 335 Klocko, A.D., Ormsby, T., Galazka, J.M., Leggett, N.A., Uesaka, M., Honda, S., Freitag, M., Selker, E.U., 2016.  
336 Normal chromosome conformation depends on subtelomeric facultative heterochromatin in *Neurospora crassa*. *PNAS*  
337 113, 15048–15053.
- 338 Lanctôt, C., Cheutin, T., Cremer, M., Cavalli, G., Cremer, T., 2007. Dynamic genome architecture in the nuclear  
339 space: regulation of gene expression in three dimensions. *Nature Reviews Genetics* 8, 104–115.
- 340 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-  
341 Delta Delta C(T)) Method. *Methods*. 2001 25:402-408.
- 342 Lo Presti, L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., Zuccaro, A., Reissmann, S., Kahmann, R.,  
343 2015. Fungal Effectors and Plant Susceptibility. *Annual Review of Plant Biology* 66, 513–545.
- 344 Lo Presti, L., Kahmann, R., 2017. How filamentous plant pathogen effectors are translocated to host cells.  
345 *Current Opinion in Plant Biology* 38, 19–24.
- 346 Möller, M., Schotanus, K., Soyer, J., Haeisen, J., Happ, K., Stralucke, M., Happel, P., Smith, K.M., Connolly, L.R.,  
347 Freitag, M., Stukenbrock, E.H., 2019. Destabilization of chromosome structure by histone H3 lysine 27 methylation.
- 348 Muller, P.Y., Janovjak, H., Miserez, A.R., Dobbie, Z., 2002. Processing of gene expression data generated by  
349 quantitative real-time RT-PCR. *BioTechniques* 32, 1372–1374, 1376, 1378–1379.
- 350 Mullins, E.D., Chen, X., Romaine, P., Raina, R., Geiser, D.M., Kang, S., 2001. *Agrobacterium*-Mediated  
351 Transformation of *Fusarium oxysporum*: An Efficient Tool for Insertional Mutagenesis and Gene Transfer.  
352 *Phytopathology* 91, 173–180.
- 353 Papadopoulos, J.S., Agarwala, R., 2007. COBALT: constraint-based alignment tool for multiple protein  
354 sequences. *Bioinformatics* 23, 1073–1079.
- 355 Petit-Houdenot, Y., Degrave, A., Meyer, M., Blaise, F., Ollivier, B., Marais, C.-L., Jauneau, A., Audran, C., Rivas,  
356 S., Veneault-Fourrey, C., Brun, H., Rouxel, T., Fudal, I., Balesdent, M.-H., 2019. A two genes - for - one gene interaction  
357 between *Leptosphaeria maculans* and *Brassica napus*. *New Phytol.* 223, 397–411.
- 358 Rouxel, T., Balesdent, M.H., 2005. The stem canker (blackleg) fungus, *Leptosphaeria maculans*, enters the  
359 genomic era. *Molecular Plant Pathology* 6, 225–241.
- 360 Rouxel, T., Grandaubert, J., Hane, J.K., Hoede, C., van de Wouw, A.P., Couloux, A., Dominguez, V., Anthouard,  
361 V., Bally, P., Bourras, S., Cozijnsen, A.J., Ciuffetti, L.M., Degrave, A., Dilmaghani, A., Duret, L., Fudal, I., Goodwin, S.B.,  
362 Gout, L., Glaser, N., Linglin, J., Kema, G.H.J., Lapalu, N., Lawrence, C.B., May, K., Meyer, M., Ollivier, B., Poulain, J.,  
363 Schoch, C.L., Simon, A., Spatafora, J.W., Stachowiak, A., Turgeon, B.G., Tyler, B.M., Vincent, D., Weissenbach, J.,  
364 Amselem, J., Quesneville, H., Oliver, R.P., Wincker, P., Balesdent, M.-H., Howlett, B.J., 2011. Effector diversification  
365 within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations. *Nature*  
366 *Communications* 2.
- 367 Rouxel, T., Balesdent, M.-H., 2017. Life, death and rebirth of avirulence effectors in a fungal pathogen of  
368 Brassica crops, *Leptosphaeria maculans*. *New Phytol.* 214, 526–532.





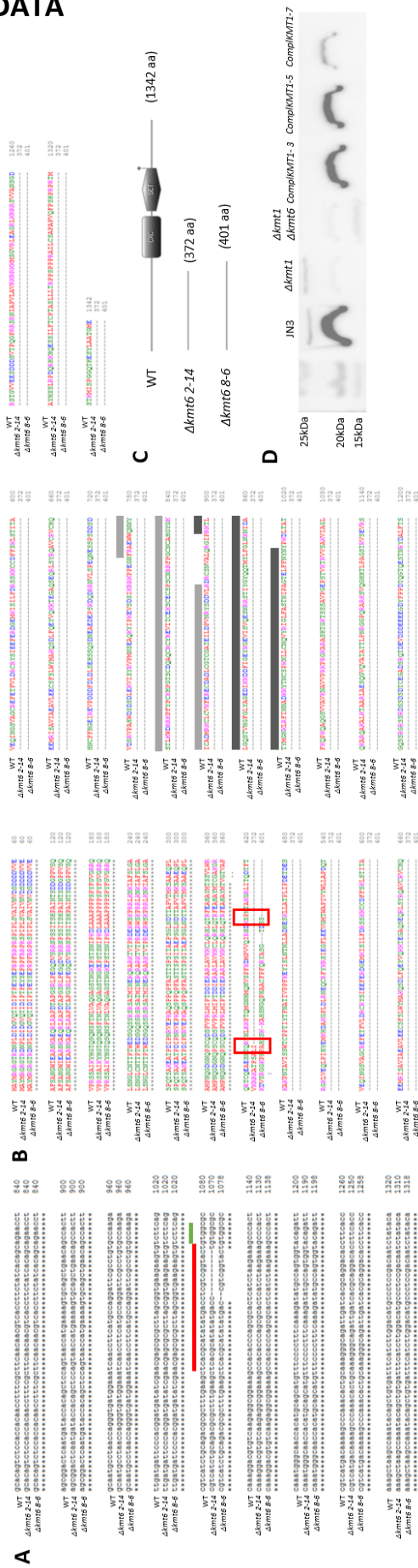
- 369 Sánchez-Vallet, A., Fouché, S., Fudal, I., Hartmann, F.E., Soyer, J.L., Tellier, A., Croll, D., 2018. The Genome  
370 Biology of Effector Gene Evolution in Filamentous Plant Pathogens. *Annual Review of Phytopathology* 56, 21–40.
- 371 Schotanus, K., Soyer, J.L., Connolly, L.R., Grandaubert, J., Happel, P., Smith, K.M., Freitag, M., Stukenbrock,  
372 E.H., 2015. Histone modifications rather than the novel regional centromeres of *Zymoseptoria tritici* distinguish core  
373 and accessory chromosomes. *Epigenetics & Chromatin* 8, 41.
- 374 Silayeva, O., Barnes, A.C., 2018. Gibson Assembly facilitates bacterial allelic exchange mutagenesis. *J.*  
375 *Microbiol. Methods* 144, 157–163.
- 376 Soyer, J.L., El Ghalid, M., Glaser, N., Ollivier, B., Linglin, J., Grandaubert, J., Balesdent, M.-H., Connolly, L.R.,  
377 Freitag, M., Rouxel, T., Fudal, I., 2014. Epigenetic Control of Effector Gene Expression in the Plant Pathogenic Fungus  
378 *Leptosphaeria maculans*. *PLoS Genetics* 10, e1004227.
- 379 Soyer, J.L., Möller, M., Schotanus, K., Connolly, L.R., Galazka, J.M., Freitag, M., Stukenbrock, E.H., 2015.  
380 Chromatin analyses of *Zymoseptoria tritici*: Methods for chromatin immunoprecipitation followed by high-throughput  
381 sequencing (ChIP-seq). *Fungal Genet. Biol.* 79, 63–70.
- 382 Stergiopoulos, I., de Wit, P.J.G.M., 2009. Fungal Effector Proteins. *Annual Review of Phytopathology* 47, 233–  
383 263. <https://doi.org/10.1146/annurev.phyto.112408.132637> Tamaru, H., Selker, E.U., 2001. A histone H3  
384 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* 414, 277.
- 385 Tan, K.-C., Oliver, R.P., 2017. Regulation of proteinaceous effector expression in phytopathogenic fungi. *PLoS*  
386 *pathogens* 13, e1006241.
- 387 Winter, D.J., Ganley, A., Young, C., Liachko, I., Schardl, C., Dupont, P., Berry, D., Ram, A., Scott, B., Cox, M.,  
388 2018. Repeat elements organize 3D genome structure and mediate transcription in the filamentous fungus *Epichloë*  
389 *festucae*. *bioRxiv* 339010.

390  
391  
392  
393  
394  
395  
396  
397  
398  
399  
400  
401  
402  
403  
404  
405



406 SUPPLEMENTARY DATA

407

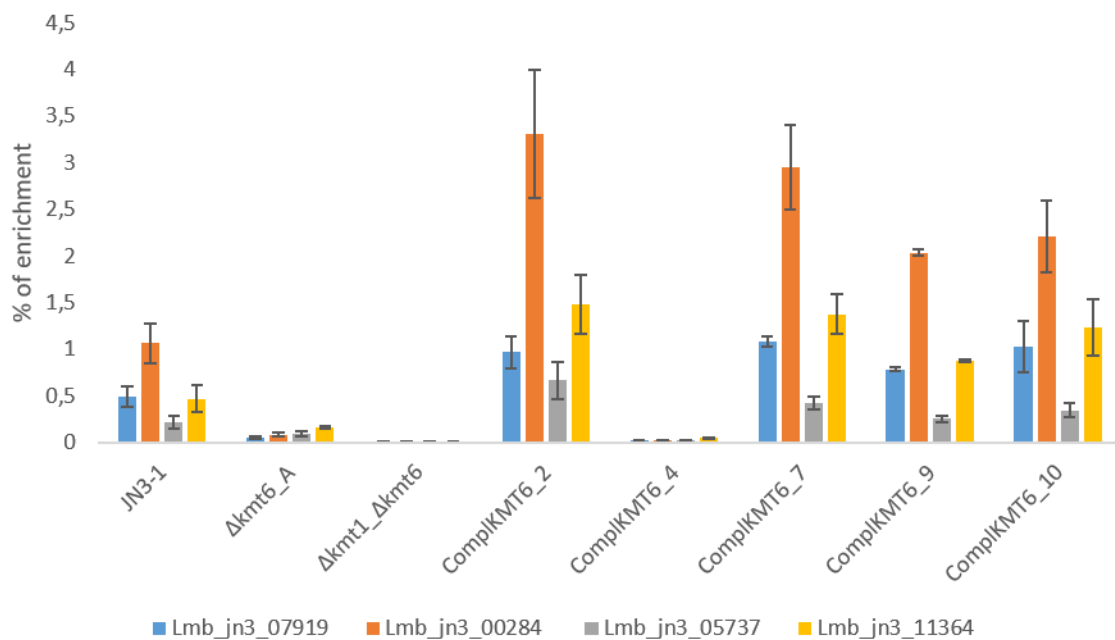


**Figure S1: Effect of  $\Delta kmt6$  mutations on KMT6 protein structure and impact of  $kmt1$  complementation on KMT6 protein structure and impact of  $kmt1$  complementation on KMT6 protein structure.** The WT isolates JN2 and in two mutants  $\Delta kmt6$  2-14 and 8-6 showing respectively a 10bp and 2bp deletion). The PAM (Protospacer Adjacent Motif) is highlighted in green and the region targeted by the guide RNA is highlighted in red **B**: Protein sequence of KMT6 in the WT isolate JN2 and in the two mutants. The red frame highlights the stop in the two mutant version of the protein. The WT version of the protein is 1342 aa long. Light grey bar and dark grey bar indicating respectively CXC and SET domains. **C**: KMT6 protein domains obtained with Pfam as described in Finn *et al.*, 2014 (<https://pfam.xfam.org/>). **D**: Western Blot analysis of H3K9 tri-methylation in JN3 (WT),  $\Delta kmt1$ ,  $\Delta kmt6$  and different  $kmt1$  complemented strains. Nuclear protein samples were extracted from mycelium, separated by SDS-PAGE, blotted and probed with antibodies against H3K9me3.



408

409



**Figure S2: Trimethylation of lysine 27 of histone H3 (H3K27me3) is abolished in  $\Delta kmt6$  and  $\Delta kmt1_{\Delta kmt6}$  backgrounds and recovered in complemented 2, 7, 9 and 10 transformants.** Chromatin immunoprecipitation analysis was performed to assess changes in the heterochromatic H3K27me3 mark in 'late' effector genes known to be in H3K27me3 environment in WT strain. We used avirulence genes as a negative control of H3K27me3 enrichment. Data were normalized using the "percent of input" method. Error bars indicate the standard deviation of two biological and technical replicates.

Table S1: Primers used in this study

Name of the fragment	DNA sequence	function	construction
gRNA against KMT1	GAACCTAATCAATCAACGGAGCAGCTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTCTGCTCCCTTCGAACGCTCGCGT TTTAGAGCTAGAAATAGC	KMT1 CRISPR-CAS9 directed inactivation	
gRNA against KMT6	GAACCTAATCAATCAACGGCGTGACTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTCTCACCACATATATGACCTCGTCTT TTAGAGCTAGAAATAGC	KMT6 CRISPR-CAS9 directed inactivation	
MAI0309	ACCTTAATCGAAACCTAATCAATCAAC	Amplification of gRNA coding sequence	CRISPR-CAS9 inactivation
MAI0310	ATTTAACTTGCTATTTCTAGCTCTAAAAC	Amplification of gRNA coding sequence	
verif_CRISPR_KMT1_F	GCCCTCCATGGTTGTGGTGC	KMT1 mutation detection	
verif_CRISPR_KMT1_R	ACCCCGAAATGCTAGCCGA	KMT1 mutation detection	
verif_CRISPR_KMT6_F	GCGCGCTTAGCGGTGAAGAA	KMT6 mutation detection	
verif_CRISPR_KMT6_R	AGAGCCGTGGCAAGCTCGTA	KMT6 mutation detection	
Compl_KMT1_F	AGTCACCTGCAGGCATGCCGAATGGAAACTTTTCTGG	KMT1 amplification	
Compl_KMT1_R	GTAAAAACGCGCCAGTGCACCTCCCTCGGATACTTTCTAC	KMT1 amplification	
Compl_2_KMT6_F	AGTCACCTGCAGGCATGCATTCCTCGGGGCCACACTAC	KMT6 amplification	
Compl_2_KMT6_R	GTAAAAACGCGCCAGTGCACCTCCCTCGGGGCCACACTAC	KMT6 amplification	
verif_KMT6_F1	CATGATCAACGAAAGACAGC	KMT6 sequence verification	gene
verif_KMT6_F2	GGCTGTCAGAAATTTGAGAC	KMT6 sequence verification	complementation
verif_KMT6_F3	CAATAGGTACAGTACGACGC	KMT6 sequence verification	
verif_KMT1_F1	TCGACTGGACAGGAGGAAGCG	KMT1 sequence verification	
verif_KMT1_F2	GTATTGCGGGAAACAAGAAATG	KMT1 sequence verification	
βTubulin_F	AAGAACTCATCTACTTCGA	qPCR	None
βTubulin_R	TGAATAGCTCCTGAATGG	qPCR	None
AvrLm4-7_F	GCCCTGCATAAATACTACCGAC	qPCR	None
AvrLm4-7_R	TCCTGGCCAAATATAAATCC	qPCR	None
AvrLm6_F	AAACGGCACTATTACGAAAA	qPCR	None
AvrLm6_R	GATTAGCCGAGAAAGCAAGT	qPCR	None
AvrLm10.A_F	ACAGGAGAACACTGACGGCT	qPCR	None
AvrLm10.A_R	GCATGGACTATCGTAGCGTTAT	qPCR	None
AvrLm11_F	TGCGTTCTTGCTCCTATATTT	qPCR	None
AvrLm11_R	CAAGTTGGATCTTTCTCATTCG	qPCR	None
Lmb_jn3_11364_F	FACTGGCTGCTTTTGTCTGC	qPCR	None
Lmb_jn3_11364_R	CCGCCACAGTGAACATAAT	qPCR	None
Lmb_jn3_00284_F	ACCTTCCACACGAAACC	qPCR	None
Lmb_jn3_00284_R	CAGCCAGCTCCCGCTTTTA	qPCR	None
Lmb_jn3_07919_F	ATCGCCATCCTCTCCCTT	qPCR	None
Lmb_jn3_07919_R	ACTCAGCGTTACCAGCCTC	qPCR	None
Lmb_jn3_05737_F	ACAAATGGAGGACAAGAAG	qPCR	None
Lmb_jn3_05737_R	AGTAGITGGAAAGAGTGACAG	qPCR	None
ActinBn_F	TCCCTCAGCACTTTC AAC	qPCR	None
ActinBn_R	TTCAATCCCCATCTCAACAA	qPCR	None



411

**Table S2: Strains used for mutant purification.**

	CRISPR-CAS9 mutation	Purification (Crosses)			
		JN2 <sup>a</sup>	JN2-GFP <sup>b</sup>	selected progeny <sup>c</sup>	Additional Transformation <sup>d</sup>
<i>Δkmt1 6-1</i>	1 bp insertion	yes	yes	<i>Δkmt1_A</i> <i>Δkmt1_A</i> -GFP	complementation
<i>Δkmt6 2-14</i>	10pb deletion	yes	yes	<i>Δkmt6_A</i> <i>Δkmt6_A</i> -GFP	complementation
<i>Δkmt6 8-6</i>	2pb deletion	no	yes	<i>Δkmt6_B</i> -GFP <i>Δkmt6_B</i> -GFP	no
<i>Δkmt1_Δkmt6 60-1</i>	2pb deletion <i>KMT1</i> and 10pb del <i>KMT6</i>	no	yes	<i>Δkmt1/Δkmt6</i>	no

a: Purification by crossing with an isolate of opposite mating type with or without GFP. b: Selected progenies purified (without *Cas9* and gRNA). c: Second round of purification by crossing with JN2-GFP to obtain purified-GFP mutants. d: Names of the selected purified (without Cas9 and gRNA) descendants. e: Transformation on purified transformants (after all purification rounds).

412

## FUNDING

413

This work was supported by the Plant Health and Environment division of the French National Institute for Agricultural Research [Chromadyn Project AAP 2018].

414

415

## CONFLICT OF INTEREST

416

The authors declare no conflict of interest.

417

418

419

420

421



422 RESULTATS COMPLEMENTAIRES

Analyse fonctionnelle de KDM8, histone déméthylase putative spécifique de H3K36me2

Certaines protéines sont capables de retirer les modifications chimiques des queues d'histones et ainsi de réguler l'expression génique (Kronholm *et al.*, 2016), et certaines déméthylases ont été identifiées comme régulatrice de processus infectieux. Il a notamment été montré que certaines déméthylases d'histones étaient impliquées dans le contrôle du métabolisme secondaire chez certains champignons filamenteux. C'est le cas de KdmA et KdmB, responsables respectivement de l'enlèvement d'une marque répressive, H3K9me3, et d'une marque activatrice, H3K4me3, jouent un rôle dans la régulation du switch entre métabolisme primaire et secondaire chez *Aspergillus nidulans* (Gacek-Matthews *et al.*, 2015 ; Gacek-Matthews *et al.*, 2016). Parmi les sept déméthylases d'histones prédites chez *L. maculans* (Pratx and Zurletto, Pers. Comm), Kmd8, est spécifiquement sur-exprimée pendant les étapes précoces de l'infection de cotylédons de colza (Figure 1). Son homologue chez *B. cinerea*, *Bckdm1* (42 % identité) a été identifiée comme impliquée dans de nombreux processus physiologiques et notamment dans la pénétration de l'hôte (Schumacher *et al.*, 2019). Pour étudier son implication dans la régulation des gènes codant des effecteurs, j'ai réalisé une mutagenèse dirigée en utilisant la technique CRISPR-Cas9 contre le gène codant cette protéine. J'ai ainsi identifié deux mutants indépendants, ayant respectivement 2 et 3 pb délétées dans la partie 5' du gène. L'une des délétions entraîne au niveau protéique la perte du domaine fonctionnel Jumonji de la protéine, alors que la seconde délétion n'entraîne qu'une perte d'un acide aminé et ne modifie que très légèrement la séquence protéique (Figure 2). J'ai ensuite croisé les mutants ainsi obtenus avec une souche sauvage pour éliminer le gène codant la Cas9 et l'ARN guide. Les mutants purifiés ont été respectivement nommés  $\Delta Kdm8\_A$  et  $\Delta Kdm8\_B$ . J'ai ensuite caractérisé les mutants pour leur croissance et leur pathogénie sur cultivar sensible de colza (Figure 3). Aucun défaut de pathogénie ou croissance

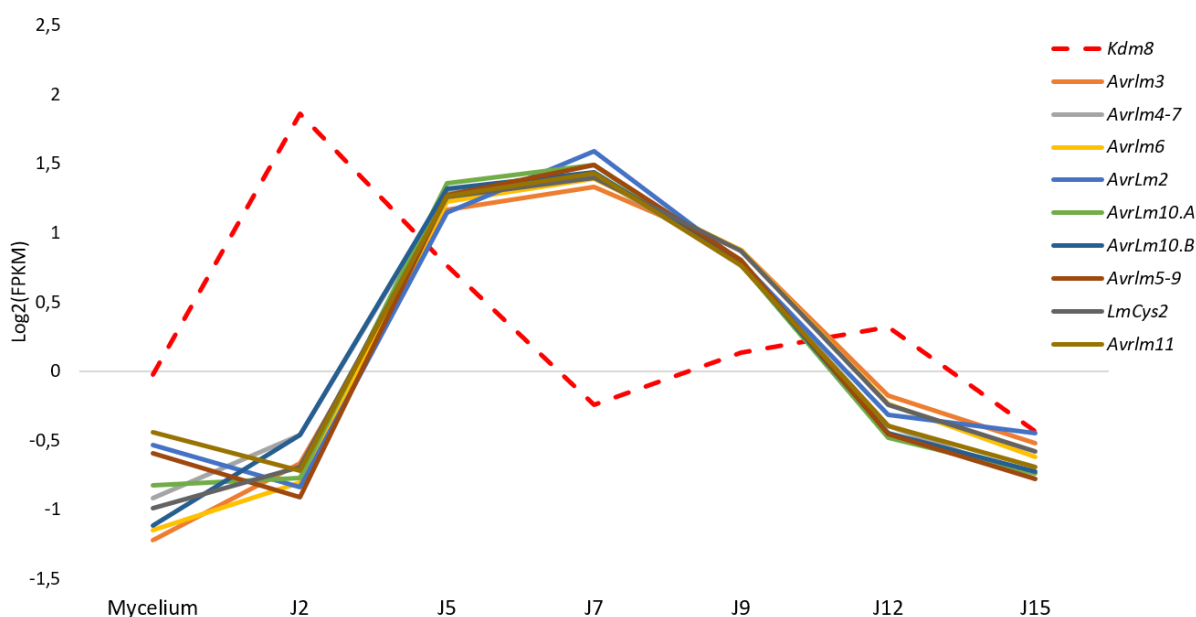
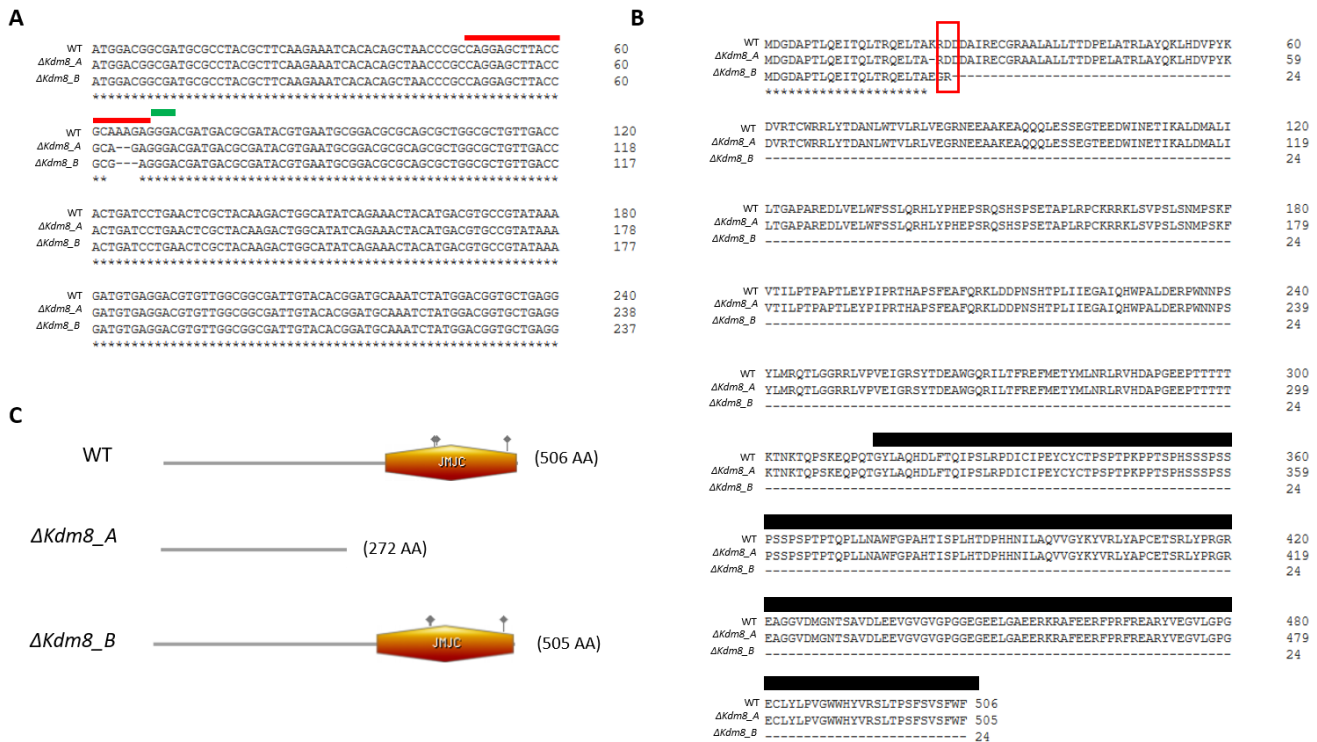


Figure 1: Expression profile of *Leptosphaeria maculans* avirulence genes and *Kdm8* during axenic growth and infection of oilseed rape cotyledons. Mycelium was obtained by growing JN2 strain on V8-agar medium for 7 days. Cotyledons of the susceptible cultivar Darmor were inoculated by JN2 and sampled at seven different time points (2, 5, 7, 9, 12 and 15 days post infection). Two biological replicates per condition were generated. Total RNA was extracted and sequenced. The expression level of eight avirulence genes and *Kdm8* is represented by the log2 of FPKM (Fragments Per Kilobase per Million mapped reads) (Gay *et al.*, ongoing thesis).

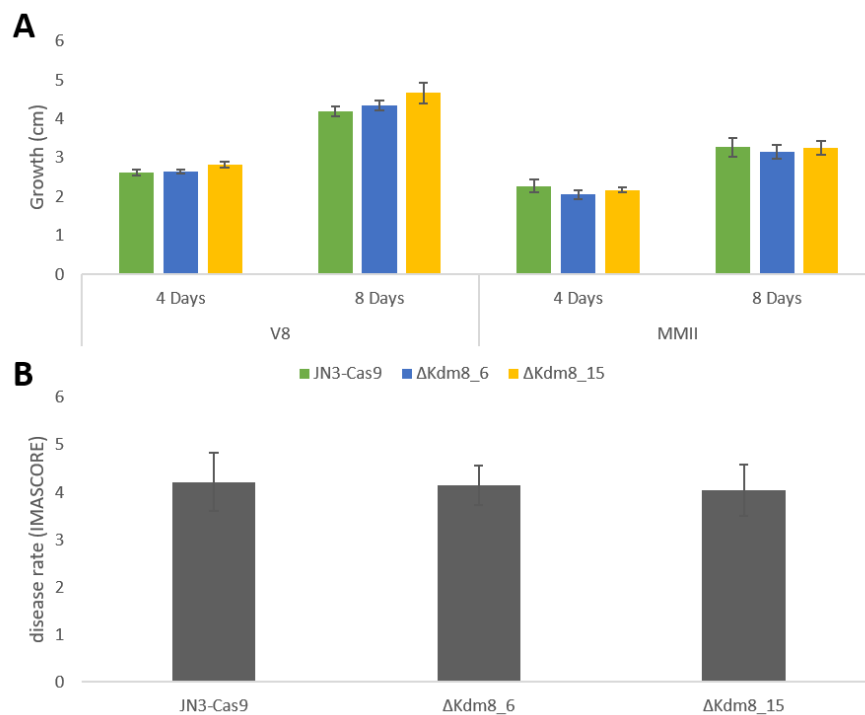


n'a été observé. Nous avons ensuite cherché à savoir si le mutant  $\Delta Kdm8\_6$  ayant perdu le domaine fonctionnel de Kdm8, était affecté pour l'expression de gènes codant des effecteurs *in vitro* et *in planta*. J'ai donc testé par RT-qPCR l'expression de différents gènes codant effecteurs pendant la croissance axénique et pendant l'infection primaire. Je n'ai mis en évidence aucune altération d'expression des effecteurs testés (Figures 4 et 5).

423

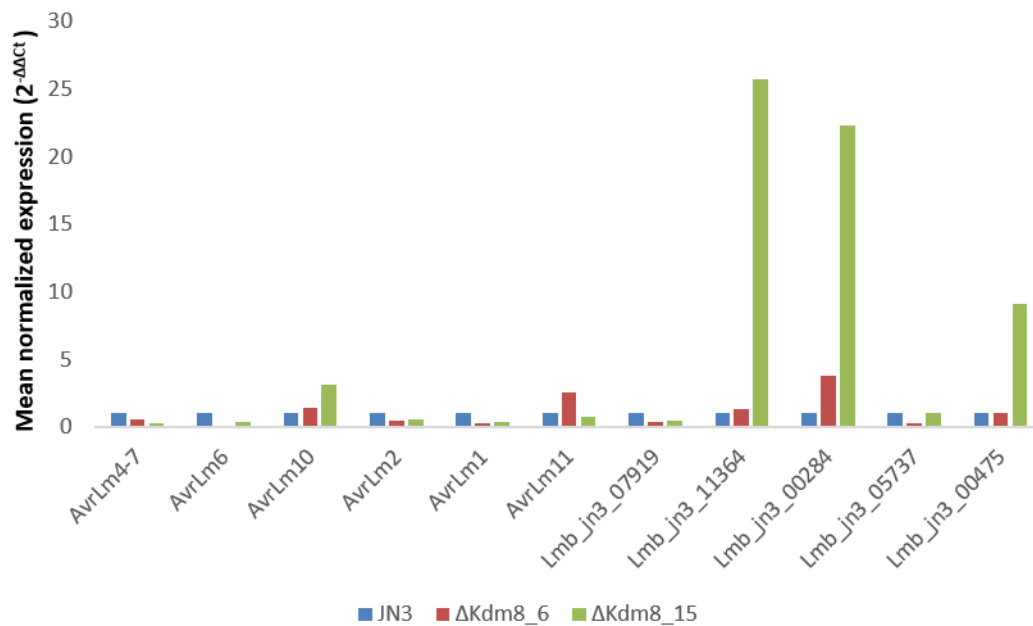


**Figure 2: Effect of the *Kdm8* mutations on *Kdm8* protein structure. A:** Sequence alignment of the *Kdm8* gene in the WT isolate JN2 and in two mutants ( $\Delta Kdm8\_A$  and  $\Delta Kdm8\_B$  showing respectively a 2-bp and 3-bp deletion). The PAM (Protospacer Adjacent Motif) is highlighted in green and the region targeted by the guide RNA is highlighted in red. **B:** Protein sequence of *Kdm8* in the WT isolate JN2 and in the two mutants. The Jumanji domain is indicated by a black bar. The red frame indicates the location of the stop codon in the mutant version of the protein. **C:** *Kdm8* protein length and domains identified with Pfam as described in Finn *et al.*, 2014 (<https://pfam.xfam.org/>).



**Figure 3: Effect of *kdm8* inactivation on mycelial growth and pathogenicity of *Leptosphaeria maculans*.** **A:** Growth assay on *kdm8* mutants and WT strain JN3 with or without Cas9 gene. Mycelial growth was measured at 4 days and 8 days post inoculation on two different agar media, V8 and MMII. **B:** Pathogenicity tests were performed on the susceptible cultivar Es-Astrid at 13DPI. Disease rate is expressed as the mean scoring using the IMAScore rating scale comprising six infection classes (IC), where IC1 to IC3 corresponded to resistance and IC4 to IC6, to susceptibility. DPI= Days Post Inoculation.

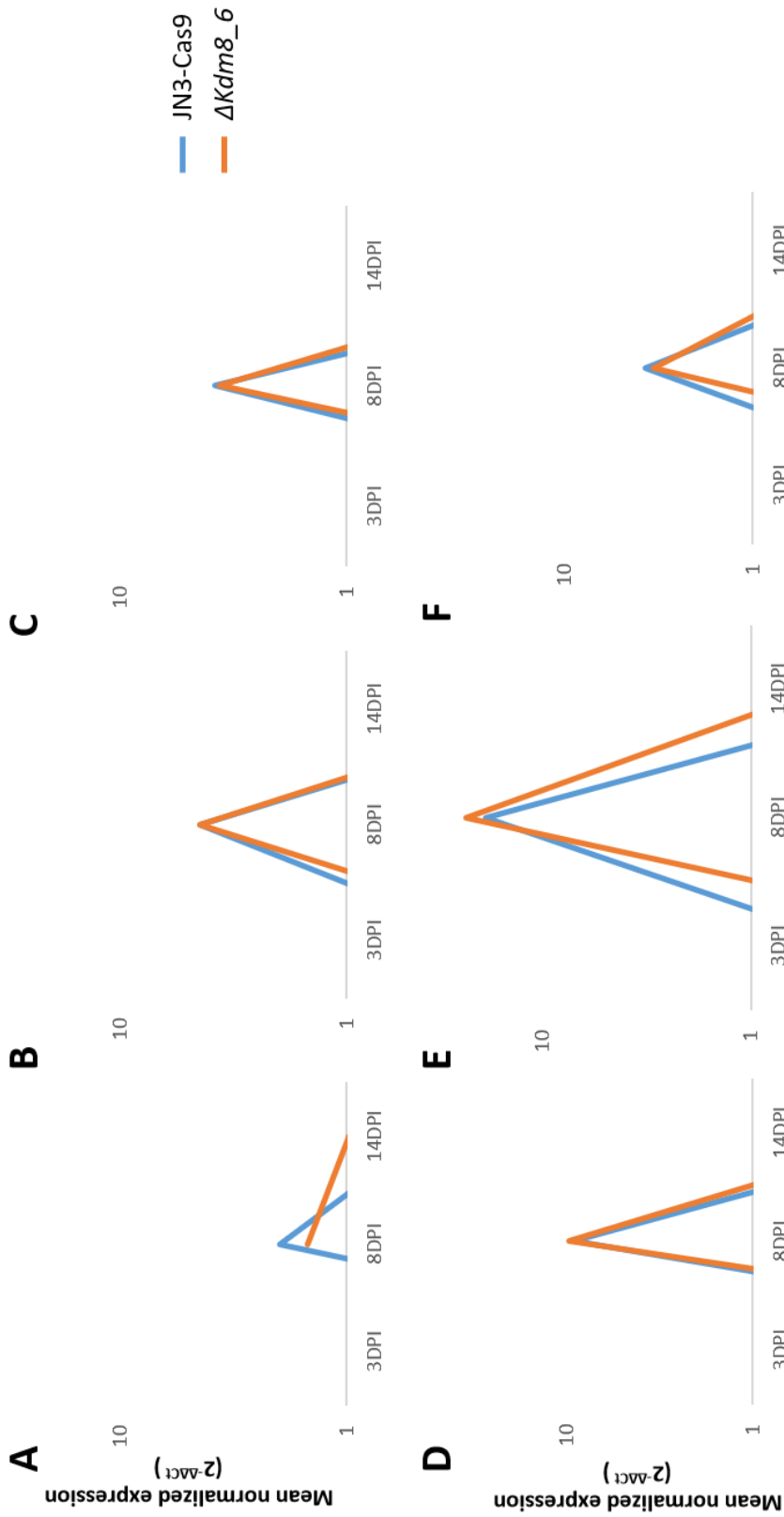




early effector genes

late effector genes

**Figure 4: Influence of *Kdm8* inactivation on expression of several *L. maculans* 'early' and 'late' effector genes *in vitro*.** Mycelium was obtained by growing JN3-Cas9 strain and transformants in Fries liquid medium for 7 days. Two biological replicates per condition were generated. Total RNA was extracted. Expression of 'early' and 'late' effector genes was measured by qRT-PCR and expressed relatively to *LmβTubulin* expression and to expression of the same genes in JN3 using the  $2^{-\Delta\Delta Ct}$  method.



