CONTROL AND FUNCTION OF FACULTATIVE HETEROCHROMATIN IN FILAMENTOUS FUNGI

by

Abigail J. Courtney

(Under the Direction of Zachary A. Lewis)

ABSTRACT

Polycomb Repressive Complex 2 is a multi-subunit complex that deposits mono- di- and tri- methyl groups on lysine 27 of histone H3, and tri-methyl H3K27 is a molecular marker of transcriptionally repressed facultative heterochromatin. Polycomb mediated silencing is important for development in metazoans and for repression of key fungal genes in certain filamentous fungi. *Neurospora crassa* and *Magnaporthe oryzae* contain minimal Polycomb repression systems which allow for genetic studies to examine the basic mechanisms of this complex in addition to its biological impact. This body of work shows Polycomb mediated repression of known effector genes in *M. oryzae*, and deletion of the H3K27 methyltransferase, KMT6, derepressed many effector genes in addition to reducing virulence of this devastating plant pathogen. In *N. crassa*, deletion of the histone variant H2A.Z downregulates the expression of a critical subunit of Polycomb Repressive Complex 2 and results in region-specific loss of H3K27 methylation. These discoveries are important in understanding how the plant pathogen *M. oryzae* regulates the genes related to infection of a host plant. Additionally, work in *N. crassa* has revealed that differential dependence on EED concentration is more critical

for establishment or maintenance of specific repressed domains in the genome. This work represents meaningful contributions to understanding the function and control of facultative heterochromatin in filamentous fungi.

INDEX WORDS: Fungal genetics; *Neurospora crassa*; Genomics; H2A.Z; Histone modifications; Chromatin; Heterochromatin; Gene expression

CONTROL AND FUNCTION OF FACULTATIVE HETEROCHROMATIN IN FILAMENTOUS FUNGI

by

ABIGAIL J. COURTNEY

BS, Ramapo College of New Jersey, 2015

AA, SUNY Rockland Community College, 2006

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2020

CONTROL AND FUNCTION OF FACULTATIVE HETEROCHROMATIN IN FILAMENTOUS FUNGI

by

ABIGAIL J. COURTNEY

Major Professor: Zachary A. Lewis
Committee: Vincent J. Starai
Shannon Quinn
Mary Goll

Electronic Version Approved:

Ron Walcott
Dean of the Graduate School
The University of Georgia
December 2020

DEDICATION

To my mother, Sara Ameri. For introducing me to science at an early age. For her constant belief in me and my abilities and supporting me in all of my endeavors. I am eternally grateful for your love and support.

ACKNOWLEDGEMENTS

I would like to acknowledge Yan Xu, Paramjeet Bagga, Thomas Owen, and Sandra Suarez for helping me achieve what they knew I could. For pushing me beyond my comfort zone and helping me discover incredible opportunities that changed the trajectory of my life. I would not be here today without their guidance and support.

To my committee members Shannon Quinn, Vinny Starai, and Mary Goll for being rock solid support, calm when I needed it, and for believing in me and my abilities beyond what I was able to see.

To Zack Lewis, for teaching me more than I can explain in words. For patience beyond patience. For helping me through some of the most difficult times in my life and seeing the scientist I am, especially when I could not.

To the undergraduate students who taught me more than I think I taught them: Sean Jackewicz, Jaqueline Nutter, Jemma Hwang, Mallory Huff, Trevor Neal, Vlad Sirbu and Daphnee Dubreuil.

There are many people who have helped me personally and scientifically throughout these past five years. They kept me sane and made me stronger throughout this experience and without them I do not know who I would be. To Masayuki Kamei, Jessica Heard, Jennifer Holihan-Colazzo, Aileen Ferraro, Eve Basenko, Jie Zhu, Katie Duval, Mariel Pfeifer, Susanna Harris, Kari Renne, Alysha Higgs, Shana Pau, Cathryn Quinn, Anastacia Parks, Alice Naftaly, and Matthew Courtney. Thank you.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS v

CHAPTER

1 INTRODUCTION AND LITERATURE REVIEW 1

Purpose 1

Chromatin structure 1

Facultative heterochromatin in fungi 3

References 8

Figures 16

2 THE HISTONE MODIFICATIONH H3K27ME3 AND TRANSCRIPTION FACTOR MOGTI1 COORDINATELY CONTROL EXPRESSION OF EFFECTOR GENES IN RICE BLAST FUNGUS 19

Abstract 20

Introduction 20

Results 24

Discussion 35

Materials and Methods 40

References 49

Figures 57

Tables 66

3 SHANNON ENTROPY AS A METRIC FOR CONDITIONAL GENE EXPRESSION IN *NEUROSPORA CRASSA*....77

Abstract....78

Introduction....78

Methods....80

Results and Discussion....81

References....86

Figures....90

4 NORMAL PATTERNS OF HISTONE H3K27 METHYLATION REQUIRE THE HISTONE VARIANT H2A.Z IN *NEUROSPORA CRASSA*....98

Abstract....99

Introduction....99

Materials and Methods....103

Results....112

Discussion....120

References....122

Figures....133

5 RTT109 IS REQUIRED FOR NORMAL PATTERNS OF H3K27ME2/3 IN *NEUROSPORA CRASSA*....143

Introduction....143

Results....144

Discussion....147

Materials and Methods....148

References.....152

Figures.....156

Tables.....162

6 DISCUSSION.....164

The histone modification H3K27me3 and transcription factor MoGti1 coordinately control expression of effector genes in rice blast fungus....165

Shannon entropy as a metric for conditional gene expression in *Neurospora crassa*.....166

Normal patterns of histone H3K27 methylation require the histone variant H2A.Z in *Neurospora crassa*.....168

RTT109 is required for normal patterns of H3K27me2/3 in *Neurospora crassa*.....168

References.....169

APPENDIX

A THE MATERNAL TO ZYGOTIC TRANSITION REGULATES GENOME-WIDE HETEROCHROMATIN ESTABLISHMENT IN THE ZEBRAFISH EMBRYO.....172

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Purpose

The purpose of this dissertation is to dissect the control and function of facultative heterochromatin in filamentous fungi. Facultative heterochromatin is important in the proper development of metazoans, but is challenging to study in higher eukaryotes (1). Model organisms such as filamentous fungi represent a rich resource to explore the complex topic of epigenetic mechanisms, which alter gene expression. Filamentous fungi are easy to propagate and have provided important insights over the years into complex biological functions in higher eukaryotes. In this dissertation, I summarize my work in analyzing how gene expression is altered by multiple epigenetic phenomena.

Chromatin structure

Chromatin was first described in the late 1800s (2). Walther Flemming coined the term chromatin, which plainly means stainable material (2, 3). Emil Heitz described the distinct staining patterns, which differentiated euchromatin from heterochromatin (4), and in the 1970s Kornberg and Thomas described the histone proteins that make up the nucleosome, the basic unit of chromatin (2, 5). Nucleosomes are composed of a histone octamer containing two copies of the core histones: H2A, H2B, H3, and H4. Approximately ~150bp of DNA is wrapped 1.5 times around the histone octamer (6) and nucleosomes are deposited following DNA replication in S phase of the cell cycle (Figure 1.1) (7). There is a wealth of knowledge about chromatin, but there are still open

questions concerning how structural changes such as histone modifications are mechanistically linked to genome function.

There are multiple ways that factors which bind or interact with this DNA-protein substrate can change both local and global chromatin structure including, DNA methylation, chromatin remodelers, replacement of core histone with histone variants, and post-translational modifications to the unstructured histone tails (8). Most importantly, the way in which nucleosomes interact with DNA is conserved between higher eukaryotes and yeast, allowing studies in fungi to shed light on functions that cannot be studied in higher eukaryotes (9). Multiple protein complexes interact with chromatin and organize the genome into functionally distinct domains (10, 11).

The eukaryotic genome is divided into separate segments of euchromatin and heterochromatin (12). Euchromatic regions of the genome are accessible, include regions that contain methylation of the residue lysine 4 of histone H3 (H3K4me) (13), enrichment of the histone variant H2A.Z at transcription start sites (14), and are transcriptionally competent (Figure 1.2A) (12). Heterochromatic regions are typically less accessible and transcriptionally silent (15). The two overarching types of heterochromatin are: constitutive heterochromatin and facultative heterochromatin (16). Marked by tri-methylated lysine 9 on the histone H3 tail (H3K9me3), constitutive heterochromatin is found in repeat-rich regions of the genome and structural features, such as the centromere and telomeric repeats (17). In fungi, facultative heterochromatin is most abundantly located at the subtelomeric regions of the chromosome, although there are domains dispersed throughout the internal regions, and is marked by H3K27me2/3 (Figure 1.2B) (18). These regions are more gene-rich and represent areas of the genome where genes

are silent typically silent but can be expressed under specific conditions and as a response to exogenous or endogenous triggers (19). Although the genomic localization of H3K27me2/3 has been well studied in *N. crassa*, the functional roles are ill understood. Many genes that reside in these H3K27me2/3 marked domains encode for hypothetical proteins with unknown functions. In addition, the general mechanisms for deposition and maintenance of this histone modification are open questions.

The Polycomb group (PcG) proteins form the complexes that establish facultative heterochromatin by depositing the histone modification, H3K27me2/3 (20). In metazoans, PcG proteins are essential in development by repressing developmental loci in a temporal and cell lineage-specific manner (11, 21, 22). There are two main PcG protein complexes, Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2), which function cooperatively to maintain stable gene repression (23, 24). Certain fungi only contain the core components of the PcG protein complex PRC2 but lack any homologs to any subunits from PRC1; however, the model yeasts *Saccharomyces cerevisiae* and *Schizosacchamyces pombe* lack any PcG protein complexes (25). Establishment of H3K27 methylation has been associated with a multitude of other phenomena including, histone modifications (23), specific DNA sequences (26, 27), and noncoding RNAs (28). None of these models have been shown to be generally applicable and highlight the complexity of the PcG system in different organisms, including fungi.

In *Neurospora*, PRC2 is composed of three core subunits, SET-7 (often referred to as KMT6), EED, SUZ12, and one accessory subunit, CAC-3. PRC2 deposits mono- di- and tri-methyl groups on lysine 27 of histone H3. These components are conserved in plants, fungi, and animals. SET-7 is the catalytic subunit homologous to EZH1/EZH2 in

humans and curly leaf, medea, or swinger in *Arabidopsis thaliana* (CLF, MEA, SWN) (29-37). EED and SUZ12 are named as such in humans, and named Esc and su(z)12 in *Drosophila*, but in *Arabidopsis* EED is named fertilization independent endosperm (FIE) and SUZ12 is named embryonic flower 2, vernalization 2 or fertilization independent seed 2 (EMF2, VER2, FIS2) (36, 38, 39). CAC-3, the accessory subunit, is named mammalian retinoblastoma binding protein 46/48 (RBAP46/48) in humans and multicopy suppressor of IRA1-5 (MSI1-5) in *Arabidopsis* (40-42).