**Figure 5: Influence of *Kdm8* inactivation on the expression of six *L. maculans* avirulence genes during infection of oilseed rape.** Oilseed rape susceptible cultivar ES Astrid was inoculated with JN3-Cas9 and ΔKdm8 mutant. Cotyledons were harvested 5, 8 and 14 dpi. Expression of *Lmb\_jn3\_11364* (A) *AvrLm10A* (B) ; *AvrLm4-7* (C) ; *AvrLm6* (D) ; *AvrLm11* (E) ; *AvrLm2* (F) was measured by qRT-PCR level using *LmβTubulin* as a constitutive reporter gene and expression in JN3-Cas9 as reference (2<sup>-ΔAct</sup> method). Each value is the average of two biological replicates (two extractions from different biological replicates) and two technical replicates (two RT-PCR).



## CHAPITRE III

# Régulation transcriptionnelle



## INTRODUCTION

Dans le chapitre III sera abordée la question de la régulation transcriptionnelle de l'expression des gènes codant des effecteurs chez *L. maculans*. Pour cette étude, nous avons identifié un facteur de transcription (*LmPf2*) impliqué dans le contrôle des effecteurs fongiques et de la pathogénie chez plusieurs Dothideomycetes phytopathogènes et conservé chez *L. maculans*. Grâce à l'exploitation de données RNA-Seq générées dans le cadre d'un autre projet et couvrant tout le cycle de vie de *L. maculans* au champ et en conditions contrôlées, nous avons pu observer que ce FT avait un profil d'expression similaire à celui des gènes d'avirulence. J'ai réalisé l'inactivation de *LmPf2* par la technique CRISPR-Cas9. J'ai également réalisé des constructions permettant la sur-expression de *LmPf2*. J'ai caractérisé les transformants ainsi obtenus pour leur croissance, leur sporulation et leur capacité à infecter le colza. J'ai également mené des analyses RNA-Seq afin d'identifier les gènes directement ou indirectement régulés par *LmPf2*. L'inactivation de *LmPf2* mène à une perte de pouvoir pathogène ainsi qu'à une absence d'expression des effecteurs exprimés durant les phases précoces de l'infection *in planta*. La sur-expression de *LmPf2* dans deux fonds génétiques distincts, un fond sauvage et un fond  $\Delta kmt1$  dans lequel la marque répressive H3K9me3 était absente, nous a permis de montrer que la sur-expression de *LmPf2* selon l'état de compaction de la chromatine va être capable de réguler différents sets de gènes selon des vagues concertées. J'ai rédigé les résultats obtenus sous forme d'un article dont je suis premier auteur (première partie du chapitre). Dans une deuxième partie, je présenterai les résultats obtenus lors de l'inactivation de deux autres FTs spécifiques de *L. maculans*, également surexprimés pendant les étapes précoces d'infection.



## RESULTATS

### ARTICLE IV

## Regulation of effector gene expression as concerted waves in *Leptosphaeria maculans*: a two-players game



# 1 Regulation of effector gene expression as concerted waves in *Leptosphaeria maculans*: a 2 two-players game

3 C. Clairet<sup>1,2</sup>, E.J. Gay<sup>1,2</sup>, A. Simon<sup>1</sup>, N. Lapalu<sup>1</sup>, F. Blaise<sup>1</sup>, C.L. Marais<sup>1</sup>, J.L. Soyer<sup>1,\*</sup> and I. Fudal<sup>1,\*</sup>

4 <sup>1</sup> UMR BIOGER, INRA, AgroParisTech, Université Paris-Saclay, Avenue Lucien Brétignières, BP 01, F-78850 Thiverval-Grignon,  
5 France

6 <sup>2</sup> Université Paris-Sud, 91405, Orsay, France

7 \* To whom correspondence should be addressed.

8 E-mail: [jessica.soyer@inra.fr](mailto:jessica.soyer@inra.fr); [isabelle.fudal@inra.fr](mailto:isabelle.fudal@inra.fr)

## 9 ABSTRACT

10 Phytopathogenic fungi secrete during infection a set of molecules collectively known as effectors, involved in  
11 overcoming the host immune defense system and in disease establishment. In *Leptosphaeria maculans*, the causal  
12 agent of oilseed rape stem canker, we established that the histone modification H3K9me3, deposited by Kmt1, was  
13 involved in the regulation of expression of genes highly expressed during infection, including effectors expressed  
14 during the early stages of infection. Effector genes are concertedly expressed as waves during infection. These  
15 concerted patterns of effector gene expression suggest a tight regulation of their expression involving both chromatin  
16 remodeling proteins and specific transcription factors (TFs). Pf2, a TF belonging to the Zn2Cys6 fungal specific TF family,  
17 was described in several Dothideomycete species as essential for pathogenicity and effector gene expression. Here,  
18 we identified the protein orthologous to Pf2 in *L. maculans*, LmPf2. To investigate the role of LmPf2 together with  
19 Kmt1 in *L. maculans*, we inactivated *LmPf2* and over-expressed *LmPf2* in a WT and a  $\Delta kmt1$  backgrounds. Functional  
20 analyses of the corresponding transformants highlighted an essential role of LmPf2 in the establishment of  
21 pathogenesis. It also showed that LmPf2 is involved in the control of effector gene expression, particularly in a  $\Delta kmt1$   
22 background. Transcriptomic analyses, using RNA-seq, during axenic growth helped deciphering the complex regulatory  
23 roles of LmPf2 and KMT1 on effector gene expression. Altogether, our findings allowed us to establish an effector gene  
24 regulatory model both *in planta* and *in vitro*.

## 25 INTRODUCTION

26 During infection, plant pathogenic fungi secrete a set of molecules, collectively known as effectors, involved  
27 in overcoming the host immune defense system, nutrient uptake and, eventually, symptom development (Lo Presti *et al.*, 2015). Effectors mainly correspond to small secreted proteins, generally cysteine-rich, with no homology to other  
28 known proteins in the databases (SSPs; Stergiopoulos and de Wit, 2009). Effector genes of filamentous plant pathogens  
29 are often located in transposable elements (TEs)-rich regions of the genomes, such as dispensable chromosomes or  
30 telomeres (Sánchez-Vallet *et al.*, 2018). Transcriptomic data generated during different stages of plant infection  
31 highlighted concerted waves of effector gene expression over the course of infection, according to the infection  
32 structure, to the host-plant infected or to the species or strain studied (Sanchez-Vallet *et al.*, 2018). Little is



34 known about how coordinated expression of effector genes is regulated. Based on observation that many effector  
35 genes are located in TE-rich regions, and up-regulated during host penetration/infection, the role of chromatin  
36 remodeling in the regulation of these genes was recently investigated. Many proteins were as involved in chromatin  
37 remodeling, and analyses in plant pathogenic fungi have shed light on two histone methyltransferases, KMT1 and  
38 KMT6. These two proteins are involved in the trimethylation of the lysine 9 of histone H3 (H3K9me3) and of the  
39 trimethylation of the lysine 27 of histone H3 (H3K27me3), typical for the tightly packed portion of the chromatin, the  
40 heterochromatin, and associated with gene repression. The role of KMT1 and KMT6 in the regulation of the expression  
41 of genes encoding pathogenicity determinants (effector and secondary metabolite-encoding genes) has been  
42 demonstrated in distantly related fungi, for instance in *Leptosphaeria maculans* and in *Epichloe festucae* (Chujo & Scott  
43 2014; Soyer *et al.*, 2014). In *L. maculans*, the causal agent of stem canker of oilseed rape (*Brassica napus*), inactivation  
44 of *KMT1*, led to an over-expression of more than 30% of the genes located in TE-rich environments, normally silenced  
45 *in vitro* in the WT strain, specifically effector genes, compared to the wild type strain. ChIP-qPCR analyses showed that  
46 over-expression of at least two effector genes was associated with a decrease of the repressive histone modification  
47 H3K9me3 in the genomic environment of these genes (Soyer *et al.*, 2014). In a recent study we constructed the map  
48 of H3K4me2 activating modification, H3K9me3 and H3K27me3 repressive modification respective repartitions on *L.*  
49 *maculans* genomes (Soyer *et al.*, 2019). We used this map to get further insight in chromatin modifications  
50 involvement in gene regulation and more specifically on effector gene regulation in *L. maculans*. Altogether, these  
51 analyses revealed that chromatin structure, via the dynamic of chromatin remodeling, was an important regulatory  
52 layer of effector genes up-regulated during host infection and located in TE-rich regions.

53 *Leptosphaeria maculans* is an ideal model to investigate the dual control of effector gene expression: (i) it  
54 displays a complex life cycle with alternating phases of saprophytism, asymptomatic growth and necrotrophy that  
55 suggest involvement of sophisticated regulation networks. It first infects oilseed rape leaves and undergoes asexual  
56 sporulation leading to leaf spots. Then, the fungus grows into the stem without visible symptom and finally provokes  
57 stem canker at the end of the oilseed rape growing season (Rouxel and Balesdent, 2005). (ii) The *L. maculans* genome  
58 displays a bipartite structure with gene-rich regions and TE-rich regions. One thousand and eighty putative effector  
59 genes are predicted in its genome, and regions enriched for histone modifications typical for heterochromatin are  
60 significantly enriched in such genes (Dutreux *et al.*, 2018; E.J. Gay, on-going PhD thesis; J.L. Soyer, unpublished). During  
61 infection of oilseed rape, several waves of effector genes are expressed (Rouxel *et al.*, 2011; Gervais *et al.*, 2017; E.J.  
62 Gay, on-going PhD thesis). Of particular interest, during the asymptomatic stages occurring on both leaves and stems,  
63 a specific set of effector genes is expressed, including all known avirulence genes (*AvrLm*) identified so far, and these  
64 genes are specifically located in TE-rich regions. Avr protein coded by these genes directly or indirectly activates plant  
65 defense upon recognition mediated by corresponding plant R proteins (Biezen and Jones, 1998). iii) Involvement of  
66 the chromatin structure in the regulation of gene expression highlighted that the chromatin folding in TE-rich regions  
67 was instrumental in the regulation of effector genes (and *AvrLm* genes) located in these regions. Nevertheless, at least  
68 in *L. maculans*, inactivation of *KMT1*, while highly inducing expression of effector genes *in vitro*, did not induce  
69 expression of these genes at the same level as during infection of oilseed rape. This suggests that, although chromatin  
70 folding plays an important role in the regulatory mechanism, a second regulator, such as a transcription factor,



might be involved. Based on the observation that loosening of the chromatin fiber alone does not result in the same expression level as during *in planta* infection, Soyer *et al.* (2015. A) proposed a model of dual control of effector gene expression: after chromatin loosening upon infection, one or several transcription factor(s) (TFs) could bind effector gene promoters resulting in their concerted expression. Synergic involvement of chromatin modifications and TF(s) remains poorly understood and an unexplored field of investigation.

In filamentous plant pathogens, only a few transcription factors influencing effector gene expression have been identified so far (see for review Tan and Oliver, 2017). Among them, AbPf2, a Zn<sub>2</sub>Cys<sub>6</sub> TF, was first described in *Alternaria brassicicola*, in which it regulates, directly or indirectly, expression of 33 genes encoding secreted proteins including eight putative effectors (Cho *et al.*, 2013). In *Parastagonospora nodorum*, the AbPf2 orthologue PnPf2 positively regulates two necrotrophic effector genes, *SnToxA* and *SnTox3* and the orthologue of *SnToxA*, *ToxA*, is regulated by PtrPf2 in *Pyrenophora tritici-repentis* (Rybak *et al.*, 2017).

Here, we investigated the involvement of the homolog of Pf2 in *L. maculans*, LmPf2, in the control of effector gene expression. We inactivated *LmPf2* using the CRISPR-Cas9 strategy and over-expressed *LmPf2* in two different genetic backgrounds: (i) a wild type strain; (ii) a strain in which the gene encoding KMT1 (responsible for the deposition of H3K9me3) is inactivated. We characterized the corresponding transformants for their growth, sporulation and pathogenicity. We performed a RNA-seq analysis in order to decipher the respective involvement of LmPf2 and KMT1 in gene regulation. We found out that KMT1 and LmPf2 are acting together to specifically control at least part of two different sets of genes expressed as concerted waves at two infection time points.

## MATERIAL AND METHODS

### Fungal culture

The reference isolates JN2 and JN3 were used as hosts for genetic transformations (Balesdent *et al.*, 2001). We also used a JN2 strain constitutively expressing *eGFP* (Sâsèk *et al.*, 2012) which was crossed with our mutants in order to follow the colonization of oilseed rape by the mutants. Fungal cultures and conidia production were performed as previously described (Ansan-Melayah *et al.*, 1995). For DNA/RNA extractions, mycelium was grown on V8-juice agar medium at 25°C in the dark for seven days and then plugs were transferred into 150 ml of Fries liquid medium in 500 ml Roux flasks. Tissues were harvested after growing for seven days at 25°C.

### Pathogenicity and growth assays

Pathogenicity assays were performed on cotyledons of 15-day-old plantlets of a susceptible cultivar of *B. napus*, Es-Astrid. Plants were incubated in a growth chamber at 16/24°C (night/day) with a 16h photoperiod. Symptoms were scored on 10-12 plants, with two biological replicates, 13 days post inoculation (dpi) using the IMAScore rating scale comprising six infection classes (IC), where IC1 to IC3 correspond to various levels of resistance of the plant and IC4 to IC6 to susceptibility (Balesdent *et al.*, 2001). Growth assays were performed by deposition of a 5 mm plug at the center of 90 mm Petri dishes (containing 20 ml of V8-juice agar medium or MMIII medium). Radial growth was measured at nine days after incubation in a growth chamber (25°C) on four biological replicates and statistical analysis were performed using Kruskal-Wallis test (Guo *et al.*, 2013).





## 106 DNA and RNA manipulation

107 For PCR, genomic DNA was extracted from conidia or from mycelium grown in Fries liquid culture with the  
108 DNAeasy 96 plant Kit (Qiagen S.A., Courtaboeuf, France). PCR amplifications were performed as previously described  
109 (Fudal *et al.*, 2007). To identify mutation arising in the sequence of the gene targeted by the CRISPR-Cas9 strategy,  
110 sequencing was performed by Eurofins Genomics (Anzinger, Ebersberg, Germany). Total RNA was extracted from  
111 mycelium grown for one week in Fries liquid medium, and from cotyledons of Es-Astrid infected by *L. maculans* seven  
112 dpi as previously described (Fudal *et al.*, 2007).

## 113 Vector construction and fungal transformation

114 Vectors pLAU2 and pLAU53 conferring respectively hygromycin and geneticin resistance were used to perform  
115 CRISPR-Cas9 gene inactivation, as described by Idnurm *et al.* (2017). DNA fragments coding for gRNA which target  
116 genes of interest were designed using the CRISPOR prediction tool and *L. maculans* as reference genome  
117 (<http://crispor.tefor.net/>; **Table S1**; Dutreux *et al.*, 2018). The gRNA were chosen not to match on any other genes.  
118 The DNA fragment coding for gRNA was amplified using primers MAI0309 and MAI0310 and then inserted into the  
119 *XhoI* site of plasmid pLAU53 using Gibson assembly (**Table S1**; Silayeva and Barnes, 2018). Hence, plasmids Plau2-Cas9  
120 and Plau53-KMT1 and Plau53-LmPf2 were generated to inactivate *KMT1* and *LmPf2*. pLAU2 containing *Cas9* gene was  
121 used without modification.

122 Over-expression plasmids were obtained using the pBht2 vector conferring resistance to hygromycin (Mullins  
123 *et al.*, 2001). *EF1 $\alpha$*  promoter was amplified using EF1aProSac1F and EF1aProKpnIR primers and genomic DNA of JN3  
124 isolate as template, and then inserted into *KpnI-SacI* digested pBht2 to obtain pBht2-promEF1a vector. *LmPf2* gene  
125 and his terminator were then amplified using surex\_*LmPf2*\_Gibs\_F and surex\_*LmPf2*\_Gibs\_R primers and inserted in  
126 3' of the *EF1 $\alpha$*  promoter in the *HindIII* digested pBht2-promEF1a vector using Gibson assembly to obtain pBht2-  
127 surexLmPf2 (**Table S1**). JN2 and  $\Delta kmt1$  were used for the transformation.

128 The constructs were introduced into the *Agrobacterium tumefaciens* strain C58-pGV2260 by electroporation  
129 (1.5 kV, 200 ohms and 25 mF). *A. tumefaciens* mediated transformation (ATMT) of *L. maculans* was performed as  
130 previously described (Gout *et al.*, 2006). Transformants were plated on minimal medium complemented with geneticin  
131 (50 mg/l) for pLAU53-gRNA or hygromycin (50 mg/l) for pBht2-surexLmPf2 and pLAU2-Cas9 and cefotaxime (250 mg/l).  
132 For the CRISPR-Cas9 gene inactivation, construct containing *Cas9* (pLAU2-Cas9) was first introduced into a JN3 strain  
133 and transformants were selected for hygromycin resistance. Transformants transformed with pLAU2-Cas9 construct  
134 were then transformed with pLAU53-gRNA directed against *KMT1* or *LmPf2*. Transformants were selected for  
135 geneticin, hygromycin and cefotaxime resistance. For the *LmPf2* over-expression, pBht2-surexLmPf2 was introduced  
136 into JN2 and  $\Delta kmt1$ . Mutations in the targeted genes were checked in the transformants by PCR amplifying with  
137 specific primers (**Table S1**) and sequencing. Correct insertions of *surexLmPf2* construction were checked in  
138 transformants by PCR amplification.

## 139 Fungal crosses



140 Purification in order to eliminate *Cas9* gene and gRNA coding gene from the CRISPR-Cas9 mutants were  
141 performed by crossing mutants with a WT isolate (expressing or not *GFP*) from the opposite mating type (**Table S2**).  
142 Crosses were performed as described by Balesdent *et al.* (2002). Progeny was harvested and plated on V8-juice agar  
143 medium. Mycelium was collected and DNA extracted. PCR and sequencing were performed to select progeny without  
144 *Cas9* and gRNA which had integrated the CRISPR-Cas9 mutation.

#### 145 **Quantitative RT-PCR**

146 Quantitative RT-PCR (qRT-PCR) was performed using a model CFX96 Real Time System (BIORAD; Hercules, CA,  
147 USA) and Absolute SYBR Green ROX dUTP Mix (ABgene, Courtaboeuf, France) as previously described (Fudal *et al.*,  
148 2007). For each condition tested, two different RNA extractions from two different biological samples and two reverse  
149 transcriptions for each biological replicate were performed. Primers used for qRT-PCR are described in (**Table S1**). Ct  
150 values were analyzed as described by Muller *et al.* (2002) or using the  $2^{-\Delta\Delta Ct}$  method (Livak and Schmittgen, 2001).  *$\beta$ -*  
151 *tubulin* was used as a constitutively expressed reference gene. Biomass of *L. maculans* in infected plants was analyzed  
152 by generating a standard curve with known concentration of *L. maculans* and *B. napus* gDNA. Total gDNA was extracted  
153 from adult 28 dpi infected plants and a  $10^{-1}$  dilution was used for qPCR. The Ct values obtained were reported on the  
154 standard curve in order to determine fungal and plant gDNA proportion.

#### 155 **Confocal microscopy and binocular observation**

156 Cotyledons of a susceptible cultivar (Es-Astrid) infected by isolates expressing *GFP* were observed at two  
157 different time points (5 and 7 dpi) using a DM5500B Leica TCS SPE laser scanning confocal microscope and a 20x HCX  
158 Fluotar Leica objective lens. GFP was excited at 488 nm and emission was captured with a 505–530 nm broad-pass  
159 filter. The detector gain was set-up between 700 and 900. Calibrations of gain settings were performed with multiple  
160 control leaves and with a range of background fluorescence. All images represent at least 4 scans. Infected cotyledons  
161 were also observed at 8 and 13 dpi using a Leica MZ16F fluorescent binocular coupled with a Leica DFC300FX camera.

#### 162 **Western Blot**

163 Total proteins were extracted from 10-100 mg lyophilized mycelium. Mycelium was ground using beads and  
164 Mixer Mill MM 400 (Retsch, Éragny, France). Total proteins were treated as described in Petit-Houdenot *et al.* (2019).  
165 Proteins were transferred on a PVDF membrane according to the manufacturer's protocol (Trans-Blot® Turbo™ Rapid  
166 Transfer System, BIORAD, Les-Ulis, France) using a small protein transfer program. The PVDF membrane was incubated  
167 in TBS 1X containing 5% milk, 0.05% tween 20 for one hour to saturate membrane. The membrane was then incubated  
168 at 4°C over-night in TBS 1% containing 0.05% Tween 20, milk 1% and H3K9me3 antibody (1:5000; 39062 ActiveMotif,  
169 La Hulpe, Belgium Germany). PVDF membrane was washed as described by Petit-Houdenot *et al.* (2019) and was then  
170 incubated in TBS 1X + 0.05% Tween 20 + milk 1% + antiRabbit IgG II<sup>R</sup> (goat anti-rabbit, Santa Cruz Biotechnology,  
171 Heidelberg, Allemagne) and washed as previously. Finally, membrane was incubated 1min in 1 mL enzyme solution  
172 (Clarity™ Western ECL, BIORAD, Les-Ulis, France) and 1 mL Luminol/enhancer solution (Clarity™ Western ECL, BIORAD,  
173 Les-Ulis, France) and observed using ChemiDoc (ChemiDoc™ Imaging Systems, BIORAD, Les-Ulis, France).



174 **Gene annotation and domain prediction**

175 The Pf2 orthologue of *L. maculans* had been previously identified by Rybak et al. (2017). As it was identified  
176 on a previous version of the *L. maculans* genome, we have performed a new search, using the protein sequence of Pf2  
177 of *P. nodorum* and the NCBI BLAST program (Altschul *et al.*, 1990). Functional domains were identified using Pfam as  
178 described in Finn *et al.*, 2014 (<https://pfam.xfam.org/>). Alignments were performed with COBALT (Papadopoulos and  
179 Agarwala, 2007).

180 **RNA-Seq and statistical analysis**

181 Eight different transformants (i.e., mutants inactivated for *KMT1* or *LmPf2* and transformants in which *LmPf2*  
182 is over-expressed in a JN2 background (JN2\_oPf2) or a  $\Delta kmt1$  background ( $\Delta kmt1$ \_oPf, **Table 1**) were grown in Fries  
183 medium during 7 days and harvested for RNA extraction. Two RNA extraction corresponding to two biological  
184 replicates were performed for each transformant. About 150 mg of mycelium was used per extraction. Libraries were  
185 prepared from all biological replicates, individually, according to the Illumina TruSeq protocol (Illumina, San Diego, CA,  
186 USA). Libraries including polyA enrichment were performed and sequenced using 150 pb paired-end strategy on an  
187 HiSeq2000 Illumina sequencer at the Genewiz sequencing facility (Leipzig, Germany) with an input of 1 $\mu$ g total RNA.  
188 Raw sequencing data are available under the SRA accession number xxxx. Quality of the reads was checked and  
189 improved using Trimmomatic (Bolger *et al.*, 2014). The resulting reads were treated to remove adaptors and reads  
190 below 30 bp and filtered reads were mapped against the *L. maculans* genome using STAR (Dobin *et al.*, 2013; Dutreux  
191 *et al.*, 2018) with default parameters. Read alignments were stored in SAM format, and indexing, sorting, and  
192 conversion to BAM format were performed using SAMtoolsv0.1.19 (Li *et al.*, 2009). Genes with a number of read > 15  
193 in at least one condition were kept for statistical analysis. Differential expression analyses were made using R (R Core  
194 Team, 2017) and the package EdgeR (Robinson *et al.*, 2010): Genes with a log2 Fold Change  $\leq -1.5$  or  $\geq 1.5$  and an  
195 associated false discovery rate  $\leq 0.05$  were considered as differentially expressed (McCarthy *et al.*, 2012).

196 **Gene Ontology enrichment analysis**

197 Gene Ontology (GO) annotation of *L. maculans* encoding genes were retrieved from Dutreux et al., (2018).  
198 Gene ontology term enrichment analysis of the differentially expressed genes in our transformants was performed  
199 with the plug-in Biological networks Gene ontology (BinGo; v3.0.3) of the cytoscape software (Shannon *et al.*, 2003).  
200 List of genes submitted to BINGO were considered as significantly enriched for a given GO term with an associated a  
201 False Discovery Rate  $\leq 0.01$  for the biological processes. All analyses were done in R ([www.r-project.org](http://www.r-project.org)).

202

203

204

205



206

207

208

209

**Table 1: Effect of *kmt1* and *LmPf2* inactivations and of *LmPf2* over-expression on mycelial growth, conidia production, pathogenicity and effector gene expression of *Leptosphaeria maculans***

Isolate / transformants <sup>a</sup>	Radial growth		Pathogenicity assay using conidia <sup>b</sup>	Pathogenicity assay using mycelium plugs <sup>c</sup>	Effector gene expression during axenic growth <sup>d</sup>	
	9 days post inoculation on V8 medium (mm)	conidia/ml			<i>AvrLm6</i>	<i>AvrLm4-7</i>
JN2	60.5 (± 2.24)	4.50E+07	4.83 (± 0.28)	pathogenic	1	1
$\Delta kmt1$	50.25 (± 1.94)	4.00E+07	3.09 (± 0.35)	pathogenic	9.70E-02 (± 0.22)	2.37E+02 (± 27)
$\Delta LmPf2\_A$	56.13 (± 0.9)	1.53E+08	1 (± 0)	non pathogenic	4.60E-02 (± 0)	0.00E+00 (± 0)
$\Delta LmPf2\_B$	80 (± 0.8)*	1.08E+08	1 (± 0)	non pathogenic	6.30E-03 (± 0)	4.10E-03 (± 0)
JN2_oPf2_14	77.38 (± 0.76)	0.00E+00	not enough conidia	pathogenic	3.05E+00 (± 0.14)	2.50E-01 (± 0.20)
JN2_oPf2_24	73.5 (± 1.19)	0.00E+00	not enough conidia	pathogenic	2.68E+00 (± 0.19)	7.00E-02 (± 0.02)
$\Delta kmt1\_oPf2\_8$	68.38 (± 1.03)	2.00E+07	4 (± 0)	pathogenic	2.61E+01 (± 3.4)	4.73E+02 (± 8.6)
$\Delta kmt1\_oPf2\_22$	53.75 (± 1.06)	1.25E+06	not enough conidia	non pathogenic	1.30E+04 (± 107.79)	2.31E+03 (± 24.64)

a: JN2 is the WT strain;  $\Delta LmPf2\_A$  and B are *LmPf2* mutants; JN2\_oPf2\_14 and JN2\_oPf2\_24 correspond to *LmPf2* over-expression transformants (with respectively 600 fold and 150 fold expression of *LmPf2* compared to JN2);  $\Delta kmt1\_oPf2\_8$  and  $\Delta kmt1\_oPf2\_22$  correspond to *LmPf2* over-expression transformants in a  $\Delta kmt1$  mutant background (with respectively 150 fold and 600 fold expression of *LmPf2* compared to the  $\Delta kmt1$  mutant). b: Pathogenicity tests were performed on the susceptible cultivar Es-Astrid. Disease rate is expressed as the mean scoring using the IMASCORE rating scale comprising six infection classes (IC), where IC1 to IC3 correspond to various levels of resistance of the plant and IC4 to IC6 to susceptibility (Balesdent *et al.*, 2001). c: ability to provoke symptoms on cotyledons; d: Mycelium was obtained by growing JN2 strain and transformants in Fries liquid medium for 7 days. Two biological replicates per condition were generated. Total RNA was extracted. Expression of *AvrLm6* and *AvrLm4-7* was measured by qRT-PCR and expressed relatively to *Lm $\beta$ Tubulin* expression and to expression of *AvrLm6* or *AvrLm4-7* in JN2 using the the  $2^{-\Delta\Delta Ct}$  method (Livak and Schmittgen, 2001). Asterisk corresponds to transformants statistically different from the WT, Kruskal Wallis test *Pvalue*<0,05.

210

211

212



213 RESULTS

214 Identification of a *Pf2* orthologue and analysis of *KMT1* and *LmPf2* expression in *Leptosphaeria maculans*

215 In order to analyze involvement of Pf2 and KMT1 in the regulation of gene expression, notably on effector  
216 genes located in TE-rich regions and highly expressed during the early stages of oilseed rape infection, we firstly  
217 identified *Pf2* in *L. maculans*. The gene encoding Pf2 in *P. nodorum*, *PnPf2*, was identified by Rybak *et al.* (2017) and  
218 phylogenetic analysis showed that the closest orthologue of *PnPf2* was found in *L. maculans* (Genbank Accession  
219 XP\_003838593.1). This Genbank Accession corresponded to gene ID *Lmb\_jn3\_06039* of *L. maculans*, located on  
220 SuperContig6 (Dutreux *et al.*, 2018). A bidirectional Best Hit with BLASTp was performed to confirm orthology between  
221 *PnPf2* and *LmPf2* (Figure 1). The *Pf2* gene of *L. maculans* (hereinafter referred to as *LmPf2*) encodes a 658 amino acids  
222 protein sharing 67% identity with PnPf2. A Zn2Cys6-type DNA-binding domain is predicted in both proteins  
223 (IPR001138; Figure 1).

```

LmPf2 MATNPTTTTPVKRACDSCHRRKVKCIGEGTAPCKNCVSAGLACTYNAVPQKKGPKGSRKVLSELRENQRNAQLAAGYPS 80
PnPf2 MSSSSTTSAPVKRACD---RRKVKCIGEGTNPCKNCLSAGLACTYNAIPQKKGPKGSRKVLSELRENQRNAQLAAGFPP 77

LmPf2 ELGFDGRALTASFARTPGLLAPGLVESCI EYFFAHVYVSEPLLRHQRAQETAMNMDRSTESYCVIVALCAYVMIKANMKV 160
PnPf2 DVGYDGRTLSTTFARAQGLLPNGLVDTCLDFFANVYVSTPVLHRQKAQELAVNMERSTEAYCLIVSLCAYVMIHANMKV 157

LmPf2 SPTMLPRPEMAQMSNVSFHGHILLEESVVRVQGYDFRENPHLTVLTSYFYSGCYFGLGRENTAWAYLRDATTQAHILGMH 240
PnPf2 PSNMFSRPEVAQMSNMTLGHALLEESVR----- 185

LmPf2 DEDTYKHDPMDISRKRVLWLLFIAERNFALHKKRPIISLYPTIHPPTLDEASSDRQFASGLELMINMYKIIDDFTINLWN 320
PnPf2 -----

LmPf2 RVHTHANPAWIAQLQTQLAEAVPAYLDCNEAQSVEIRVTQHHLRAQAWQLCVTQGLVSSVTSDSPLTFKYPIEIIARDLLT 400
PnPf2 -----LSEAVPAYLECTEAQGV EIRITQQWLKAMAWQLCVQGLVSSVTNDNCMTFKYPIEISRDLIT 248

LmPf2 ATHQFSQQSMEVHGAGLI-----EKLFDVACCLTDVVAVTS 436
PnPf2 MTHQFSQQAMEVHGAELILRSRNPCHLDEGPPSSVFSNFLARRILGNVLLILDHLLTWLQIEKLFDIACCLADVAVTS 328

LmPf2 FSPDAFALGPRDYVSRFLTLISTLRGGQSRYLPLLLAKLSEVLPNLPPLPRSLNLPQGVSTSSIGLSGTGSPVPSNVGDD 516
PnPf2 FSPDAFALGPRDYVSRFLTLISTLRGGHSRYLPLLLAKLSEVLPNLPPLPRSLNLPQTLPASTISMSGTG--TVPSNITDD 406

LmPf2 FSAMGGQTSPSYPSNDLIRRLAAQTG---AQLPFNPQHSSYPAPTS-----HVEDLSLYDTSHTTHSSGSVP 582
PnPf2 YSAMPATSSPSYPSSELIRRLAAQTGTFESRRRSVFPVRYRTQPQRIAFFLICAKEQLHYTRTIRVYHVTTELPNSLSPVR 486

LmPf2 RSTSATPG--PYEPSMSQSRQIVGHSTMQIPTSHPQHNMQSHHVEESAGTAVGAIGWV LHKGNVEGCKLPDSPAQ 658
PnPf2 ADSNLTPTATQPYAIAPHFSRPDCL----- 510

```

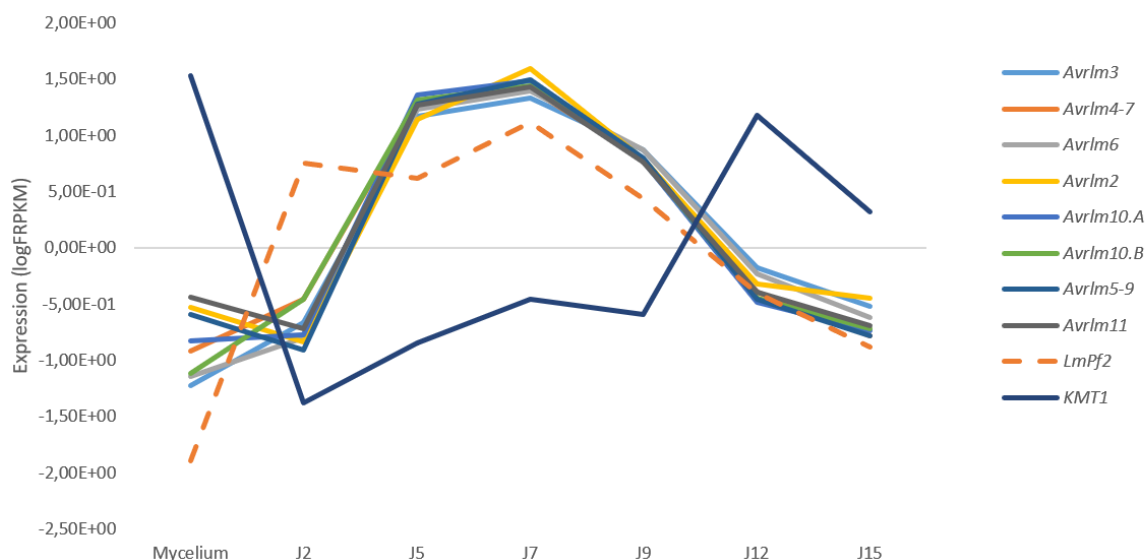
224 **Figure 1: Identification of a PnPf2 orthologue in *Leptosphaeria maculans*.** Alignment between LmPf2 of *L. maculans* and *Parastagonospora nodorum* PnPf2 was performed using COBALT (Papadopoulos and Agarwala, 2007). The black bar corresponds to the Zn2Cys6-type DNA binding domain identified using Pfam as described in Finn *et al.*, 2014 (<https://pfam.xfam.org/>). LmPf2 and PnPf2 share 67 % identity.



225 We then investigated the expression profile of *Lmpf2* and *KMT1* during axenic growth and at different stages  
226 of primary infection of oilseed rape and compared their expression profiles to that of the avirulence genes of *L.*  
227 *maculans* (*AvrLm2*, *AvrLm3*, *AvrLm4-7*, *AvrLm5-9*, *AvrLm6*, *AvrLm10.A*, *AvrLm10.B* and *AvrLm11*; Fudal *et al.*, 2007;  
228 Parlange *et al.*, 2009; Balesdent *et al.*, 2013; van de Wouw *et al.*, 2014; Plissonneau *et al.*, 2016; Ghanbarnia *et al.*,  
229 2018; Petit-Houdenot *et al.*, 2019) (**Figure 2**). As already described, expression of the avirulence genes of *L. maculans*  
230 was repressed during axenic. Expression strongly increased during the asymptomatic infection of oilseed rape  
231 cotyledons, between two and nine dpi, reaching a peak seven dpi and then, expression slowly decreased up to 12-15  
232 dpi, when necrotic symptoms were visible (**Figure 2**). Likewise, *Lmpf2* was not expressed *in vitro* while its expression  
233 increased between 5 and 7 dpi and decreased again until 15 dpi. *KMT1* was inversely expressed with high expression  
234 during axenic growth, no expression during early infection (J2 to J9), and was up-regulated at J12 when *AvrLm* gene  
235 expression decreased. To summarize, *Lmpf2* showed a similar expression profile as *AvrLm* genes described so far in *L.*  
236 *maculans*, and *KMT1* was inversely correlated with *AvrLm* genes expression (**Figure 2**).

237

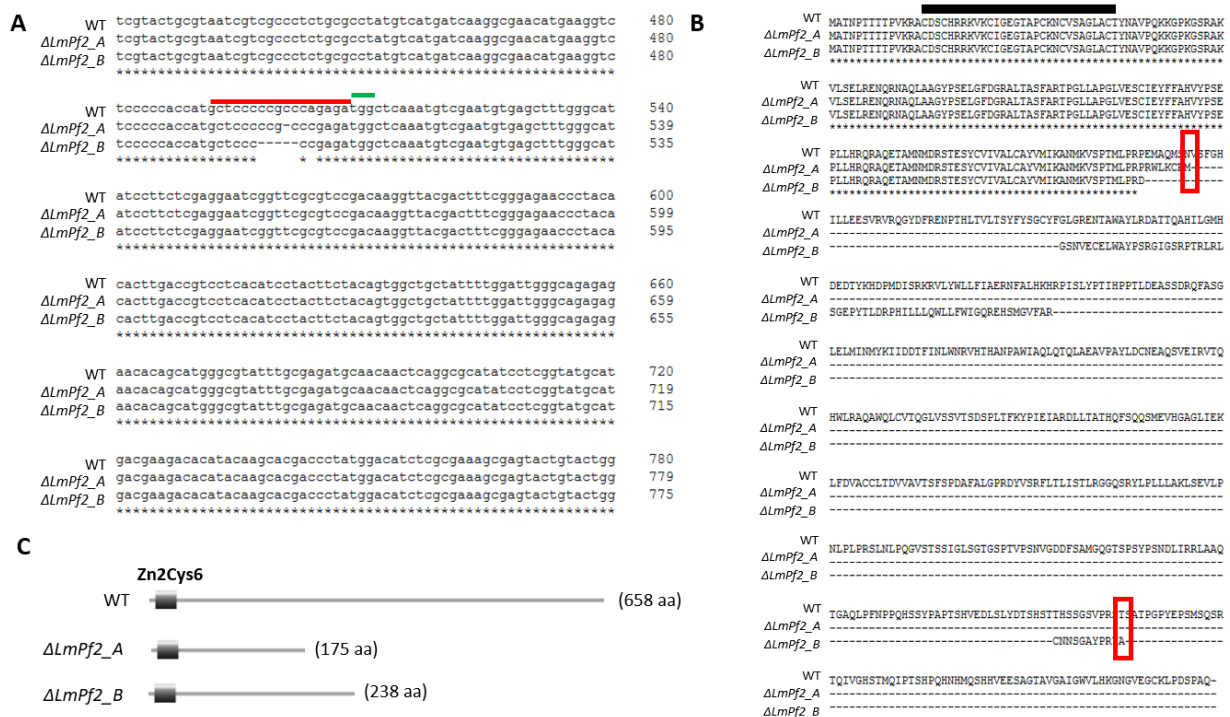
238



**Figure 2: Expression profile of *Leptosphaeria maculans* avirulence genes, *Lmpf2* transcription factor and *KMT1* during axenic growth and infection of oilseed rape cotyledons.** Mycelium was obtained by growing JN2 strain on V8 agar medium for 7 days. Cotyledons of the susceptible cultivar Darmor-bzh were inoculated by JN2 and sampled at 7 different time points (2, 5, 7, 9, 12 and 15 days post infection). Two biological replicates per condition were generated. Total RNA was extracted and sequenced. The expression level of nine avirulence genes, *Lmpf2* and *Kmt1* in RNA-seq data is represented by the log<sub>2</sub> of FPKM (Reads Per Kilobase Per Million mapped reads) centered and reduced (Gay *et al.*, in prep.).

239 ***Lmpf2* and *KMT1* inactivations do not induce morphology, conidiation or pigmentation defects while *Lmpf2* over-**  
240 **expression induces developmental defects in *L. maculans***

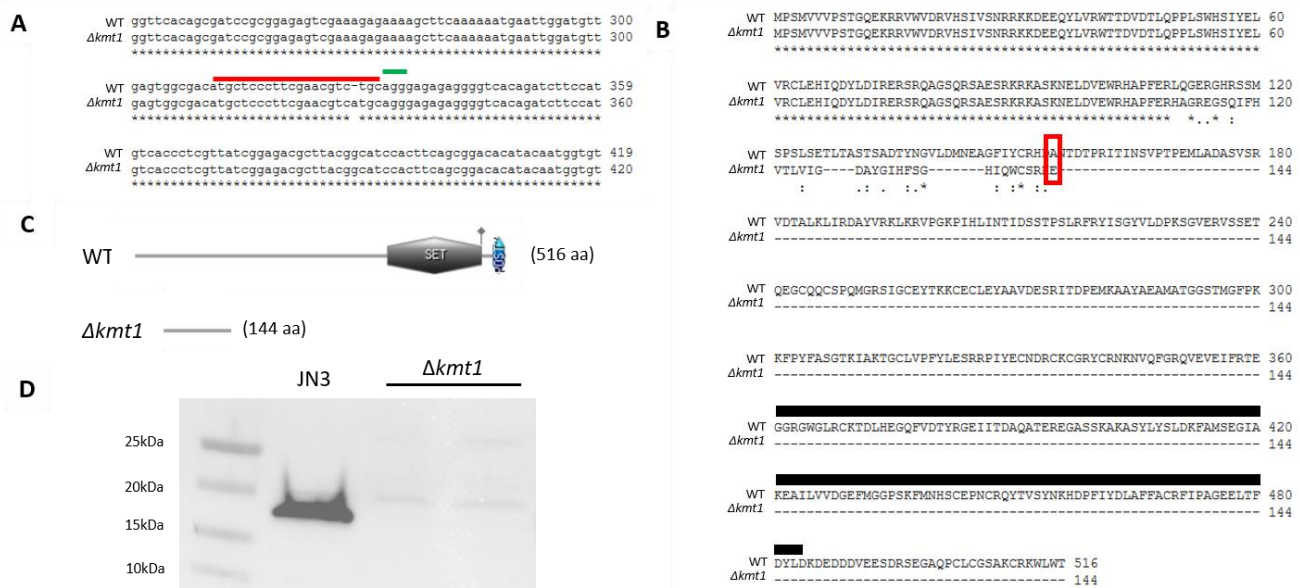
241 We inactivated *Lmpf2* in order to evaluate its involvement in the regulation of *L. maculans* avirulence genes.  
242 Inactivation of *Lmpf2* was performed using the CRISPR-Cas9 strategy as detailed in the Materials and Methods section.  
243 Seventeen transformants resistant to hygromycin and geneticin were obtained and sequenced for the *Lmpf2* gene. Of  
244 the 17 transformants, 12 had no mutations in *Lmpf2* compared to the WT, three had a 1-bp and one had a 5-bp deletion  
245 near the cleavage site (**Figure 3A**). These mutations resulted, at the protein level, in frame-shifts leading to two  
246 different truncated proteins: 175 amino-acids and 238 amino-acids respectively for the 1-bp and the 5-bp deletions  
247 compared to the 658 amino-acids length of the WT protein (**Figure 3B**). The two mutants were crossed with JN2-GFP  
248 in order to obtain purified  $\Delta Lmpf2$  and  $\Delta Lmpf2$ -GFP mutants without *Cas9* and CRISPR guide RNA (**Table S2**). The four  
249 corresponding mutants are hereinafter referred to as  $\Delta Lmpf2\_A$ ,  $\Delta Lmpf2\_A$ -GFP,  $\Delta Lmpf2\_B$  and  $\Delta Lmpf2\_B$ -GFP (A  
250 corresponding to the 1-bp and B to the 5-bp deletion). The four  $\Delta Lmpf2$  mutants were able to produce conidia and  
251 showed dark coloration of the hyphae, corresponding to melanin production, as the WT after 14 days of growth on  
252 V8-juice agar medium (**Figure S1**; **Table 1**). After nine days of growth on V8-agar plate, only  $\Delta Lmpf2\_B$  exhibited  
253 statistically more growth than the WT indicating that *Lmpf2* deletion does not lead to growth defect. To summarize,  
254 no major defect in conidia production, growth rate or morphology was associated with the inactivation of *Lmpf2*.



**Figure 3: Effect of the *Lmpf2* mutations on *Lmpf2* protein structure. A:** Sequence alignment of the *Lmpf2* gene in the WT isolates JN2 and in two mutants  $\Delta Lmpf2\_A$  and B showing respectively a 1bp and 5bp deletion. The PAM (Protospacer Adjacent Motif) is highlighted in green and the region targeted by the guide RNA is highlighted in red. **B:** Protein sequence of *Lmpf2* in the WT isolate JN2 and in the two mutants. The DNA-binding Domain is indicated by a black bar. The red frames indicate the location of the stop codons in the mutant versions of the protein. **C:** *Lmpf2* protein length and domains identified with Pfam as described in Finn *et al.* (2014)(<https://pfam.xfam.org/>).



255 In a previous study, Soyer *et al.* (2014) silenced expression of *KMT1* (with a residual expression of 16%  
256 compared to the WT). Silencing of *KMT1* led to an over-expression of effector genes located in TE-rich regions, notably  
257 avirulence genes, during axenic growth. This over-expression was associated, at least for two avirulence genes, with a  
258 decrease of H3K9me3 at their loci. Here, we took advantage of the availability of the CRISPR-Cas9 strategy to better  
259 investigate involvement of *KMT1* in the regulation of *L. maculans* gene expression. Twenty-five transformants resistant  
260 to hygromycin and geneticin were obtained and sequenced for the *KMT1* gene. Of the 25 transformants, 24 had no  
261 mutations in *KMT1* compared to the WT and one had a 1-bp insertion resulting, at the protein level, in a truncated  
262 protein of 144 aa (while the WT protein had a length of 516 amino-acids; **Figure 4A and B**). Both functional domains  
263 of *KMT1* (Pre-SET and SET) are absent from the truncated protein (Pfam analysis; Finn *et al.*, 2014;  
264 <https://pfam.xfam.org/>; **Figure 4C**) and loss of H3K9me3 was confirmed in the  $\Delta kmt1$  mutant strain by Western blot  
265 analysis (**Figure 4D**; **Figure S2**). The  $\Delta kmt1$  mutant was crossed with JN2-GFP in order to obtain purified  $\Delta kmt1$  and  
266  $\Delta kmt1$ -GFP mutants without *Cas9* and CRISPR guide RNA (**Table S2**). The two corresponding mutants are hereinafter  
267 referred to as  $\Delta kmt1$  and  $\Delta kmt1$ -GFP mutants. We tested the effect of *KMT1* inactivation on growth and conidiation  
268 of *L. maculans*.  $\Delta kmt1$  was not significantly altered in its mycelial growth, conidia production and production of  
269 melanin (**Table 1**; **Figure S1**).



**Figure 4: Effect of the  $\Delta kmt1$  mutation on the *KMT1* protein structure and function. A:** Nucleotide sequence alignment of the *KMT1* gene in the WT isolate and in the  $\Delta kmt1$  mutant (showing a 1bp insertion). The PAM (Protospacer Adjacent Motif) is highlighted in green and the region targeted by the guide RNA is highlighted in red. **B:** Protein sequence of *KMT1* in the WT isolate JN2 and in the  $\Delta kmt1$  mutant. The red frame indicates the stop in the mutant version of the protein. The black bar indicates the SET domain which has been lost in the truncated protein. **C:** *KMT1* protein length and domains identified with Pfam as described in Finn *et al.* (2014) (<https://pfam.xfam.org/>). **D:** Western Blot analysis of H3K9 tri-methylation in JN3 (WT) and  $\Delta kmt1$ . Nuclear proteins were extracted from mycelium, separated by SDS-PAGE, blotted and probed with antibodies against H3K9me3.  $\Delta kmt1$  proteins replicates have been used for this test.

270



271 We then expressed *LmPf2* under the control of the *EF1 $\alpha$*  promoter. The construct was introduced in two  
272 different genetic backgrounds: the WT isolate JN2 and the  $\Delta kmt1$  mutant. Twenty-five transformants were recovered  
273 for each transformation and named respectively JN2\_oPf2 and  $\Delta kmt1\_oPf2$ . We measured expression of *LmPf2* during  
274 axenic growth by RT-qPCR. The JN2\_oPf2 transformants showed a 15 to 18,000-fold increase of expression of *LmPf2*  
275 compared to the WT while *LmPf2* expression increased 15 to 700-fold in the  $\Delta kmt1\_oPf2$  transformants (**Figure S3**).  
276 For further analyses, we selected two transformants from each genetic background with similar *LmPf2* expression  
277 levels (150 and 600 more expression than in the JN2 or the  $\Delta kmt1$  strains; **Figure S3**): JN2\_oPf2\_14, JN2\_oPf2\_24,  
278  $\Delta kmt1\_oPf2\_8$  and  $\Delta kmt1\_oPf2\_22$ . Over-expression of the *LmPf2* gene in the WT strain or the  $\Delta kmt1$  strains did not  
279 induce any growth defect even though the thallus was denser and harbored a white coloration indicating that the  
280 mutants were not able to produce melanin. In addition, over-expression of *LmPf2* had critical impact on *L. maculans*  
281 conidiation (**Table 1**; **Figure S1**).

### 282 **LmPf2 and KMT1 are involved in the control of effector gene expression and pathogenicity of *L. maculans***

283 We inoculated  $\Delta LmPf2$ ,  $\Delta LmPf2$ -GFP,  $\Delta kmt1$ ,  $\Delta kmt1$ -GFP, *LmPf2*-overexpressing transformants and a WT  
284 strain on cotyledons of the susceptible cultivar Es-Astrid.  $\Delta Kmt1$  and  $\Delta Kmt1$ -GFP showed reduced symptoms compared  
285 to the WT, indicating a decrease of pathogenicity (**Figure 5A**; **Figure S4**; **Table 1**). Cotyledons infected with  $\Delta Kmt1$ -GFP  
286 were observed from 5 to 13 dpi, which allowed us to distinguish living plant cells and GFP fungal hyphae, and  
287 production of pycnidia. The  $\Delta kmt1$ -GFP transformant was able to colonize the plant and to produce pycnidia but  
288 induced lesser symptoms than the WT (**Figure 5C**; **Figure S4**). The  $\Delta LmPf2$  mutants were not able to invade the  
289 cotyledon further than the inoculation site, and consequently, did not induce any visible symptom (**Figure 5**; **Figure**  
290 **S4**; **Table 1**).  $\Delta LmPf2$  mutants started producing pycnidia at 5 DPI at the inoculation site whereas the WT started  
291 producing pycnidia at 7 DPI (**Figure 5A, B**). These results indicate that both KMT1 and *LmPf2* are involved in infection  
292 establishment. Finally, the transformants over-expressing *LmPf2* were not able to produce enough conidia, which are  
293 normally used to perform pathogenicity tests. We thus performed pathogenicity tests using plugs of mycelium and all  
294 the transformants, excepted  $\Delta kmt1\_oPf2\_22$ , were able to produce symptoms on oilseed rape cotyledons (**Table 1**).



295 We investigated expression of two avirulence genes (*AvrLm4-7* and *AvrLm6*) in the  $\Delta Kmt1$  mutant, the  $\Delta LmPf2$   
296 mutant, transformants over-expressing *LmPf2* and the WT strain during axenic growth by qRT-PCR (**Figure 6; Table 1**).  
297 While expression of *AvrLm6* was comparable to that of the WT, *AvrLm4-7* was over-expressed in  $\Delta Kmt1$  compared to  
298 the WT strain (237 times more expressed; **Table 1**). *AvrLm4-7* and *AvrLm6* expression was fully abolished in both  
299  $\Delta LmPf2$  mutant (**Table 1**). In both genetic backgrounds of *LmPf2* over-expression, we observed an up-regulation of  
300 these genes, with a higher level of up-regulation in the  $\Delta kmt1$  background. *AvrLm4-7* expression level reach up to 172-  
301 fold change in JN2\_oPf2 mutants while *AvrLm6* expression level reach up to 1,000-fold change in the JN2\_oPf2 mutants  
302 and up to 20,000-fold change in the  $\Delta kmt1$ \_oPf2 mutants; **Figure 6**). To investigate whether a stronger over-expression  
303 of *LmPf2* in a WT background would result in a higher increase of expression of the two avirulence genes, we also  
304 investigated expression of these avirulence genes in two JN2\_oPf2 transformants in which *LmPf2* was over-expressed  
305 10,000 and 18,000 times compared to the WT strain (**Figure S5**). The strong *LmPf2* over-expression induced an up-  
306 regulation of *AvrLm4-7* and *AvrLm6* expression. Nevertheless, it did not reach the same level as in a  $\Delta kmt1$  background  
307 (*AvrLm4-7* was induced 50-1,100-fold and *AvrLm6* 700-4,700-fold compared to the WT). We then investigated  
308 expression of four avirulence genes (*AvrLm4-7*, *AvrLm6*, *AvrLm10.A* and *AvrLm11*) in the  $\Delta LmPf2$  and  $\Delta kmt1$  mutants  
309 and the WT strain during infection of the susceptible cultivar Es-Astrid at 7 dpi (when expression of *L. maculans*  
310 effectors reaches a peak in the WT strain). Expression of the four avirulence genes was very low in  $\Delta LmPf2$  compared  
311 to the WT (1000 to 2000-fold less expressed than in the WT) or even not detected (**Table 2**). In  $\Delta kmt1$ , avirulence gene  
312 expression was not altered *in planta* (**Table 2**). Altogether, our results suggest that KMT1 is able to repress avirulence  
313 gene expression whereas *LmPf2* induces their expression.

314

**Table 2: Effect of *LmPf2* and *KMT1* inactivation on avirulence gene expression during primary oilseed rape infection**

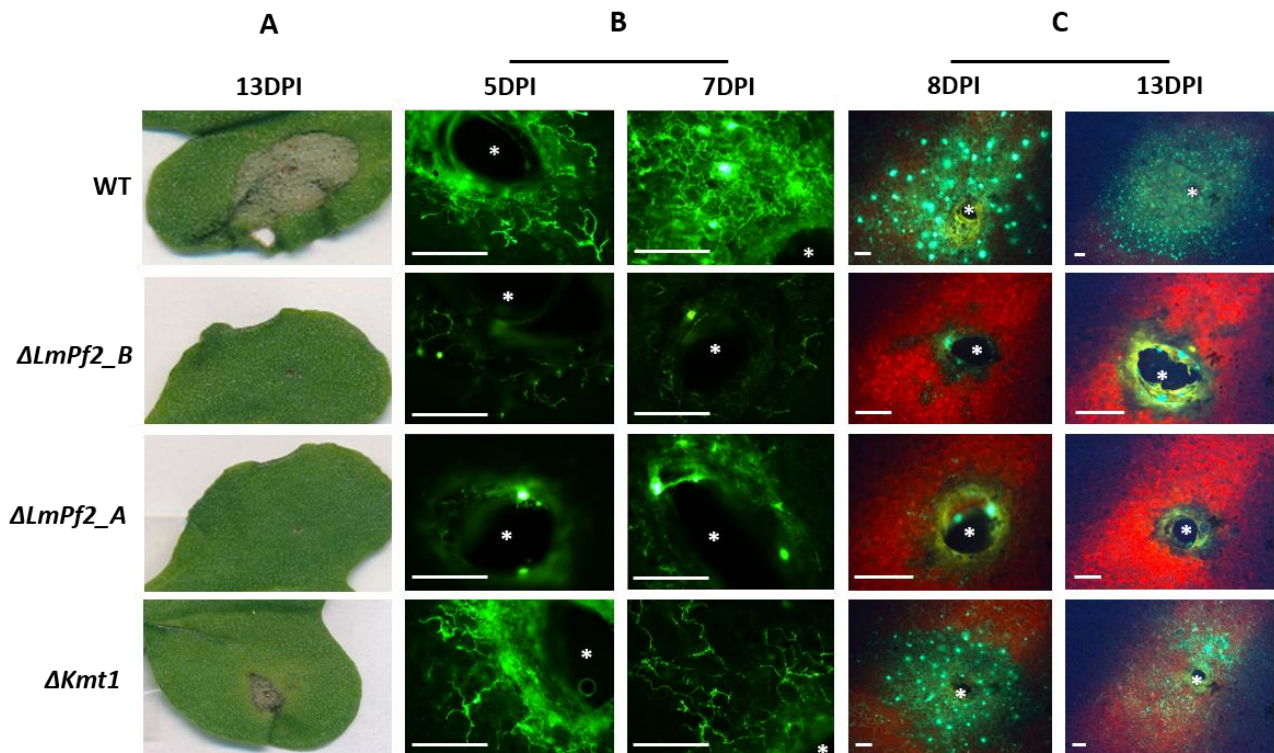
Isolate/ transformants	expression <i>in planta</i> 7DPI <sup>a</sup>			
	<i>AvrLm4-7</i>	<i>AvrLm6</i>	<i>AvrLm10A</i>	<i>AvrLm11</i>
WT	3.25E+01	7.41E+01	2.92E+01	1.96E+02
$\Delta LmPf2\_A$	nd <sup>b</sup>	4.58E-02	2.49E-02	7.64E-02
$\Delta LmPf2\_B$	4.08E-03	6.27E-03	3.25E-02	9.55E-02
$\Delta kmt1$	1,03E+00	4,23E+00	1,93E+00	5,23E+00

Expression of *AvrLm4-7*, *AvrLm6*, *AvrLm10.A* and *AvrLm11* at 7dpi on cotyledons of a susceptible cultivar (Es-Astrid) in  $\Delta LmPf2$ ,  $\Delta kmt1$  mutants and a WT isolate. Legends: <sup>a</sup>Gene expression levels are relative to *Lmβtubulin* and calculated as described by Muller *et al.* (2002). Each value is the average of two biological replicates (two extractions from different biological replicates) and two technical replicates (two RT-PCR). <sup>b</sup>nd:



315 **KMT1 and LmPf2 differentially control effector genes expression**

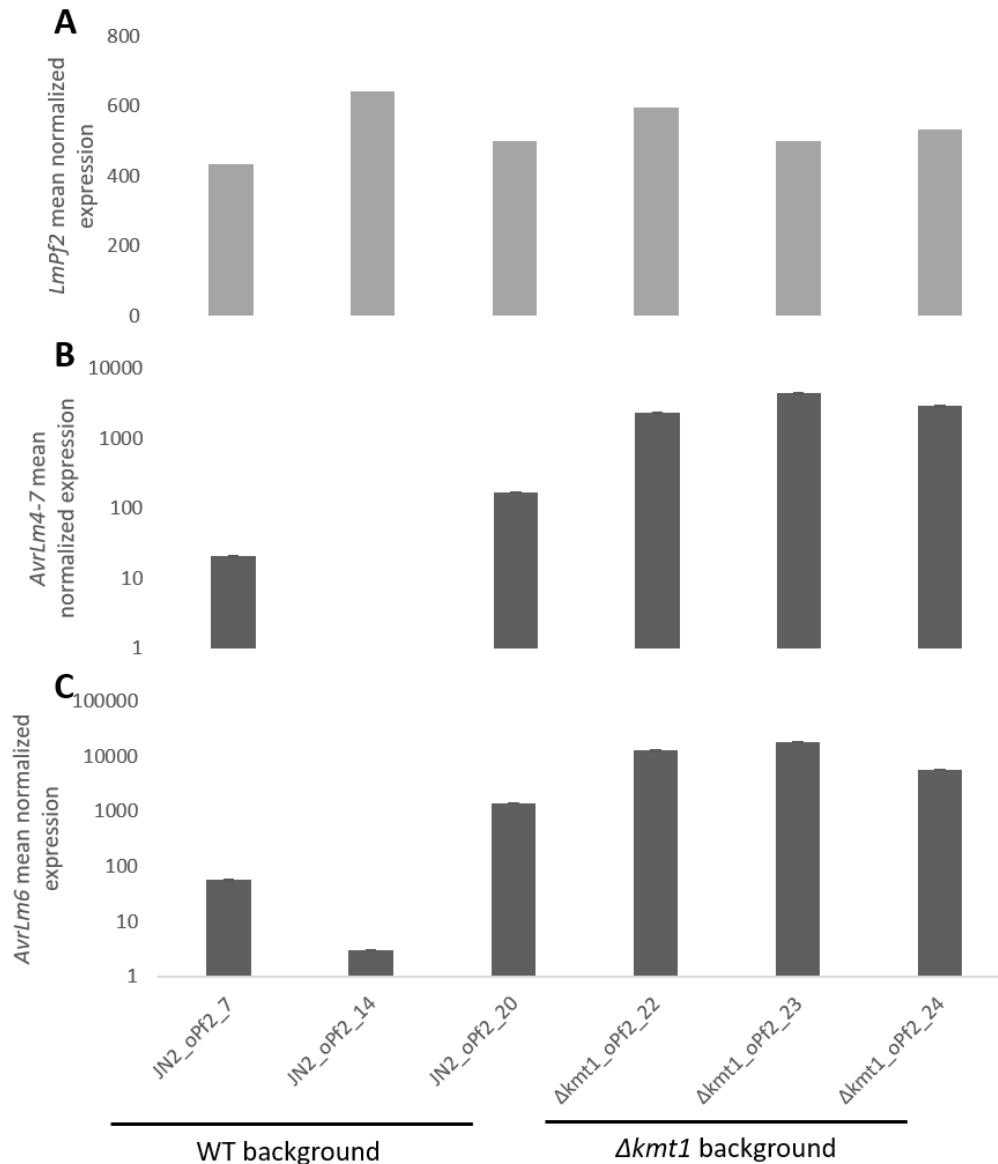
316 Inactivation of *LmPf2* abolished expression of the avirulence genes tested *in planta* while the up-regulation of  
317 *LmPf2* strongly induced expression of these genes *in vitro*, and the effect was stronger in a  $\Delta kmt1$  background than in  
318 a WT background. Based on these observations, we generated RNA-seq data during axenic growth of all types of  
319 transformants generated, and of the WT strain (i.e., JN2,  $\Delta kmt1$ ,  $\Delta LmPf2\_A$ ,  $\Delta LmPf2\_B$ , JN2\_0Pf2\_14, JN2\_0Pf2\_24,  
320  $\Delta kmt1\_oPf2\_8$  and  $\Delta kmt1\_oPf2\_22$ ; **Table 1**). Eighteen million reads were obtained, on average, for each sample, that  
321 were filtered based on their quality using Trimmomatic and mapped on the *L. maculans* genome. According to the  
322 sample considered, between 79.5% and 88% paired-end reads were uniquely mapped, except for one technical  
323 replicate of the  $\Delta LmPf2\_A$  mutant for which only 39% of paired-end reads were mapped (**Table S3**). We then plotted  
324 the log<sub>2</sub>(RPKM) of each sample and observed that biological replicates of the same condition were correlated, except  
325 for  $\Delta kmt1\_oPf2\_8$  and  $\Delta kmt1\_oPf2\_22$  transformants (**Figure 7**; **Figure S6**). Based on that observation, data from  
326 biological replicates were merged, except that of  $\Delta kmt1\_oPf2\_8$  and  $\Delta kmt1\_oPf2\_22$  that were considered separately  
327 for subsequent statistical analyses.



**Figure 5: Effect of *LmPf2* and *LmKmt1* inactivations on pathogenicity of *L. maculans*.** The WT isolate JN3 and mutants inactivated for *LmPf2* or *kmt1* were inoculated on cotyledons of a susceptible cultivar of oilseed rape (Es-Astrid).  $\Delta LmPf2\_A$  has a 1 bp deletion leading to a truncated protein of 175 aa.  $\Delta LmPf2\_B$  has a 5 bp deletion leading to a truncated protein of 238 aa.  $\Delta Kmt1$  has a 1 bp insertion leading to a truncated protein of 144 aa. **A:** Pictures of symptoms at 13 DPI. **B:** Fluorescence images of cotyledons infected by GFP mutants and WT strains at 5 and 7 DPI using a confocal microscope. **C:** Fluorescence images of cotyledons infected by GFP mutants and WT strains at 8 and 13 DPI using a fluorescence binocular. Infection site (needle hole) is indicated by a white asterisk. Scale bars represent 1 mm. DPI = Days Post Inoculation.

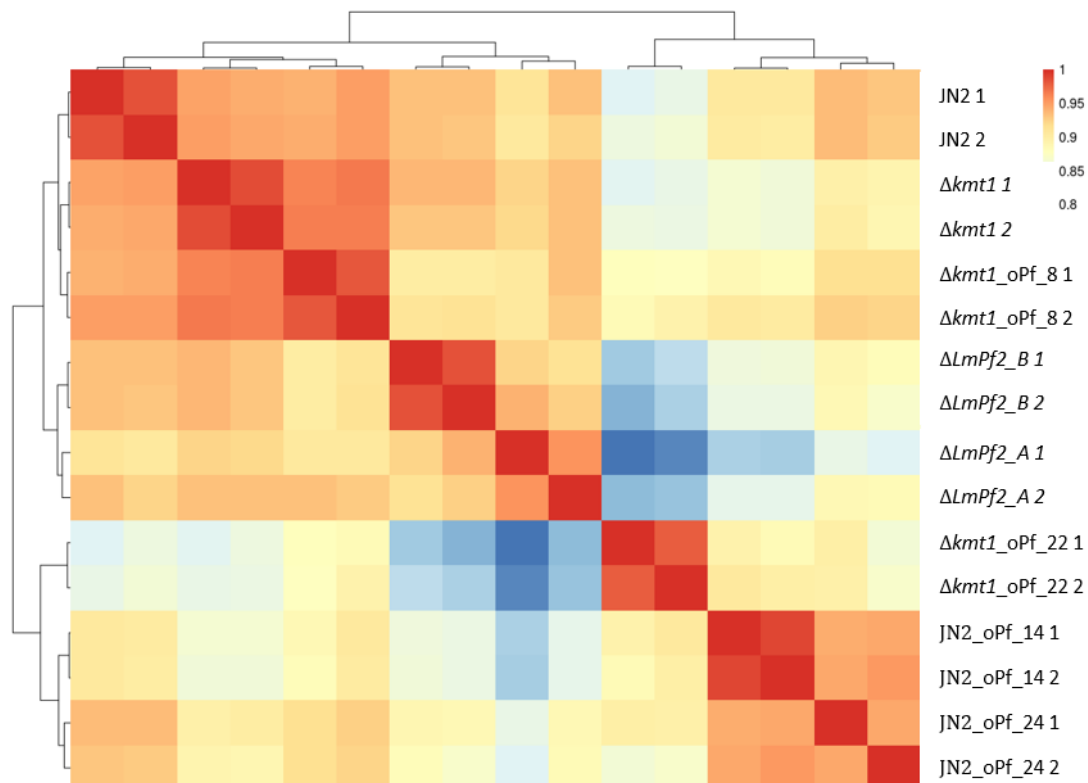
328

329



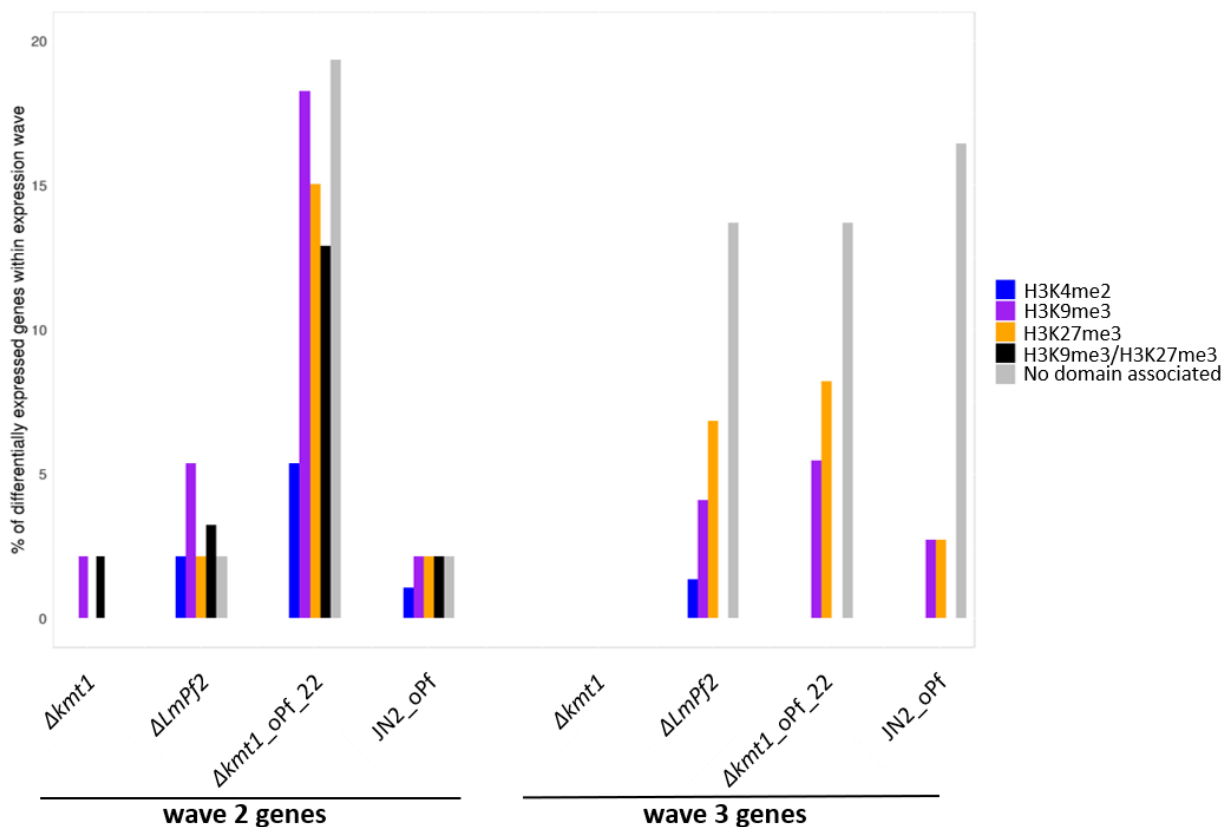
**Figure 6: Influence of *LmPf2* overexpression and *KMT1* inactivation on the expression of two *L. maculans* avirulence genes during axenic growth.** Mycelium was grown in Fries liquid medium during seven days. Expression of *LmPf2* (A), *AvrLm4-7* (B) and *AvrLm6* (C) was measured by qRT-PCR level using *Lm $\beta$ Tubulin* as a constitutive reporter gene and expression in JN2 or  $\Delta kmt1$  as reference ( $2^{-\Delta\Delta Ct}$  method). Each value is the average of two biological replicates (two extractions from different biological replicates) and two technical replicates (two RT-PCR).

332 We identified respectively 272, 427, 524, 433 and 1792 genes differentially expressed in  $\Delta kmt1$ ,  $\Delta LmPf2$ ,  
333  $JN2\_oPf2$ ,  $\Delta kmt1\_oPf\_8$ ,  $\Delta kmt1\_oPf\_22$ . We observed respectively 129, 247, 275, 931 genes up-regulated in  $\Delta kmt1$ ,  
334  $JN2\_oPf2$ ,  $\Delta kmt1\_oPf\_8$ ,  $\Delta kmt1\_oPf\_22$  and 129 genes down-regulated in  $\Delta LmPf2$ . In  $JN2\_oPf2$  we observed 247 up-  
335 regulated genes which were statistically enriched in H3K27me3 and H3K9me3/H3K27me3 genomic compartments,  
336 and statistically involved in carbohydrate metabolic process (19 genes;  $Pvalue=4.94E^{-4}$ ) and polysaccharide metabolic  
337 process (6 genes;  $Pvalue=2.27E^{-2}$ ; **Table S4**). Interestingly down-regulated genes in  $\Delta LmPf2$  were statistically enriched  
338 in H3K9me3, H3K9me3/H3K27me3 and SSP-encoding genes and GOs analysis revealed that 33 of them were involved  
339 in oxidation-reduction process ( $Pvalue=8.86E^{-4}$ ). Finally,  $\Delta kmt1$ ,  $\Delta kmt1\_oPf\_8$  and  $\Delta kmt1\_oPf\_22$  up-regulated genes  
340 were specifically located in H3K9me3, H3K27me3 and H3K9me3/H3K27me3 domains and were statistically enriched  
341 in SSP-encoding genes indicating that H3K9me3 and  $LmPf2$  are able to control expression of facultative and  
342 constitutive heterochromatinian located genes and specifically SSP coding genes (**Table S4**). Up-regulated genes in  
343  $\Delta kmt1\_oPf2\_8$  were involved in starch, sucrose, disaccharide, glucan, cellular glucan, cellular polysaccharide and  
344 oligosacchide metabolic processes with 5 genes involved in each process. Finally, Up-regulated genes in  $\Delta kmt1\_oPf2\_22$   
345 were involved in carbohydrate metabolic processes nucleic acid phosphodiester bond hydrolysis, cell wall modification  
346 and transmembrane transport. Altogether these results indicate that  $KMT1$  inactivation and  $LmPf2$  over-expression in  
347 two different background lead to the same metabolic process enrichment and that the more  $LmPf2$  is expressed the  
348 more we get differentially regulated metabolic processes.



**Figure 7: Hierarchical clustering of the pairwise Pearson correlation between the 18 RNA-Seq conditions.** Heatmap of  $\log_2(\text{RPKM})$  of each RNA-Seq biological and technical replicate plotted using Pearson correlation. JN2 is the WT strain;  $\Delta LmPf2\_A$  and B are  $LmPf2$  mutants; JN2\_oPf2\_14 and JN2\_oPf2\_24 correspond to  $LmPf2$  over-expression transformants (with respectively 600 fold and 150 fold expression of  $LmPf2$  compared to JN2);  $\Delta kmt1\_oPf2\_8$  and  $\Delta kmt1\_oPf2\_22$  correspond to  $LmPf2$  over-expression transformants in a  $\Delta kmt1$  mutant background (with respectively 150 fold and 600 fold expression of  $LmPf2$  compared to the  $\Delta kmt1$  mutant).