In mouse embryonic stem cells (mESCs) and plants, the histone variant H2A.Z has been implicated directly in the regulation of H3K27 methylation (43-45). H2A.Z is one of the most extensively studied and conserved histone variants and is enriched at transcription start sites (TSS) and vertebrate enhancers (46-53). H2A.Z has been functionally linked to a multitude of seemingly discordant functions including, gene activation and gene repression (46, 54-61). Although much is known about H2A.Z itself and its genomic location, mechanisms behind its various actions are not well understood. In mESCs, there is a strong correlation between the activity of PRC2, and colocalization of H2A.Z and H3K27me3 (62). This colocalization is found not only between the modification H3K27me3, but also with the PRC2 subunit SUZ12 in mESCs at developmentally important loci, such as the HOX clusters (48). The N- and C-terminal tails of H2A.Z can be acetylated or ubiquitylated, respectively, and these modifications either repress or stimulate the action of PRC2 through BRD2, a transcriptional activator or through Polycomb Repressive Complex 1 (PRC1) (43). In *Arabidopsis thaliana*, there is a genetic interaction between the chromatin remodeler PKL, which promotes H3K27me3, and PIE-1, the homologous protein to SWR-1, the histone variant exchanger

that deposits H2A.Z (45). In repressed genes in rice, H2A.Z is also found to be colocalized with H3K27me3 (44). The relationship between the histone variant H2A.Z and H3K27me3 is not fully understood in any eukaryote.

Facultative heterochromatin in fungi

Facultative heterochromatin has been studied in a multitude of organisms and in general functions as a system that controls conditional gene repression, but how this is accomplished is not fully understood (63). The classical or sequential model states that PRC2 methylates H3K27 and then PRC1 ubiquitylates H2A residues promoting compaction (11). This model has been refuted by studies which have found that PRC1 can be recruited without PRC2 activity in PRC2-deficient mESCs (64). Knockdown of PRC1 has also been shown to lead to decreased binding of PRC2 indicating that PRC1 is required for PRC2 stable association (65). In addition, facultative heterochromatin exists in organisms that lack PRC1, namely fungi. This represents a gap in knowledge for how PRC1 and PRC2 function with each other in higher eukaryotes, and more importantly how PRC2 achieves gene repression in systems that lack PRC1.

The percentage of the genome covered in H3K27me2/3 varies in different filamentous fungi that have been studied. It ranges from 5-7% in *Neurospora crassa* and *Cryptococcus neoformans* to 30% in *Fusarium graminearum* (25, 66, 67). In addition, in all the above and *Magnaporthe oryzae*, *Fusarium fujikuroi, Zymoseptoria tritici*, and *Epichloë festucae* the bulk of the H3K27 methylation found is close to the telomeres (68-70). Even in Fusarium where 30% of the genome is covered in this modification, the regions of the chromosomes that are H3K27me marked may be ancestral telomeric sites that fused to form the chromosomal makeup that we now know (67). Although there are

still many genes marked by H3K27me2/3 of unknown function, secondary metabolites and effector genes are some of the better understood genes that can be located in such regions. Subtelomeric regions in filamentous fungi are generally enriched for genes specific to that species, such as these secondary metabolite clusters or effector genes (71, 72).

Secondary metabolites (SM) are part of defense or survival mechanisms and not required for primary metabolism or growth and development (73). Genes encoding SM are typically found in clusters, expressed when the cell is under nonoptimal conditions and include important compounds such as, cyclosporin, penicillin, and lovastatin (73). These gene clusters have been found in subterminal regions of chromosomes and in certain filamentous fungi their expression is controlled by epigenetic regulation (67-69, 74). Filamentous fungal plant pathogens secrete effectors during infection, which hijack the plant's immune system; these effectors are small secreted proteins and the genes that encode them have been found to be under epigenetic control (75, 76). Under laboratory conditions in *Magnaporthe oryzae,* these effector genes are normally silent and the removal of KMT6 activity (H3K27 methylation) alters expression for many, but not all. It is interesting to note that in organisms with genes that are enriched for H3K27me2/3, removal of this modification is not sufficient for upregulation (66, 67, 77), indicating that there is another tier of regulation for subsets of these genes.

Abolishment of H3K27me2/3 in higher eukaryotes is lethal, which makes fungi good models to probe for insights into control and function of this modification. In organisms where the methyltransferase for H3K27 has been knocked-down or deleted there are wide-ranging effects including: upregulation of secondary metabolite and

effector genes, fruiting body formation defects (67, 77, 78), hyphal growth defects (69), conidiation defects (79), and loss of virulence (67, 79). When H3K27 methylation is removed or significantly depleted, in *E. festucae*, the alkaloid biosynthetic gene cluster exhibited derepression (68), 20 out of 47 predicted SM key enzymes were misregulated in *F. fujikuroi* (69), 14% of all genes became upregulated in *F. graminearum* (67), 75 subtelomeric transcripts were derepressed in *C. neoformans* (66), and 130 out of ~770 genes were upregulated in *N. crassa* (25). Although H3K27me2/3 itself is clearly important for gene repression of certain subsets of genes, these results clearly point to a complex system of regulation with multiple layers and likely more players involved.

In *Neurospora*, we know there are at least two categories of H3K27 methylation, telomere-dependent and telomere-independent (Figure 1.3); placement of repetitive telomere repeat sequences (5'-TTAGGG-3') in euchromatic loci induces de novo H3K27 methylation across large regions (80). ASH-1 catalyzed H3K36me2 was recently found to mark inactive or silent genes and 95% of H3K27me2/3 domains are also marked by this modification (81). A forward genetics screen identified a new protein required for subtelomeric H3K27 methylation, PRC2 accessory subunit (PAS) (78). Strains lacking PAS have upregulated expression of telomere-proximal genes as well as specific loss of H3K27me2/3 in subtelomeric regions (78). Prior to this study, CAC-3/NPF, another accessory subunit of PRC2 was found to have a similar pattern of H3K27me2/3 loss at subtelomeric regions (25). In addition, my research has led to the knowledge that H2A.Z is required for normal internal patterns of H3K27 methylation. Deletion of this histone variant results in downregulation of EED and region-specific loss of H3K27me2/3 genome wide.

References

1. Li B, Carey M, Workman JL. 2007. The role of chromatin during transcription. Cell 128:707-19.10.1016/j.cell.2007.01.015

2. Trojer P, Reinberg D. 2007. Facultative heterochromatin: Is there a distinctive molecular signature? Molecular Cell 28:1-13.10.1016/j.molcel.2007.09.011

3. Paweletz N. 2001. Walther Flemming: pioneer of mitosis research., http://www.nature.com/articles/35048077.

4. Passarge E. 1979. Emil Heitz and the concept of heterochromatin: longitudinal chromosome differentiation was recognized fifty years ago. Am J Hum Genet 31:106-115

5. May 24 1974. 4139. Chromatin structure; oligomers of the histones. Science, 184.865-868. http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=4825888&retmode=ref&cmd=prlinks.

6. Luger K, Mader AW, Richmond RK, Sargent DF, Richmond TJ. 1997. Crystal structure of the nucleosome core particle at 2.8 angstrom resolution. Nature 389:251-260

7. Gunjan A, Paik J, Verreault A. 2005. Regulation of histone synthesis and nucleosome assembly. Biochimie 87:625-635.10.1016/j.biochi.2005.02.008

8. Dulac C. 2010. Brain function and chromatin plasticity. Nature 465:728-35.10.1038/nature09231

9. Lewis ZA. 2017. Polycomb Group Systems in Fungi: New Models for Understanding Polycomb Repressive Complex 2. Trends Genet 33:220-231

10. Janssen A, Colmenares SU, Karpen GH. 2018. Heterochromatin: Guardian of the Genome. Annu Rev Cell Dev Biol 34:265-288.10.1146/annurev-cellbio-100617-062653

11. Schuettengruber B, Bourbon HM, Di Croce L, Cavalli G. 2017. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. Cell 171:34-57.10.1016/j.cell.2017.08.002

12. Tamaru H. 2010. Confining euchromatin/heterochromatin territory: jumonji crosses the line. Genes Dev 24:1465-78.10.1101/gad.1941010

13. Noma K, Allis CD, Grewal SI. 2001. Transitions in distinct histone H3 methylation patterns at the heterochromatin domain boundaries. Science 293:1150-5.10.1126/science.1064150

14. Giaimo BD, Ferrante F, Herchenrother A, Hake SB, Borggrefe T. 2019. The histone variant H2A.Z in gene regulation. Epigenetics Chromatin 12:37.10.1186/s13072-019-0274-9

15. Berlowitz L. 1965. Correlation of Genetic Activity, Heterochromatization, and Rna Metabolism. Proc Natl Acad Sci U S A 53:68-73

16. Strahl BD, Allis CD. 2000. The language of covalent histone modifications. Nature 403:41-45

17. Saksouk N, Simboeck E, Dejardin J. 2015. Constitutive heterochromatin formation and transcription in mammals. Epigenetics Chromatin 8:3.10.1186/1756-8935-8-3

18. Wiles ET, Selker EU. 2016. H3K27 methylation: a promiscuous repressive chromatin mark. Curr Opin Genet Dev 43:31-37.10.1016/j.gde.2016.11.001

19. Blackledge NP, Rose NR, Klose RJ. 2015. Targeting Polycomb systems to regulate gene expression: modifications to a complex story. Nat Rev Mol Cell Biol 16:643-649.10.1038/nrm4067

20. Margueron R, Reinberg D. 2011. The Polycomb complex PRC2 and its mark in life. Nature 469:343-9.10.1038/nature09784

21. Muller J. 1995. Transcriptional silencing by the Polycomb protein in Drosophila embryos. EMBO J 14:1209-20

22. Hennig L, Derkacheva M. 2009. Diversity of Polycomb group complexes in plants: same rules, different players? Trends Genet 25:414-23.10.1016/j.tig.2009.07.002

23. Simon JA, Kingston RE. 2009. Mechanisms of polycomb gene silencing: knowns and unknowns. Nat Rev Mol Cell Biol 10:697-708.10.1038/nrm2763

24. Kuroda MI, Kang H, De S, Kassis JA. 2020. Dynamic Competition of Polycomb and Trithorax in Transcriptional Programming. Annu Rev Biochem doi:10.1146/annurev-biochem-120219-103641.10.1146/annurev-biochem-120219-103641

25. Jamieson K, Rountree MR, Lewis ZA, Stajich JE, Selker EU. 2013. Regional control of histone H3 lysine 27 methylation in Neurospora. Proc Natl Acad Sci U S A 110:6027-32.10.1073/pnas.1303750110

26. Li H, Liefke R, Jiang J, Kurland JV, Tian W, Deng P, Zhang W, He Q, Patel DJ, Bulyk ML, Shi Y, Wang Z. 2017. Polycomb-like proteins link the PRC2 complex to CpG islands. Nature 549:287-291.10.1038/nature23881