349 Recent transcriptomic analyses throughout the life cycle of *L. maculans* have highlighted that synchronous  
350 waves of genes, including effector genes, are expressed (Gay et al., in prep.). Among these waves, four were specifically  
351 expressed during oilseed rape infection. Here, we investigated two of the four waves of genes concertedly expressed  
352 during infection and determined whether they were up-regulated in  $\Delta kmt1$ ,  $\Delta kmt1\_oPf2\_22$  or JN2\_oPf2 or down-  
353 regulated in  $\Delta LmPf2$  to investigate the respective effect of H3K9me3 removal and LmPf2 involvement in gene  
354 regulation. Respectively 4, 9 and 14 genes of wave 2 were differentially regulated in  $\Delta kmt1$ , JN2\_oPf2 and  $\Delta LmPf2$   
355 indicating that H3K9me3 removal or *LmPf2* over-expression alone lead to limited wave 2 effector genes up-regulation.  
356 In contrast, sixty-six genes from wave 2 (71%) were specifically up-regulated in  $\Delta kmt1\_oPf2\_22$  on 931 genes up-  
357 regulated in total representing a 9.94-time enrichment in wave 2 effector genes ( $Khi^2$ ;  $pValue=2.20E^{-16}$ ). Interestingly,  
358 half of these genes were located in H3K27me3 or in no particular compartment indicating that H3K9me3 removal and  
359 *LmPf2* over-expression do not only target H3K9me3 enriched genes. We did not identify any genes of wave 3 up-  
360 regulated in  $\Delta kmt1$  indicating that H3K9me3 do not actively repress these genes. Surprisingly, we identified 16, 19 and  
361 20 of wave 3 genes respectively up-regulated in JN2\_oPf2,  $\Delta LmPf2$  and  $\Delta kmt1\_oPf2\_22$  confirming the limited impact  
362 of H3K9me3 removal on this set of genes. Furthermore, the 16 and 19 genes specifically up-regulated in JN2\_oPf2 and  
363  $\Delta LmPf2$  transformants represented respectively an 11.6-time enrichment and an 15-time enrichment ( $Khi^2$ ;  
364  $pValue=2.20E^{-16}$ ) whereas the 20 genes of wave 3 up-regulated in  $\Delta kmt1\_oPf2\_22$  represented a 3.8-time enrichment

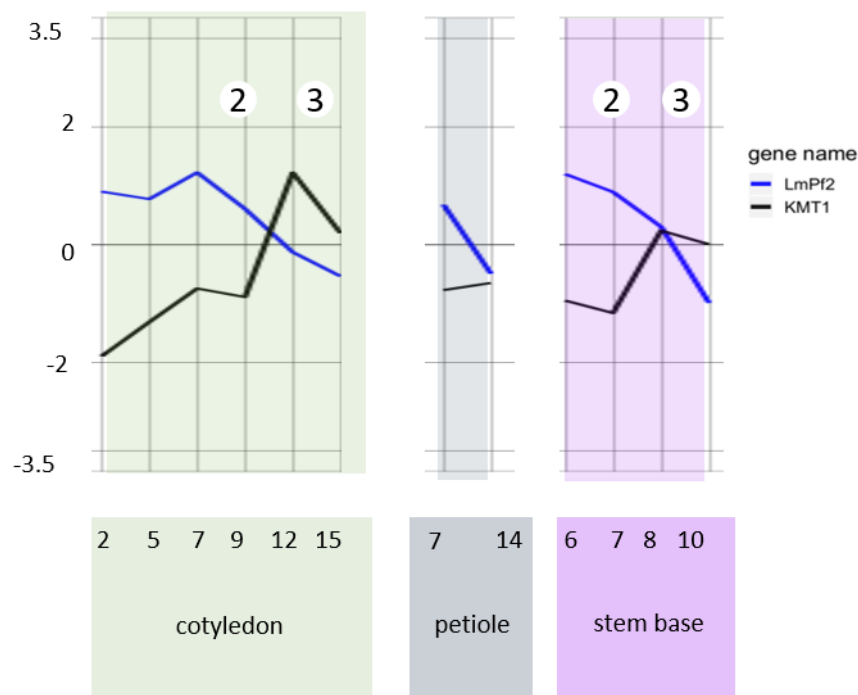


**Figure 8: Influence of KMT1 and LmPf2 on expression of genes specifically expressed during infection.** RNA-Seq analysis of differentially up-regulated genes in  $\Delta kmt1$ ,  $\Delta kmt1\_oPf2\_22$  and JN2\_oPf2 transformants and down-regulated in  $\Delta LmPf2$  depending on their chromatinian environnement and compared to genes concertedly expressed as wave 2 and wave 3 during *Leptosphaeria maculans* life-cycle. Histograms representing % of genes belonging to each wave of concertedly expressed genes in each transformant depending on their chromatin environment. On the left, genes concertedly expressed at 7 DPI during early infection on cotyledons of susceptible cultivar Darmor (including avirulence genes; wave 2). On the right, genes concertedly expressed between 9 and 12 DPI (wave 3) during early infection on cotyledons of a susceptible cultivar Darmor-bzh. DPI = Days Post Inoculation.

365 ( $Khi^2$ ;  $pValue= 3,68E^{-10}$  ; **Figure 8**). Altogether these results indicate that genes of wave 2 are under H3K9me3 and  
366 LmPf2 double control whereas genes of wave 3 are under LmPf2 control only.

367 In order to get insight on how concerted effector gene expression could be controlled by H3K9me3 and LmPf2  
368 in *planta*, we plotted the expression of *KMT1* and *LmPf2* during different stages of *L. maculans* life cycle. We observed  
369 that at least at three time points (Cotyledons 9-12DPI; Petioles 7-14DPI; and Stems between May and October) *KMT1*  
370 and *LmPf2* are conversely expressed indicating an inverted control exerted by these two genes (**Figure 9**). Interestingly,  
371 each time the *LmPf2* and *KMT1* expression was inverted, we observed a switch between sets of concertedly expressed  
372 genes.

373 Finally, we investigated whether TF-encoding genes were up-regulated in case of *LmPf2* over-expression.  
374 Surprisingly, we identified 21 TF-encoding genes up-regulated in  $\Delta kmt1\_oPf2\_22$  and 6 in JN2\_oPf2 with only 2 of  
375 them in common (**Table S5**). We conclude that LmPf2 could act as a master regulator and activate different sets of TF  
376 depending on the chromatin context.



**Figure 9: Inverted regulation of *KMT1* and *LmPf2* during *Leptosphaeria maculans* life cycle.** Expression obtained by RNA-Seq in different *in planta* conditions. Expression corresponding to scaled  $\log_2(\text{FPKM})$ . The black line corresponds to *KMT1* expression and the blue line to *LmPf2* expression. Numbers at the top corresponding to concerted effector waves expressed depending on the infection time. Green part of the graph corresponding to infection of Darmor-bzh cotyledons between 2 and 15DPI. Grey part of the graph corresponding to petioles infected by the JN2 in controlled conditions at two time points 7 and 14DPI. The purple part of the graph corresponding to stem harvested in the field between June (6) and October (10).

377

## 378 DISCUSSION

379 Here we demonstrated that LmPf2, a Zn2Cys6 TF, is involved in the establishment of infection and in the  
380 control of effector gene expression in *L. maculans*. We also determined that this TF is able to differentially regulate  
381 two sets of effector genes, expressed at different time points, depending on the chromatinin condensation state  
382 regulated by H3K9me3 deposition via KMT1. Furthermore, *LmPf2* presents an inverted expression profile compared  
383 to *KMT1* which suggests an opposite role on effector gene expression. To our knowledge, this is the first evidence of  
384 a double control involving both the chromatin structure and specific TF on expression of effector genes in a fungal  
385 species. Altogether, these results allowed us to draw a model for the double control of effector gene expression  
386 mediated by KMT1 and LmPf2 in *L. maculans*. The next step forward will be to investigate whether LmPf2 directly  
387 binds on effector gene promoters.

### 388 ***LmPf2* and *KMT1* inactivations lead to pathogenicity defects in *L. maculans***

389 In this study, we investigated function of *LmPf2* and *KMT1* in *L. maculans*. CRISPR-Cas9 inactivation of *LmPf2*  
390 and *KMT1* induced pathogenicity defects while the mutants were not altered in their growth or conidiation. We  
391 conclude that, in *L. maculans*, LmPf2 and KMT1 are specifically involved in the establishment of oilseed rape infection  
392 by *L. maculans*. This is in accordance with previous results obtained for three other Pleosporales species *A. brassicicola*,  
393 *P. tritici repentis* and *P. nodorum* in which Pf2 is also essential for establishment of infection (Cho *et al.*, 2013; Rybak  
394 *et al.*, 2017). Other fungal TF were also reported as involved in the establishment of infection (Michielse *et al.*, 2009,  
395 Ipcho *et al.*, 2010; Zahiri *et al.*, 2010; Soyer *et al.*, 2015. B). KMT1 was also demonstrated to be involved in several  
396 cellular processes in fungi. Histone modifications enzymes are more and more documented for their ability to express  
397 concertedly, large portion of genomic compartments (Gacek and Strauss, 2012). In a fungal endophyte, *Epichloë*  
398 *festucea*, KMT1 is able to regulate symbiosis-specific alkaloids bioprotective metabolites synthesis and to be crucial  
399 for mutualistic interaction establishment (Chujo and Scott, 2014). Despite the fact that KMT1 has been demonstrated  
400 to be crucial for pathogenicity establishment and effector and pathogenicity gene regulation in *B. cinerea* and *L.*  
401 *maculans* (Soyer *et al.*, 2014; Zhang *et al.*, 2016). The role of KMT1 in phytopathogenic fungi during infection remain  
402 poorly understood. H3K9me3 has also been described to regulate secondary metabolism, via the inactivation of the  
403 KdmA demethylase responsible for H3K9me3 removal in *A. nidulans* (Gacek-Matthews *et al.*, 2015).

### 404 ***LmPf2* and *KMT1* inactivations lead to changes in effector gene expression**

405  $\Delta kmt1$  mutant exhibited *AvrLm6* up-regulation *in vitro*. This is consistent with our previous finding that *DIM5*  
406 (*KMT1*) partial inactivation via RNAi led to avirulence genes up-regulation and H3K9me3 depletion. H3K9me3 is known  
407 to regulate secondary metabolism and alkaloid production in several fungi and even in a species related to *L. maculans*,  
408 *Epichloe festuceae*, harboring a bipartite genome structure as well (Gacek and Strauss, 2012; Chujo et Scott, 2014).  
409 Soyer *et al.* (2014) determined that the silencing of KMT1 allowed up-regulation of 30% of the genes specifically  
410 located in H3K9me3 whereas in this study we only identified five genes located in a H3K9me3 context specifically up-  
411 regulated in the  $\Delta kmt1$  mutant. Soyer *et al.* (2014) were using DNA oligoarrays for their transcriptomic analysis and a  
412 previous annotation of *L. maculans* genome. Furthermore, the complete inactivation of *KMT1* could have led to





413 H3K27me3 relocation at H3K9me3 previous emplacements, as previously reported for *Neurospora crassa* or  
414 *Zymoseptoria tritici*, that phenomenon being less pronounced after silencing of *KMT1* (Basenko *et al.*, 2015; Möller *et*  
415 *al.*, 2019). The inactivation of *KMT1* also induced expression of genes which were located in H3K4me2 and in no  
416 particular environment indicating that H3K9me3 removal not only modify gene expression in its environment but also  
417 in other compartments due to a structural role of H3K9me3 in genome organization as previously described in *N.*  
418 *crassa* (Winter *et al.*, 2018).

419 In this study, we also determined that *LmPf2* inactivation led to effector gene expression defect *in planta*.  
420 Furthermore, *LmPf2* over-expression led to effector genes up-regulation. Altogether these results indicate that *LmPf2*  
421 specifically regulates effector gene expression in *L. maculans*. Once again, these results are in accordance with  
422 previous studies that described *Pf2* as a regulator of effector gene expression in three Pleosporales species. *Pf2*, was  
423 reported to regulates, expression of 33 genes encoding secreted proteins including eight putative effectors in  
424 *Alternaria brassicicola* (Cho *et al.*, 2013). In *Parastagonospora nodorum*, *PnPf2* positively regulates two necrotrophic  
425 effector genes, *SnToxA* and *SnTox3* and the orthologue of *SnToxA*, *ToxA*, is regulated by *PtrPf2* in *Pyrenophora tritici-*  
426 *repentis* (Rybak *et al.*, 2017). This can confirm the conserved function of *Pf2* on effector gene regulation in several  
427 phytopathogenic fungi.

428 ***LmPf2* over-expression leads to an up-regulation of effector-encoding genes and genes in an heterochromatin**  
429 **context**

430 In this study, we observed that up-regulated genes in JN2\_o*Pf2* transformants were statistically enriched in  
431 SSP-encoding genes, and genes located in a heterochromatin context (enriched in H3K9me3 or H3K27me3).  
432 Furthermore, we determined that the inactivation of *KMT1* combined with an over-expression of *LmPf2* induced  
433 expression of avirulence genes. We thus conclude that *LmPf2* specifically regulate effector genes which are mostly  
434 located in H3K9me3 and H3K27me3 environments. In addition, avirulence genes up-regulation was much higher in a  
435  $\Delta kmt1$  than in a WT background. Finally, we observed that *LmPf2* and *KMT1* were inversely expressed during infection.  
436 This indicate that *AvrLm* genes in *L. maculans* are under double control of both H3K9me3 and *LmPf2*. Histone  
437 modifications H3K9me3 and H3K27me3 have been identified as important regulators of effector and secondary  
438 metabolism gene expression in some fungi (Strauss and Reyes-Dominguez, 2011; Chujo and Scott, 2014; Connolly *et*  
439 *al.*, 2013; Soyer *et al.*, 2015. A). But to our knowledge this is the first study focusing on both a transcription factor and  
440 a chromatin remodeler. We now need to determine if that model of double control of effector gene expression  
441 involving a specific TF and a chromatin remodeler could be generalized to other pathogenic fungi with two speed  
442 genome structures or if it remains very specific of *L. maculans*. In *F. oxysporum*, two specific transcription factors, *FTF1*  
443 and *FTF2*, have been described for their ability to regulate all effector genes located on a particular pathogenicity  
444 dispensable chromosome (van der Does *et al.*, 2016). Such differences in effector genes regulation could indicate that  
445 phytopathogenic fungi use their incredible genome plasticity in very different ways to regulates effector genes during  
446 infection.

447



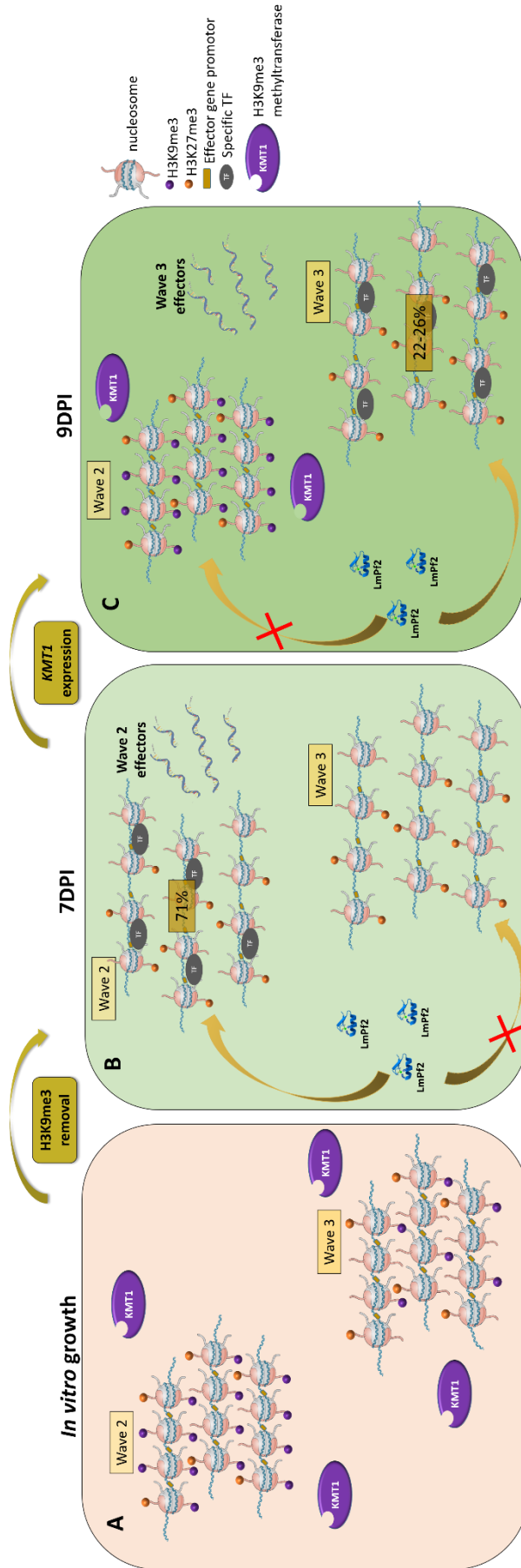
448 **LmPf2 controls different sets of genes according to their chromatin context**

449 Effector genes are known to need a tight regulation to permit fungus escaping plant recognition and  
450 manipulating host cell (Raffaele et Kamoun, 2012; Hacquard *et al.*, 2012; O'Connell *et al.*, 2012; Gervais *et al.*, 2017;  
451 Haueisen *et al.*, 2018). In this study, we determined the respective impact of chromatin remodeling and LmPf2 on  
452 effector gene expression. We determined that, in a WT background, *LmPf2* over-expression leads to over-expression  
453 of wave 3 effector genes. In contrast, in a  $\Delta kmt1$  background, *LmPf2* over-expression leads to over-expression of wave  
454 2 effector genes including six of the nine cloned avirulence genes. Since *KMT1* is lowly expressed at the peak of  
455 expression of wave 2 effector genes (7DPI), we can hypothesize that H3K9me3 is removed at these effector gene loci  
456 during early infection and allow direct or indirect control by LmPf2. Furthermore, the inverted expression of *KMT1* and  
457 *LmPf2* during *L. maculans* life cycle supports our model of the double regulation. Other specific TFs could act in addition  
458 to LmPf2 to concertedly express all genes of wave 2. We propose the following regulation model: (i) *in vitro*, wave 2  
459 and 3 genes are tightly repressed by H3K9me3 deposited by KMT1. (ii) at the onset of infection, H3K9me3 is removed  
460 and *LmPf2* is highly expressed, specifically activating wave 2 genes possibly because of a better direct or indirect  
461 affinity for wave 2 promoters. (iii) At 9 DPI, *KMT1* starts being re-expressed leading to a deposition of H3K9me3 on  
462 wave 2 genes first, possibly because of their high level of H3K9me3 *in vitro* or their specific localization in TE-rich  
463 compartments. At this stage, LmPf2 is not involved anymore on the control of wave 2 gene expression and starts  
464 activating wave 3 genes (**Figure 10**). These different sets of genes could be differentially controlled by LmPf2 indirectly  
465 taking into account the diversity of TF-encoding genes differentially expressed depending on the chromatin  
466 condensation state.

467

468





**Figure 10: Model of key players of effector gene regulation as concerted waves during the early stage of *Leptosphaeria maculans* colonization of oilseed rape and *in vitro* growth.** During *Leptosphaeria maculans*' life cycle, LmPf2 and KMT1 are able to regulate effector gene expression respectively acting on effector gene expression activation and chromatin condensation. **A:** During axenic growth effector waves repertoire are tightly repressed by H3K9me3 maintained by KMT1. **B:** At 7 DPI H3K9me3 is fully removed and heterochromatin is loosened. At this stage *LmPf2* is highly expressed and is able to regulate specifically 71% of Wave 2 effector genes probably because of a better affinity for wave 2 effector promoters. Between 7 DPI and 9 DPI KMT1 is being expressed and start heterochromatin recondensation by H3K9me3 redeposition. **C:** Wave 2 effector genes are being recondense and LmPf2 start acting on (22-26%) of Wave 3 effector genes, provoking their up-regulation. DPI = Days Post Inoculation.

471 **ACCESSION NUMBERS**

472 All sequenced reads have been deposited with the Short Read Archive under accession numbers XXXX.

473 **REFERENCES**

474 Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. *Journal*  
475 *of Molecular Biology* 215, 403–410.

476 Ansan-Melayah, D. (inra, Balesdent, M.H., Buee, M., Rouxel, T., 1995. Genetic characterization of AvrLm1, the  
477 first avirulence gene of *Leptosphaeria maculans*. *Phytopathology (USA)*.

478 Balesdent, M.H., Attard, A., Ansan-Melayah, D., Delourme, R., Renard, M., Rouxel, T., 2001. Genetic Control  
479 and Host Range of Avirulence Toward *Brassica napus* Cultivars Quinta and Jet Neuf in *Leptosphaeria maculans*.  
480 *Phytopathology* 91, 70–76.

481 Balesdent, M.H., Attard, A., Kühn, M.L., Rouxel, T., 2002. New Avirulence Genes in the Phytopathogenic Fungus  
482 *Leptosphaeria maculans*. *Phytopathology* 92, 1122–1133.

483 Balesdent, M.-H., Fudal, I., Ollivier, B., Bally, P., Grandaubert, J., Eber, F., Chèvre, A.-M., Leflon, M., Rouxel, T.,  
484 2013. The dispensable chromosome of *Leptosphaeria maculans* shelters an effector gene conferring avirulence  
485 towards *Brassica rapa*. *New Phytologist* 198, 887–898.

486 Basenko, E.Y., Sasaki, T., Ji, L., Prybol, C.J., Burckhardt, R.M., Schmitz, R.J., Lewis, Z.A., 2015. Genome-wide  
487 redistribution of H3K27me3 is linked to genotoxic stress and defective growth. *Proceedings of the National Academy*  
488 *of Sciences* 112, E6339–E6348.

489 Bolger, A.M., Lohse, M., Usadel, B., 2014. Trimmomatic: a flexible trimmer for Illumina sequence data.  
490 *Bioinformatics* 30, 2114–2120.

491 Van der Biezen, E.A., Jones, J.D., 1998. Plant disease-resistance proteins and the gene-for-gene concept.  
492 *Trends Biochem. Sci.* 23, 454–456.

493 Cho, Y., Ohm, R.A., Grigoriev, I.V., Srivastava, A., 2013. Fungal-specific transcription factor AbPf2 activates  
494 pathogenicity in *Alternaria brassicicola*. *The Plant Journal* 75, 498–514.

495 Chujo, T., Scott, B., 2014. Histone H3K9 and H3K27 methylation regulates fungal alkaloid biosynthesis in a  
496 fungal endophyte-plant symbiosis: K9 and K27 methylation regulates symbiosis. *Molecular Microbiology* 92, 413–434.

497 Connolly, L.R., Smith, K.M., Freitag, M., 2013. The *Fusarium graminearum* Histone H3 K27 Methyltransferase  
498 KMT6 Regulates Development and Expression of Secondary Metabolite Gene Clusters. *PLoS Genetics* 9, e1003916.

499 Dobin, A., Davis, C.A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., Batut, P., Chaisson, M., Gingeras, T.R.,  
500 2013. STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21.

501 van der Does, H.C., Fokkens, L., Yang, A., Schmidt, S.M., Langereis, L., Lukasiewicz, J.M., Hughes, T.R., Rep, M.,  
502 2016. Transcription factors encoded on core and accessory chromosomes of *Fusarium oxysporum* induce expression  
503 of effector genes. *PLoS genetics* 12, e1006401.



- 504 Dutreux, F., Da Silva, C., d'Agata, L., Couloux, A., Gay, E.J., Istace, B., Lapalu, N., Lemainque, A., Linglin, J., Noel,  
505 B., Wincker, P., Cruaud, C., Rouxel, T., Balesdent, M.-H., Aury, J.-M., 2018. De novo assembly and annotation of three  
506 Leptosphaeria genomes using Oxford Nanopore MinION sequencing. *Scientific Data* 5.
- 507 Finn, R.D., Bateman, A., Clements, J., Coggill, P., Eberhardt, R.Y., Eddy, S.R., Heger, A., Hetherington, K., Holm,  
508 L., Mistry, J., Sonnhammer, E.L.L., Tate, J., Punta, M., 2014. Pfam: the protein families database. *Nucleic Acids Res.* 42,  
509 D222-230.
- 510 Fudal, I., Ross, S., Gout, L., Blaise, F., Kuhn, M.L., Eckert, M.R., Cattolico, L., Bernard-Samain, S., Balesdent,  
511 M.H., Rouxel, T., 2007. Heterochromatin-Like Regions as Ecological Niches for Avirulence Genes in the *Leptosphaeria*  
512 *maculans* Genome: Map-Based Cloning of *AvrLm6*. *Molecular Plant-Microbe Interactions* 20, 459–470.
- 513 Gacek, A., Strauss, J., 2012. The chromatin code of fungal secondary metabolite gene clusters. *Appl Microbiol*  
514 *Biotechnol* 95, 1389–1404.
- 515 Gacek-Matthews, A., Noble, L.M., Gruber, C., Berger, H., Sulyok, M., Marcos, A.T., Strauss, J., Andrianopoulos,  
516 A., 2015. KdmA, a histone H3 demethylase with bipartite function, differentially regulates primary and secondary  
517 metabolism in *Aspergillus nidulans*: Histone H3K9 demethylase function in growth. *Molecular Microbiology* 96, 839–  
518 860.
- 519 Ghanbarnia, K., Ma, L., Larkan, N.J., Haddadi, P., Fernando, W.G.D., Borhan, M.H., 2018. *Leptosphaeria*  
520 *maculans* AvrLm9: a new player in the game of hide and seek with AvrLm4-7. *Molecular Plant Pathology* 19, 1754–  
521 1764.
- 522 Gervais, J., Plissonneau, C., Linglin, J., Meyer, M., Labadie, K., Cruaud, C., Fudal, I., Rouxel, T., Balesdent, M.-  
523 H., 2017. Different waves of effector genes with contrasted genomic location are expressed by *Leptosphaeria*  
524 *maculans* during cotyledon and stem colonization of oilseed rape. *Molecular plant pathology* 18, 1113–1126.
- 525 Gout, L., Fudal, I., Kuhn, M.-L., Blaise, F., Eckert, M., Cattolico, L., Balesdent, M.-H., Rouxel, T., 2006. Lost in the  
526 middle of nowhere: the AvrLm1 avirulence gene of the Dothideomycete *Leptosphaeria maculans*. *Molecular*  
527 *Microbiology* 60, 67–80.
- 528 Guo, S., Zhong, S., Zhang, A., 2013. Privacy-preserving Kruskal-Wallis test. *Comput Methods Programs Biomed*  
529 112, 135–145.
- 530 Hacquard, S., Joly, D.L., Lin, Y.-C., Tisserant, E., Feau, N., Delaruelle, C., Legué, V., Kohler, A., Tanguay, P., Petre,  
531 B., Frey, P., Van de Peer, Y., Rouzé, P., Martin, F., Hamelin, R.C., Duplessis, S., 2012. A comprehensive analysis of genes  
532 encoding small secreted proteins identifies candidate effectors in *Melampsora larici-populina* (poplar leaf rust). *Mol.*  
533 *Plant Microbe Interact.* 25, 279–293.
- 534 Haueisen, J., Möller, M., Eschenbrenner, C.J., Grandaubert, J., Seybold, H., Adamiak, H., Stukenbrock, E.H.,  
535 2019. Highly flexible infection programs in a specialized wheat pathogen. *Ecology and Evolution* 9, 275–294.
- 536 Idnurm, A., Urquhart, A.S., Vummadi, D.R., Chang, S., Van de Wouw, A.P., López-Ruiz, F.J., 2017. Spontaneous  
537 and CRISPR/Cas9-induced mutation of the osmosensor histidine kinase of the canola pathogen *Leptosphaeria*  
538 *maculans*. *Fungal Biology and Biotechnology* 4.



- 539 IpCho, S.V.S., Tan, K.-C., Koh, G., Gummer, J., Oliver, R.P., Trengove, R.D., Solomon, P.S., 2010. The  
540 Transcription Factor StuA Regulates Central Carbon Metabolism, Mycotoxin Production, and Effector Gene Expression  
541 in the Wheat Pathogen *Stagonospora nodorum*. *Eukaryotic Cell* 9, 1100–1108.
- 542 Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., 2009.  
543 The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079.
- 544 Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-  
545 Delta Delta C(T)) Method. *Methods*. 2001 25:402-408.
- 546 Lo Presti, L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., Zuccaro, A., Reissmann, S., Kahmann, R.,  
547 2015. Fungal Effectors and Plant Susceptibility. *Annual Review of Plant Biology* 66, 513–545.
- 548 McCarthy, D.J., Chen, Y., Smyth, G.K., 2012. Differential expression analysis of multifactor RNA-Seq  
549 experiments with respect to biological variation. *Nucleic Acids Res* 40, 4288–4297.
- 550 Michielse, C.B., van Wijk, R., Reijnen, L., Manders, E.M.M., Boas, S., Olivain, C., Alabouvette, C., Rep, M., 2009.  
551 The Nuclear Protein Sge1 of *Fusarium oxysporum* Is Required for Parasitic Growth. *PLoS Pathogens* 5, e1000637.
- 552 Möller, M., Schotanus, K., Soyer, J., Haeisen, J., Happ, K., Stralucke, M., Happel, P., Smith, K.M., Connolly, L.R.,  
553 Freitag, M., Stukenbrock, E.H., 2019. Destabilization of chromosome structure by histone H3 lysine 27 methylation.
- 554 Muller, P.Y., Janovjak, H., Miserez, A.R., Dobbie, Z., 2002. Processing of gene expression data generated by  
555 quantitative real-time RT-PCR. *BioTechniques* 32, 1372–1374, 1376, 1378–1379.
- 556 Mullins, E.D., Chen, X., Romaine, P., Raina, R., Geiser, D.M., Kang, S., 2001. *Agrobacterium*-Mediated  
557 Transformation of *Fusarium oxysporum*: An Efficient Tool for Insertional Mutagenesis and Gene Transfer.  
558 *Phytopathology* 91, 173–180.
- 559 O'Connell, R.J., Thon, M.R., Hacquard, S., Amyotte, S.G., Kleemann, J., 2012. Lifestyle transitions in plant  
560 pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nature Genetics* 44, 1060–1065.
- 561 Papadopoulos, J.S., Agarwala, R., 2007. COBALT: constraint-based alignment tool for multiple protein  
562 sequences. *Bioinformatics* 23, 1073–1079.
- 563 Parlange, F., Daverdin, G., Fudal, I., Kuhn, M.-L., Balesdent, M.-H., Blaise, F., Grezes-Besset, B., Rouxel, T., 2009.  
564 *Leptosphaeria maculans* avirulence gene *AvrLm4-7* confers a dual recognition specificity by the *Rlm4* and *Rlm7*  
565 resistance genes of oilseed rape, and circumvents *Rlm4*-mediated recognition through a single amino acid change.  
566 *Mol. Microbiol.* 71, 851–863.
- 567 Petit-Houdenot, Y., Degrave, A., Meyer, M., Blaise, F., Ollivier, B., Marais, C.-L., Jauneau, A., Audran, C., Rivas,  
568 S., Veneault-Fourrey, C., Brun, H., Rouxel, T., Fudal, I., Balesdent, M.-H., 2019. A two genes - for - one gene interaction  
569 between *Leptosphaeria maculans* and *Brassica napus*. *New Phytol.* 223, 397–411.
- 570 Plissonneau, C., Rouxel, T., Chèvre, A.-M., Van De Wouw, A.P., Balesdent, M.-H., 2018. One gene-one name:  
571 the *AvrLmJ1* avirulence gene of *Leptosphaeria maculans* is *AvrLm5*. *Molecular plant pathology* 19, 1012–1016.
- 572 Robinson, M.D., McCarthy, D.J., Smyth, G.K., 2010. edgeR: a Bioconductor package for differential expression  
573 analysis of digital gene expression data. *Bioinformatics* 26, 139–140.
- 574 Rouxel, T., Balesdent, M.H., 2005. The stem canker (blackleg) fungus, *Leptosphaeria maculans*, enters the  
575 genomic era. *Molecular Plant Pathology* 6, 225–241.



- 576 Rouxel, T., Grandaubert, J., Hane, J.K., Hoede, C., van de Wouw, A.P., Couloux, A., Dominguez, V., Anthouard,  
577 V., Bally, P., Bourras, S., Cozijnsen, A.J., Ciuffetti, L.M., Degrave, A., Dilmaghani, A., Duret, L., Fudal, I., Goodwin, S.B.,  
578 Gout, L., Glaser, N., Linglin, J., Kema, G.H.J., Lapalu, N., Lawrence, C.B., May, K., Meyer, M., Ollivier, B., Poulain, J.,  
579 Schoch, C.L., Simon, A., Spatafora, J.W., Stachowiak, A., Turgeon, B.G., Tyler, B.M., Vincent, D., Weissenbach, J.,  
580 Amselem, J., Quesneville, H., Oliver, R.P., Wincker, P., Balesdent, M.-H., Howlett, B.J., 2011. Effector diversification  
581 within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations. *Nature*  
582 *Communications* 2.
- 583 Rybak, K., See, P.T., Phan, H.T., Syme, R.A., Moffat, C.S., Oliver, R.P., Tan, K.-C., 2017. A functionally conserved  
584 Zn<sub>2</sub>Cys<sub>6</sub> binuclear cluster transcription factor class regulates necrotrophic effector gene expression and host-specific  
585 virulence of two major Pleosporales fungal pathogens of wheat. *Molecular plant pathology* 18, 420–434.
- 586 Sánchez-Vallet, A., Fouché, S., Fudal, I., Hartmann, F.E., Soyer, J.L., Tellier, A., Croll, D., 2018. The Genome  
587 Biology of Effector Gene Evolution in Filamentous Plant Pathogens. *Annual Review of Phytopathology* 56, 21–40.
- 588 Šašek, V., Nováková, M., Dobrev, P.I., Valentová, O., and Burketová, L. (2012). β-aminobutyric acid protects  
589 Brassica napus plants from infection by *Leptosphaeria maculans*. Resistance induction or a direct antifungal effect?  
590 *Eur. J. Plant Pathol.* 133, 279–289.
- 591 Shannon, P., Markiel, A., Ozier, O., Baliga, N.S., Wang, J.T., Ramage, D., Amin, N., Schwikowski, B., Ideker, T.,  
592 2003. Cytoscape: a software environment for integrated models of biomolecular interaction networks. *Genome Res.*  
593 13, 2498–2504.
- 594 Silayeva, O., Barnes, A.C., 2018. Gibson Assembly facilitates bacterial allelic exchange mutagenesis. *J.*  
595 *Microbiol. Methods* 144, 157–163.
- 596 Soyer, J.L., El Ghalid, M., Glaser, N., Ollivier, B., Linglin, J., Grandaubert, J., Balesdent, M.-H., Connolly, L.R.,  
597 Freitag, M., Rouxel, T., Fudal, I., 2014. Epigenetic Control of Effector Gene Expression in the Plant Pathogenic Fungus  
598 *Leptosphaeria maculans*. *PLoS Genetics* 10, e1004227.
- 599 Soyer, J.L., Rouxel, T., Fudal, I., 2015. A Chromatin-based control of effector gene expression in plant-  
600 associated fungi. *Current Opinion in Plant Biology* 26, 51–56.
- 601 Soyer, J.L., Hamiot, A., Ollivier, B., Balesdent, M.-H., Rouxel, T., Fudal, I., 2015. B The APSES transcription factor  
602 LmStuA is required for sporulation, pathogenic development and effector gene expression in *Leptosphaeria maculans*:  
603 Functional analysis of LmStuA from *L. maculans*. *Molecular Plant Pathology* 16, 1000–1005.
- 604 Soyer, J.L., Grandaubert, J., Haueisen, J., Schotanus, K., Stukenbrock, E.H., 2019. In planta chromatin  
605 immunoprecipitation in *Zymoseptoria tritici* reveals chromatin-based regulation of putative effector gene expression.  
606 *bioRxiv* 544627.
- 607 Stergiopoulos, I., de Wit, P.J.G.M., 2009. Fungal Effector Proteins. *Annual Review of Phytopathology* 47, 233–  
608 263.
- 609 Strauss, J., Reyes-Dominguez, Y., 2011. Regulation of secondary metabolism by chromatin structure and  
610 epigenetic codes. *Fungal Genetics and Biology* 48, 62–69.
- 611 Tan, K.-C., Oliver, R.P., 2017. Regulation of proteinaceous effector expression in phytopathogenic fungi. *PLoS*  
612 *pathogens* 13, e1006241.



613 Winter, D.J., Ganley, A., Young, C., Liachko, I., Schardl, C., Dupont, P., Berry, D., Ram, A., Scott, B., Cox, M.,  
614 2018. Repeat elements organize 3D genome structure and mediate transcription in the filamentous fungus *Epichloë*  
615 *festucae*. bioRxiv 339010.

616 Wouw, A.P.V. de, Lowe, R.G.T., Elliott, C.E., Dubois, D.J., Howlett, B.J., 2014. An avirulence gene, *AvrLmJ1*,  
617 from the blackleg fungus, *Leptosphaeria maculans*, confers avirulence to *Brassica juncea* cultivars. *Molecular Plant*  
618 *Pathology* 15, 523–530.