27. Wang L, Brown JL, Cao R, Zhang Y, Kassis JA, Jones RS. 2004. Hierarchical recruitment of polycomb group silencing complexes. Mol Cell 14:637-46.10.1016/j.molcel.2004.05.009

28. Kohlmaier A, Savarese F, Lachner M, Martens J, Jenuwein T, Wutz A. 2004. A chromosomal memory triggered by Xist regulates histone methylation in X inactivation. PLoS Biol 2:E171.10.1371/journal.pbio.0020171

29. Jones RS, Gelbart WM. 1993. The Drosophila Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax. Mol Cell Biol 13:6357-66.10.1128/mcb.13.10.6357

30. Abel KJ, Brody LC, Valdes JM, Erdos MR, McKinley DR, Castilla LH, Merajver SD, Couch FJ, Friedman LS, Ostermeyer EA, Lynch ED, King MC, Welcsh PL, Osborne-Lawrence S, Spillman M, Bowcock AM, Collins FS, Weber BL. 1996. Characterization of EZH1, a human homolog of Drosophila Enhancer of zeste near BRCA1. Genomics 37:161-71.10.1006/geno.1996.0537

31. Chen H, Rossier C, Antonarakis SE. 1996. Cloning of a human homolog of the Drosophila enhancer of zeste gene (EZH2) that maps to chromosome 21q22.2. Genomics 38:30-7.10.1006/geno.1996.0588

32. Goodrich J, Puangsomlee P, Martin M, Long D, Meyerowitz EM, Coupland G. 1997. A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. Nature 386:44-51.10.1038/386044a0

33. Grossniklaus U, Vielle-Calzada JP, Hoeppner MA, Gagliano WB. 1998. Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis. Science 280:446-50.10.1126/science.280.5362.446

34. Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. 2002. Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal polycomb sites. Cell 111:185-196.Doi 10.1016/S0092-8674(02)00975-3

35. Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D. 2002. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev 16:2893-905.10.1101/gad.1035902

36. Chanvivattana Y, Bishopp A, Schubert D, Stock C, Moon YH, Sung ZR, Goodrich J. 2004. Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. Development 131:5263-76.10.1242/dev.01400

37. Schwartz YB, Pirrotta V. 2007. Polycomb silencing mechanisms and the management of genomic programmes. Nat Rev Genet 8:9-22.10.1038/nrg1981

38. Schumacher A, Lichtarge O, Schwartz S, Magnuson T. 1998. The murine Polycomb-group gene eed and its human orthologue: functional implications of evolutionary conservation. Genomics 54:79-88.10.1006/geno.1998.5509

39. Birve A, Sengupta AK, Beuchle D, Larsson J, Kennison JA, Rasmuson-Lestander A, Muller J. 2001. Su(z)12, a novel Drosophila Polycomb group gene that is conserved in vertebrates and plants. Development 128:3371-9

40. Huang S, Lee WH, Lee EY. 1991. A cellular protein that competes with SV40 T antigen for binding to the retinoblastoma gene product. Nature 350:160-2.10.1038/350160a0

41. Dong J, Gao Z, Liu S, Li G, Yang Z, Huang H, Xu L. 2013. SLIDE, the protein interacting domain of Imitation Switch remodelers, binds DDT-domain proteins of different subfamilies in chromatin remodeling complexes. J Integr Plant Biol 55:928-37.10.1111/jipb.12069

42. Derkacheva M, Steinbach Y, Wildhaber T, Mozgova I, Mahrez W, Nanni P, Bischof S, Gruissem W, Hennig L. 2013. Arabidopsis MSI1 connects LHP1 to PRC2 complexes. EMBO J 32:2073-85.10.1038/emboj.2013.145

43. Surface LE, Fields PA, Subramanian V, Behmer R, Udeshi N, Peach SE, Carr SA, Jaffe JD, Boyer LA. 2016. H2A.Z.1 Monoubiquitylation Antagonizes BRD2 to Maintain Poised Chromatin in ESCs. Cell Rep 14:1142-1155.10.1016/j.celrep.2015.12.100

44. Zhang K, Xu W, Wang C, Yi X, Zhang W, Su Z. 2017. Differential deposition of H2A.Z in combination with histone modifications within related genes in Oryza sativa callus and seedling. Plant J 89:264-277.10.1111/tpj.13381

45. Carter B, Bishop B, Ho KK, Huang R, Jia W, Zhang H, Pascuzzi PE, Deal RB, Ogas J. 2018. The Chromatin Remodelers PKL and PIE1 Act in an Epigenetic Pathway That Determines H3K27me3 Homeostasis in Arabidopsis. Plant Cell 30:1337-1352.10.1105/tpc.17.00867

46. Guillemette B, Bataille AR, Gevry N, Adam M, Blanchette M, Robert F, Gaudreau L. 2005. Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. Plos Biology 3:2100-2110.ARTN e38410.1371/journal.pbio.0030384

47. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. Cell 129:823-37.10.1016/j.cell.2007.05.009

48. Creyghton MP, Markoulaki S, Levine SS, Hanna J, Lodato MA, Sha K, Young RA, Jaenisch R, Boyer LA. 2008. H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment. Cell 135:649-61.10.1016/j.cell.2008.09.056

49. Bargaje R, Alam MP, Patowary A, Sarkar M, Ali T, Gupta S, Garg M, Singh M, Purkanti R, Scaria V, Sivasubbu S, Brahmachari V, Pillai B. 2012. Proximity of H2A.Z containing nucleosome to the transcription start site influences gene expression levels in the mammalian liver and brain. Nucleic Acids Res 40:8965-78.10.1093/nar/gks665

50. Weber CM, Ramachandran S, Henikoff S. 2014. Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. Mol Cell 53:819-30.10.1016/j.molcel.2014.02.014

51. Latorre I, Chesney MA, Garrigues JM, Stempor P, Appert A, Francesconi M, Strome S, Ahringer J. 2015. The DREAM complex promotes gene body H2A.Z for target repression. Genes Dev 29:495-500.10.1101/gad.255810.114

52. Dai X, Bai Y, Zhao L, Dou X, Liu Y, Wang L, Li Y, Li W, Hui Y, Huang X, Wang Z, Qin Y. 2017. H2A.Z Represses Gene Expression by Modulating Promoter Nucleosome Structure and Enhancer Histone Modifications in Arabidopsis. Mol Plant 10:1274-1292.10.1016/j.molp.2017.09.007

53. Gomez-Zambrano A, Merini W, Calonje M. 2019. The repressive role of Arabidopsis H2A.Z in transcriptional regulation depends on AtBMI1 activity. Nat Commun 10:2828.10.1038/s41467-019-10773-1

54. Adam M, Robert F, Larochelle M, Gaudreau L. 2001. H2A.Z is required for global chromatin integrity and for recruitment of RNA polymerase II under specific conditions. Molecular and Cellular Biology 21:6270-6279

55. Meneghini MD, Wu M, Madhani HD. 2003. Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. Cell 112:725-36.10.1016/s0092-8674(03)00123-5

56. Rangasamy D, Greaves I, Tremethick DJ. 2004. RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. Nat Struct Mol Biol 11:650-5.10.1038/nsmb786

57. Bruce K, Myers FA, Mantouvalou E, Lefevre P, Greaves I, Bonifer C, Tremethick DJ, Thorne AW, Crane-Robinson C. 2005. The replacement histone H2A.Z in a

hyperacetylated form is a feature of active genes in the chicken. Nucleic Acids Res 33:5633-9.10.1093/nar/gki874

58. Dhillon N, Oki M, Szyjka SJ, Aparicio OM, Kamakaka RT. 2006. H2A.Z functions to regulate progression through the cell cycle. Mol Cell Biol 26:489-501.10.1128/MCB.26.2.489-501.2006

59. Xu Y, Ayrapetov MK, Xu C, Gursoy-Yuzugullu O, Hu Y, Price BD. 2012. Histone H2A.Z controls a critical chromatin remodeling step required for DNA double-strand break repair. Mol Cell 48:723-33.10.1016/j.molcel.2012.09.026

60. Hu G, Cui K, Northrup D, Liu C, Wang C, Tang Q, Ge K, Levens D, Crane-Robinson C, Zhao K. 2013. H2A.Z facilitates access of active and repressive complexes to chromatin in embryonic stem cell self-renewal and differentiation. Cell Stem Cell 12:180-92.10.1016/j.stem.2012.11.003

61. Neves LT, Douglass S, Spreafico R, Venkataramanan S, Kress TL, Johnson TL. 2017. The histone variant H2A.Z promotes efficient cotranscriptional splicing in S. cerevisiae. Genes & Development 31:702-717

62. Wang Y, Long H, Yu J, Dong L, Wassef M, Zhuo B, Li X, Zhao J, Wang M, Liu C, Wen Z, Chang L, Chen P, Wang QF, Xu X, Margueron R, Li G. 2018. Histone variants H2A.Z and H3.3 coordinately regulate PRC2-dependent H3K27me3 deposition and gene expression regulation in mES cells. BMC Biol 16:107.10.1186/s12915-018-0568-6

63. Golbabapour S, Majid NA, Hassandarvish P, Hajrezaie M, Abdulla MA, Hadi AH. 2013. Gene silencing and Polycomb group proteins: an overview of their structure, mechanisms and phylogenetics. OMICS 17:283-96.10.1089/omi.2012.0105

64. Tavares L, Dimitrova E, Oxley D, Webster J, Poot R, Demmers J, Bezstarosti K, Taylor S, Ura H, Koide H, Wutz A, Vidal M, Elderkin S, Brockdorff N. 2012. RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. Cell 148:664-78.10.1016/j.cell.2011.12.029

65. Blackledge NP, Farcas AM, Kondo T, King HW, McGouran JF, Hanssen LL, Ito S, Cooper S, Kondo K, Koseki Y, Ishikura T, Long HK, Sheahan TW, Brockdorff N, Kessler BM, Koseki H, Klose RJ. 2014. Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. Cell 157:1445-59.10.1016/j.cell.2014.05.004

66. Dumesic PA, Homer CM, Moresco JJ, Pack LR, Shanle EK, Coyle SM, Strahl BD, Fujimori DG, Yates JR, Madhani HD. 2015. Product Binding Enforces the

Genomic Specificity of a Yeast Polycomb Repressive Complex. Cell 160:204-218.10.1016/j.cell.2014.11.039

67. Connolly LR, Smith KM, Freitag M. 2013. The Fusarium graminearum histone H3 K27 methyltransferase KMT6 regulates development and expression of secondary metabolite gene clusters. PLoS Genet 9:e1003916.10.1371/journal.pgen.1003916

68. Chujo T, Scott B. 2014. Histone H3K9 and H3K27 methylation regulates fungal alkaloid biosynthesis in a fungal endophyte-plant symbiosis. Mol Microbiol 92:413-34.10.1111/mmi.12567

69. Studt L, Rosler SM, Burkhardt I, Arndt B, Freitag M, Humpf HU, Dickschat JS, Tudzynski B. 2016. Knock-down of the methyltransferase Kmt6 relieves H3K27me3 and results in induction of cryptic and otherwise silent secondary metabolite gene clusters in Fusarium fujikuroi. Environmental Microbiology 18:4037-4054.10.1111/1462-2920.13427

70. Schotanus K, Soyer JL, Connolly LR, Grandaubert J, Happel P, Smith KM, Freitag M, Stukenbrock EH. 2015. Histone modifications rather than the novel regional centromeres of Zymoseptoria tritici distinguish core and accessory chromosomes. Epigenetics & Chromatin 8.ARTN 4110.1186/s13072-015-0033-5

71. Dean RA, Talbot NJ, Ebbole DJ, Farman ML, Mitchell TK, Orbach MJ, Thon M, Kulkarni R, Xu JR, Pan H, Read ND, Lee YH, Carbone I, Brown D, Oh YY, Donofrio N, Jeong JS, Soanes DM, Djonovic S, Kolomiets E, Rehmeyer C, Li W, Harding M, Kim S, Lebrun MH, Bohnert H, Coughlan S, Butler J, Calvo S, Ma LJ, Nicol R, Purcell S, Nusbaum C, Galagan JE, Birren BW. 2005. The genome sequence of the rice blast fungus Magnaporthe grisea. Nature 434:980-6.10.1038/nature03449

72. Fedorova ND, Khaldi N, Joardar VS, Maiti R, Amedeo P, Anderson MJ, Crabtree J, Silva JC, Badger JH, Albarraq A, Angiuoli S, Bussey H, Bowyer P, Cotty PJ, Dyer PS, Egan A, Galens K, Fraser-Liggett CM, Haas BJ, Inman JM, Kent R, Lemieux S, Malavazi I, Orvis J, Roemer T, Ronning CM, Sundaram JP, Sutton G, Turner G, Venter JC, White OR, Whitty BR, Youngman P, Wolfe KH, Goldman GH, Wortman JR, Jiang B, Denning DW, Nierman WC. 2008. Genomic islands in the pathogenic filamentous fungus Aspergillus fumigatus. PLoS Genet 4:e1000046.10.1371/journal.pgen.1000046

73. Boruta T. 2018. Uncovering the repertoire of fungal secondary metabolites: From Fleming's laboratory to the International Space Station. Bioengineered 9:12-16.10.1080/21655979.2017.1341022

74. Niehaus EM, Studt L, von Bargen KW, Kummer W, Humpf HU, Reuter G, Tudzynski B. 2016. Sound of silence: the beauvericin cluster in Fusarium

fujikuroi is controlled by cluster-specific and global regulators mediated by H3K27 modification. Environ Microbiol 18:4282-4302.10.1111/1462-2920.13576

75. Oliva R, Win J, Raffaele S, Boutemy L, Bozkurt TO, Chaparro-Garcia A, Segretin ME, Stam R, Schornack S, Cano LM, van Damme M, Huitema E, Thines M, Banfield MJ, Kamoun S. 2010. Recent developments in effector biology of filamentous plant pathogens. Cell Microbiol 12:705-15.10.1111/j.1462-5822.2010.01471.x

76. Soyer JL, El Ghalid M, Glaser N, Ollivier B, Linglin J, Grandaubert J, Balesdent MH, Connolly LR, Freitag M, Rouxel T, Fudal I. 2014. Epigenetic control of effector gene expression in the plant pathogenic fungus Leptosphaeria maculans. PLoS Genet 10:e1004227.10.1371/journal.pgen.1004227

77. Basenko EY, Sasaki T, Ji LX, Prybol CJ, Burckhardt RM, Schmitz RJ, Lewis ZA. 2015. Genome-wide redistribution of H3K27me3 is linked to genotoxic stress and defective growth. Proceedings of the National Academy of Sciences of the United States of America 112:E6339-E6348.10.1073/pnas.1511377112

78. McNaught KJ, Wiles ET, Selker EU. 2020. Identification of a PRC2 accessory subunit required for subtelomeric H3K27 methylation in Neurospora. Mol Cell Biol doi:10.1128/MCB.00003-20.10.1128/MCB.00003-20

79. Pham KT, Inoue Y, Vu BV, Nguyen HH, Nakayashiki T, Ikeda K, Nakayashiki H. 2015. Correction: MoSET1 (Histone H3K4 Methyltransferase in Magnaporthe oryzae) Regulates Global Gene Expression during Infection-Related Morphogenesis. PLoS Genet 11:e1005752.10.1371/journal.pgen.1005752

80. Jamieson K, McNaught KJ, Ormsby T, Leggett NA, Honda S, Selker EU. 2018. Telomere repeats induce domains of H3K27 methylation in Neurospora. Elife 7.10.7554/eLife.31216

81. Bicocca VT, Ormsby T, Adhvaryu KK, Honda S, Selker EU. 2018. ASH1-catalyzed H3K36 methylation drives gene repression and marks H3K27me2/3-competent chromatin. Elife 7.10.7554/eLife.41497

Figures

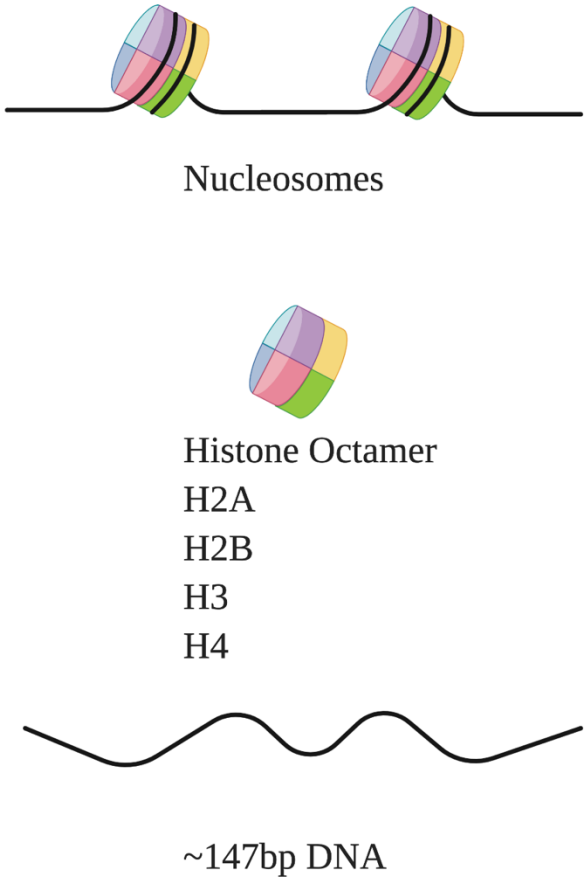


Figure 1.1: Nucleosome composition and core histones

A) Nucleosomes are composed of ~147 bp of DNA wrapped 1.5 times around an octamer of histone subunits. There are two copies of the core histone subunits H2A, H2B, H3, and H4 in the histone octamer. Image created with BioRender.com. Adapted from Takatori *et al* 2013.

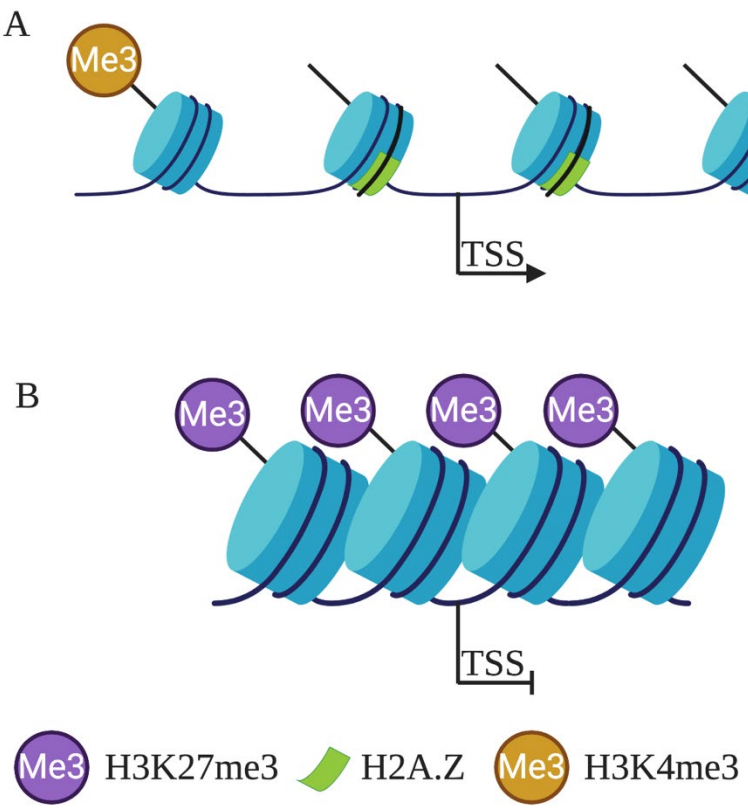


Figure 1.2: Euchromatin and Heterochromatin

A) Euchromatin is loosely packed, transcriptionally active, contains regions of H3K4me3, and the histone variant H2A.Z flanks the transcription start site (TSS). B) Heterochromatin is densely packed, transcriptionally silent, and is marked by H3K27me3. Image created with BioRender.com.

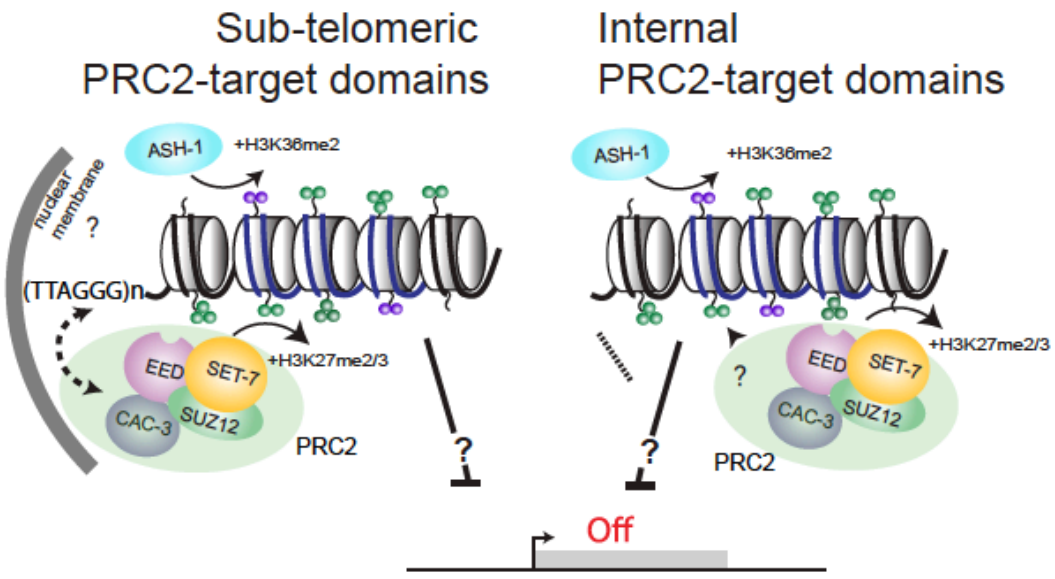


Figure 1.3: Two mechanisms of Polycomb mediated silencing in *Neurospora*

Sub-telomeric and internal regions have two mechanisms by which H3K27me2/3 is established. PRC2 catalyzes methylation of H3K27 through the action of SET-7, the catalytic subunit, which leads to repression of genes in these domains. The telomere repeat is able to recruit PRC2 where PRC2 can then deposit and H3K27me2/3. An unknown mechanism recruits PRC2 to internal domains. Many genes found in H3K27me2/3 domains are also enriched for H3K36me2.