619 Zahiri, A., Heibel, K., Wahl, R., Rath, M., Kämper, J., 2010. The *Ustilago maydis* Forkhead Transcription  
620 Factor Fox1 Is Involved in the Regulation of Genes Required for the Attenuation of Plant Defenses During Pathogenic  
621 Development. *Molecular Plant-Microbe Interactions* 23, 1118–1129.

622 Zhang, X., Liu, X., Zhao, Y., Cheng, J., Xie, J., Fu, Y., Jiang, D., Chen, T., 2016. Histone H3 Lysine 9  
623 Methyltransferase DIM5 Is Required for the Development and Virulence of *Botrytis cinerea*. *Frontiers in*  
624 *Microbiology* 7.

625

626

627

628

629

630

631

632

633

634

635

636





637 SUPPLEMENTARY DATA

638 Supplementary Data are available at NAR online.

639

640

641

642

643

644

645

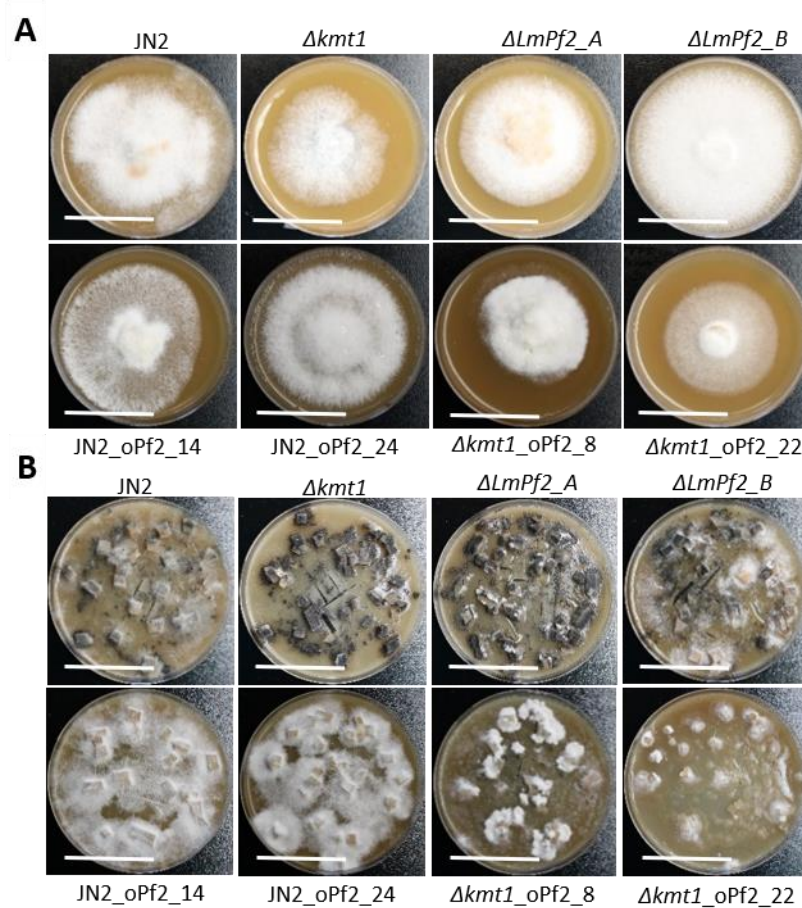
646

647

648

649

650



**Figure S1: Morphological characterization of *LmPf2* mutants and over-expression transformants.** The over-expression transformants have been obtained in a WT background (JN2) and in a  $\Delta kmt1$  background. **A:** Transformant thallus morphology on V8 agar medium at 6 days after inoculation. Scale bars corresponding to 50 mm. **B:** Transformant thallus morphology during conidiation on V8 agar 14 days after inoculation. Scale bars corresponding to 90 mm.

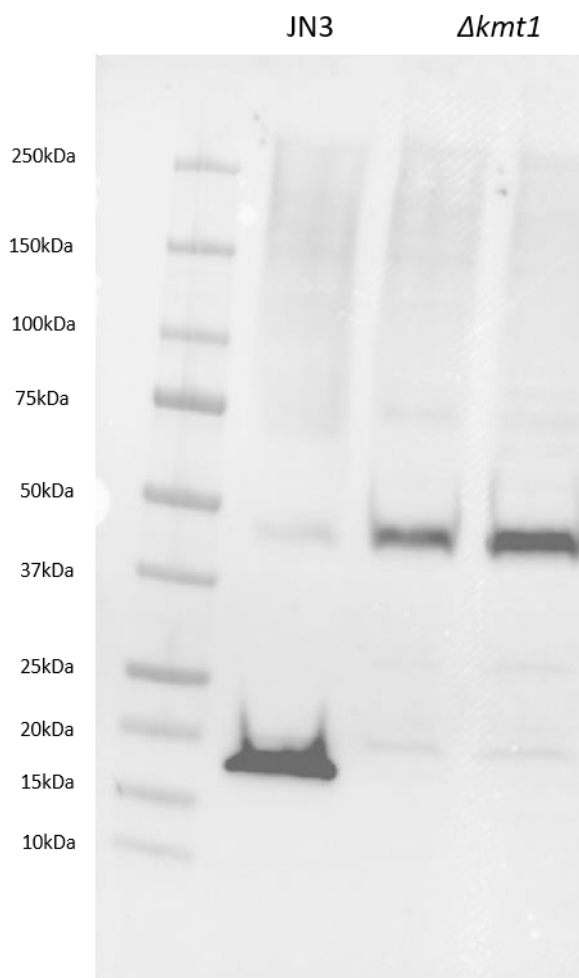
651

652

653

654

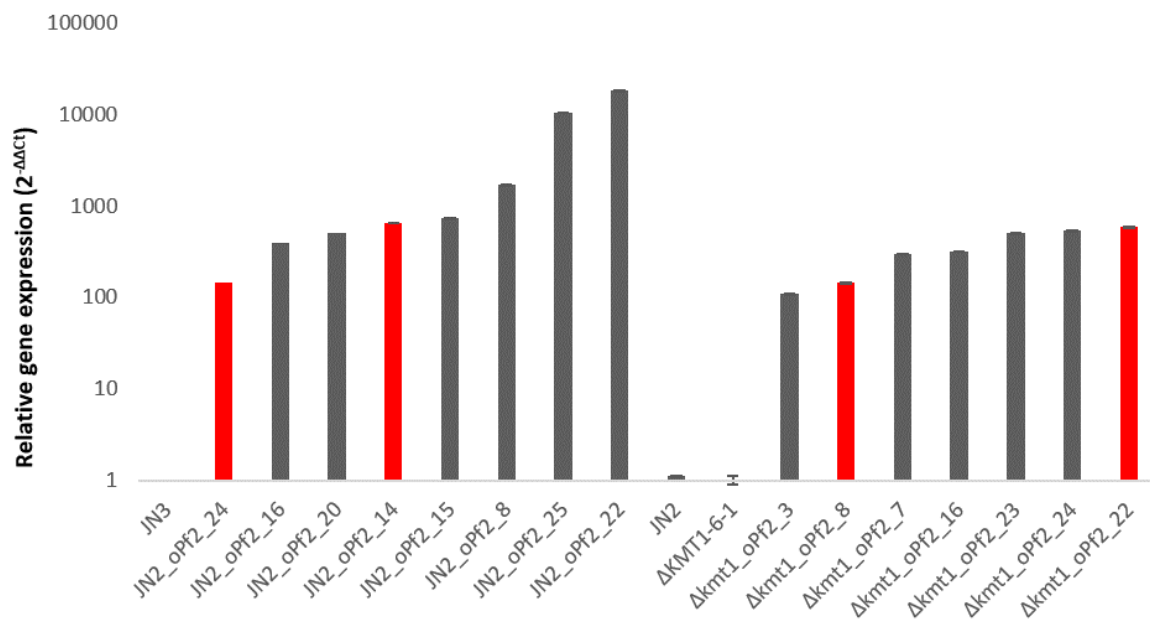
655



**Figure S2: Immunoblot analysis showing presence of H3K9me3.** Protein extracts were analysed by immunoblotting with anti-H3K9me3 antibodies.  $\Delta kmt1$  have been tested using two replicates of total protein extract.

656

657



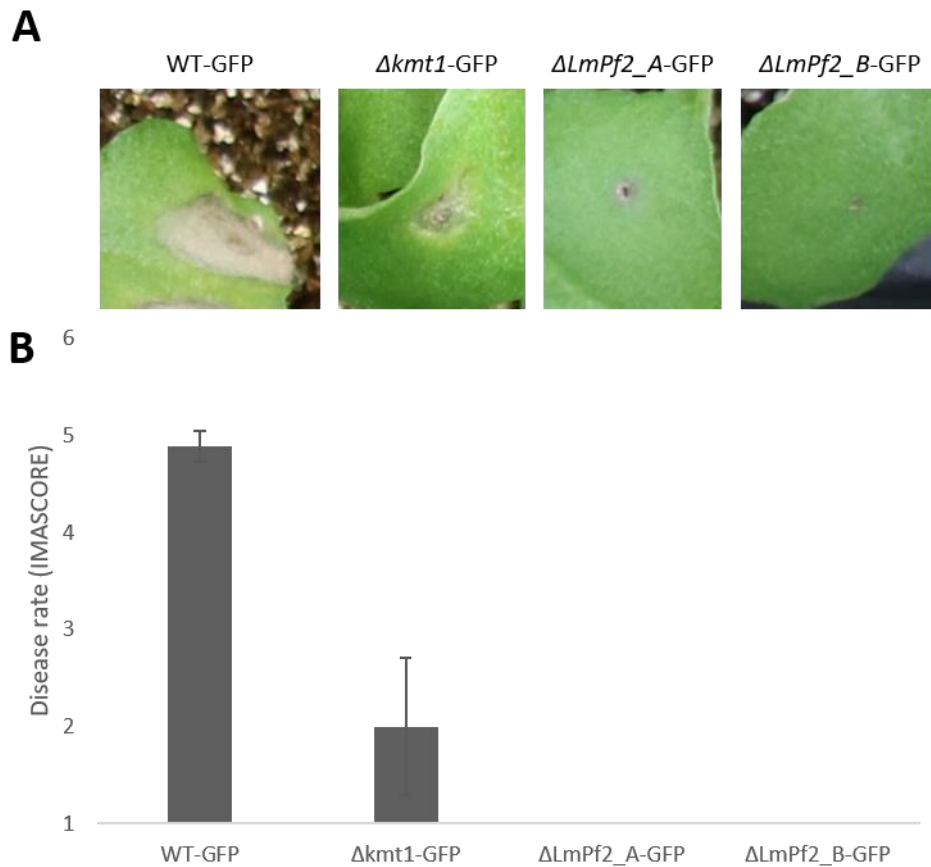
**Figure S3: Level of *Lmpf2* expression in the *L. maculans* transformants over-expressing *Lmpf2* during axenic growth.** Mycelium was obtained by growing JN2 and JN3 strains and transformants in Fries liquid medium for 7 days. Two biological replicates per condition were generated. Total RNA was extracted. Expression of *Lmpf2* was measured by qRT-PCR and expressed relatively to *LmβTubulin* expression and to expression of *Lmpf2* in the strain before transformation using the  $2^{-\Delta\Delta C_t}$  method (Livak and Schmittgen, 2001). Red bars correspond to transformants which were selected for further characterization (JN2\_opf2\_14, JN2\_opf2\_24,  $\Delta kmt1\_opf2\_8$  and  $\Delta kmt1\_opf2\_22$ ).

659

660

661

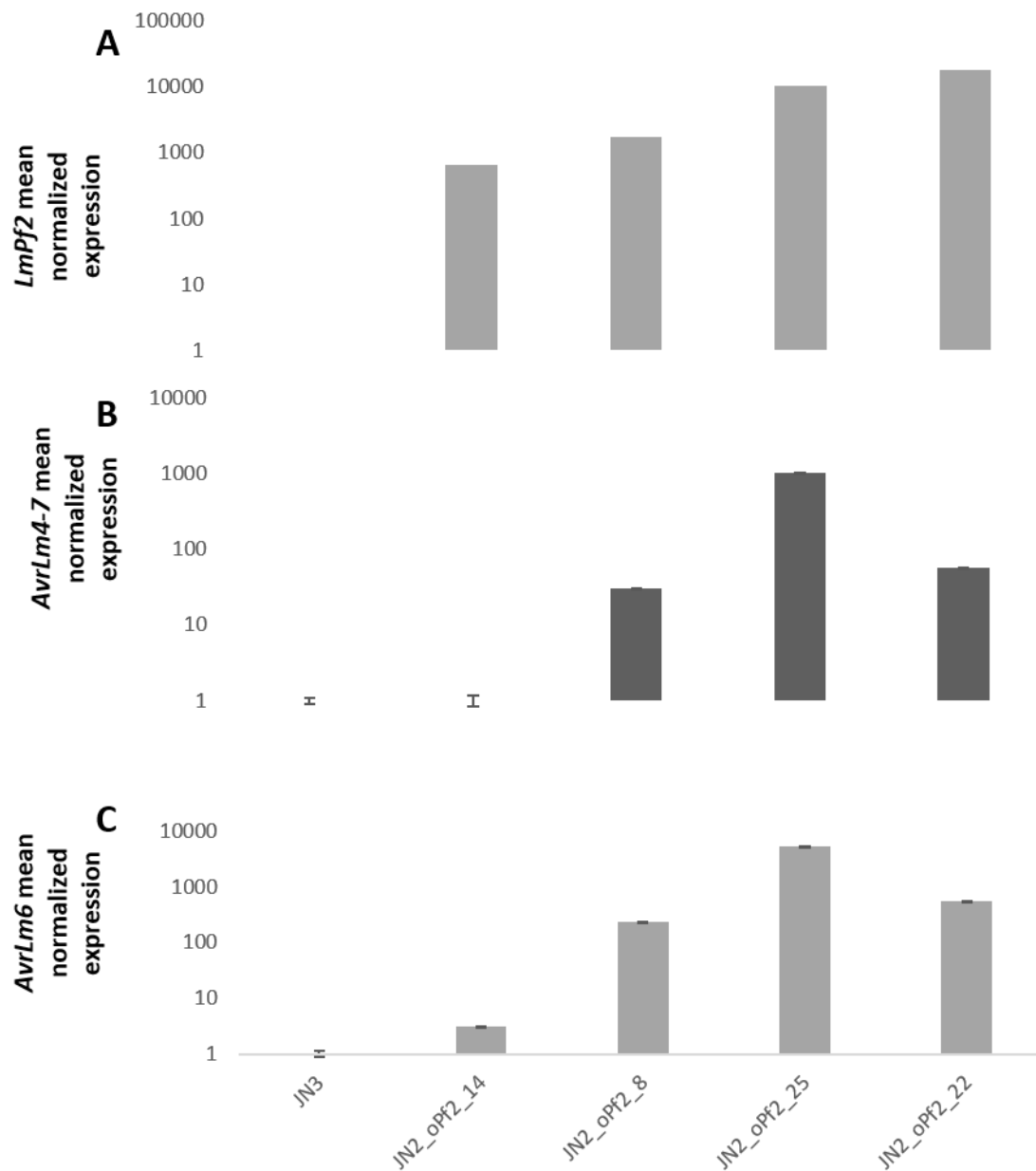
662



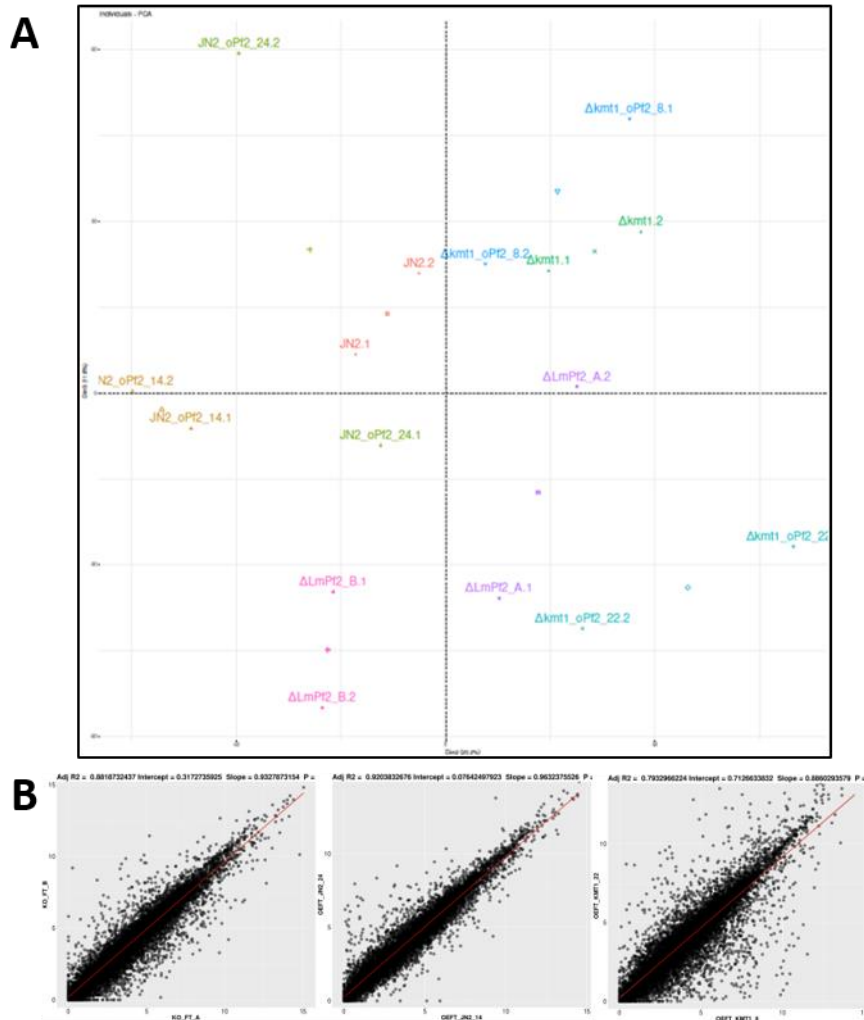
**Figure S4: Pathogenicity assays on WT,  $\Delta kmt1$  and  $\Delta LmPf2$  transformants after crossing with JN2-GFP. A:** Photos of symptoms on Es-Astrid at 14 DPI. **B:** Pathogenicity test of three different CRISPR-Cas9 mutants on oilseed rape cotyledons of a susceptible cultivar (Es-Astrid) at 13 DPI. Disease rate is expressed as the mean scoring using the IMASCORE rating scale comprising six infection classes (IC), where IC1 to IC3 correspond to various levels of resistance of the plant and IC4 to IC6 to susceptibility (Balesdent *et al.*, 2001).

663

664



**Figure S5: Influence of *LmPf2* over-expression on the expression of two *L. maculans* avirulence genes during axenic growth.** Mycelium was grown in Fries liquid medium during seven days. Expression of *LmPf2* (A), *AvrLm4-7* (B) and *AvrLm6* (C) was measured by qRT-PCR level using *LmβTubulin* as a constitutive reporter gene and expression in JN2 as reference ( $2^{-\Delta\Delta Ct}$  method). Each value is the average of two biological replicates (two extractions from different biological replicates) and two technical replicates (two RT-PCR).



**Figure S6: Reproducibility of RNA-Seq replicates. A:** log2(RPKM) PCA of all samples. Legend labels: WT = JN2;  $\Delta kmt1\_X$  =  $\Delta kmt1\_A$  and B;  $\Delta LmPf2\_X$  =  $\Delta LmPf2\_A$  and B; JN2\_oPf2\_X = JN2\_oPf2\_14 and 24 ; and  $\Delta kmt1\_oPf2\_X$  =  $\Delta kmt1\_oPf2\_8$  and 22. **B:** Scatterplot obtained with Log2(RPKM) of three biological replicates. Left panel =  $\Delta LmPf2\_A$  vs  $\Delta LmPf2\_B$ ; middle panel = JN2\_oPf2\_14 vs JN2\_oPf2\_24; and right panel =  $\Delta kmt1\_oPf2\_8$  vs  $\Delta kmt1\_oPf2\_22$ .

Table S1: Primers used in the study

Name of the fragment	DNA sequence	function	construction
gRNA against <i>Lmpf2</i>	GAAACCTAATCAATCAACAGCATGCTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTCCATGCTCCCGCCAG AGAGTTTTAGAGCTAGAAATAGC	<i>Lmpf2</i> CRISPR-CAS9 directed inactivation	
gRNA against <i>KMT1</i>	GAAACCTAATCAATCAACGAGCACTGATGAGTCCGTGAGGACGAAACGAGTAAGCTCGTCCCGTCCGAAACGTC TGCGTTTTAGAGCTAGAAATAGC	<i>KMT1</i> CRISPR-CAS9 directed inactivation	
MAI0309	ACCTCTAATCGAAACCTAATCAATCAAC	Amplification of gRNA coding sequence	CRISPR-Cas9
MAI0310	ATTTAACTTGCTATTTCTAGCTCTAAAAAC	Amplification of gRNA coding sequence	CRISPR-Cas9
verif_CRISPR_Lmpf2_F	CCACTCCAGTCAAGCGGGC	<i>Lmpf2</i> mutation detection	
verif_CRISPR_Lmpf2_R	AATTGGCGGTGGACGATGC	<i>Lmpf2</i> mutation detection	
verif_CRISPR_KMT1_F	GCCCTCCATGGTTGTGGTGC	<i>KMT1</i> mutation detection	
verif_CRISPR_KMT1_R	ACCCCGAAATGCTAGCCGA	<i>KMT1</i> mutation detection	
EF1aProSac1F	GAGAGAGGCTGGCACTCGTAAGGAGTATGAGC	<i>Ef1a</i> promoter amplification and cloning	
EF1aProKpnIR	GAGAGAGGTACCGTTGGGGTGTGTGATGG	<i>Ef1a</i> promoter amplification and cloning	
surex_Lmpf2_Gibs_F	AGTCGACCTGCAGGCATGCAATGGCTACCAATCCACCACC	<i>Lmpf2</i> amplification and cloning	<i>Lmpf2</i> overexpression
surex_Lmpf2_Gibs_R	GTAAAAACGACGGCCAGTGCACCAAGTTGGGGGAGAGTACATACTG	<i>Lmpf2</i> amplification and cloning	
Lmpf2_primer_Up	ATGGCTACCAATCCACCACC	<i>Lmpf2</i> sequence verification	
Lmpf2_Seq1_F	CCGCCAATTCGGTCTGGCC	<i>Lmpf2</i> sequence verification	
Lmpf2_primer_Low	CAGTATGTACTCTCCCCCAAC	<i>Lmpf2</i> sequence verification	
$\beta$ Tubulin_F	AAGAACTCATCCTACTTCGA	qPCR	None
$\beta$ Tubulin_R	TGAATAGCTCCTGAATGG	qPCR	None
AvrLm4-7_F	GCCCTGCATAACTACCGAC	qPCR	None
AvrLm4-7_R	TCCTGGCCAAATATAACTCC	qPCR	None
AvrLm6_F	AAACGGCACTATTACGAAAA	qPCR	None
AvrLm6_R	GATTAGGGGAGAAAGCAAGT	qPCR	None
AvrLm10.A_F	ACAGGAAACACTGACGGCT	qPCR	None
AvrLm10.A_R	GCATGGACTATCGTAGCGTTAT	qPCR	None
AvrLm11_F	TGCGTTTCTTGCTTCCCTATATT	qPCR	None
AvrLm11_R	CAAGTTGGATCTTCTCATTCG	qPCR	None



**Table S2: Crosses performed for mutant purification**

	CRISPR-Cas9 mutation	Compatible isolate used for crosses	Name of the selected progenies <sup>b</sup>	Additional Transformation
<i>Δkmt1 6-1</i>	1 bp insertion	JN2 <sup>a</sup>	<i>ΔKMT1</i>	<i>LmPf2</i> over-expression
		JN2-GFP	<i>ΔKMT1-GFP</i>	no
<i>ΔLmPf2 1-2</i>	1 bp deletion	JN2-GFP	<i>ΔLmPf2_A</i>	no
		JN2-GFP	<i>ΔLmPf2_A-GFP</i>	no
<i>ΔLmPf2 4-2</i>	5 bp deletion	JN2-GFP	<i>ΔLmPf2_B</i>	no
		JN2-GFP	<i>ΔLmPf2_B-GFP</i>	no

<sup>a</sup>Purification by crossing with an isolate of opposite mating type with or without GFP. <sup>b</sup> Selected progenies purified (without *Cas9* and gRNA).





668

669

Table S3: RNA-Seq read information.

	Total numbers of paired reads	Paired reads after pre- processing <sup>a</sup>	% paired kept reads	STAR mapping Uniquely mapped reads <sup>b</sup>
JN2-1	1.49 E+07	1.24E+07	83,28%	81.76%
JN2-2	2.35E+07	1.95E+07	83,14%	83.75%
Δkmt1-1	1.71E+07	1.42E+07	83,23%	83.12%
Δkmt1-2	1.72E+07	1.43E+07	83,06%	85.79%
ΔLmPf2_A-1	1.66E+07	1.37E+07	82,76%	83.92%
ΔLmPf2_A-2	1.59E+07	1.31E+07	82,51%	38.76%
ΔLmPf2_B-1	1.59E+07	1.32E+07	82,84%	81.68%
ΔLmPf2_B-2	1.62E+07	1.33E+07	82,17%	81.20%
JN2_oPf_14-1	2.43E+07	2.01E+07	82,73%	79.73%
JN2_oPf_14-2	2.32E+07	1.92E+07	82,86%	79.11%
JN2_oPf_24-1	2.48E+07	2.06E+07	83,02%	80.51%
JN2_oPf_24-1	2.42E+07	2.01E+07	83,17%	82.96%
Δkmt1_oPf_8-1	2.23E+07	1.85E+07	82,81%	84.44%
Δkmt1_oPf_8-2	2.23E+07	1.85E+07	83,04%	82.61%
Δkmt1_oPf_22-1	2.48E+07	2.08E+07	83,98%	75.19%
Δkmt1_oPf_22-2	2.11E+07	1.75E+07	83,00%	71.57%

<sup>a</sup> Paired read was treated with Trimmomatic. Sequencing adaptors and too short reads (<30 bp) were removed. <sup>b</sup> Reads were then mapped against *L. maculans* genome using STAR (Dobin *et al.*, 2013; Dutreux *et al.*, 2018).



670

671

672

**Table S4: Differentially expressed genes encoding small secreted proteins or located in particular chromatin domains.**

		Total genes	SSP encoding genes	H3K4me2 genes	H3K9me3 genes	H3K27me3 genes	H3K9me3/H3K27me3 genes
Up-regulated	<i>Δkmt1</i>	129	22*	33 <sup>#</sup>	5*	22*	2*
	<i>ΔLmPf2</i>	202	19	85 <sup>#</sup>	4	23	5*
	JN2_oPf2	247	45	77 <sup>#</sup>	5	41*	2*
	<i>Δkmt1_oPf_8</i>	275	54*	61 <sup>#</sup>	16*	56*	10*
	<i>Δkmt1_oPf_22</i>	931	154*	218 <sup>#</sup>	34*	152*	15*
Down-regulated	<i>Δkmt1</i>	143	25*	36 <sup>#</sup>	13*	13	6*
	<i>ΔLmPf2</i>	225	36*	75 <sup>#</sup>	10*	26	4*
	JN2_oPf2	277	19	126 <sup>#</sup>	3	26	0
	<i>Δkmt1_oPf_8</i>	158	23*	41 <sup>#</sup>	9*	16	8*
	<i>Δkmt1_oPf_22</i>	861	83	399 <sup>#</sup>	16	80	9*

Differential expression is obtained by comparing WT expression to each transformant. We studied genes specifically coding for Small Secreted Proteins (SSPs) or specifically located in euchromatin environment marked by di-methylation of lysine 4 of histone H3 (H3K4me2) and heterochromatin environment marked by tri-methylation of lysine 9 of histone H3 (H3K9me3), tri-methylation of lysine 27 of histone H3 (H3K27me3) and both (H3K9me3/H3K27me3). # corresponds to statistically under-represented genes and \* corresponds to statistically enriched genes. Statistical analysis performed using  $\chi^2$ , Pvalue < 0.05.



**Table S5: Differentially expressed transcription factor coding genes in *LmPf2* over-expressor harbouring different chromatinian context.**

Upregulated transcription factor genes			
<i>Δkmt1_oPf_22</i>		JN2_oPf2	
ID	Fonction	ID	Fonction
Lmb_jn3_00338	zinc_finger-containing	Lmb_jn3_00032	zinc_finger_740
Lmb_jn3_00666	AN1_zinc_finger	Lmb_jn3_04461	C2H2_type_zinc_finger_domain-containing
Lmb_jn3_01078	C2H2_type_zinc_finger-containing	Lmb_jn3_04938	translation_factor_mitochondrial
Lmb_jn3_01759	zinc_knuckle	Lmb_jn3_05737	Zn(2)-C6_fungal-type_D-binding_domain
Lmb_jn3_02114	zinc_finger_ring-type	Lmb_jn3_07527	transcription_factor_jumonji_aspartyl_beta-hydroxylase
Lmb_jn3_03373	C6_transcription_factor	Lmb_jn3_11428	C-x8-C-x5-C-x3-H_type_zinc_finger
Lmb_jn3_04191	C2H2_conidiation_transcription_factor		
Lmb_jn3_04849	C2H2_finger_domain-containing		
Lmb_jn3_05461	zinc_ion_binding		
Lmb_jn3_05737	Zn(2)-C6_fungal-type_D-binding_domain		
Lmb_jn3_06942	C6_finger_domain		
Lmb_jn3_06984	transcriptiol_co-activator		
Lmb_jn3_07222	C2H2_finger_domain-containing		
Lmb_jn3_07297	C2H2_type_zinc_finger_domain-containing		
Lmb_jn3_07916	C6_Zinc_finger_domain-containing		
Lmb_jn3_08804	zinc_ion_binding		
Lmb_jn3_10400	zinc_metalloprotease		
Lmb_jn3_11010	transcription_factor_tfiiic_complex_subunit_sfc4		
Lmb_jn3_11399	zinc_ion_binding		
Lmb_jn3_11428	C-x8-C-x5-C-x3-H_type_zinc_finger		
Lmb_jn3_12660	C2H2_finger_domain		

## 673 ACKNOWLEDGEMENT

674 We are grateful to the Genewiz bioinformatics platform Leipzig Germany for providing computing resources.

## 675 FUNDING

## 676 CONFLICT OF INTEREST

677 The authors declare no conflict of interest



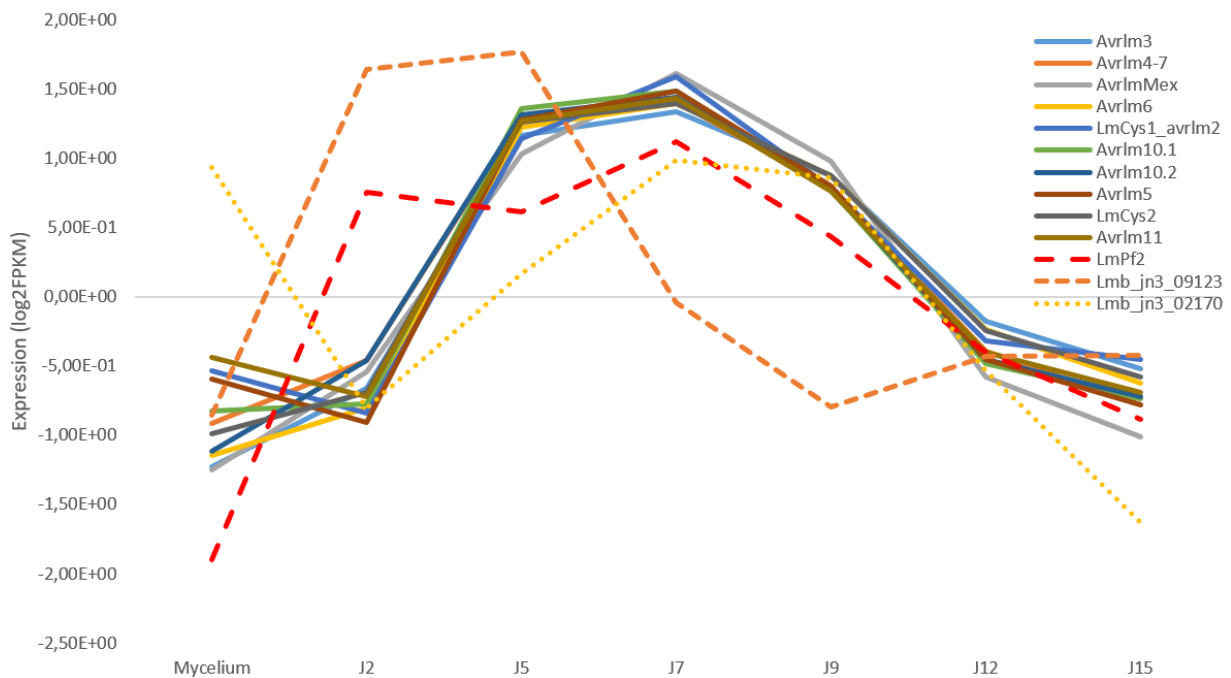
## RESULTATS COMPLEMENTAIRES

### Analyse fonctionnelle de facteurs de transcription putatifs pour leur implication dans la régulation de l'expression des gènes codant des effecteurs de *Leptosphaeria maculans*

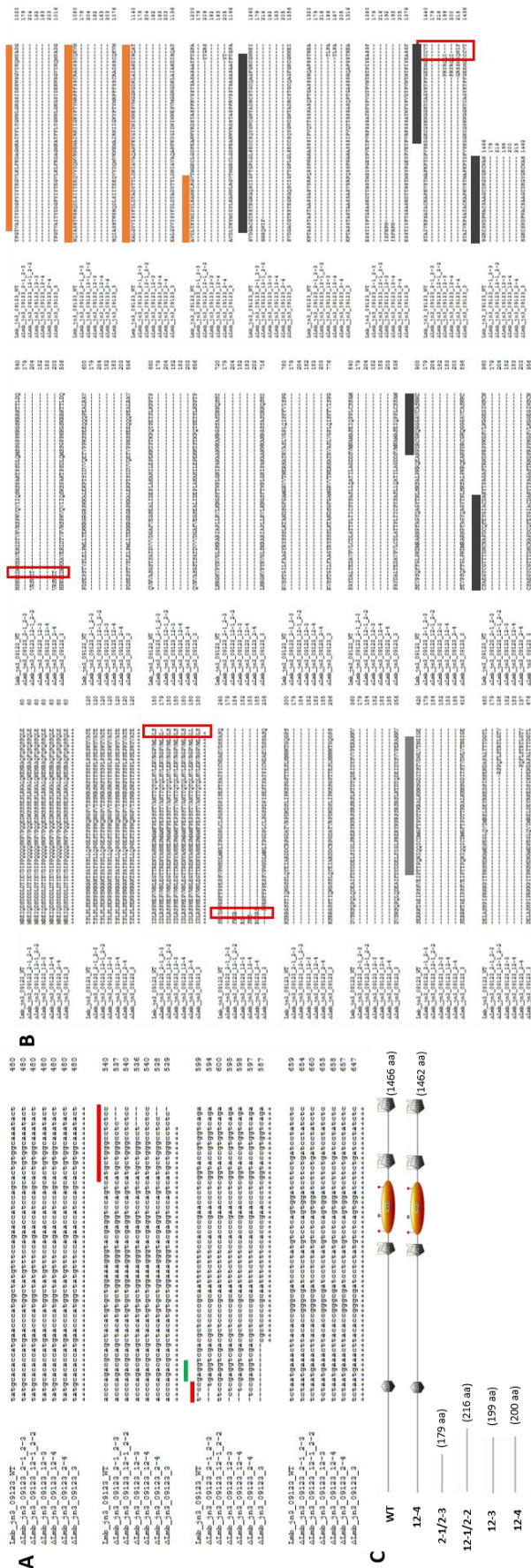
Certains FTs sont capables de réguler la mise en place de différents processus cellulaires chez les champignons filamenteux tels que le métabolisme primaire, secondaire, la reproduction sexuée ou encore l'expression des gènes codant des effecteurs (Cho *et al.*, 2013 ; Soyer *et al.*, 2015. B; Niño-Sánchez *et al.*, 2016; Rybak *et al.*, 2017). Lors de cette thèse, j'ai pu mettre en évidence le rôle majeur du FT LmPf2 dans le contrôle de l'expression des gènes codant des effecteurs de *L. maculans* (Voire **chapitre 3**). Afin d'identifier un ou des FTs contrôlant l'expression des gènes codant des effecteurs de *L. maculans*, deux approches avaient été choisies : (i) une approche avec *a priori* consistant à identifier les gènes orthologues de FTs décrits chez les champignons comme impliqués dans le contrôle de l'expression des effecteurs (e.g. *Sge1*, *StuA*, cf. **partie 5.3**) et (ii) une approche sans *a priori* consistant à prédire le répertoire de FTs de *L. maculans* et à rechercher les FTs surexprimés pendant l'infection ou spécifiques de *L. maculans* (J. Soyer, thèse). Parmi eux, j'ai sélectionné deux FTs de *L. maculans*, codant des AT-Hook (ciblant spécifiquement les régions riches en AT) et différentiellement exprimés *in planta* pour une étude fonctionnelle. Le premier gène codant un FT est spécifique de *L. maculans*, *Lmb\_jn3\_09123*. Le second, *Lmb\_jn3\_02170*, est présent chez toutes les espèces du complexe d'espèce de *L. maculans* (**Figure 1**).