CHAPTER 2

THE HISTONE MODIFICATION H3K27ME3 AND TRANSCRIPTION FACTOR MOGTI1 COORDINATELY CONTROL EXPRESSION OF EFFECTOR GENES IN RICE BLAST FUNGUS[1][2]

[1] Abigail J. Courtney, Jie Zhu, Peng Qi, Katrien Devos, Zachary A. Lewis and Chang Hyun Khang. To be submitted to *Plos Genetics*.

[2] Abigail J. Courtney and Jie Zhu are equal co-authors. AJC Experimental Contributions: Aided in experimental design of ChIP-seq and RNA-seq experiments. Constructed ChIP- and RNA-seq libraries. Performed bioinformatic analyses of ChIP-seq, RNA-seq, and microarray data and generated visualizations for all sequencing related experiments.

Abstract

Fungal pathogens are a major threat to global food production. Plant infection by fungi involves the secretion of effector proteins, which is imperative in facilitating invasion of plant cells. In *Magnaporthe oryzae*, the expression of effector genes is repressed in mycelial culture, but becomes induced during plant infection; however, the mechanisms governing this transcriptional reprogramming are not yet understood. We found through chromatin immunoprecipitation followed by Illumina sequencing that 40% of a total of 178 known and predicted effector genes possess enrichment of the repressive histone modification H3K27me3 during mycelial growth. Deletion of the H3K27me3 methyltransferase, KMT6, derepressed many of these effector genes. In addition, the overexpression of the transcription factor *MoGti1* also upregulated the expression of these genes, and in a KMT6 deletion background, the effect is synergistic for 21 effector genes. The synergistically upregulated genes were found to be upregulated in previous microarray studies that examined post infection expression in *M. oryzae*. Together, these data suggest that there are two components, H3K27me3 and *MoGti1,* which are important in the transcriptional regulation of effector genes needed for plant colonization.

Introduction

The emergence of filamentous fungal pathogens is a major threat to global food production. During plant infection, fungal pathogens secrete effector molecules to exploit host plants and to facilitate invasion of plant cells. Effectors typically function as essential pathogenicity determinants of plant-pathogen interactions. For instance, avirulence (Avr) effector proteins can interfere with the plant immune response to promote colonization in the absence of the cognate host resistant (R) proteins, referred to

as effector-triggered susceptibility (ETS). In the presence of corresponding R proteins, however, plant defense response is triggered to suppress colonization (effector-triggered immunity, ETI). While the function of fungal effectors has been extensively studied (1, 2), the mechanisms regulating effector gene expression remain largely unknown.

During plant infection, filamentous fungi undergo transcriptional reprogramming to mediate colonization. Regulated expression of effector genes is one of the most critical reprogramming actions. Expression of effector genes is typically suppressed during mycelial growth in axenic culture but is highly induced during plant infection. This is consistent with *in planta*-specific functions of effectors to facilitate fungal invasion. Transcriptomic studies in both oomycetes and fungi have revealed concerted waves of effector gene expression during different infection stages, suggesting that globally induced effector gene expression during plant infection is a shared property of filamentous pathogens (3-8). Thus, this expression pattern raises the question of how effector genes are globally repressed during mycelial growth and activated in waves during plant infection.

Filamentous fungi are multicellular organisms. All cells are genetically homogeneous but can be structurally and functionally heterogeneous at different development stages, due to differential gene expression. As mentioned above, expression of effector genes is typically repressed in mycelia and conidia, but highly induced in appressoria and invasive hyphae produced during plant infection. Heterochromatic regions possess less accessibility than euchromatin and are transcriptionally silent (9). The two overarching types of heterochromatin are constitutive heterochromatin and facultative heterochromatin (10). Constitutive heterochromatin is marked by tri-

methylated lysine 9 on the histone H3 tail (H3K9me3) and found in repeat-rich regions of the genome (11). In fungi, facultative heterochromatin is most abundantly located at the subtelomeric regions of the chromosome and is marked by H3K27me2/3 (12). Accumulating evidence suggests that epigenetic regulation in the form of these repressive chromatin modifications might be important for global repression of effector genes in a number of filamentous pathogens. Indeed, in *Leptosphaeria maculans,* expression of small secreted protein (SSP)-encoding genes was induced in axenic culture by silencing epigenetic regulators including *Heterochromatin Protein 1* (*HP1*) or *DIM5,* which catalyzes H3K9me3 (13). It still remains to be investigated whether other epigenetic modifications (e.g., H3K27me3) are involved in regulating effector gene expression. Nonetheless, absence of H3K27me3 derepresses expression of secondary metabolite (SM) biosynthetic genes as well as putative secreted pathogenicity factors in *Fusarium graminearum* (14), possibly indicating that H3K27me3 also plays roles in effector gene expression. Additionally, a recent study reports that deletion of *KMT1* and *KMT6* catalyzing trimethylation of H3K9 and H3K27, respectively, deregulates expression of pathogenicity-related genes, including some putative effector genes, in *Zymoseptoria tritici* (15).

H3K27 methylation is found mostly in subtelomeric regions on all chromosomes in the filamentous fungi *Neurospora crassa*, *F. graminearum* and *Cryptococcus neoformans* (16). Interestingly, the subtelomeric regions in *Magnaporthe oryzae* and *Phytophthora infestans* are characterized by a high density of Transposable Elements (TEs) and known effector genes (17-19). This is consistent with the idea that

H3K27me3, a well-characterized transcriptionally repressive mark, might repress gene expression during growth outside the host.

Conversely, the concerted waves of effector gene expression that occur during plant infection suggest that one or more global transcriptional regulators must direct coordinated activation of effector genes *in planta*. One good candidate is the WOPR transcription factor, which was initially characterized as a master regulator of morphological switching and virulence in the human fungal pathogen *Candida albicans* (white-opaque regulator 1, Wor1; (20)). Wor1 homologs were subsequently identified in various plant pathogenic fungi; all deletion mutants show both up- and downregulation of many effector genes (21-27). However, it is still unknown whether the WOPR transcription factor is under epigenetic control or can regulate effector gene expression in a cooperative manner with epigenetic machinery.

In this study, we profiled genome-wide distributions of histone modifications (H3K4me2, H3K9me3, H3K27me3 and H3K36me3) during mycelial growth of rice blast fungus *Magnaporthe oryzae* using chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq). We found that a large proportion of known and predicted effector genes in *M. oryzae* are enriched with H3K27me3. Then we demonstrated that H3K27me3 is involved in epigenetic repression of effector gene expression during mycelial growth using RNA-sequencing and quantitative PCR (qRT-PCR). Interestingly, many effector genes that exhibited de-repression due to loss of H3K27me3 by *MoKMT6* knockout during mycelial growth were also induced during plant infection of a wild-type strain at 36 hours post inoculation (hpi). To investigate transcriptional activation of effector genes, we overexpressed MoGti1, an *M. oryzae*

Wor1 homolog, and found expression of some effector genes were significantly upregulated even during mycelial growth. Then, we investigated the possible interplay between epigenetic (H3K27me3) and transcriptional (MoGti1) regulation on effector gene expression and uncovered a synergistic effect when *MoGti1* was overexpressed in the mutant strain lacking H3K27me3. Our data demonstrate H3K27me3-mediated repression of effector genes during mycelial growth and MoGti1-mediated activation of effector genes during plant infection. Taken together, this strongly suggests synchronized control of effector gene expression by epigenetic and transcriptional regulation in a temporal manner.

Results

Histone modifications of known and predicted effector genes in *M. oryzae*

To investigate the role of histone modifications in regulating effector gene expression, we carried out ChIP-seq to determine the genome-wide distributions of four well-conserved histone H3 modifications: H3K4me2, H3K9me3, H3K27me3 and H3K36me3. Chromatin was extracted from *M. oryzae* mycelia cultured in complete medium (hereafter referred to as "axenic culture"), in which most *M. oryzae* effector genes are transcriptionally repressed (28). Our ChIP-seq results revealed that two repressive histone marks, H3K9me3 and H3K27me3, were predominantly enriched in presumed heterochromatic regions, such as chromosomal ends, which contrasts to the distribution of the active H3K4me2 mark (Fig 2.1). Distributions of H3K9me3 and H3K27me3 were not colocalized but neighboring. Each mark was deposited at distinct genomic regions with little to no overlap (Fig 2.1). As reported for other fungi, most

chromosomes exhibit enrichment of H3K27me3 in subtelomeric regions with additional domains of H3K27me3 found at internal regions.

Based on existing literature, we curated a list of 24 *M. oryzae* effector genes that are repressed during mycelial growth in axenic culture but activated during various infection stages (2.1 Table). These effector genes include avirulence genes and pathogenicity genes. Loci for the 23 of the 24 previously identified effector genes are located in or near H3K27me3 domains (Fig 2.1). We constructed heatmaps to depict enrichment of H3K27me3, H3K36me3, H3K9me3, and H3K4me2 across the transcription start sites (TSS) of all 24 known effector genes. This confirmed high enrichment of H3K27me3 as well as H3K36me3 spanning both the promoters and gene bodies, whereas H3K4me2 and H3K9me3 were not enriched in these effector genes (Table 2.1). This striking association of the repressed effector genes with silencing H3K27me3 marks prompted us to expand our analysis to a total of 448 known and predicted effector genes that we identified in *M. oryzae* isolate O-137. We found that ~30% of these effector genes (132/448) are enriched with H3K27me3. In contrast, only ~6.2% (819/13,144) of the entire *M. oryzae* genome is enriched with H3K27me3. Then, we constructed heatmaps to depict enrichment of H3K27me3, H3K36me3, H3K9me3, and H3K4me2 across the TSS of all 448 known and predicted effector genes (Fig 2.2). We performed k-means clustering with k=2 based on H3K27 methylation enrichment and this method placed 40% of the effector genes into cluster 1, based on higher levels of H3K27me3 (40%, 178/448). These genes also displayed higher levels of H3K36me3 enrichment in both the promoter and the gene body, mirroring our previous results (Fig 2.2). However, the remaining 60% (268/448) of effector genes were grouped into cluster

2 based on their lower enrichment of H3K27me3 (Fig 2.2), including a constitutively expressed effector gene *MC69* (29). In addition, H3K36me3 enrichment at effector gene loci in cluster 2 is lower than that of cluster 1 (Fig 2.2). H3K9me3 enrichment at both clusters is comparable to the level of H3K9me3 throughout the entire genome, showing no bias for H3K9me3 at these loci. The patterns exhibited by both clusters of all modifications are strikingly different than the pattern of modifications when viewed on the TSSs for all genes (Fig 2.2B). Our clusters have little to no H3K4me3, as opposed to all genes where there is the characteristic spike of H3K4me3 immediately after the TSS (Fig 2.2B).

Deletion of the histone methyltransferase *MoKMT6* eliminates H3K27 trimethylation in *M. oryzae*

To further characterize the role of H3K27me3 in regulating effector gene expression, we decided to focus on *MoKMT6* (MGG_00152) in *M. oryzae*. *MoKMT6* was previously cloned from a wheat-pathogenic strain of *M. oryzae* and shown to encode a histone methyltransferase responsible for catalyzing methylation of H3K27 (30). We first identified the same gene in the rice-pathogenic *M. oryzae* strain O-137 by BLAST analysis and used a homologous recombination strategy to generate *MoKMT6* deletion mutants. Compared to wild-type and ectopic strains, all deletion mutants (n=8) produced noticeably lighter mycelia when grown on conidia-inducing OMA media, indicating defects in conidiation. Indeed, the Δ*mokmt6* deletion mutant strain CKF3472 (chosen for further studies) showed reduced conidiation and virulence on rice (Fig 2.3A and Table 2.2). Complementation of the deletion strain with the wild-type allele restored virulence defects but not conidiation (Fig 2.3A and Table 2.2). These phenotypes were consistent

with previous observations where *MoKMT6* was disrupted in a wheat-pathogenic strain of *M. oryzae* (30). ChIP-seq analysis clearly showed that the enrichment of H3K27me3 was lost in the Δ*mokmt6* mutant but was restored in the complementation strain (Fig 2.3B), confirming that *MoKMT6* is responsible for deposition of H3K27me3 in *M. oryzae*.

Loss of H3K27me3 results in de-repression of effector genes

We used qRT-PCR to compare the expression of nine effector genes during mycelial growth in axenic culture of wild type and Δ*mokmt6*. These effector genes include five AVR genes (*AVR-Pik*, *AVRPiz-t*, *AVR-Pi9*, *ACE1*, and *PWL2*), two biotrophy-associated genes (*BAS3* and *BAS4*), and two virulence genes (*Slp1* and *MC69*). Seven out of nine genes were enriched with H3K27me3 in wild type but not in in the Δ*mokmt6* mutant during mycelial growth (Fig 2.4 and Table 2.3). We found that the transcript levels of all seven effector genes in the Δ*mokmt6* mutant were significantly increased, ranging from 3.5-fold (i.e, *PWL2*) to over 150-fold (i.e, *AVRPiz-t* and *BAS4*), compared with those in wild type (Fig 2.4A, 2.4B, and Table 2.3). These results indicate that H3K27me3 is involved in repressing expression of these effector genes during mycelial growth in axenic culture and that, consequently, loss of H3K27me3 results in de-repression. Furthermore, we found that the expression of *MC69*, which has no enrichment of H3K27me3 during mycelial growth of wild type, was not changed in the Δ*mokmt6* mutant (Fig 2.4C). Interestingly, *AVR-Pi9* was clustered into the group of lower H3K27me3 enrichment in our hierarchical clustering, but the genome-wide loss of H3K27me3 led to its de-repression during mycelial growth (Fig 2.4A), indicating the

H3K27me3 regulates effector gene expression through both direct and indirect mechanisms.

Regulation of effector gene expression by the transcription factor MoGti1

Gene expression is increasingly shown to be regulated by both histone modifications and transcription factors (TFs) (31, 32). To investigate the interrelationship between H3K27me3 and TFs in regulation of effector gene expression, we focused on the *M. oryzae* TF MoGti1 and effector gene *BAS4*. MoGti1 is a homolog of the Wor1 TF, and deletion of *MoGti1* in *M. oryzae* results in altered expression of several effector genes, including *BAS4* (26). *BAS4* expression was strongly induced immediately after appressorium-mediated host penetration at 25 hpi (Fig 2.5A and 5B). We found that the *BAS4* promoter contains the core motif (TTAAAGTTT), recognized by Wor1 (33), suggesting that the expression of *BAS4* is directly regulated by MoGti1. This is further supported by qRT-PCR results revealing low expression of *MoGti1* and *BAS4* during axenic growth and increased expression during the course of plant infection, with earlier induction of *MoGti1* (Fig 2.5B).

To further confirm that MoGti1 positively regulates *BAS4* expression, we transformed *M. oryzae* ectopically with the *MoGti1* coding sequence under control of the constitutive promoter of the ribosomal protein *RP27* gene. Four randomly selected transformants showed more than 10-fold increase of *MoGti1* expression, compared to wild type, in axenic culture. One transformant (MoGti1oe; CKF3790AB) that exhibited the highest fold increase, ~55-fold, was chosen for further experiments. We found that the transcript level of *BAS4* was significantly increased, ~250-fold, in MoGti1oe relative to wild type in axenic culture (Fig 2.5C).

We then examined another eight effector genes (*AVR-Pik*, *AVR-Pi9*, *AVRPiz-t*, *PWL2*, *ACE1*, *MC69*, *Slp1,* and *BAS3*) in both wild type and MoGti1oe during mycelial growth in axenic culture (Fig 2.5C and Table 2.3). Similar to *BAS4*, the expression of *AVR-Pik*, *PWL2,* and *BAS3* was significantly upregulated in MoGti1oe compared to wild type. Interestingly, we also found that the expression of some effector genes was downregulated (i.e., *AVR-Pi9*), or not altered (i.e., *AVRPiz-t*, *ACE1*, *Slp1*, and *MC69*) in MoGti1oe compared to wild type. Taken together, we suggest that MoGti1 plays both a positive and a negative role in regulating expression of a subset of effector genes in *M. oryzae*.

Double control of effector gene expression by H3K27me3 and MoGti1

Since deletion of *MoKMT6* (*Δmokmt6*) or overexpression of *MoGti1* (MoGti1oe) leads to increased expression of *BAS4* and other effector genes (Fig 2.4 and 2.5), we asked whether the combination of *Δmokmt6* and MoGti1oe would lead to an additive or synergistic effect on effector gene expression. To answer this, we generated *M. oryzae* strains, ectopically overexpressing *MoGti1* in the *Δmokmt6* mutant background and subsequently identified one strain, CKF4034 (*Δmokmt6*-MoGti1oe), in which the level of *MoGti1* overexpression was comparable to that in the MoGti1oe strain. The qRT-PCR assays with RNA isolated from axenic cultures showed that *Δmokmt6* or MoGti1oe alone increased the *BAS4* transcripts by ~150-fold or ~250-fold, respectively. Strikingly, the *Δmokmt6*-MoGti1oe strain increased *BAS4* expression by ~20,000-fold (Fig 2.6A). Similar results were also observed in another independent *Δmokmt6*-MoGti1oe transformant. This synergistic effect was consistently observed for other effector genes,

such as *AVR-Pik*, *BAS3* and *PWL2*, whose expression, just like *BAS4*, was increased by *Δmokmt6* or MoGti1oe alone (Table 2.3).

To gain insight into the synergistic effect of *Δmokmt6* and MoGti1oe on effector gene expression, we first tested if H3K27me3 plays a role in regulating *MoGti1* expression by measuring the *MoGti1* transcript in *Δmokmt6* using qRT-PCR. We found that *MoGti1* expression was significantly induced (~3-fold) in *Δmokmt6* compared to wild type, suggesting that H3K27me3 is somehow involved in repression of *MoGti1* expression (Fig 2.6B). In *F. graminearum* and *N. crassa*, H3K27me3 loss can upregulate both H3K27me3-marked and non-H3K27me3-marked genes (14, 34, 35). To determine if *MoGti1* is a H3K27me3-marked gene, we examined H3K27me3 distribution at the *MoGti1* locus in axenic culture. Interestingly, our ChIP-seq analyses demonstrate lack of H3K27me3 at the *MoGti1* locus, even though removal of H3K27me3 upregulates this gene (Fig 2.6B). Taken together, these results suggest that H3K27me3 is likely repressing an upstream activator of *MoGti1*.

Four effector genes (*AVR-Pik*, *BAS3*, *BAS4,* and *PWL2*) in our qRT-PCR analysis showed increased expression in MoGti1oe, as well as enrichment of H3K27me3 in wild type (Table 2.3). We then tested if *MoGti1* overexpression can alter H3K27me3 patterns at these loci. To this end, we first measured the transcripts of *MoKMT6* in both wild type and MoGti1oe and found no significant difference of *MoKMT6* expression between them (Fig 2.6C). ChIP-seq results showed no noticeable difference of the H3K27me3 patterns from wild type at individual loci (Fig 2.6C). Similarly, we found almost indistinguishable differences of H3K27me3 in MoGti1oe from wild type when viewed genome-wide using heatmaps of all wild-type peak locations (Fig 2.6D). These data suggest that altered

expression of effector genes due to MoGti1oe does not involve changes in the histone modification and silencing mark H3K27me3.

Previously, we showed that MoGti1 overexpression upregulated expression of a subset of effector genes (Fig 2.5C). The deletion of *MoKMT6* was able to induce expression of both *MoGti1* and effector genes in axenic culture (Fig 2.4 and 2.6B). This raises a question whether induced effector gene expression by *MoKMT6* deletion is solely due to upregulated *MoGti1* expression. Due to the comparable expression of *MoGti1* in both MoGti1oe and *Δmokmt6*-MoGti1oe strains, any differences of effector gene expression between these two strains is likely due to the presence or absence of H3K27me3. We indeed found that H3K27me3 loss led to the remarkable upregulation of effector gene expression (*AVR-Pik*, *BAS3*, *BAS4,* and *PWL2*) in *Δmokmt6*-MoGti1oe compared to MoGti1oe. This suggests that induced effector gene expression (*AVR-Pik*, *BAS3*, *BAS4,* and *PWL2*) by *MoKMT6* deletion is not solely due to upregulated *MoGti1* expression but also because of the absence of H3K27me3. This was further confirmed by induced expression of effector genes *ACE1*, *AVRPiz-t* and *AVR-Pi9* in *Δmokmt6*. Because their expression was unchanged or negatively regulated by MoGti1 (Fig 2.5), the upregulation of these effector genes in *Δmokmt6* is likely due to the loss of H3K27me3.