J'ai inactivé ces deux facteurs de transcription par la technique CRISPR-Cas9. Le premier gène, *Lmb\_jn3\_09123*, est exprimé dès 2 jpi. J'ai obtenu 10 mutants indépendants présentant six types de mutations différentes menant pour cinq d'entre eux à la perte des domaines fonctionnels (**Figure 2**). Les mutants ont été caractérisés pour leur croissance et leur capacité à infecter un cultivar sensible de colza. Aucun des mutants ne présentait d'altération de croissance ou de pathogénie (**Figure 3**). Il reste cependant possible que ce FT soit impliqué dans la mise en place de l'infection étant donné son profil d'expression et les domaines prédits (deux domaines de liaison à la chitine, un domaine de liaison à l'ubiquitine et un domaine déacétylase de chitine ; **Figures 1 et 2**).



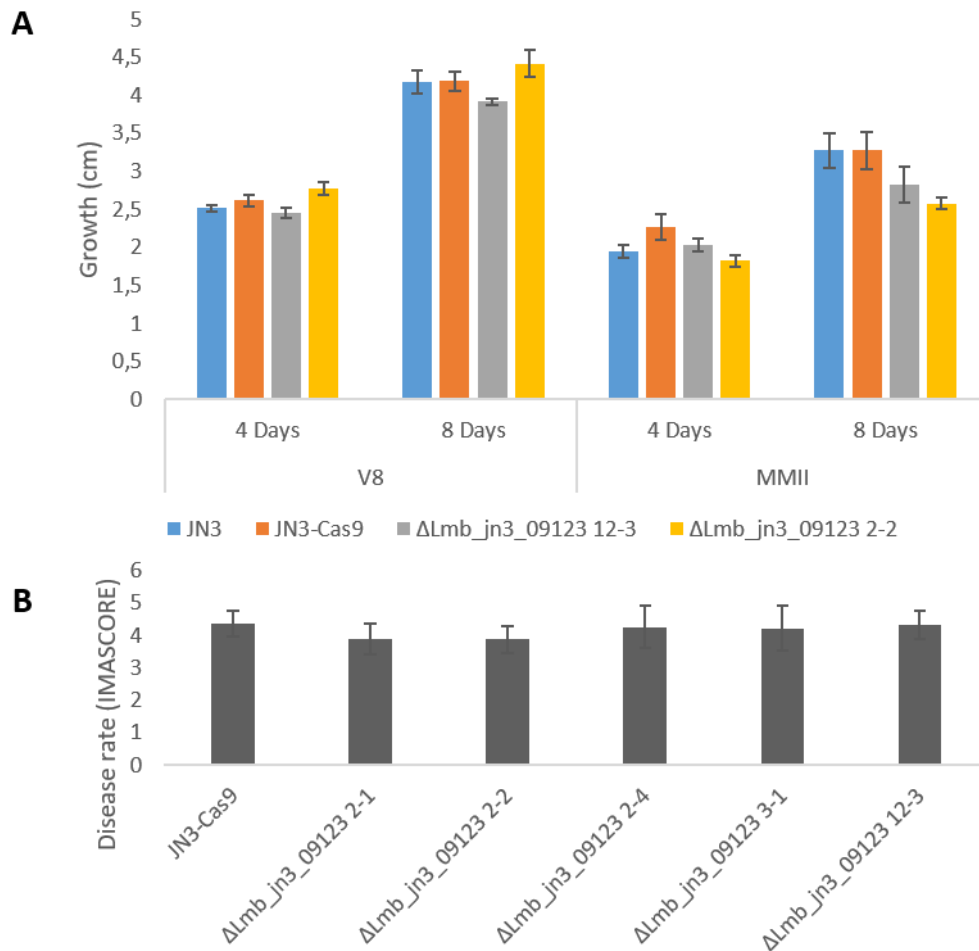


**Figure 1: Expression profile of cloned effector genes and three specific transcription factors during infection and axenic growth obtained by RNA-seq.** Mycelium were obtained by growing JN2 on V8 for 7 days. Cotyledons were sampled at 7 different time points (2, 5, 7, 9, 12, 15 days post infection). All samples were sequenced in two replicats. Cloned effector genes (*AvrLm4-7*, *AvrLm6*, *AvrLm2*, *AvrLm3*, *AvrLm10.1*, *AvrLm10.2*, *AvrLmCys2*, *AvrLm11*, *AvrLm5*) relative expression is plotted depending on the conditions. Transcription factors (*LmPf2*, *Lmb\_jn3\_09123* and *Lmb\_jn3\_02170*) relative expression is plotted depending on the same conditions (E.J. Gay, on-going PhD thesis, Data extracted from Leptolife project).



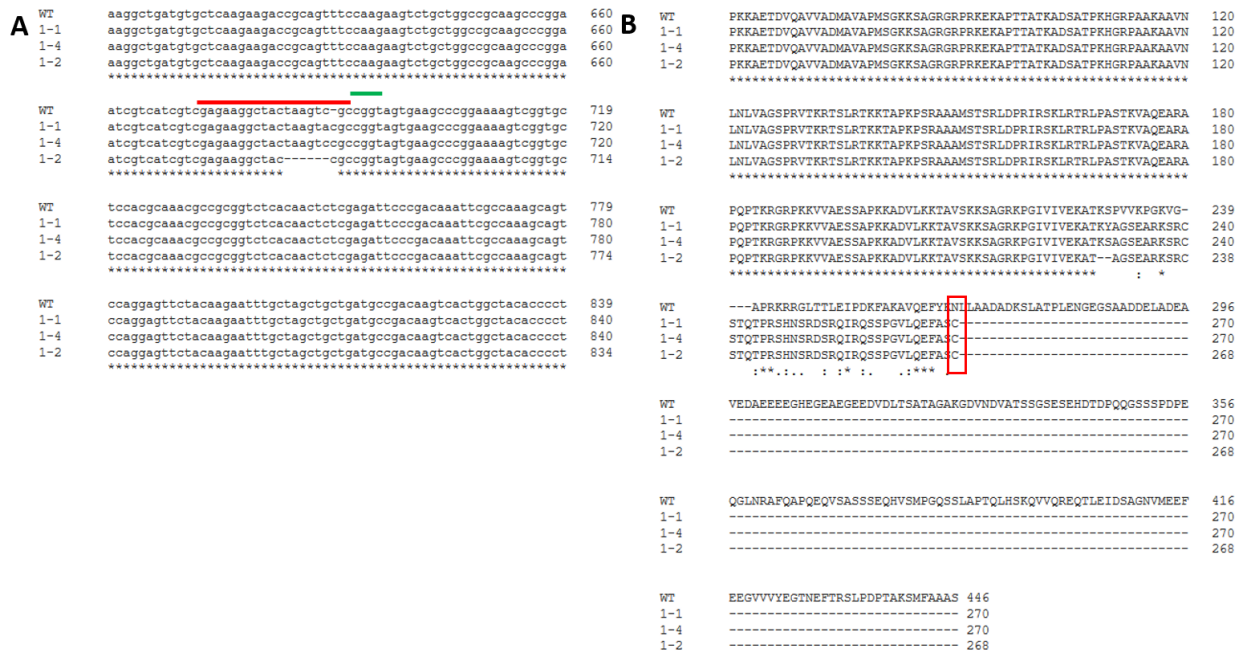
**Figure 2: Effect of the *Lmb\_jn3\_09123* mutations on *Lmb\_jn3\_09123* protein structure.** A: Sequence alignment of the *Lmb\_jn3\_09123* gene in the WT isolate JN2 and in six mutants. The PAM (Protospacer Adjacent Motif) is highlighted in green and the region targeted by the guide RNA is highlighted in red. B: Protein sequence of *Lmb\_jn3\_09123* in the WT isolate JN2 and in the six mutants. Light grey, dark greys and orange domains corresponding respectively to UBA, chitin binding domains and NODB domain. C: *Lmb\_jn3\_09123* protein length and domains identified with Pfam (UBA(Ubiquitin-associated domain), three chitin binding domains and NODB domain (chitin deacetylase domain)); <https://pfam.xfam.org/>.





**Figure 3: Effect of *Lmb\_jn3\_09123* inactivation on mycelial growth and pathogenicity of *Leptosphaeria maculans*.** **A:** Growth assay on *Lmb\_jn3\_09123* mutants and WT strain JN3 with or without *Cas9* gene. Mycelial growth was measured at 4 days and 8 days post inoculation on two different agar media, V8 and MMII. **B:** Pathogenicity tests were performed on the susceptible cultivar Es-Astrid at 13 DPI. Disease rate is expressed as the mean scoring using the IMAScore rating scale comprising six infection classes (IC), where IC1 to IC3 corresponded to resistance and IC4 to IC6, to susceptibility. DPI = Days Post Inoculation.

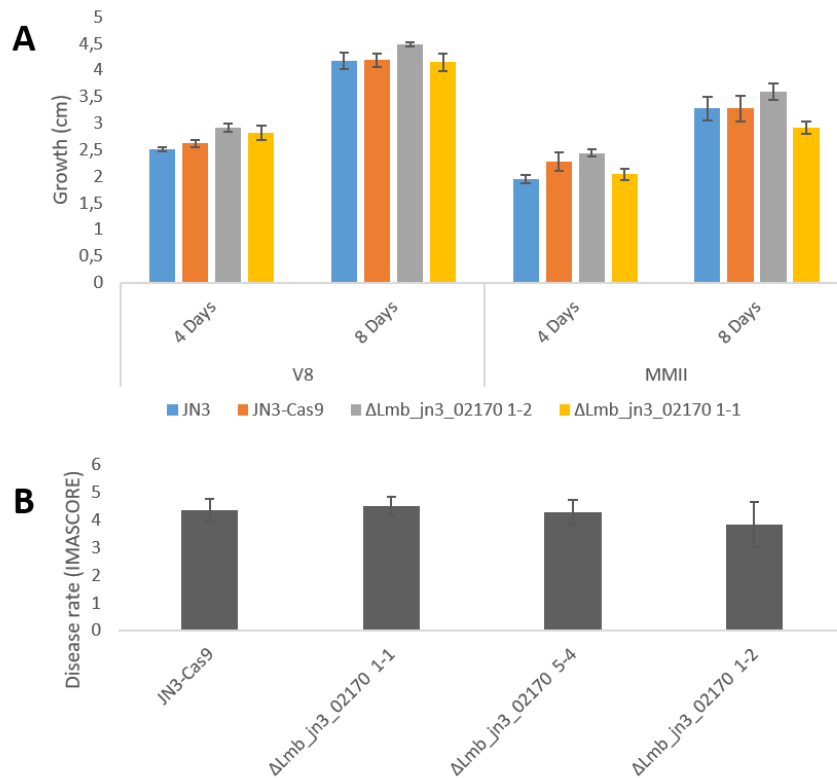
L'autre FT que j'ai étudié, *Lmb\_jn3\_02170*, possède un profil d'expression corrélé avec celui des effecteurs précoces (**Figure 1**). L'inactivation de *Lmb\_jn3\_02170* par CRISPR-Cas9 a permis d'obtenir trois types de mutations différentes et quatre mutants indépendants. Ces trois mutations mènent à protéine tronquée de la moitié de sa taille. Ce FT ne possède pas de domaine fonctionnel prédit et il est donc difficile d'identifier sa fonction (**Figure 4**). L'inactivation de *Lmb\_jn3\_02170* ne conduit à aucun défaut de croissance ou de pathogénie (**Figure 5**).



**Figure 4: Effect of the *Lmb\_jn3\_02170* mutations on *Lmb\_jn3\_02170* protein structure.** **A:** Sequence alignment of the *Lmb\_jn3\_02170* gene in the WT isolate JN2 and in three mutants. The PAM (Protospacer Adjacent Motif) is highlighted in green and the region targeted by the guide RNA is highlighted in red. **B:** Protein sequence of *Lmb\_jn3\_02170* in the WT isolate JN2 and in the three mutants. The red frame indicates the stop in the mutant versions of the protein. No domains have been predicted using Pfam (<https://pfam.xfam.org/>).







**Figure 5: Effect of *Lmb\_jn3\_02170* inactivation on mycelial growth and pathogenicity of *Leptosphaeria maculans*.** **A:** Growth assay on *Lmb\_jn3\_02170* mutants and WT strain JN3 with or without *Cas9* gene. Mycelial growth was measured at 4 days and 8 days post inoculation on two different agar media, V8 and MMII. **B:** Pathogenicity tests were performed on the susceptible cultivar Es-Astrid at 13 DPI. Disease rate is expressed as the mean scoring using the IMASCORE rating scale comprising six infection classes (IC), where IC1 to IC3 corresponded to resistance and IC4 to IC6, to susceptibility. DPI = Days Post Inoculation.

## DISCUSSION

Ce travail de thèse a été réalisé en forte interaction avec d'autres membres de l'UMR Bioger, dont l'équipe de bio-informatique de l'unité, le plateau de cytologie ainsi qu'avec Elise Gay qui réalise une thèse de bio-informatique dans notre équipe.

L'étude de l'action de protéines effectrices dans la relation entre un organisme phytopathogène et sa plante hôte est un sujet brûlant (Lo Presti *et al.*, 2015 ; Lo Presti et Kahmann, 2017 ; Irieda *et al.*, 2019). Les gènes codant des effecteurs sont souvent localisés dans des régions particulières des génomes fongiques et sont souvent corégulés (Schardl *et al.*, 2013 ; Soyer *et al.*, 2015,A ; van Dam *et al.*, 2017 ; Sánchez-Vallet *et al.*, 2018). Précédemment à ma thèse, l'étude de la régulation chromatinienne et transcriptionnelle de l'expression des effecteurs localisés dans les régions riches en ETs chez *L. maculans* avait été initiée par I. Fudal et J. Soyer. Notre équipe avait pu identifier que les régions riches en ETs étaient enrichies en effecteurs, notamment en gènes d'avirulence, sur-exprimés pendant l'infection primaire du colza. L'inactivation par ARNi de gènes codant des protéines impliquées de la mise en place de l'hétérochromatine (KMT1, HP1) avait permis de montrer un contrôle chromatinien de l'expression d'au moins deux gènes codant des effecteurs, via l'implication de la modification H3K9me3 qui réprime l'expression de ces gènes *in vitro* (Soyer *et al.*, 2014). L'analyse d'un contrôle transcriptionnel de l'expression des effecteurs précoces via des FTs spécifiques avait également été initiée. Ainsi, l'analyse fonctionnelle de StuA, connu pour réguler de nombreuses fonctions chez les champignons filamenteux (IpCho *et al.*, 2010 ; Ohara and Tsuge, 2004 ; Tong *et al.*, 2007), avait aussi été menée par ARNi. Il avait été montré que ce FT régule, de façon directe ou indirecte, l'expression d'au moins trois gènes codant des effecteurs localisés en régions riches en ETs (Soyer *et al.*, 2015. B). Depuis les premières analyses transcriptomiques qui avaient permis de montrer une expression spécifique des gènes codant des effecteurs pendant l'infection des cotylédons de colza (Rouxel *et al.*, 2011), d'autres analyses transcriptomiques ont été menées à d'autres étapes du cycle infectieux, et ainsi, d'autres vagues d'expression d'effecteurs ont également été identifiées (Gervais *et al.*, 2017, Gay *et al.*, en préparation).

L'objectif de ma thèse était de comprendre le rôle de la structure de la chromatine (à plusieurs échelles : position des nucléosomes, modification des histones et organisation intra-nucléaire des chromosomes) et de FTs dans le contrôle de l'expression concertée des gènes codant des effecteurs chez *L. maculans*, et plus globalement dans le contrôle de l'expression génique et de la pathogénie.

Pour comprendre comment ces mécanismes intervenaient dans le contrôle de l'expression génique, différentes approches ont été choisies. Une étude « genome-wide » de trois modifications des histones, de la localisation des nucléosomes, des interactions physiques du génome de *L. maculans* ont été entreprises. Une analyse *in vitro* des modifications chimiques des histones et de la localisation des nucléosomes a aussi été menée chez une espèce proche de *L. maculans* 'brassicae', *Leptosphaeria maculans* 'lepidii' (Lml), ainsi que de l'expression des gènes associés. Une étude fonctionnelle de remodeleurs chromatinien via une inactivation par CRISPR-Cas9 chez *L. maculans* et enfin une étude fonctionnelle de FTs possiblement impliqués dans cette régulation encore une fois via



une inactivation par CRISPR-Cas9 et une sur-expression dans différents fonds génétiques de *L. maculans*. Cette étude m'a permis d'identifier un FT essentiel à la mise en place de la pathogénie chez *L. maculans*. De plus la surexpression de ce FT dans différents contextes chromatinien m'a permis de mettre en évidence un double contrôle via la structure chromatinienne et ce FT spécifique de l'expression de différents sets de gènes codant notamment des effecteurs et exprimés de façon concertée tout au long du cycle de vie de *L. maculans*.

### Contrôle chromatinien de l'expression des gènes codant des effecteurs

Le contrôle chromatinien de l'expression des gènes codant des effecteurs n'a commencé à être investigué que récemment. Ainsi, seuls, quelques cas de régulation des gènes codant des effecteurs protéiques ou impliqués dans des clusters de métabolisme secondaire ont été analysés jusqu'à présent. Ces études tendent à montrer que les gènes codant des effecteurs fongiques, souvent localisés dans des régions particulières des génomes, sont souvent régulés par des modifications chimiques des queues d'histones (Strauss et Reyes-Dominguez, 2011 ; Connolly *et al.*, 2013 ; Chujo and Scott, 2014 ; Soyer *et al.*, 2015. A). La cartographie des territoires nucléosomiques est également un domaine de recherche innovant pour essayer de comprendre les dynamiques régissant l'expression des gènes (Ponts *et al.*, 2010 ; Locke *et al.*, 2013 ; Chereji *et al.*, 2016), qui n'avait encore pas été développé chez les champignons phytopathogènes. L'étude réalisée dans le cadre de ma thèse est, à ce jour, le seul exemple d'analyse comparative de distribution nucléosomique chez des champignons phytopathogènes à notre connaissance (**Chapitre 1 ; Article 1**). L'étude, par ChIP-Seq, de marques chromatinien à l'échelle du génome de *L. maculans* a permis de mettre en évidence la localisation des gènes codant des effecteurs, exprimés pendant les phases biotrophes d'infection, en environnement hétérochromatinien H3K9me3, H3K27me3 ou H3K9me3/H3K27me3. Ainsi nous avons pu déterminer que les gènes spécifiquement exprimés à 7 jpi sur cotylédons, notamment les gènes d'aviorulence, sont spécifiquement localisés en environnement H3K9me3 et H3K27me3 alors que les autres gènes codant des effecteurs et exprimés durant la phase de colonisation systémique de la tige de colza sont plutôt enrichis en domaines H3K27me3 (**Chapitre 1 ; Article 2**). De plus l'inactivation de *KMT1*, responsable du dépôt de la marque H3K9me3, a permis de déterminer que *KMT1* est impliqué dans la mise en place de l'infection car le mutant  $\Delta kmt1$  induit des symptômes plus réduits que le WT. *KMT1* est également impliqué dans l'expression des gènes codant des effecteurs puisque son inactivation induit l'expression de *Lmb\_jn3\_11364* à 7 jpi alors qu'il est normalement exprimé pendant les phases de colonisation tardives de la tige. *KMT1* est également impliqué dans la répression de l'expression des gènes d'aviorulence *AvrLm4-7*, *AvrLm6*, *AvrLm10.A* et d'*AvrLm11* pendant l'infection de la tige à 14 jpi. En résumé *KMT1* est essentiel à l'expression du set d'effecteurs adéquat à plusieurs étapes de l'infection. *KMT6* via la mise en place de la marque H3K27me3, joue un rôle plus limité dans la mise en place de l'infection et dans le contrôle de l'expression des effecteurs. Son inactivation ne mène à aucun défaut de pathogénie sur cotylédon mais le développement de la biomasse fongique dans la tige (28 jpi) est altéré. Enfin l'inactivation de *KMT6* n'a pas permis de détecter de défaut de régulation des effecteurs testés. *KMT6* est donc impliqué dans la réalisation totale du cycle infectieux mais n'influe pas directement sur l'expression des effecteurs testés / à moins qu'en cas d'inactivation de *KMT6*, une autre histone méthyltransferase prenne le relais (**Chapitre 2 ; Article 3**). Ces résultats confirment ceux précédemment obtenus par l'équipe concernant la régulation des gènes d'aviorulence par *KMT1* et *HP1* (Soyer *et al.*, 2014). L'inactivation d'une déméthylase d'histone,



Kdm8, présentant un profil d'expression similaire à celui des gènes d'avirulence n'a pas permis de mettre en évidence son implication dans le contrôle de l'expression de ces gènes. Il demeure néanmoins possible qu'elle soit impliquée dans le contrôle de l'expression de set de gènes non indispensables à la mise en place de l'infection. En effet l'implication de déméthylases d'histones dans la régulation de processus développementaux et dans la production de métabolites secondaires a été décrite chez *A. nidulans* (Gacek-Matthews *et al.*, 2015 ; Gacek-Matthews *et al.*, 2016).

Le contrôle de l'expression des gènes codant des effecteurs exprimés spécifiquement dans les tiges de colza semble plus complexe. En effet ces gènes sont statistiquement plus présents dans les régions marquées par la marque H3K27me3. Cependant, l'inactivation de KMT6 responsable du dépôt de cette marque n'entraîne pas de surexpression spécifique de ces gènes. Il semble donc que des régulateurs agissent en plus ou même sans qu'il n'y ait besoin d'éliminer H3K27me3 pour activer l'expression concertée des gènes codant les effecteurs exprimés pendant les phases tardives de l'infection. Il est également possible que le retrait des marques H3K27me3 mène à la relocalisation de modification H3K9me3 sur les sites normalement occupés par du H3K9me3 (Möller *et al.*, 2019). L'identification de FTs présentant le même profil d'expression que les effecteurs exprimés pendant les premières étapes de l'infection a permis d'identifier *LmPf2* lors de la thèse de J. Soyer puis de réaliser son analyse fonctionnelle durant ma thèse, permettant de mettre en évidence sa capacité à réguler la pathogénie via le contrôle successif de l'expression de sets de gènes exprimés sous forme de vagues concertées. Il serait donc intéressant de réaliser le même type d'analyse pour identifier des FTs co-régulés avec les gènes codant des effecteurs exprimés pendant les phases tardives de l'infection.

### Contrôle transcriptionnel de l'expression des gènes codant des effecteurs

Les FTs, représentent de 0.5 à 8% du génome des Eucaryotes et sont indispensables à la régulation du cycle de vie chez la plupart des organismes complexes (Weirauch et Hughes, 2011). Certains FTs sont généraux et régulent plusieurs mécanismes cellulaires (Ohara et Tsuge, 2004 ; Tong *et al.*, 2007 ; IpCho *et al.*, 2010). D'autres ont été identifiés comme directement impliqués dans la régulation de la pathogénie chez les Mycètes (Michielse *et al.*, 2009 ; Niño-Sánchez *et al.*, 2016). Chez les champignons phytopathogènes, la régulation du métabolisme secondaire ou de l'expression des effecteurs via des FTs spécifiques a déjà été relatée (Tollot *et al.*, 2016 ; van der Does *et al.*, 2016).

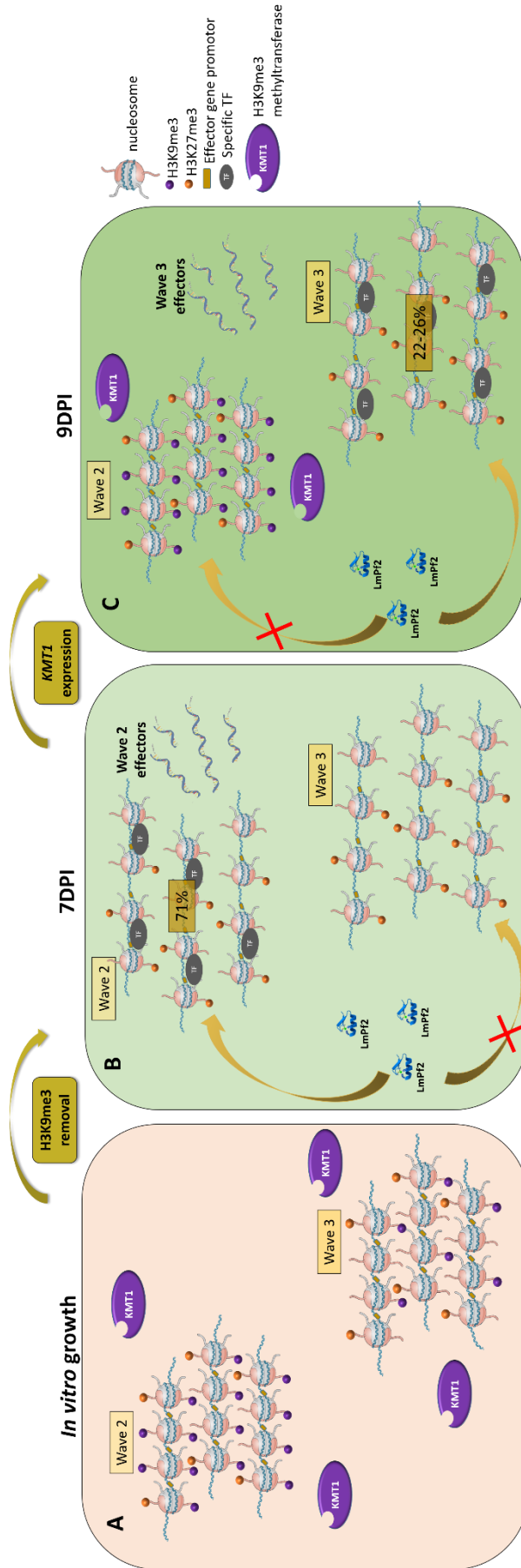
Pendant ma thèse j'ai menée l'inactivation par CRISPR-Cas9 et la surexpression de *LmPf2* décrit comme régulant la pathogénie et l'expression des effecteurs chez des Pleosporales proches de *L. maculans* (Cho *et al.*, 2013 ; Rybak *et al.*, 2017). Cette étude a permis de déterminer que ce FT est également impliqué dans la pathogénie chez *L. maculans*, son inactivation menant à une absence totale de développement de symptômes sur cotylédons de colza. De plus, cette étude a également permis de déterminer que ce FT contrôle l'expression de plusieurs sets d'effecteurs impliqués dans la phase de colonisation biotrophe du colza (**Figure 1 ; Chapitre 3 ; Article 4**). L'inactivation de deux autres FTs ayant des profils d'expression similaires à ceux des gènes d'avirulence de *L. maculans* n'a mené à aucun défaut de pathogénie. Il reste néanmoins possible que ces FTs contrôlent l'expression d'effecteurs non-essentiels à l'établissement de la pathogénie.



## Double contrôle, chromatinien et transcriptionnel, de l'expression des gènes codant des effecteurs précoces

L'étude menée pendant ma thèse est à notre connaissance la seule mettant en évidence l'implication de protéines impliquées dans le remodelage de la chromatine et de FTs dans le contrôle de l'expression des gènes codant des effecteurs protéiques chez les champignons phytopathogènes. Nous avons déterminé que la surexpression de *LmPf2* chez une souche sauvage contrôle spécifiquement les gènes de la vague 3 exprimés spécifiquement entre 9 et 12 jpi sur cotylédons alors que sa sur-expression dans un contexte  $\Delta kmt1$ , présentant une structure chromatinienne relâchée, va spécifiquement activer l'expression des gènes de la vague 2 composée de 91 gènes comportant les huit gènes d'avirulence clonés et spécifiquement exprimés à 7 jpi. Nous avons à partir de ces résultats proposé un modèle de régulation de l'expression des effecteurs chez *L. maculans* dans lequel le champignon pénètre dans la plante hôte et perçoit un signal végétal ou environnemental (demeurant inconnu). Le champignon va alors retirer les marques H3K9me3 au niveau des gènes codant des effecteurs en particulier ceux localisés dans les régions riches en ET, rendant les promoteurs de ces gènes accessibles à des FTs spécifiques tels que *LmPf2*. *LmPf2* est alors capable d'activer successivement 71% des gènes de la vague 2. Entre 9 et 12 jours après infection, *KMT1*, codant la méthylase d'histone responsable du dépôt des marques H3K9me3, est exprimée ce qui induirait une recondensation des régions hétérochromatiniennes et bloquant ainsi l'accessibilité aux promoteurs des gènes de la vague 2. *LmPf2* activerait alors 22-26 % des gènes de la vague 3. La production de ces effecteurs va permettre à *L. maculans* de coloniser les tissus de la plante en échappant à sa reconnaissance (**Figure 1 ; Chapitre 3 ; Article 4**).





**Figure 1: Model of key players of effector gene regulation as concerted waves during the early stage of *Leptosphaeria maculans* colonization of oilseed rape and *in vitro* growth.** During *Leptosphaeria maculans*' life cycle, LmPf2 and KMT1 are able to regulate effector gene expression respectively acting on effector gene expression and chromatin condensation. **A:** During axenic growth effector waves repertoire are tightly repressed by H3K9me3 maintained by KMT1. **B:** At 7 DPI H3K9me3 is fully removed and heterochromatin is loosened. At this stage *LmPf2* is highly expressed and is able to regulate specifically 71% of Wave 2 effector genes probably because of a better affinity for wave 2 effector promoters. Between 7 DPI and 9 DPI *KMT1* is being recondensed and start heterochromatin recondensation by H3K9me3 redeposition. **C:** Wave 2 effector genes are being recondensed and *LmPf2* start acting on (22 -26%) of Wave 3 effector genes, provoking their up-regulation. DPI = Days Post Inoculation.

## PERSPECTIVES

### Conformation 3D des génomes et contrôle de l'expression des gènes codant des effecteurs chez *L. maculans*

La compréhension de l'organisation 3D des génomes et le decryptage de la régulation génique par des assemblages protéiques complexes sont en pleine expansion. Il a été décrit que des parties entières de chromosomes peuvent être co-régulées en interagissant avec différents partenaires protéiques (Lanctôt *et al.*, 2007 ; Ulianov *et al.*, 2016). Chez les champignons phytopathogènes, une étude récente a permis de mettre en évidence une structuration du génome d'*Epichloë festucae* (ayant, comme *L. maculans*, une structure de génome en isochores) par les éléments répétés constituant les Isochores AT. On retrouve les gènes différentiellement exprimés *in planta* près des régions riches en ETs, cette structuration 3D pourrait donc être impliquée dans leur expression concertée *in planta* (Winter *et al.*, 2018). Ma thèse a permis d'initier l'étude de l'impact de la structure chromosomique sur la régulation génique à travers la création de mutants HMEs et une expériences de Hi-C chez ces mutants et chez une souche sauvage (échantillons sont en cours de séquençage) afin de définir les régions génomiques interagissant physiquement. Chez d'autres champignons filamenteux, les protéines KMT1 et KMT6 ont été montrées comme impliquées dans l'organisation des génomes. En effet, chez *N. crassa* et *E. festucae*, les méthylations H3K9me3 et H3K27me3 jouent un rôle crucial dans l'organisation du noyau en permettant les interactions entre régions hétérochromatiniennes (Klocko *et al.*, 2016 ; Winter *et al.*, 2018). Chez *N. crassa*, l'inactivation de KMT1 mène à la relocalisation des marques H3K27me3 à l'emplacement des modifications H3K9me3 ce qui induit de larges réarrangements chromosomiques. L'inactivation de KMT6 qui dépose normalement la marque H3K27me3 sur les régions subtélomériques induit une diminution des liaisons hétérochromatine-hétérochromatine, une augmentation des liaisons euchromatine-hétérochromatine et une destabilisation des liaisons entre télomères et enveloppe nucléaire (Klocko *et al.*, 2016). Chez *Z. tritici* également, on retrouve des modifications H3K27me3 sur les télomères ce qui permet leur interaction avec la membrane nucléaire. L'inactivation de KMT1 chez ce champignon induit de la même façon une relocalisation de H3K27me3 à l'emplacement des modifications H3K9me3 ce qui induit de larges réarrangements chromosomiques (Möller *et al.*, 2019). Au contraire l'inactivation de KMT6 n'induit que des réarrangements chromosomiques mineurs et en faible nombre. Chez *Fusarium oxysporum*, il existe des chromosomes dispensables, fortement enrichies en gènes différentiellement exprimés pendant l'infection et fortement enrichis en modifications H3K27me3 (Ma *et al.*, 2010). Les régions enrichies en H3K27me3 étant soumises à une évolution plus rapide (« fast » génome). Enfin, une étude récente a permis d'identifier que le core génome de cette espèce peut être divisé en deux catégories, le vrai « core » génome, enrichie en gènes de ménages et en modification H3K4me2 et le « fast core » génome enrichie en gènes différentiellement exprimés pendant l'infection (transport et défense). Les chromosomes dispensables, au contraire, sont spécifiquement enrichis en gènes codant des SSPs différentiellement exprimés pendant l'infection. Le génome à deux vitesses serait donc composé du vrai « core » génome et du « fast » génome constitué lui-même de plusieurs régions enrichies en H3K27me3 (chromosomes dispensables et « fast core » génome). Il serait également possible que des réarrangements entre « fast core » génome et chromosomes dispensables se produisent augmentant grandement la diversité génétique du pathogène (Fokkens *et al.*, 2018). Chez *L. maculans* qui possède aussi des chromosomes dispensables (Balesdent *et al.*, 2013). le « fast » genome pourrait être composé du « fast core » genome



composé des isochores AT enrichies en H3K9me3 et des régions GC-riche enrichies en H3K27me3 ce qui pourrait constituer en plus du chromosome dispensable deux types distincts de « fast » génome.

### Différents répertoires d'effecteurs pour différentes phases d'infection

Lors de ce travail de thèse j'ai pu mettre en évidence que les gènes codant des effecteurs exprimés pendant les premières étapes de l'infection du colza par *L. maculans* 'brassicae' (Lmb), étaient, au moins partiellement, contrôlés par LmPf2. Les effecteurs exprimés précocement sont décrits comme impliqués dans le contournement des réactions de défense de la plante et comme facilitant le développement fongique dans la plante hôte. De plus, il a été montré que la compartimentalisation des génomes d'organismes phytopathogènes filamenteux, avec des régions génomiques enrichies en effecteurs et en TEs évoluant plus rapidement, apporte un avantage sélectif à ces derniers (Raffaele and Kamoun, 2012 ; van Dam *et al.*, 2017). Je fais donc l'hypothèse que les effecteurs exprimés durant les phases plus tardive de l'infection, localisés dans les régions riches en gènes du génome de *L. maculans* 'brassicae', constituent le répertoire d'origine de l'ancêtre commun entre *L. maculans* 'brassicae', *L. biglobosa* et 'lepidii'(Lml) mais qu'une invasion du génome de Lmb par des TEs lui a permis de diversifier son répertoire d'effecteurs et d'acquérir, progressivement, une nouvelle phase d'infection, la phase biotrophe, permettant une colonisation primaire de la plante plus efficace avant de passer à une phase nécrotrophe comme celle des autres espèces du complexe. Cette capacité aurait été obtenue en acquérant progressivement un (ou plusieurs) set d'effecteurs permettant d'échapper à la reconnaissance de la plante hôte. Cette hypothèse est soutenue par l'existence d'espèces proches ayant la même structure genomique et chromatinienne telles qu'*Epichloë festucae* qui présente un mode de vie endophyte (Winter *et al.*, 2018). Il semble donc que les Isochores AT enrichis en gènes codant des effecteurs servent au champignon à la colonisation biotrophe de leur plante hôte. Cependant, l'utilisation du modèle Lmb/Lml est limité pour comprendre ces mécanismes car nous ne possédons pas d'hôte pour étudier l'infection de Lml et celui-ci n'est pas capable d'infecter le colza. Il serait donc à l'avenir plus intéressant d'étudier la spécialisation à l'hôte et la régulation du cycle infectieux en comparant *L. maculans* 'brassicae'/*Leptosphaeria biglobosa* 'brassicae' qui est aussi capable d'infecter le colza mais ne possède pas de grands isochores AT comme chez Lmb et n'initie pas son cycle infectieux par une phase de colonisation biotrophe. Il serait donc intéressant de comprendre si ce sont bien les isochores AT marqués par du H3K9me3 et enrichis en gènes codant des effecteurs qui sont responsables de la colonisation biotrophe initiale de l'hôte.