Trimethylation of H3K27 and the transcription factor MoGti1 globally control expression of effector genes

To investigate expression changes in strains lacking H3K27me3 and/or overexpressing MoGti1, we performed RNA-seq on wild-type, *Δmokmt6*, MoGti1oe, and *Δmokmt6*-MoGti1oe strains during mycelial growth in axenic culture. We found that expression of seven of nine qRT-PCR tested effector genes showed consistent expression

patterns in the RNA-seq data in *Δmokmt6* and MoGti1oe respectively (Table 2.4). The remaining two effector genes *AVR-Pik* and *PWL2* lacked read coverage in the wild-type strain but had sufficient read coverage in *Δmokmt6* and MoGti1oe based on RNA-seq, suggesting possible upregulated expression, which is also consistent with qRT-PCR results. The consistency between qRT-PCR and RNA-seq results serves as validation of these RNA-seq data. Overall, the expression levels of 32.1% (144/448) of known and predicted effector genes were differentially regulated by the removal of H3K27me3 including 125 upregulated and 19 downregulated genes and 42% (75/178) of the cluster 1 genes (enriched for H3K27me3) were upregulated upon removal of this modification (Table 2.5). Similarly, the overexpression of MoGti1 in mycelia led to an altered expression of ~20% (89/448) of this effector gene set, where we found 63 upregulated effector genes including *BAS3*, *BAS4* and *SPD5*, and 26 downregulated effector genes, including *AVR-Pi9* (Table 2.5). Considering that effector genes only represent 3.4% (448/13,144) of the genes in the entire *M. oryzae* genome, we suggest that both H3K27me3 and MoGti1 globally control expression of effector genes in *M. oryzae*.

H3K27me3 is involved in repressing expression of genome-wide effector genes

We next asked if expression of the 448 known and predicted effector genes changes upon removal of H3K27me3. We compared expression levels of H3K27me3-enriched (cluster 1) and -unenriched effector genes (cluster 2) in wild type and *Δmokmt6*. As expected, expression levels of cluster 1 effector genes in the wild-type strain were less than that of cluster 2 effector genes (Fig 2.7). However, the loss of H3K27me3 in *Δmokmt6* mutant resulted in significantly induced expression of cluster 1 effector genes, but not for cluster 2 effector genes (Fig 2.7). These data suggest that H3K27me3 plays an

important role in repressing expression of effector genes in *M. oryzae*. This is also supported by the higher percentage of upregulated effector genes (87%, 125/144) in *Δmokmt6* than that of downregulated genes (13%, 19/144). In particular, many of the known effector genes (15/24 genes), of which expression is repressed during mycelial growth (Table 2.1), were detected to be upregulated, while none were downregulated (Table 2.5). For the remaining 424 effector genes, the number of upregulated effector genes was approximately six times higher than that of downregulated genes (110 vs. 19).

Synergistic effect of H3K27me3 and MoGti1 on effector gene expression

Because effector genes that are typically silent in culture were upregulated in the absence of H3K27me3 and these genes are known to be upregulated during plant infection [24], we sought to determine if H3K27me3 and MoGti1 had a synergistic effect on their induction. Using the RNA-seq dataset we compared the *Δmokmt6*-MoGti1oe strain with *Δmokmt6* and MoGti1oe. We found that the *Δmokmt6*-MoGti1oe strain had a similar number of upregulated effector genes to the total number for both *Δmokmt6* and MoGti1oe strains combined (Table 2.4). Notably, the expression of 36 effector genes was significantly upregulated in both *Δmokmt6* mutant and MoGti1oe (Table 2.7). Out of the 36 genes, 58% (21/36) exhibit a fold change increase in the *Δmokmt6*-MoGti1oe strain compared to the individual strains. For example, *BAS4* is upregulated in *Δmokmt6* and MoGti1oe by 14.8-fold, and 37.6-fold, respectively; however, in *Δmokmt6*-MoGti1oe *BAS4* is upregulated by 1,531.8-fold (Table 2.7). Taken together, these data suggest that H3K27me3 and MoGti1 synergistically control expression of a subset of effector genes in *M. oryzae.* We analyzed the transcriptomic data generated from infected rice sheath by *M. oryzae* O-137 at 36 hpi from a previous study [24]. Interestingly, 17/21 (81%) of

these synergistically upregulated effector genes in *Δmokmt6*-MoGti1oe were also highly induced during biotrophic invasion of the rice sheath at 36 hpi by a wild-type *M. oryzae* strain (Table 2.7), indicating a temporal correlation between H3K27me3-MoGti1 regulatory mechanisms on effector gene expression and the biotrophic invasion stage.

Effector genes upregulated during biotrophic invasion are enriched with H3K27me3 during mycelial growth

Since expression of 15/24 (62.5%) known effector genes were derepressed in *Δmokmt6* (Table 2.4) and these genes are known to be induced during plant infection (Table 2.1), we asked if other predicted effector genes enriched for H3K27me3 are induced during plant infection. As expected, many predicted effector genes (33.7%, 151/448) were upregulated by more than two-fold at 36 hpi, including 75% (18/24) of known effector genes. Conversely, only 10.3% of all genes in the genome were induced at 36 hpi (Table 2.5). Almost 50% (86/178) of predicted effector genes enriched for H3K27me3 (cluster 1 genes) were induced by two-fold or more at 36 hpi. Conversely, only ~25% (65/268) of the remaining effector genes (cluster 2) were induced (Table 2.5). This suggests that, indeed, many effector genes induced during plant infection are enriched for H3K27me3 during mycelial growth. Considering that effector genes exhibit stage-specific expression across the infection cycle (3, 6), identification of 86 induced effector genes that are upregulated at 36 hpi and also enriched by H3K27me3 during mycelial growth is notable, highlighting a key role of H3K27me3 in regulation of pathogenicity genes. It is possible that other predicted effector genes marked with H3K27me3 during axenic growth will also exhibit increased expression *in planta* during earlier or later infection stages.

We next compared the gene expression changes that occur during plant infection at 36 hpi to the expression changes we observed in the Δ*mokmt6* strain. The overall percentage of genes upregulated in the Δ*mokmt6* strain and the *in planta* 36 hpi strain were similar at 8.0% (1053) and 10.3% (1351), respectively. In addition, the cluster 1 effector genes, showed the same upregulation pattern between the two, where Δ*mokmt6* had 42.1% (75/178) and 36 hpi had 48.3% (86/178) (Table 2.5). However, this was not true for the comparison of cluster 2 effector genes (31% [11/268]) and 36 hpi (24.3% [65/268]) (Table 2.5). For a more direct comparison of effector genes that are upregulated in both Δ*mokmt6* and 36 hpi, we restricted our search to genes that were upregulated by at least 2-fold in each. We identified 46 genes in common. Of these, 80% (37/46) are enriched for H3K27me3 and these include seven known effector genes such as, *AVRPiz-t* and *BAS4* (Table 2.6). Taken together, these results suggest there exists a mechanism during plant infection at 36 hpi that is able to reverse silencing mediated by KMT6 and H3K27me3.

Discussion

In this study, we have found that expression of effector genes in *M. oryzae* is controlled by two layers of regulatory control: epigenetic (H3K27me3) and transcriptional (MoGti1) regulation (Fig 2.8). Specifically, histone H3K27 trimethylation, deposited by MoKMT6, plays a role in repressing expression of effector genes that are enriched for H3K27me3 during mycelial growth (Fig 2.1, 2.2, 2.3, and 2.4). Our data suggests a possible reduction of H3K27me3 at effector gene loci during plant infection (Table 2.6). It is not clear yet if highly and widely upregulated effector gene expression during plant infection is due to altered histone modifications. Previously, reduced

H3K9me3 derepressed expression of effector genes in axenic culture but does not change expression pattern of effector genes during primary infection of *L. maculans* (13), indicating that the repressive histone modification (H3K9me3) might be reduced at effector gene loci during plant infection. In our study, absence of H3K27me3 derepresses expression of H3K27me3-enriched effector genes under non-inducible conditions (Fig 2.4). Interestingly, expression of many known and predicted effector genes that are enriched by H3K27me3 during mycelial growth is also highly upregulated during plant infection (Table 2.1, 2.6, and 2.8). Particularly, the transcript level of *BAS4* during plant infection of the wild-type strain at 25 hpi is massively induced and is significantly higher than that even during mycelial growth of Δ*mokmt6*-MoGti1oe strain. These results indicate that the highly induced expression of H3K27me3-enriched effector genes during plant infection may be correlated with H3K27me3 removal at corresponding effector gene loci. Indeed, ChIP-qPCR *in vitro* and *in planta* analyses of a putative effector gene in *Z. tritici* reveals that reduced H3K27me3 levels at this putative effector gene locus is correlated with its upregulated expression *in planta* (15). Additionally, *in planta* exclusively induced SM genes *lolitrems* (*ltm*) and *ergot alkaloids* (*eas*) are also correlated with reduced H3K27me3 levels at each locus when compared to axenic culture in *Epichloe festucae* (36). These results suggest that to allow efficient transcription, fungal H3K27me3 must be removed from individual H3K27me3-enriched gene loci during plant infection. A similar observation has been reported in mouse AtT-20 cells, for instance, high H3K9me2 levels are strongly depleted but H3K4me1 and H3K27ac levels are increased after the pioneer factor Pax7 action at pioneered sites, which facilitates the binding of other TFs and coactivators to promote transcription (e.g. p300) (32). Thus, it is

possible that the genome-wide or local removal of H3K27me3 during plant infection results in open chromatin and gain of active histone modifications (e.g. H3K4me2 and H3K27ac) at effector gene loci to accelerate transcription. If this is the case, one or more histone demethylases may be critical for plant infection. Alternatively, it is also possible that H3K27me3 persists *in planta*, but H3K27me3-mediated repression is lost due to altered activity of a yet-unknown H3K27me3-binding protein. Future studies will distinguish between these and other possible mechanisms for reactivating H3K27me3-repressed *M. oryzae* effector genes during plant infection.

Effector genes are regulated by multiple mechanisms

Our study indicates that expression of effector genes is also under control of other regulations besides H3K27me3 and MoGti1 in *M. oryzae*. We observed unaltered expression of *MC69* either by removal of H3K27me3 or by overexpression of MoGti1 (Fig 2.4C). One possibility is regulation by other histone modifications in addition to H3K27me3. Our ChIP-seq result reveals that 40% (178/446) of effector genes have relatively high enrichment of H3K27me3 during mycelial growth of *M. oryzae* (Fig 2.2). That only a subset of H3K27me3-enriched effector genes was derepressed indicates that activating signals may be required to upregulate expression of other H3K27me3-enriched effector genes in addition to H3K27me3 loss. However, there are still 60% (268/446) of effector genes with low or no H3K27me3 enrichment at the same growth stage and condition, such as *MC69*, of which expression is not controlled by H3K27me3 (Fig 2.2 and Table 2.3). These results indicate that H3K27me3 plays a role, but is not the only factor in repressing effector gene expression during mycelial growth of *M. oryzae*. As discussed in *L. maculans* and *Z. tritici*, expression of effector or putative effector genes is

upregulated due to the absence of H3K9me3 in axenic culture (13, 15). Interestingly, we also found an enrichment of H3K9me3 in a small subset of effector genes, although future investigation is needed to confirm the role of H3K9me3 on effector gene expression in *M. oryzae*.