### Applications liées à la compréhension des mécanismes épigénétiques et transcriptionnels

Bien que très fondamental, le projet de thèse que j'ai mené a tout de même des applications concrètes. En effet, la compréhension des mécanismes épigénétiques régissant l'expression des gènes codant des effecteurs et ainsi la colonisation de la plante hôte pourrait permettre de développer de nouveaux moyens de lutte contre les champignons phytopathogènes. On peut par exemple penser à des traitement chimiques agissant sur les protéines HMEs et ainsi empêcher la décondensation de la chromatine et l'expression des effecteurs. La limitation de ce type de traitements est que les HMEs sont très conservées au sein des Eucaryote et qu'il sera difficile de cibler les HMEs de l'agent pathogène sans inactiver celles de l'hôte. Cependant, il a été prouvé récemment que les plantes et leurs agents





pathogènes sont capables de manipuler le système de défense adverse à l'aide de siRNA (Weiberg *et al.*, 2013 ; Wang *et al.*, 2016). Il est ainsi possible de créer des plantes OGM produisant des siRNA dirigés contre les domaines fonctionnels très conservés de protéines essentielles à la régulation des effecteurs (e.g. le site Zn2Cys6 de LmPf2 ou les domaines catalytiques de KMT1). Du fait du fort niveau de conservation de ce type de domaine, la possibilité de mutation et de conservation de fonction sera limitée et le contournement de la résistance sera, du même coup, difficile. Il serait également possible de traiter le champ directement avec des siRNA dirigés contre ce type de domaines fonctionnels.

### Valorisation du matériel produit pendant la thèse

A mon arrivée au laboratoire, nous utilisons la technique RNAi pour étudier fonctionnellement un gène. Cette technique a prouvé son efficacité, notamment pour l'identification de gènes d'avirulence, mais également ses limites quand il s'agit de l'étude de FTs ou de protéines modificatrices d'histones (HMEs) qui sont très spécifiques, parfois vitales pour la cellule, avec des régulations complexes et des fenêtres d'expression limitées. J'ai donc au cours de ma thèse transféré l'outil de mutagenèse dirigée, CRISPR-Cas9, au laboratoire, technique qui m'a permis d'inactiver les gènes codant les FTs et HMEs que nous voulions étudier sans rencontrer de soucis relatifs au niveau d'expression du transgène. Cette technique a nécessité la transformation successive de *L. maculans* par la construction portant la Cas9 puis par la construction contenant l'ARN guide. Les transformants ont ensuite été cultivés jusqu'à ce qu'une mutation apparaisse puis croisés avec une souche sauvage (GFP ou non) pour éliminer la Cas9 et l'ARN guide. Cette méthode sera par la suite très utile à l'équipe puisque l'inactivation des gènes par recombinaison homologue au locus chez *L. maculans* est très difficile (de l'ordre de 1/300) comme chez beaucoup de champignons du fait de la présence de gènes KU70 et KU80 (Impliqués dans la recombinaison non-homologue ; Gardiner et Howlett, 2004). L'inactivation par CRISPR-Cas9 de Ku70/Ku80 pourra même par la suite permettre de réaliser des recombinaison homologues chez *L. maculans* et faciliter grandement les études fonctionnelles.

Notre équipe à l'origine développait des approches de génétique formelle pour identifier des gènes d'avirulence de *L. maculans* et les gènes de résistance associés chez le colza. Pendant ma thèse j'ai pu obtenir des transformants surexprimant *LmPf2* dans un contexte sauvage ou dans un mutant  $\Delta kmt1$ . En fonction du contexte chromatinien et du niveau d'expression de *LmPf2*, différents sets de gènes codant des effecteurs sont sur-exprimés, parfois à de très hauts niveaux (20 000 fois plus). On pourrait donc faire croître ces souches en milieu liquide (Fries) et utiliser le surnageant de cultures pour obtenir des mélanges protéiques pouvant permettre de cribler des génotypes de colza pour la présence de gènes de résistance non encore identifiés. Nous pourrions également découvrir de nouvelles molécules d'intérêt, par ex. des métabolites secondaires, en inactivant et/ou surexprimant des HMEs/FTs connus pour déréguler le métabolisme secondaire chez d'autres champignons et produire des molécules habituellement non produites en condition de culture axénique.



## CONCLUSION

Lorsque j'ai commencé ma thèse, J. Soyer avait déjà démontré l'implication de H3K9me3 déposée par KMT1 dans la répression de l'expression d'au moins deux gènes codant des effecteurs. Mais leur surexpression *in planta* n'était pas uniquement dû au retrait de modifications chromatinienne, ce qui suggérait l'intervention de FTs spécifiques. Au cours de cette thèse, nous avons pu cartographier à l'échelle du génome : (i) trois modifications chromatinienne dans les génomes de *L. maculans* 'brassicae' et d'une espèce proche *L. maculans* 'lepidii' et (ii) la répartition nucléosomale chez quatre espèces de champignons phytopathogènes. J'ai également pu mettre en évidence un double contrôle par KMT1 et LmPf2 de différents sets de gènes très enrichies en effecteurs et exprimés de façon concertée au cours de l'infection. C'est à ce jour la première étude mettant en lumière une double régulation génique par une HME et un FT spécifique. De plus nous avons pu déterminer l'implication de deux méthylase d'histone, KMT1 et KMT6 dans la mise en place de la pathogénèse et le contrôle des gènes codant des effecteurs. Enfin nous avons initié au cours de ma thèse des expériences de Hi-C Seq couplées à du ChIP-Seq et RNA-Seq *in vitro* qui nous renseigneront grandement à la fois sur l'implication des modifications chromatinienne sur la structure du génome et sur l'expression des gènes.

## BIBLIOGRAPHIE

Andrulis, E.D., Neiman, A.M., Zappulla, D.C., Sternglanz, R., 1998. Perinuclear localization of chromatin facilitates transcriptional silencing. *Nature* 394, 592–595.

Antequera, F., Tamame, M., Villanueva, J.R., Santos, T., 1984. DNA methylation in the fungi. *J. Biol. Chem.* 259, 8033–8036.

Aramayo R., Peleg Y, Addison R, Metzenberg R., 1996. *Asm-1+*, a *Neurospora crassa* gene related to transcriptional regulators of fungal development. *Genetics* 144: 991-1003.

Balesdent, M.H., Fudal, I., Ollivier, B., Bally, P., Grandaubert, J., Eber, F., Chèvre, A.-M., Leflon, M., Rouxel, T., 2013. The dispensable chromosome of *Leptosphaeria maculans* shelters an effector gene conferring avirulence towards *Brassica rapa*. *New Phytologist* 198, 887–898.

Bannister, A.J., Kouzarides, T., 2005. Reversing histone methylation. *Nature* 436, 1103–1106.

Bantignies, F., Grimaud, C., Lavrov, S., Gabut, M., Cavalli, G., 2003. Inheritance of Polycomb-dependent chromosomal interactions in *Drosophila*. *Genes Dev.* 17, 2406–2420.

Basenko, E.Y., Sasaki, T., Ji, L., Prybol, C.J., Burckhardt, R.M., Schmitz, R.J., Lewis, Z.A., 2015. Genome-wide redistribution of H3K27me3 is linked to genotoxic stress and defective growth. *Proceedings of the National Academy of Sciences* 112, E6339–E6348.

Belaghzal, H., Dekker, J., Gibcus, J.H., 2017. HI-C 2.0: AN OPTIMIZED HI-C PROCEDURE FOR HIGH-RESOLUTION GENOME-WIDE MAPPING OF CHROMOSOME CONFORMATION. *Methods* 123, 56–65.

Bewick, A.J., Hofmeister, B.T., Powers, R.A., Mondo, S.J., Grigoriev, I.V., James, T.Y., Stajich, J.E., Schmitz, R.J., 2019. Diversity of cytosine methylation across the fungal tree of life. *Nat Ecol Evol* 3, 479–490.



Bird, A., 2002. DNA methylation patterns and epigenetic memory. *Genes Dev.* 16, 6–21.

Bonev, B., Cavalli, G., 2016. Organization and function of the 3D genome. *Nature Reviews Genetics* 17, 661–678.

Bonneuil, C. and Thomas, F., 2012. Semences, une histoire politique : amélioration des plantes, agriculture et alimentation en France depuis la Seconde Guerre mondiale, ECLM.

Brundrett, M.C., 2002. Coevolution of roots and mycorrhizas of land plants. *New phytologist* 154, 275–304.

Bunnik, E.M., Polishko, A., Prudhomme, J., Ponts, N., Gill, S.S., Lonardi, S., Le Roch, K.G., 2014. DNA-encoded nucleosome occupancy is associated with transcription levels in the human malaria parasite *Plasmodium falciparum*. *BMC Genomics* 15, 347.

Carhart-Harris, R.L., Roseman, L., Bolstridge, M., Demetriou, L., Pannekoek, J.N., 2017. Psilocybin for treatment-resistant depression: fMRI-measured brain mechanisms. *Scientific Reports* 7.

Carrillo, A.J., Schacht, P., Cabrera, I.E., Blahut, J., Prudhomme, L., Dietrich, S., Bekman, T., Mei, J., Carrera, C., Chen, V., Clark, I., Fierro, G., Ganzen, L., Orellana, J., Wise, S., Yang, K., Zhong, H., Borkovich, K.A., 2017. Functional Profiling of Transcription Factor Genes in *Neurospora crassa*. *G3 (Bethesda)* 7, 2945–2956.

Casolari, J.M., Brown, C.R., Komili, S., West, J., Hieronymus, H., Silver, P.A., 2004. Genome-Wide Localization of the Nuclear Transport Machinery Couples Transcriptional Status and Nuclear Organization. *Cell* 117, 427–439.

Cedar, H., Bergman, Y., 2009. Linking DNA methylation and histone modification: patterns and paradigms. *Nature Reviews Genetics* 10, 295–304.

Chalhoub, B., Denoeud, F., Liu, S., et al. 2014. Early allopolyploid evolution in the post-Neolithic *Brassica napus* oilseed genome. *Science* 345, 950–953.

Chereji, R.V., Kan, T.-W., Grudniewska, M.K., Romashchenko, A.V., Berezikov, E., Zhimulev, I.F., Guryev, V., Morozov, A.V., Moshkin, Y.M., 2016. Genome-wide profiling of nucleosome sensitivity and chromatin accessibility in *Drosophila melanogaster*. *Nucleic Acids Research* 44, 1036–1051.

Cho, Y., Ohm, R.A., Grigoriev, I.V., Srivastava, A., 2013. Fungal-specific transcription factor AbPf2 activates pathogenicity in *Alternaria brassicicola*. *The Plant Journal* 75, 498–514.

Chujo, T., Scott, B., 2014. Histone H3K9 and H3K27 methylation regulates fungal alkaloid biosynthesis in a fungal endophyte-plant symbiosis: K9 and K27 methylation regulates symbiosis. *Molecular Microbiology* 92, 413–434.

Coleman, J.J., Rounsley, S.D., Rodriguez-Carres, M., Kuo, A., Wasmann, C.C., Grimwood, J., Schmutz, J., Taga, M., White, G.J., Zhou, S., Schwartz, D.C., Freitag, M., Ma, L., Danchin, E.G.J., Henrissat, B., Coutinho, P.M., Nelson, D.R., Straney, D., Napoli, C.A., Barker, B.M., Gribskov, M., Rep, M., Kroken, S., Molnár, I., Rensing, C., Kennell, J.C., Zamora, J., Farman, M.L., Selker, E.U., Salamov, A., Shapiro, H., Pangilinan, J., Lindquist, E., Lamers, C., Grigoriev, I.V., Geiser, D.M., Covert, S.F., Temporini, E., VanEtten, H.D., 2009. The Genome of *Nectria haematococca*: Contribution of Supernumerary Chromosomes to Gene Expansion. *PLOS Genetics* 5, e1000618.

Collemare, J., Seidl, M.F., 2019. Chromatin-dependent regulation of secondary metabolite biosynthesis in fungi: is the picture complete? *FEMS Microbiol. Rev.*



Colot, H.V., Park, G., Turner, G.E., Ringelberg, C., Crew, C.M., Litvinkova L., Weiss, R.L., Borkovich, K.A., and Dunlap, J.C., 2006. A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. *PNas*, 103, 10352–10357.

Connolly, L.R., Smith, K.M., Freitag, M., 2013. The *Fusarium graminearum* Histone H3 K27 Methyltransferase KMT6 Regulates Development and Expression of Secondary Metabolite Gene Clusters. *PLoS Genetics* 9, e1003916.

Cooke, R.C., Whipps, J.M., 1980. The evolution of modes of nutrition in fungi parasitic on terrestrial plants. *Biological Reviews* 55, 341–362.

Cremer, T., Cremer, C., 2001. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. *Nature Reviews Genetics* 2, 292–301.

van Dam, S., Vösa, U., van der Graaf, A., Franke, L., de Magalhães, J.P., 2018. Gene co-expression analysis for functional classification and gene–disease predictions. *Brief Bioinform* 19, 575–592.

Deising, H.B., Werner, S., Wernitz, M., 2000. The role of fungal appressoria in plant infection. *Microbes and infection* 2, 1631–1641.

Dhillon, B., Cavaletto, J.R., Wood, K.V., Goodwin, S.B., 2010. Accidental amplification and inactivation of a methyltransferase gene eliminates cytosine methylation in *Mycosphaerella graminicola*. *Genetics* 186, 67–77.

Dilmaghani, A., Balesdent, M.H., Didier, J.P., Wu, C., Davey, J., Barbetti, M.J., Li, H., Moreno-Rico, O., Phillips, D., Despeghel, J.P., Vincenot, L., Gout, L., Rouxel, T., 2009. The *Leptosphaeria maculans* – *Leptosphaeria biglobosa* species complex in the American continent. *Plant Pathology* 58, 1044–1058.

Dilmaghani, A., Gladieux, P., Gout, L., Giraud, T., Brunner, P.C., Stachowiak, A., Balesdent, M.-H., Rouxel, T., 2012. Migration patterns and changes in population biology associated with the worldwide spread of the oilseed rape pathogen *Leptosphaeria maculans*. *Molecular Ecology* 21, 2519–2533.

van der Does, H.C., Fokkens, L., Yang, A., Schmidt, S.M., Langereis, L., Lukasiewicz, J.M., Hughes, T.R., Rep, M., 2016. Transcription factors encoded on core and accessory chromosomes of *Fusarium oxysporum* induce expression of effector genes. *PLoS genetics* 12, e1006401.

Dunlap, J.C., Borkovich, K.A., Henn, M.R., Turner, G.E., Sachs, M.S., Glass, N.L., McCluskey, K., Plamann, M., Galagan, J.E., Birren, B.W., Weiss, R.L., Townsend, J.P., Loros, J.J., Nelson, M.A., Lambregts, R., Colot, H.V., Park, G., Collopy, P., Ringelberg, C., Crew, C., Litvinkova, L., DeCaprio, D., Hood, H.M., Curilla, S., Shi, M., Crawford, M., Koerhsen, M., Montgomery, P., Larson, L., Pearson, M., Kasuga, T., Tian, C., Baştürkmen, M., Altamirano, L., Xu, J., 2007. Enabling a Community to Dissect an Organism: Overview of the *Neurospora* Functional Genomics Project, in: *Advances in Genetics*. Elsevier, pp. 49–96.

Duplessis, S., Cuomo, C.A., Lin, Y.-C., Aerts, A., Tisserant, E., Veneault-Fourrey, C., Joly, D.L., Hacquard, S., Amselem, J., Cantarel, B.L., Chiu, R., Coutinho, P.M., Feu, N., Field, M., Frey, P., Gelhaye, E., Goldberg, J., Grabherr, M.G., Kodira, C.D., Kohler, A., Kues, U., Lindquist, E.A., Lucas, S.M., Mago, R., Mauceli, E., Morin, E., Murat, C., Pangilinan, J.L., Park, R., Pearson, M., Quesneville, H., Rouhier, N., Sakthikumar, S., Salamov, A.A., Schmutz, J., Selles, B., Shapiro, H., Tanguay, P., Tuskan, G.A., Henrissat, B., Van de Peer, Y., Rouze, P., Ellis, J.G., Dodds, P.N., Schein, J.E., Zhong, S., Hamelin, R.C., Grigoriev, I.V., Szabo, L.J., Martin, F., 2011. Obligate biotrophy features unraveled by the genomic analysis of rust fungi. *Proceedings of the National Academy of Sciences* 108, 9166–9171.



Dutreux, F., Da Silva, C., d'Agata, L., Couloux, A., Gay, E.J., Istace, B., Lapalu, N., Lemainque, A., Linglin, J., Noel, B., Wincker, P., Cruaud, C., Rouxel, T., Balesdent, M.-H., Aury, J.-M., 2018. De novo assembly and annotation of three *Leptosphaeria* genomes using Oxford Nanopore MinION sequencing. *Scientific Data* 5, 180235.

Eberharter, A., 2002. Histone acetylation: a switch between repressive and permissive chromatin: Second in review series on chromatin dynamics. *EMBO Reports* 3, 224–229

Farnham, P.J., 2009. Insights from genomic profiling of transcription factors. *Nature Reviews Genetics* 10, 605–616.

Fisher, M.C., Henk, D.A., Briggs, C.J., Brownstein, J.S., Madoff, L.C., McCraw, S.L., Gurr, S.J., 2012. Emerging fungal threats to animal, plant and ecosystem health. *Nature* 484, 186–194.

Fitt, B.D.L., Brun, H., Barbetti, M.J., Rimmer, S.R., 2006. World-Wide Importance of Phoma Stem Canker (*Leptosphaeria maculans* and *L. biglobosa*) on Oilseed Rape (*Brassica napus*). *European Journal of Plant Pathology* 114, 3–15.

Fokkens, L., Shahi, S., Connolly, L.R., Stam, R., Schmidt, S.M., Smith, K.M., Freitag, M., Rep, M., 2018. The multi-speed genome of *Fusarium oxysporum* reveals association of histone modifications with sequence divergence and footprints of past horizontal chromosome transfer events. *bioRxiv*.

Foss, H.M., Roberts, C.J., Claeys, K.M., Selker, E.U., 1993. Abnormal chromosome behavior in *Neurospora* mutants defective in DNA methylation. *Science* 262, 1737–1741.

Francis, N.J., Kingston, R.E., 2001. Mechanisms of transcriptional memory. *Nat. Rev. Mol. Cell Biol.* 2, 409–421.

Freitag, M., 2017. Histone Methylation by SET Domain Proteins in Fungi. *Annual Review of Microbiology* 71, 413–439.

Gacek, A., Strauss, J., 2012. The chromatin code of fungal secondary metabolite gene clusters. *Appl Microbiol Biotechnol* 95, 1389–1404.

Gacek-Matthews, A., Noble, L.M., Gruber, C., Berger, H., Sulyok, M., Marcos, A.T., Strauss, J., Andrianopoulos, A., 2015. KdmA, a histone H3 demethylase with bipartite function, differentially regulates primary and secondary metabolism in *Aspergillus nidulans*: Histone H3K9 demethylase function in growth. *Molecular Microbiology* 96, 839–860.

Gacek-Matthews, A., Berger, H., Sasaki, T., Wittstein, K., Gruber, C., Lewis, Z.A., Strauss, J., 2016. KdmB, a Jumonji Histone H3 Demethylase, Regulates Genome-Wide H3K4 Trimethylation and Is Required for Normal Induction of Secondary Metabolism in *Aspergillus nidulans*. *PLOS Genetics* 12, e1006222.

Galagan, J.E., Selker, E.U., 2004. RIP: the evolutionary cost of genome defense. *Trends in Genetics* 20, 417–423.

Galazka, J.M., Klocko, A.D., Uesaka, M., Honda, S., Selker, E.U., Freitag, M., 2016. *Neurospora* chromosomes are organized by blocks of importin alpha-dependent heterochromatin that are largely independent of H3K9me3. *Genome Research* 26, 1069–1080.

Garcia-Diaz, M., Bebenek, K., 2007. Multiple Functions of DNA Polymerases. *Critical Reviews in Plant Sciences* 26, 105–122.



- Gardiner, D.M., Howlett, B.J., 2004. Negative selection using thymidine kinase increases the efficiency of recovery of transformants with targeted genes in the filamentous fungus *Leptosphaeria maculans*. *Curr. Genet.* 45, 249–255.
- Gerke, J., Braus, G.H., 2014. Manipulation of fungal development as source of novel secondary metabolites for biotechnology. *Applied Microbiology and Biotechnology* 98, 8443–8455.
- Gervais, J., Plissonneau, C., Linglin, J., Meyer, M., Labadie, K., 2017. Different waves of effector genes with contrasted genomic location are expressed by *Leptosphaeria maculans* during cotyledon and stem colonization of oilseed rape. *Molecular plant pathology* 18, 1113–1126.
- Gimeno, C.J., Fink GR., 1994. Induction of pseudohyphal growth by overexpression of PHD1, a *Saccharomyces cerevisiae* gene related to transcriptional regulators of fungal development. *Mol Cell Biol* 14: 2100-2112.
- Goyon, C., Barry, C., Grégoire, A., Faugeron, G., Rossignol, J.L., 1996. Methylation of DNA repeats of decreasing sizes in *Ascobolus immersus*. *Mol Cell Biol* 16, 3054–3065.
- Gowher, H., Ehrlich, K.C., Jeltsch, A., 2001. DNA from *Aspergillus flavus* contains 5-methylcytosine. *FEMS Microbiol Lett* 205, 151–155.
- Grandaubert, J., Lowe, R.G., Soyer, J.L., Schoch, C.L., Van de Wouw, A.P., 2014. Transposable element-assisted evolution and adaptation to host plant within the *Leptosphaeria maculans*-*Leptosphaeria biglobosa* species complex of fungal pathogens. *BMC genomics* 15, 891.
- Grewal, S.I.S., Jia, S., 2007. Heterochromatin revisited. *Nature Reviews Genetics* 8, 35–46.
- Hou, C., Li, L., Qin, Z.S., Corces, V.G., 2012. Gene Density, Transcription, and Insulators Contribute to the Partition of the *Drosophila* Genome into Physical Domains. *Molecular Cell* 48, 471–484.
- Harrison, M.J., 2012. Cellular programs for arbuscular mycorrhizal symbiosis. *Current Opinion in Plant Biology* 15, 691–698.
- Hassing, B., Winter, D., Becker, Y., Mesarich, C.H., Eaton, C.J., Scott, B., 2019. Analysis of *Epichloë festucae* small secreted proteins in the interaction with *Lolium perenne*. *PLoS ONE* 14, e0209463.
- Hebinger, H. 2013. *Le Colza*, Editions France Agricole.
- Heitz E., 1928. Das heterochromatin der moose. *Jahrb Wiss Botanik* 69: 762-818.
- Horbach, R., Navarro-Quesada, A.R., Knogge, W., Deising, H.B., 2011. When and how to kill a plant cell: Infection strategies of plant pathogenic fungi. *Journal of Plant Physiology* 168, 51–62.
- Huisinga, K.L., Brower-Toland, B., Elgin, S.C.R., 2006. The contradictory definitions of heterochromatin: transcription and silencing. *Chromosoma* 115, 110–122.
- IpCho, S.V.S., Tan, K.-C., Koh, G., Gummer, J., Oliver, R.P., Trengove, R.D., Solomon, P.S., 2010. The Transcription Factor StuA Regulates Central Carbon Metabolism, Mycotoxin Production, and Effector Gene Expression in the Wheat Pathogen *Stagonospora nodorum*. *Eukaryotic Cell* 9, 1100–1108.
- Irieda, H., Inoue, Y., Mori, M., Yamada, K., Oshikawa, Y., Saitoh, H., Uemura, A., Terauchi, R., Kitakura, S., Kosaka, A., Singkaravanit-Ogawa, S., Takano, Y., 2019. Conserved fungal effector suppresses PAMP-triggered immunity by targeting plant immune kinases. *Proceedings of the National Academy of Sciences* 116, 496–505.



- Jackson, J.P., Lindroth, A.M., Cao, X., Jacobsen, S.E., 2002. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. *Nature* 416, 556–560.
- Jamieson, K., Rountree, M.R., Lewis, Z.A., Stajich, J.E., Selker, E.U., 2013. Regional control of histone H3 lysine 27 methylation in *Neurospora*. *Proceedings of the National Academy of Sciences* 110, 6027–6032.
- Janevska, S., Baumann, L., Sieber, C.M.K., Münsterkötter, M., Ulrich, J., Kämper, J., Güldener, U., Tudzynski, B., 2018. Elucidation of the Two H3K36me3 Histone Methyltransferases Set2 and Ash1 in *Fusarium fujikuroi* Unravels Their Different Chromosomal Targets and a Major Impact of Ash1 on Genome Stability. *Genetics* 208, 153–171.
- Jenuwein, T., Allis, C.D., 2001. Translating the histone code. *Science* 293, 1074–1080.
- Jeon, J., Choi, J., Lee, G.-W., Park, S.-Y., Huh, A., Dean, R.A., Lee, Y.-H., 2015. Genome-wide profiling of DNA methylation provides insights into epigenetic regulation of fungal development in a plant pathogenic fungus, *Magnaporthe oryzae*. *Scientific Reports* 5, 8567.
- Jupe, J., Stam, R., Howden A.J.M., Morris J.A., Zhang R., 2013. *Phytophthora capsici*-tomato interaction features dramatic shifts in gene expression associated with a hemi-biotrophic lifestyle. *Genome Biology*,14:R63
- van Kan, J.A.L., 2006. Licensed to kill: the lifestyle of a necrotrophic plant pathogen. *Trends in Plant Science* 11, 247–253.
- Khatter, H., Vorländer, M.K., Müller, C.W., 2017. RNA polymerase I and III: similar yet unique. *Current Opinion in Structural Biology* 47, 88–94.
- Kim, J.-A., Kwon, M., Kim, J., 2019. Allosteric Regulation of Chromatin-Modifying Enzymes. *Biochemistry* 58, 15–23.
- Klocko, A.D., Ormsby, T., Galazka, J.M., Leggett, N.A., Uesaka, M., Honda, S., Freitag, M., Selker, E.U., 2016. Normal chromosome conformation depends on subtelomeric facultative heterochromatin in *Neurospora crassa*. *PNAS* 113, 15048–15053.
- Kouzarides, T., 2007. Chromatin Modifications and Their Function. *Cell* 128, 693–705.
- Kouzminova, E., Selker, E.U., 2001. *dim-2* encodes a DNA methyltransferase responsible for all known cytosine methylation in *Neurospora*. *EMBO J.* 20, 4309–4323.
- Kronholm, I., Johannesson, H., Ketola, T., 2016. Epigenetic Control of Phenotypic Plasticity in the Filamentous Fungus *Neurospora crassa*; *Genes|Genomes|Genetics*.
- Kubartová, A., Ranger, J., Berthelin, J., Beguiristain, T., 2009. Diversity and Decomposing Ability of Saprophytic Fungi from Temperate Forest Litter. *Microbial Ecology* 58, 98–107.
- Lancôt, C., Cheutin, T., Cremer, M., Cavalli, G., Cremer, T., 2007. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. *Nature Reviews Genetics* 8, 104–115.
- Latchman, D.S., 1997. Transcription factors: An overview. *The International Journal of Biochemistry & Cell Biology* 29, 1305–1312.
- Law, J.A., Jacobsen, S.E., 2010. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. *Nature Reviews Genetics* 11, 204–220.
- Levine, M., Tjian, R., 2003. Transcription regulation and animal diversity. *Nature* 424, 147–151.



- Li, E., Beard, C., Forster, A.C., Bestor, T.H., Jaenisch, R., 1993. DNA Methylation, Genomic Imprinting, and Mammalian Development. *Cold Spring Harb Symp Quant Biol* 58, 297–305.
- Lieberman-Aiden, E., van Berkum, N.L., Williams, L., Imakaev, M., Ragooczy, T., Telling, A., Amit, I., Lajoie, B.R., Sabo, P.J., Dorschner, M.O., Sandstrom, R., Bernstein, B., Bender, M.A., Groudine, M., Gnirke, A., Stamatoyannopoulos, J., Mirny, L.A., Lander, E.S., Dekker, J., 2009. Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science* 326, 289–293.
- Lindequist, U., Niedermeyer, T.H.J., Jülich, W.-D., 2005. The Pharmacological Potential of Mushrooms. *Evidence-Based Complementary and Alternative Medicine* 2, 285–299.
- Locke, G., Haberman, D., Johnson, S.M., Morozov, A.V., 2013. Global remodeling of nucleosome positions in *C. elegans*. *BMC genomics* 14, 284.
- Lomvardas, S., Barnea, G., Pisapia, D.J., Mendelsohn, M., Kirkland, J., Axel, R., 2006. Interchromosomal Interactions and Olfactory Receptor Choice. *Cell* 126, 403–413.
- Lo Presti, L., Lanver, D., Schweizer, G., Tanaka, S., Liang, L., Tollot, M., Zuccaro, A., Reissmann, S., Kahmann, R., 2015. Fungal Effectors and Plant Susceptibility. *Annual Review of Plant Biology* 66, 513–545.
- Lo Presti, L., Kahmann, R., 2017. How filamentous plant pathogen effectors are translocated to host cells. *Current Opinion in Plant Biology* 38, 19–24.
- Luger, K., Mäder, A.W., Richmond, R.K., Sargent, D.F., Richmond, T.J., 1997. Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251.
- Luger, K., Richmond, T.J., 1998. The histone tails of the nucleosome. *Current Opinion in Genetics & Development* 8, 140–146.
- Ma, L.-J., van der Does, H.C., Borkovich, K.A., Coleman, J.J., Daboussi, M.-J., Di Pietro, A., Dufresne, M., Freitag, M., Grabherr, M., Henrissat, B., Houterman, P.M., Kang, S., Shim, W.-B., Woloshuk, C., Xie, X., Xu, J.-R., Antoniw, J., Baker, S.E., Bluhm, B.H., Breakspear, A., Brown, D.W., Butchko, R.A.E., Chapman, S., Coulson, R., Coutinho, P.M., Danchin, E.G.J., Diener, A., Gale, L.R., Gardiner, D.M., Goff, S., Hammond-Kosack, K.E., Hilburn, K., Hua-Van, A., Jonkers, W., Kazan, K., Kodira, C.D., Koehrsen, M., Kumar, L., Lee, Y.-H., Li, L., Manners, J.M., Miranda-Saavedra, D., Mukherjee, M., Park, G., Park, J., Park, S.-Y., Proctor, R.H., Regev, A., Ruiz-Roldan, M.C., Sain, D., Sakthikumar, S., Sykes, S., Schwartz, D.C., Turgeon, B.G., Wapinski, I., Yoder, O., Young, S., Zeng, Q., Zhou, S., Galagan, J., Cuomo, C.A., Kistler, H.C., Rep, M., 2010. Comparative genomics reveals mobile pathogenicity chromosomes in *Fusarium*. *Nature* 464, 367–373.
- Marino-Ramirez, L., Levine, K.M., Morales, M., Zhang, S., Moreland, R.T., Baxevanis, A.D., Landsman, D., 2011. The Histone Database: an integrated resource for histones and histone fold-containing proteins. *Database* 2011, bar048-bar048.
- Mausser, R., Kungulovski, G., Keup, C., Reinhardt, R., Jeltsch, A., 2017. Application of dual reading domains as novel reagents in chromatin biology reveals a new H3K9me3 and H3K36me2/3 bivalent chromatin state. *Epigenetics & Chromatin* 10.
- McGee DC & Petrie GA. 1978. Variability of *Leptosphaeria maculans* in relation to blackleg of oilseed rape. *Phytopathology* 68: 47-52.





- Mendes-Pereira E., Balesdent M.H., Brun H., Rouxel T., 2003. Molecular phylogeny of the *Leptosphaeria maculans*-*L. biglobosa* species complex. *Mycol Res* 107:1287-1304.
- Michielse, C.B., van Wijk, R., Reijnen, L., Manders, E.M.M., Boas, S., Olivain, C., Alabouvette, C., Rep, M., 2009. The Nuclear Protein Sge1 of *Fusarium oxysporum* Is Required for Parasitic Growth. *PLoS Pathogens* 5, e1000637.
- Möller, M., Schotanus, K., Soyer, J.L., Haueisen, J., Happ, K., Stralucke, M., Happel, P., Smith, K.M., Connolly, L.R., Freitag, M., Stukenbrock, E.H., 2019. Destabilization of chromosome structure by histone H3 lysine 27 methylation. *PLOS Genetics* 15, e1008093.
- Motlagh, H.N., Wrabl, J.O., Li, J., Hilser, V.J., 2014. The ensemble nature of allostery. *Nature* 508, 331–339.
- Muggia, L., Grube, M., 2018. Fungal Diversity in Lichens: From Extremotolerance to Interactions with Algae. *Life* 8, 15.
- Niño-Sánchez, J., Castillo, C.-D., Tello, V., Vega-Bartol, D., José, J., Ramos, B., Sukno, S.A., Díaz Mínguez, J.M., 2016. The FTF gene family regulates virulence and expression of SIX effectors in *Fusarium oxysporum*. *Molecular plant pathology* 17, 1124–1139.
- Nishida, H., Motoyama, T., Yamamoto, S., Aburatani, H., Osada, H., 2009. Genome-wide maps of mono- and di-nucleosomes of *Aspergillus fumigatus*. *Bioinformatics* 25, 2295–2297.
- O'Connell, R.J., Panstruga, R., 2006. Tete a tete inside a plant cell: establishing compatibility between plants and biotrophic fungi and oomycetes. *New Phytologist* 171, 699–718.
- O'Connell, R.J., Thon, M.R., Hacquard, S., Amyotte, S.G., Kleemann, J., 2012. Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. *Nature Genetics* 44, 1060–1065.
- Ohara, T., Tsuge, T., 2004. FoSTUA, Encoding a Basic Helix-Loop-Helix Protein, Differentially Regulates Development of Three Kinds of Asexual Spores, Macroconidia, Microconidia, and Chlamydospores, in the Fungal Plant Pathogen *Fusarium oxysporum*. *Eukaryot Cell* 3, 1412–1422.
- Ohm, R.A., Feau, N., Henrissat, B., Schoch, C.L., Horwitz, B.A., Barry, K.W., Condon, B.J., Copeland, A.C., Dhillon, B., Glaser, F., Hesse, C.N., Kosti, I., LaButti, K., Lindquist, E.A., Lucas, S., Salamov, A.A., Bradshaw, R.E., Ciuffetti, L., Hamelin, R.C., Kema, G.H.J., Lawrence, C., Scott, J.A., Spatafora, J.W., Turgeon, B.G., Wit, P.J.G.M. de, Zhong, S., Goodwin, S.B., Grigoriev, I.V., 2012. Diverse Lifestyles and Strategies of Plant Pathogenesis Encoded in the Genomes of Eighteen Dothideomycetes Fungi. *PLOS Pathogens* 8, e1003037.
- Osborne, C.S., Chakalova, L., Brown, K.E., Carter, D., Horton, A., Debrand, E., Goyenechea, B., Mitchell, J.A., Lopes, S., Reik, W., Fraser, P., 2004. Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat. Genet.* 36, 1065–1071.
- Pabo, C.O., Sauer, R.T., 1992. *Transcription Factors: Structural Families and Principles of DNA Recognition* 43.
- Pais, P., Costa, C., Cavalheiro, M., Romão, D., Teixeira, M.C., 2016. Transcriptional Control of Drug Resistance, Virulence and Immune System Evasion in Pathogenic Fungi: A Cross-Species Comparison. *Front Cell Infect Microbiol* 6, 131.
- Park, J., Park, J., Jang, S., Kim, S., Kong, S., Choi, J., Ahn, K., Kim, J., Lee, S., Kim, S., Park, B., Jung, K., Kim, S., Kang, S., Lee, Y.-H., 2008. FTFD: an informatics pipeline supporting phylogenomic analysis of fungal transcription factors. *Bioinformatics* 24, 1024–1025.



- Papamichos-Chronakis, M., Peterson, C.L., 2013. Chromatin and the genome integrity network. *Nature Reviews Genetics* 14, 62–75.
- Patikoglou, G., Burley, S.K., 1997. Eukaryotic transcription factor-DNA complexes. *Annu Rev Biophys Biomol Struct* 26, 289–325.
- Ponts, N., Harris, E.Y., Prudhomme, J., Wick, I., Eckhardt-Ludka, C., Hicks, G.R., Hardiman, G., Lonardi, S., Le Roch, K.G., 2010. Nucleosome landscape and control of transcription in the human malaria parasite. *Genome Research* 20, 228–238.
- Radman-Livaja M., Rando Q.J., 2010. Nucleosome positioning: how is it established and why does it matter? *Dev. Biol.* 339: 258-266.
- Raffaele, S., Kamoun, S., 2012. Genome evolution in filamentous plant pathogens: why bigger can be better. *Nature Reviews Microbiology* 10, 417–430.
- Remy, W., Taylor, T.N., Hass, H., Kerp, H., 1994. Four hundred-million-year-old vesicular arbuscular mycorrhizae. *Proceedings of the National Academy of Sciences* 91, 11841–11843.
- Richards, E.J., Elgin, S.C.R., 2002. Epigenetic codes for heterochromatin formation and silencing: rounding up the usual suspects. *Cell* 108, 489–500.
- Rohe, M., Gierlich, A., Hermann, H., Hahn, M., Schmidt, B., 1995. The race-specific elicitor, NIP1, from the barley pathogen, *Rhynchosporium secalis*, determines avirulence on host plants of the Rrs1 resistance genotype. *The EMBO Journal* 14, 4168–4177.
- Rossetto, D., Avvakumov, N., Côté, J., 2012. Histone phosphorylation: A chromatin modification involved in diverse nuclear events. *Epigenetics* 7, 1098–1108.
- Rouxel T, Mendes-Pereira E, Brun H, Balesdent MH. 2004. Species complex of fungal phytopathogens: the *Leptosphaeria maculans*-*L. biglobosa* case study. In: *Plant Genome: Diversity and Evolution. Vol. 2: Cryptogams*. AK Sharma & A Sharma (Eds.), Science Publishers, Inc., Enfield, USA, pp.33-75.
- Rouxel, T., Grandaubert, J., Hane, J.K., Hoede, C., van de Wouw, A.P., Couloux, A., Dominguez, V., Anthouard, V., Bally, P., Bourras, S., Cozijnsen, A.J., Ciuffetti, L.M., Degrave, A., Dilmaghani, A., Duret, L., Fudal, I., Goodwin, S.B., Gout, L., Glaser, N., Linglin, J., Kema, G.H.J., Lapalu, N., Lawrence, C.B., May, K., Meyer, M., Ollivier, B., Poulain, J., Schoch, C.L., Simon, A., Spatafora, J.W., Stachowiak, A., Turgeon, B.G., Tyler, B.M., Vincent, D., Weissenbach, J., Amselem, J., Quesneville, H., Oliver, R.P., Wincker, P., Balesdent, M.-H., Howlett, B.J., 2011. Effector diversification within compartments of the *Leptosphaeria maculans* genome affected by Repeat-Induced Point mutations. *Nature Communications* 2.
- Ruthenburg, A.J., Allis, C.D., Wysocka, J., 2007. Methylation of Lysine 4 on Histone H3: Intricacy of Writing and Reading a Single Epigenetic Mark. *Molecular Cell* 25, 15–30.
- Rybak, K., See, P.T., Phan, H.T., Syme, R.A., Moffat, C.S., 2017. A functionally conserved Zn<sub>2</sub>Cys<sub>6</sub> binuclear cluster transcription factor class regulates necrotrophic effector gene expression and host-specific virulence of two major Pleosporales fungal pathogens of wheat. *Molecular plant pathology* 18, 420–434.
- Saksouk, N., Simboeck, E., Déjardin, J., 2015. Constitutive heterochromatin formation and transcription in mammals. *Epigenetics & Chromatin* 8, 3.



- Sánchez-Vallet, A., Fouché, S., Fudal, I., Hartmann, F.E., Soyer, J.L., Tellier, A., Croll, D., 2018. The Genome Biology of Effector Gene Evolution in Filamentous Plant Pathogens. *Annual Review of Phytopathology* 56, 21–40.
- Savitsky, P., Krojer, T., Fujisawa, T., Lambert, J.-P., Picaud, S., Wang, C.-Y., Shanle, E.K., Krajewski, K., Friedrichsen, H., Kanapin, A., Goding, C., Schapira, M., Samsonova, A., Strahl, B.D., Gingras, A.-C., Filippakopoulos, P., 2016. Multivalent Histone and DNA Engagement by a PHD/BRD/PWWP Triple Reader Cassette Recruits ZMYND8 to K14ac-Rich Chromatin. *Cell Reports* 17, 2724–2737.
- Selker, E.U., Stevens, J.N., 1985. DNA methylation at asymmetric sites is associated with numerous transition mutations. *Proc. Natl. Acad. Sci. U.S.A.* 82, 8114–8118.
- Selker, E.U., Richardson, G.A., Garrett-Engle, P.W., Singer, M.J., Miao, V., 1993. Dissection of the Signal for DNA Methylation in the  $\zeta$ - $\eta$  Region of *Neurospora*. *Cold Spring Harb Symp Quant Biol* 58, 323–329.
- Selker, E.U. 2002. Repeat-induced gene silencing in fungi. *Advances in genetics* 46:439-450.
- Selker, E.U., Tountas, N.A., Cross, S.H., Margolin, B.S., Murphy, J.G., Bird, A.P., Freitag, M., 2003. The methylated component of the *Neurospora crassa* genome. *Nature* 422, 893–897.
- Shelest, E., 2008. Transcription factors in fungi. *FEMS Microbiology Letters* 286, 145–151.
- Sheppard, D.C., Doedt, T., Chiang, L.Y., Kim, H.S., Chen, D., Nierman, W.C., Filler, S.G., 2005. The *Aspergillus fumigatus* StuA Protein Governs the Up- Regulation of a Discrete Transcriptional Program during the Acquisition of Developmental Competence. *Molecular Biology of the Cell* 16, 14.
- Schardl, C.L., Young, C.A., Hesse, U., Amyotte, S.G., Andreeva, K., Calie, P.J., Fleetwood, D.J., Haws, D.C., Moore, N., Oeser, B., Panaccione, D.G., Schweri, K.K., Voisey, C.R., Farman, M.L., Jaromczyk, J.W., Roe, B.A., O'Sullivan, D.M., Scott, B., Tudzynski, P., An, Z., Arnaoudova, E.G., Bullock, C.T., Charlton, N.D., Chen, L., Cox, M., Dinkins, R.D., Florea, S., Glenn, A.E., Gordon, A., Güldener, U., Harris, D.R., Hollin, W., Jaromczyk, J., Johnson, R.D., Khan, A.K., Leistner, E., Leuchtman, A., Li, C., Liu, JinGe, Liu, Jinze, Liu, M., Mace, W., Machado, C., Nagabhyru, P., Pan, J., Schmid, J., Sugawara, K., Steiner, U., Takach, J.E., Tanaka, E., Webb, J.S., Wilson, E.V., Wiseman, J.L., Yoshida, R., Zeng, Z., 2013. Plant-Symbiotic Fungi as Chemical Engineers: Multi-Genome Analysis of the Clavicipitaceae Reveals Dynamics of Alkaloid Loci. *PLOS Genetics* 9, e1003323.
- Schmid, M., Arib, G., Laemmli, C., Nishikawa, J., Durussel, T., Laemmli, U.K., 2006. Nup-PI: The Nucleopore-Promoter Interaction of Genes in Yeast. *Molecular Cell* 21, 379–391.
- Schotanus, K., Soyer, J.L., Connolly, L.R., Grandaubert, J., Happel, P., Smith, K.M., Freitag, M., Stukenbrock, E.H., 2015. Histone modifications rather than the novel regional centromeres of *Zygomycetes tritici* distinguish core and accessory chromosomes. *Epigenetics & Chromatin* 8, 41.
- Schumacher, J., Studt, L., Tudzynski, P., 2019. The putative H3K36 demethylase BcKDM1 affects virulence, stress responses and photomorphogenesis in *Botrytis cinerea*. *Fungal Genetics and Biology* 123, 14–24.
- Shoemaker R.A. & Brun H., 2001. The telemorph of the weakly aggressive segregate of *Leptosphaeria maculans*. *Can J Bot* 79: 412-419.
- Silar et Malagnac, Les champignons redécouverts, Paris, Belin, Hors collection, 2013, 0-231.
- Skibbens, R.V., 2019. Condensins and cohesins - one of these things is not like the other! *J. Cell. Sci.* 132.
- Smith, M.M., 1991. Histone structure and function. *Current opinion in cell biology* 3, 429–437.



- Smith, K.M., Galazka, J.M., Phatale, P.A., Connolly, L.R., Freitag, M., 2012. Centromeres of filamentous fungi. *Chromosome Res* 20, 635–656.
- Soyer, J.L., El Ghalid, M., Glaser, N., Ollivier, B., Linglin, J., Grandaubert, J., Balesdent, M.-H., Connolly, L.R., Freitag, M., Rouxel, T., Fudal, I., 2014. Epigenetic Control of Effector Gene Expression in the Plant Pathogenic Fungus *Leptosphaeria maculans*. *PLoS Genetics* 10, e1004227.
- Soyer, J.L., Rouxel, T., Fudal, I., 2015. A. Chromatin-based control of effector gene expression in plant-associated fungi. *Current Opinion in Plant Biology* 26, 51–56.
- Soyer, J.L., Hamiot, A., Ollivier, B., Balesdent, M.-H., Rouxel, T., Fudal, I., 2015. B. The APSES transcription factor LmStuA is required for sporulation, pathogenic development and effector gene expression in *Leptosphaeria maculans*: Functional analysis of LmStuA from *L. maculans*. *Molecular Plant Pathology* 16, 1000–1005.
- Soyer, J.L., Grandaubert, J., Haueisen, J., Schotanus, K., Stukenbrock, E.H., 2019. In planta chromatin immunoprecipitation in *Zymoseptoria tritici* reveals chromatin-based regulation of putative effector gene expression. *bioRxiv* 544627.
- Spanu, P.D., Abbott, J.C., Amselem, J., Burgis, T.A., Soanes, D.M., Stüber, K., Themaat, E.V.L. van, Brown, J.K.M., Butcher, S.A., Gurr, S.J., Lebrun, M.-H., Ridout, C.J., Schulze-Lefert, P., Talbot, N.J., Ahmadinejad, N., Ametz, C., Barton, G.R., Benjdia, M., Bidzinski, P., Bindschedler, L.V., Both, M., Brewer, M.T., Cadle-Davidson, L., Cadle-Davidson, M.M., Collemare, J., Cramer, R., Frenkel, O., Godfrey, D., Harriman, J., Hoede, C., King, B.C., Klages, S., Kleemann, J., Knoll, D., Koti, P.S., Kreplak, J., López-Ruiz, F.J., Lu, X., Maekawa, T., Mahanil, S., Micali, C., Milgroom, M.G., Montana, G., Noir, S., O'Connell, R.J., Oberhaensli, S., Parlange, F., Pedersen, C., Quesneville, H., Reinhardt, R., Rott, M., Sacristán, S., Schmidt, S.M., Schön, M., Skamnioti, P., Sommer, H., Stephens, A., Takahara, H., Thordal-Christensen, H., Vigouroux, M., Weßling, R., Wicker, T., Panstruga, R., 2010. Genome Expansion and Gene Loss in Powdery Mildew Fungi Reveal Tradeoffs in Extreme Parasitism. *Science* 330, 1543–1546.
- Strauss, J., Reyes-Dominguez, Y., 2011. Regulation of secondary metabolism by chromatin structure and epigenetic codes. *Fungal Genetics and Biology* 48, 62–69.
- Struhl, K., Segal, E., 2013. Determinants of nucleosome positioning. *Nature Structural & Molecular Biology* 20, 267–273.
- Stukenbrock, E.H., Banke, S., Javan-Nikkhah, M., McDonald, B.A., 2007. Origin and domestication of the fungal wheat pathogen *Mycosphaerella graminicola* via sympatric speciation. *Mol. Biol. Evol.* 24, 398–411.
- Tamaru, H., Selker, E.U., 2001. A histone H3 methyltransferase controls DNA methylation in *Neurospora crassa*. *Nature* 414, 277.
- Tan, K.-C., Oliver, R.P., 2017. Regulation of proteinaceous effector expression in phytopathogenic fungi. *PLoS pathogens* 13, e1006241.
- Tanizawa, H., Kim, K.-D., Iwasaki, O., Noma, K., 2017. Architectural alterations of the fission yeast genome during the cell cycle. *Nature Structural & Molecular Biology* 24, 965–976.
- Taylor, T.N., Osborn, J.M., 1996. The importance of fungi in shaping the paleoecosystem. *Review of Palaeobotany and Palynology* 90, 249–262.



- Thomas, M.C., Chiang, C.-M., 2006. The General Transcription Machinery and General Cofactors. *Critical Reviews in Biochemistry and Molecular Biology* 41, 105–178.
- Tong, X., Zhang, X., Plummer, K.M., Stowell, K.M., Sullivan, P.A., Farley, P.C., 2007. GcSTUA, an APSES Transcription Factor, Is Required for Generation of Appressorial Turgor Pressure and Full Pathogenicity of *Glomerella cingulata*. *Molecular Plant-Microbe Interactions* 20, 1102–1111.
- Tollot, M., Assmann, D., Becker, C., Altmüller, J., Dutheil, J.Y., Wegner, C.-E., Kahmann, R., 2016. The WOPR Protein Ros1 Is a Master Regulator of Sporogenesis and Late Effector Gene Expression in the Maize Pathogen *Ustilago maydis*. *PLOS Pathogens* 12, e1005697.
- Trojer, P., Reinberg, D., 2007. Facultative Heterochromatin: Is There a Distinctive Molecular Signature? *Molecular Cell* 28, 1–13.
- Tsankov, A.M., Thompson, D.A., Socha, A., Regev, A., Rando, O.J., 2010. The Role of Nucleosome Positioning in the Evolution of Gene Regulation. *PLoS Biology* 8, e1000414.
- Ulianov, S.V., Khrameeva, E.E., Gavrillov, A.A., Flyamer, I.M., Kos, P., Mikhaleva, E.A., Penin, A.A., Logacheva, M.D., Imakaev, M.V., Chertovich, A., Gelfand, M.S., Shevelyov, Y.Y., Razin, S.V., 2016. Active chromatin and transcription play a key role in chromosome partitioning into topologically associating domains. *Genome Res.* 26, 70–84.
- Vincenot, L., Balesdent, M.H., Li, H., Barbetti, M.J., Sivasithamparam, K., Gout, L., Rouxel, T., 2008. Occurrence of a New Subclade of *Leptosphaeria biglobosa* in Western Australia. *Phytopathology* 98, 321–329.
- Voigt, K., Cozijnsen, A.J., Kroymann, J., Pöggeler, S., Howlett, B.J., 2005. Phylogenetic relationships between members of the crucifer pathogenic *Leptosphaeria maculans* species complex as shown by mating type (MAT1-2), actin, and beta-tubulin sequences. *Mol. Phylogenet. Evol.* 37, 541–557.
- De Wit, P.J.G.M., Mehrabi, R., Van Den Burg, H.A., Stergiopoulos, I., 2009. Fungal effector proteins: past, present and future. *Molecular Plant Pathology* 10, 735–747.
- Wang, M., Weiberg, A., Jin, H., 2015. Pathogen small RNAs: a new class of effectors for pathogen attacks: Pathogen small RNAs at attack. *Molecular Plant Pathology* 16, 219–223.
- Wang, M., Weiberg, A., Lin, F.-M., Thomma, B.P.H.J., Huang, H.-D., Jin, H., 2016. Bidirectional cross-kingdom RNAi and fungal uptake of external RNAs confer plant protection. *Nature Plants* 2, 16151.
- Weiberg, A., Wang, M., Lin, F.-M., Zhao, H., Zhang, Z., Kaloshian, I., Huang, H.-D., Jin, H., 2013. Fungal Small RNAs Suppress Plant Immunity by Hijacking Host RNA Interference Pathways. *Science* 342, 118–123.
- Weirauch, M.T., Hughes, T.R., 2011. A Catalogue of Eukaryotic Transcription Factor Types, Their Evolutionary Origin, and Species Distribution, in: Hughes, Timothy R. (Ed.), *A Handbook of Transcription Factors, Subcellular Biochemistry*. Springer Netherlands, Dordrecht, pp. 25–73.
- West, J.S., Fitt, B.D.L., Leech, P.K., Biddulph, J.E., Huang, Y.-J., Balesdent, M.-H., 2002. Effects of timing of *Leptosphaeria maculans* ascospore release and fungicide regime on phoma leaf spot and phoma stem canker development on winter oilseed rape (*Brassica napus*) in southern England. *Plant Pathology* 51, 454–463.
- Williams RH & Fitt BDL. 1999. Differentiating A and B groups of *Leptosphaeria maculans*, causal agent of stem canker (blackleg) of oilseed rape. *Plant Pathology* 48: 161-175.



Wipf, D., Krajinski, F., Tuinen, D. van, Recorbet, G., Courty, P.-E., 2019. Trading on the arbuscular mycorrhiza market: from arbuscules to common mycorrhizal networks. *New Phytologist* 223, 1127–1142.

Winter, D.J., Ganley, A., Young, C., Liachko, I., Schardl, C., Dupont, P., Berry, D., Ram, A., Scott, B., Cox, M., 2018. Repeat elements organize 3D genome structure and mediate transcription in the filamentous fungus *Epichloë festucae*. *bioRxiv* 339010.

Wolter, F., Schindele, P., Puchta, H., 2019. Plant breeding at the speed of light: the power of CRISPR/Cas to generate directed genetic diversity at multiple sites. *BMC Plant Biology* 19.

Wouw, A.P.V. de, Marcroft, S.J., Barbetti, M.J., Hua, L., Salisbury, P.A., Gout, L., Rouxel, T., Howlett, B.J., Balesdent, M.H., 2009. Dual control of avirulence in *Leptosphaeria maculans* towards a *Brassica napus* cultivar with 'sylvestris-derived' resistance suggests involvement of two resistance genes. *Plant Pathology* 58, 305–313.

Wunderlich, Z., Mirny, L.A., 2009. Different gene regulation strategies revealed by analysis of binding motifs. *Trends in Genetics* 25, 434–440.

Yadav, T., Quivy, J.-P., Almouzni, G., 2018. Chromatin plasticity: A versatile landscape that underlies cell fate and identity. *Science* 361, 1332–1336.

Zolan ME, Pukkila PJ (1985) DNA methylation in *Coprinus cinereus*, pp. 333-344 in *Molecular Genetics of Filamentous Fungi*, edited by W. TIMBERLAKE. Alan R. Liss, New-York.

Zuccaro, A., Lahrmann, U., Langen, G., 2014. Broad compatibility in fungal root symbioses. *Current Opinion in Plant Biology* 20, 135–145.



## ANNEXES

### Poster

#### Investigating the involvement of the chromatin state in the control of gene expression in *Leptosphaeria maculans*

Colin Clairet<sup>1,2</sup>, Jessica L. Soyer<sup>1</sup>, Elise J. Gay<sup>1,2</sup>, Nicolas Lapalu<sup>1</sup>, Adeline Simon<sup>1</sup>, Françoise Blaise<sup>1</sup>, Eva H. Stukenbrock<sup>3</sup> and Isabelle Fudal<sup>1</sup>

<sup>1</sup>UMR BIOGER, INRA, AgroParisTech, Université Paris-Saclay, Thiverval-Grignon, France

<sup>2</sup>Université Paris-Sud, Orsay, France

<sup>3</sup>Max Planck Institute for Evolutionary Biology, Plön, and Christian-Albrechts University of Kiel, Germany

*Leptosphaeria maculans*, a hemibiotrophic fungus responsible for stem canker, colonises oilseed rape in two stages: an early stage of cotyledon or leaf colonisation, and a late colonisation stage during which the fungus colonises systemically without visible symptom the plant before stem canker appears. *L. maculans* presents a bipartite genome structure alternating gene-rich and transposable element (TE)-rich region. While TE-rich regions are enriched in putative effector-encoding genes strongly over-expressed during early infection ('early' effectors), gene-rich regions contain putative effector-encoding genes specifically expressed during late infection ('late' effectors). Here, we analyzed nucleosome positioning, chromatin structure and gene expression at the genome scale combining MAINE-seq, ChIP-seq and RNA-seq data during axenic growth and performed functional analysis of two players of heterochromatin assembly (*KMT1* and *KMT6*). We analyzed *in vitro* ChIP-seq data targeting two heterochromatin modifications, H3K9me3 and H3K27me3, and a euchromatin modification, H3K4me2, and found that gene-rich regions are associated with H3K4me2 and H3K27me3 while TE-rich regions are associated with H3K9me3. Analysis of *in vitro* MAINE-seq data showed distinct nucleosome organization for genes located in TE-rich or gene-rich regions, and according to gene expression level. CRISPR-Cas9 mutants of *KMT1* and *KMT6* have been generated and are currently characterized for their ability to infect leaves and stems of oilseed rape, and to express effector genes.

Presented at:

**40th Symposium of New phytologist** (Vienne, Austria)

**19th Fungal Genetics conference** (Asilomar, Ca, USA)





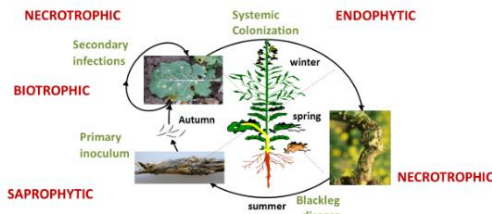
# Investigating the involvement of the chromatin state in the control of effector gene expression in *Leptosphaeria maculans*

Colin Clairot<sup>1,2</sup>, Jessica L. Soyer<sup>1,3</sup>, Elise J. Gay<sup>1,2</sup>, Nicolas Lapalu<sup>1</sup>, Adeline Simon<sup>1</sup>, Françoise Blaise<sup>1</sup>, Eva H. Stukenbrock<sup>3</sup> and Isabelle Fudal<sup>1</sup>  
<sup>1</sup>INRA BIOGER<sup>2</sup>, Thiverval-Grignon, Université Paris Saclay and Université Paris-Sud<sup>1</sup>, Orsay France  
<sup>2</sup>Université Paris-Sud, Orsay, France  
<sup>3</sup>Max Planck Institute for Evolutionary Biology, Plön, and Christian-Albrechts University of Kiel, Germany

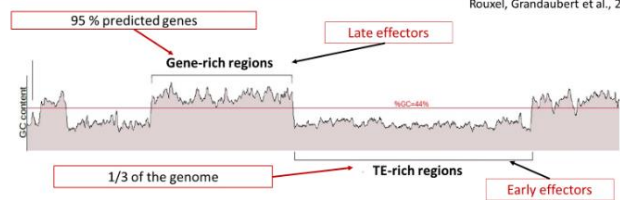
Poster Number: 648T



## 1/Context: *Leptosphaeria* life cycle and control of effector gene expression



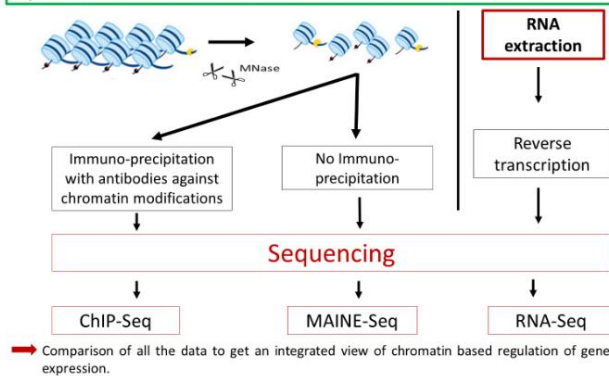
*Leptosphaeria maculans*, a hemibiotrophic fungus responsible for stem canker, colonizes oilseed rape in two stages: an early stage of cotyledon or leaf colonization, and a late colonization stage during which the fungus colonizes, systemically, without visible symptom, the plant before stem canker appears.



Rouxel, Grandaubert et al., 2011

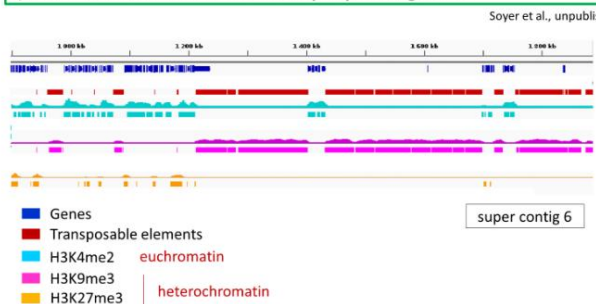
*L. maculans* presents a bipartite genome structure alternating gene-rich and transposable element (TE)-rich regions. TE-rich regions are enriched in putative effector-encoding genes that are expressed specifically in an early stage of infection ('early effectors'). In contrast, gene-rich regions were recently reported to contain putative effector-encoding genes specifically expressed during the late stages of stem infection ('late' effectors).

## 2/Strategy to decipher the role of chromatin structure on effector gene expression *in vitro*



Comparison of all the data to get an integrated view of chromatin based regulation of gene expression.

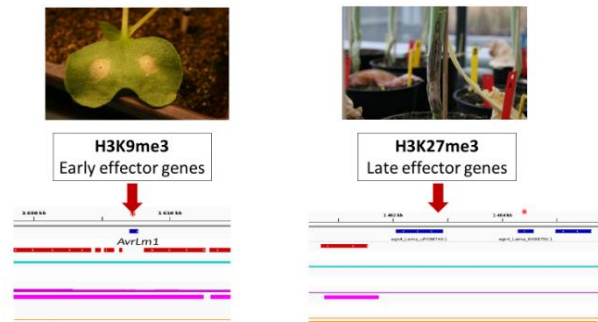
## 3/Localization of chromatin marks on *Leptosphaeria* genome



Soyer et al., unpublished

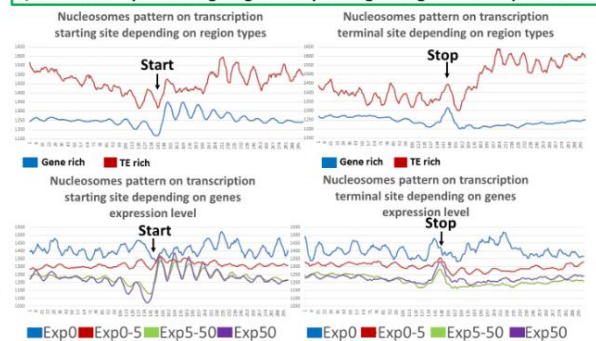
Gene-rich regions are enriched in an euchromatin mark, H3K4me2. Both TE-rich and gene-rich regions are enriched in heterochromatin marks, but these marks are distinct. While TE-rich regions are enriched in H3K9me3, gene-rich regions are enriched in H3K27me3.

## 4/Chromatin marks at late and early effector loci



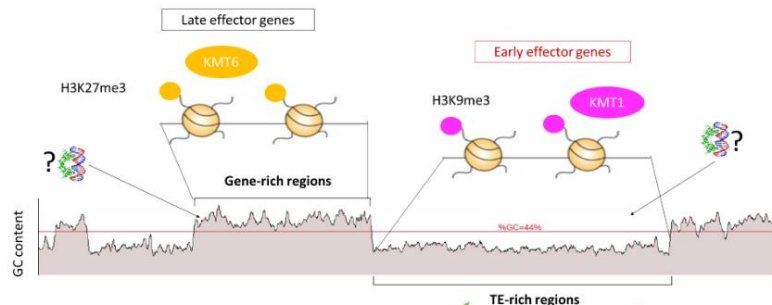
Early effector genes are statistically associated to H3K9me3 compartments. Late effector genes on the contrary are specifically associated to H3K27me3 compartments. The differential time of expression of these genes sets can be led by these different type of marks.

## 5/Nucleosome positioning on genes depending on regions and expression level



Nucleosomes have well defined positions on transcription starting site and stop of gene-rich regions unlike nucleosomes on TE-rich regions. Nucleosomes have also better positioning on well expressed genes than on poorly expressed genes. These observations are clearer on nucleosomes localized just after the start which are crucial for transcription efficiency. With Exp corresponding to expression level in TPM (transcript per million).

## 6/Effector concerted expression model and perspectives



Repression of Early effector is maintained by H3K9me3 methylation deposited by KMT1 (Soyer et al. 2014). Chromatin remodeling leads to a decondensation at effector gene loci, rendering their promoters accessible to specific transcription factors. The late effector expression mechanism would be similar. But the repression should be maintained by H3K27me3 deposited by KMT6.

**Next stages:** CRISPR-CAS9 inactivation of *KMT1* and *KMT6* to assess the specific regulation of early and late effector genes.



Contacts  
 Colin Clairot: colin.clairot@inra.fr  
 Isabelle Fudal: isabelle.fudal@inra.fr  
 Jessica Soyer: Jessica.soyer@inra.fr





## Bilan d'activités :

### Catégorie : Communication et médiation scientifique

**Congrès Doctorant SPE** (Montpellier, France) du 20 au 22 Juin 2017; Doctorant SPE: "Investigating the involvement of the chromatin state in the control of effector gene expression in *Leptosphaeria maculans*"

**Congrès EpiAgro** (Bordeaux, France) 18 et 19 juin 2018; réseau REACTION : "Investigating the involvement of the chromatin state in the control of effector gene expression in *Leptosphaeria maculans*"

**19th Fungal Genetics conference** (Asilomar, Ca, USA) 12 au 17 mars 2019; Genetic society of America: "Transcriptional and chromatin-based control of gene expression and pathogenicity in the plant pathogenic fungus *Leptosphaeria maculans*"

**Effectome Montpellier** (Montpellier, France) du 01 au 04 Novembre 2016: "Investigating the involvement of the heterochromatic-associated histone modification, H3K27me3, in the control of effector gene expression in *Leptosphaeria maculans*"

**Journées Jean-Chevauchon** (Aussois, France) 16 au 19 janvier 2018: "Investigating the involvement of the chromatin state in the control of effector gene expression in *Leptosphaeria maculans*"

Total du nombre d'heures pour la catégorie Communication et médiation scientifique : 76 h

### Catégorie : Enseignement et pédagogie

**Encadrement étudiant de 1ère année d'AgroParisTech** (6 mois/ 2h par semaine) INRA BIOGER

Total du nombre d'heures pour la catégorie Enseignement et pédagogie : 48 h

### Catégorie : Ethique et intégrité scientifique

**Introduction à l'éthique de la recherche et intégrité scientifique** (CentraleSupélec, Orsay, France) 29 mai

Total du nombre d'heures pour la catégorie Ethique et intégrité scientifique : 6 h

### Catégorie : Parcours ' Conseil et expertise en innovation '

**Rencontre Doctorants-Entreprises** (ENSTA ParisTech, Palaiseau, France) 17 mai 2018.

Total du nombre d'heures pour la catégorie Parcours ' Conseil et expertise en innovation ' : 3 h

### Catégorie : Définir son projet professionnel

**DOC'AVENIR 2018** - (AgroParisTech, Paris, France) 04 juillet 2018

**Total heures de formation : 140 heures / 10 modules**



**Title: Chromatin and transcriptional-based control of the expression of effector-encoding genes in *Leptosphaeria maculans***

**Keywords:** effectors, chromatin, transcription factors, histone modifying proteins

**Summary:** *Leptosphaeria maculans* is a pathogenic fungus responsible for the stem canker of oilseed rape (*Brassica napus*). *L. maculans* exhibits a complex life cycle with successive biotrophic and necrotrophic phases. In addition, it has a bipartite genome alternating gene-rich regions (GC isochores) and regions enriched in transposable elements and in effector-encoding genes expressed during the early stages of infection (AT isochores). Very little is known on how the expression of effectors genes is controlled. A previous study had established that the expression of at least two effector-encoding genes was repressed by a chromatin-based control involving the repressive mark H3K9me3 deposited by the histone methyltransferase KMT1, *in vitro*. During infection, this control was postulated to be lifted allowing the action of specific transcription factors (TFs).

The objective of my thesis was to study the influence of chromatin and transcriptional control on gene expression and to determine the role of histone modifying proteins (HMEs) and TFs in the control of effector-encoding gene expression. To this end, two approaches have been developed: (i) a "genome-wide" approach combining analysis of nucleosome positioning (MAINE-Seq), distribution of heterochromatin and euchromatin modifications (ChIP-Seq) and transcriptomic analyses (RNA-Seq) in *L. maculans* and a closely related fungal species not pathogenic of oilseed rape and (ii) a functional approach based on the inactivation by CRISPR-Cas9 of TFs-encoding genes and HMEs to study the impact of these inactivations on growth, pathogenicity and gene expression.

MAINE-Seq analyses allowed us establishing for the first time a nucleosome positioning map in four phytopathogenic species, including *L. maculans*. We

also determined a chromatin modification map along the genome of two closely related species within the *L. maculans* species complex. We established that the AT isochores are largely enriched in the heterochromatin mark H3K9me3 while the GC isochores are enriched in the Euchromatin mark H3K4me2 and the heterochromatin mark H3K27me3.

The genes encoding the HME KMT1 and LmPf2, a TF with an expression profile similar to that of effectors genes and described in several Pleosporales species as controlling the expression of toxins and effectors, have been inactivated in *L. maculans*. We also overexpressed *LmPf2* in a  $\Delta kmt1$  mutant in which the chromatin is decondensed and in a wild type strain of *L. maculans*. These experiments allowed us to determine that LmPf2 controls the expression of effector waves in a context where the chromatin is open. This is the first example of a double control of the expression of effector-encoding genes by an HME and a specific TF.

Finally, I initiated a study of the involvement of KMT1 and KMT6, the protein responsible for the deposition of the repressive modification H3K27me3, in the control of effector-encoding gene expression. For this purpose, I generated mutants  $\Delta kmt1$ ,  $\Delta kmt6$ , and a double mutant  $\Delta kmt1 \Delta kmt6$  and initiated their characterization. I also generated material for ChIP-Seq, RNA-Seq and Hi-C-Seq analyses that will inform us about the impact of the repressive marks H3K9me3 and H3K27me3 on genome structure, organization within nuclei and gene expression.

In conclusion, the work carried out during this thesis has allowed us to better understand the involvement of chromatin structure and TFs in the control of effector gene expression.



**Titre : Contrôle chromatinien et transcriptionnel de l'expression des gènes codant des effecteurs chez *Leptosphaeria maculans***

**Mots clés :** effecteurs, chromatine, facteurs de transcription, protéines modificatrices d'histones

**Résumé:** *Leptosphaeria maculans* est un champignon pathogène, responsable de la nécrose du collet du colza (*Brassica napus*). *L. maculans* présente un cycle de vie complexe où se succèdent des phases biotrophes et nécrotrophes. De plus il possède un génome bipartite alternant régions riches en gènes (isochores GC) et régions riches en élément transposable et en gènes codant des effecteurs exprimés pendant les étapes précoces de l'infection (isochores AT). Très peu de connaissances sont disponibles concernant la façon dont l'expression des effecteurs est contrôlée. Une précédente étude avait établi que l'expression d'au moins deux gènes codant des effecteurs étaient sous contrôle d'une modification chromatinienne via la mise en place de la marque répressive H3K9me3 par l'histone méthyltransférase KMT1, *in vitro*. Au cours de l'infection, ce contrôle serait levé permettant l'action de facteurs de transcription (FTs) spécifiques.

L'objectif de ma thèse était d'étudier l'influence d'un contrôle chromatinien et transcriptionnel sur l'expression génique et de déterminer le rôle de protéines modificatrices d'histone (HMEs) ainsi que de FTs dans le contrôle de l'expression des gènes codant des effecteurs. Pour cela, deux approches ont été développées: (i) une approche «genome-wide» combinant analyse du positionnement des nucléosomes (MAINE-Seq), de la répartition de marques hétérochromatiniennes et euchromatiniennes (ChIP-Seq) et analyses transcriptomiques (RNA-Seq) chez *L. maculans* et une espèce fongique proche, non pathogène du colza et (ii) une approche fonctionnelle basée sur l'inactivation par CRISPR-Cas9 de gènes codant des FTs et des HMEs afin d'étudier l'impact de ces inactivations sur la croissance, la pathogénie et l'expression génique.

Grâce aux analyses de MAINE-Seq, nous avons pu établir pour la première fois une carte de positionnement des nucléosomes chez quatre espèces

phytopathogènes, dont *L. maculans*. Nous avons également construit une carte de modifications chromatinienne le long du génome de deux espèces très proches au sein du complexe d'espèces de *L. maculans*. Nous avons établi que les isochores AT sont largement enrichis en marque hétérochromatinienne H3K9me3 alors que les isochores GC sont enrichis en marque euchromatinienne H3K4me2 et en marque hétérochromatinienne H3K27me3.

La HME KMT1, et LmPf2, un FT ayant un profil d'expression similaire à celui des effecteurs et décrit chez plusieurs espèces de Pleosporales comme contrôlant l'expression de toxines et d'effecteurs, ont été inactivés chez *L. maculans*. Nous avons également réalisé une surexpression de *LmPf2* dans une souche  $\Delta kmt1$  dans laquelle la chromatine est décondensée et dans une souche sauvage de *L. maculans*. Ces expériences nous ont permis de déterminer que LmPf2 contrôle l'expression des effecteurs AvrLm dans un contexte où la chromatine est ouverte ( $\Delta kmt1$ ). Il s'agit du premier exemple d'un double contrôle de l'expression de gènes codant des effecteurs par une HME et un FT.

Enfin, j'ai pu initier l'étude de l'implication de KMT1 et KMT6, la protéine responsable du dépôt de la marque répressive H3K27me3, dans l'expression des gènes codant des effecteurs. Pour cela j'ai généré les mutants  $\Delta kmt1$ ,  $\Delta kmt6$ , et double mutant  $\Delta kmt1_{\Delta kmt6}$  et initié leur caractérisation. J'ai par ailleurs généré du matériel pour des analyses de ChIP-Seq, RNA-Seq et Hi-C-Seq qui nous renseigneront sur l'impact des marques répressives H3K9me3 et H3K27me3 sur la structure du génome, l'organisation au sein des noyaux et l'expression génique.

En conclusion, le travail réalisé au cours de cette thèse nous a permis de comprendre l'implication de la structure de la chromatine et de FTs dans l'expression des effecteurs précoces en condition de vie axénique.