Another possibility is regulation by various transcription factors. Indeed, MoGti1 has been previously shown to control expression of a subset of, but not all, effector genes in *M. oryzae* (26). Our MoGti1 overexpression study confirms this observation and includes new observations that provide a more complete understanding of the complexities of MoGti1 regulation of effector gene expression (Fig 2.5 and Table 2.3). Similarly, in *Ustilago maydis*, MoGti1 ortholog Ros1 has been reported as a transcriptional regulator for downregulation of effector genes that are essential during early infection but upregulation of effector genes during late infection (8, 27). In addition to Ros1, multiple transcriptional regulators have been identified in *U. maydis*, like Rbf1 and Fox1, to control expression of effector genes at different infection stages (8). Interestingly, we have also noticed that *MoGti1* expression during plant infection coincides with *BAS4* and *AVR-Pik* upregulation and *AVR-Pi9* downregulation around the appressorium-mediated penetration stage (Fig 2.5B, Zhu et al., unpublished and (37)). This temporal association between expression of *MoGti1* and various effector genes suggests that MoGti1 is only one of the transcriptional regulators that controls expression of effector genes at a specific stage. Although no other TFs have been demonstrated to regulate effector gene expression in *M. oryzae*, so far 13 TFs from different families in various fungi have been found to control expression of effector or candidate effector

genes (38). Thus, it is possible that, in addition to MoGti1, multiple other TFs are also involved in effector gene regulation during plant infection of *M. oryzae*.

MoGti1 overexpression is not able to reprogram H3K27me3 patterns but might override repressive effects of H3K27me3 on effector gene expression

We did not find any evidence that ectopic overexpression of the transcription factor MoGti1 changed H3K27me3 patterns during mycelial growth in *M. oryzae* (Fig 2.6C, 2.6D), although ectopic expression of transcription factors mediating epigenetic reprogramming has been reported in animal studies. For example, in human cell lines having the physiologic phosphatidylinositol 3-kinase (PI3K) pathway, overexpression of the master transcription factor Nrf2 that plays roles in cellular detoxification decreases levels of H3K27me3 but increases H3K4me3 in promoters of the Nrf2's targets *mTOR* and *NQO1* (39). Additionally, the ectopic expression of four transcription factors — OCT4, SOX2, KLF4 and MYC (OSKM) in mouse embryonic fibroblasts leads to the genome-wide change of H3K4me2, but H3K27me3 remains largely unchanged (40). Alternatively, ectopic overexpression of MoGti1 could result in epigenetic reprogramming (e.g. H3K27ac) that we did not investigate in this study.

Despite the unaltered H3K27me3 pattern both globally and locally when MoGti1 is ectopically overexpressed during mycelial growth of *M. oryzae*, expression of many H3K27me3-enriched effector genes is still significantly induced (Fig 2.5C). It appears that repression of effector genes by H3K27me3 can be overridden by overexpression of MoGti1 under our experimental condition. How is a transcription factor able to access the condensed chromatin to activate gene expression? We propose two possible explanations. First, overexpression of the transcription factor MoGti1 might recruit chromatin

modifiers to increase the chromatin accessibility leading to the transcription of effector genes. For instance, chromatin modifiers BRG1/BRM can be recruited to promoters of pro-inflammatory genes by a transcription co-factor MRTF-A to steer the transcription of downstream genes in animal cells (41). Second, Petruck *et al.* (2017) found the delayed H3K27me3 accumulation on nascent DNA after DNA replication following induction of embryonic stem cell differentiation, thus providing a "window of opportunity" for recruitment of induced transcription factors to their binding sites (42). Therefore, it is possible that MoGti1 that is constitutively overexpressed during mycelial growth gains access to its binding sites on nascent DNA immediately after DNA replication to lead to expression of downstream effector genes.

Taken together, our work has uncovered a complex interplay between chromatin modifications and sequence-specific transcription factors in orchestrating stage-specific expression of pathogenicity genes in *M. oryzae.*

Materials and Methods

Strains and growth conditions

M. oryzae field isolate O-137 was used as the wild-type strain and the recipient of fungal transformations. *M. oryzae* strains are the fungal transformants. The fungi were maintained in frozen storage (-20°C) and cultured on oatmeal agar (OMA) plates at 25°C under continuous light (43).

Vector construction and fungal transformation

To obtain a *MoKMT6* knock-out mutant of *M. oryzae*, a homologous gene replacement strategy was applied as previously described (44). Briefly, the 5'- (1.4 kb) and 3'-(1.3kb) flanking regions of *MoKMT6* were amplified by PCR from genomic DNA

of *M. oryzae* isolate O-137. The *Neomycin phosphotransferase–II* (*NTPII*) gene was cloned from pBV141 (45). Each primer was designed with a restriction enzyme site at 5' end. PCR was performed using Phusion High-Fidelity PCR Master Mix with HF Buffer (Thermo Scientific™). The PCR cycling program consisted of an initial denaturation for 2 min at 98°C, two cycles of 30 s denaturation at 98°C, 30 s annealing at 56°C, 1 min extension at 72°C and 25 cycles of 30 s denaturation at 98°C, 30 s annealing at 62°C, 1 min extension at 72°C, followed by a final extension for 10 min at 72°C. PCR products were isolated from gels using E.Z.N.A.® Gel Extraction Kit (Omega Bio-tek). The three fragments were first cloned in pJET1.2 using CloneJET PCR Cloning Kit (Thermo Fisher Scientific) for sequence analysis and later into binary vector pBV108 (pGKO2)(44). *NTPII* gene was constructed between the two flanking regions using a restriction ligation strategy. Fungal transformation was performed using *Agrobacterium tumefaciens* mediated transformation (ATMT) according to previous description (44). After two rounds of selections on TB3 (0.3% yeast extract, 0.3% casamino acid, 20% sucrose) and V8 (8% V8 vegetable juice (Campbell's), pH6.97) media containing 800 µg/ml of G418 (Fisher BioReagents) and 200 µM of Cerfotaxime (Gold Biotechnology), 72 independent fungal transformants were screened by negative selection on V8 media containing 800 µg/ml of G418, 200 µM of Cerfotaxime and 5 µM of (+)-5-Fluoro-2"-deoxyuridine (F2dU, Acros organics). Eleven fungal transformants after negative selection were analyzed for gene replacement events by two different PCR amplification strategies that were used previously (46-48). Then the selected knock-out transformant was further confirmed by qRT-PCR assay.

Genetic complementation of *MoKMT6* deletion mutant was performed by introducing a wild-type allele of *MoKMT6* ectopically into *Δmokmt6* genome, which resulted in a complemented strain carrying the wild-type allele of *MoKMT6* at a random locus. Genomic sequences containing *MoKMT6* coding sequence with its 5'- and 3' flanking regions were amplified from genomic DNA of *M. oryzae* isolate O-137 with primers shown in Table S12 and first cloned into pJET1.2 and later into binary vector pCK1806. pCK1806 was generated by replacing *Xho*I-*EcoR*I fragment of pBV141 with *Nourseothricin acetyltransferase* gene (*Nat1*) amplified from pDONR207 (generously shared by Ane Sesma at Universidad Politécnica de Madrid, Madrid, Spain). Twenty independent transformants were selected on V8 media containing 400 μg/ml nourseothricin (Gold Biotechnology) and analyzed by PCR amplification. Then, two transformants were further confirmed by qRT-PCR analysis of *BAS4* expression and ChIP-seq analysis of H3K27me3.

To construct MoGti1 overexpression cassette, *MoGti1* coding sequence with 300-bp of the 3'-flanking region after stop codon was amplified from genomic DNA of *M. oryzae* O-137 with primers in Table S12. 1-kb of the strong constitutive *M. oryzae* ribosomal protein (P27) promoter was isolated from *EcoR*I-*BamH*I fragment of pBV126 (49). Two fragments were first cloned into pJET1.2 and later into binary vector pCK1806. Twelve independent transformants were selected and behaved similarly to wild-type strain under microscopy examination. Then four transformants with wild-type and recipient strains were inoculated into liquid complete medium (CM, 10 mg/ml of sucrose, 6 mg/ml of casamino acids, and 6 mg/ml of yeast extract) for 5 days to isolate RNA for determining *MoGti1* expression level.

To obtain dual transformants of *Δmokmt6*-MoGti1oe, we introduced *MoGti1* overexpression construct into a *MoKMT6* deletion mutant. Twelve independent transformants were selected on V8 media containing 400 μg/ml nourseothricin. Then four transformants with wild-type and recipient strains were inoculated into liquid CM for 5 days to isolate RNA for determining *MoGti1* expression level.

To monitor *BAS4* expression at signal cell resolution, 1-kb 5'- and 0.5-kb 3'-flanking regions of *BAS4* coding sequence were amplified, respectively, from genomic DNA of *M. oryzae* isolate O-137. EGFP was isolated from *BamH*I-*BsrG*I fragment of pGXT (50), and the protein degradation signal peptide PEST was isolated from *BsrG*I-*Not*I fragment of pBV118 (pd2EGFP-1)(51). EGFP:PEST was constructed between the two flanking regions using a restriction ligation strategy. All four fragments were fused into binary vector pBV1(pBHt2) (52). After fungal transformation, ten independent transformants were selected on V8 media containing 200 μg/ml hygromycin (Fisher) and purified by single spore isolation.

Genomic DNA isolation

Fungal conidia were harvested from ~10-day-old cultures on OMA plates. $1x10^5$ spores/ml in distilled water were inoculated into liquid complete medium and shaken at 25°C, 100 rpm for 5 days under dark environment. Then fungal mycelia were collected and washed by filtration to remove extra water, and frozen immediately in liquid nitrogen and stored at −80°C for subsequent DNA extraction. The CTAB (cetyltrimethyl ammonium bromide) DNA extraction method was used to isolate genomic DNA from the mycelia samples (53).

Preparation of infected rice sheath and mycelia samples for gene expression analysis

To examine expression of effector genes and the transcription factor *MoGti1* during plant infection, a time course qRT-PCR assay was performed. Briefly, rice sheath inoculations were performed by inoculating fungal spores at concentration of 1×10^5 spores/ml in distilled water as described (54). 8cm-long sheath pieces from ~20-day-old plants were used. Fifteen infected rice sheath samples at each time point of 18 hours post inoculation (hpi), 24hpi, 33hpi and 40hpi were collected as described (28), and frozen immediately in liquid nitrogen and stored at −80°C for subsequent RNA extraction. Preparation of mycelia samples for RNA extraction was performed exactly as mycelia samples for genomic DNA extraction and ChIP-seq assay.

ChIP-seq, ChIP-seq library construction and data analysis

For preparation of mycelia for ChIP-seq samples, fungal mycelia after 5 days growth in CM were washed once by 1% phosphate-buffered saline (PBS) and transferred into 50 ml flasks containing 10 ml of PBS with 1% formaldehyde to perform chemical cross-linking at room temperature on a rotating platform for 30 min. The reaction was quenched with 125 mM glycine. ChIP methods were performed as described previously (55) and 1 μL of the relevant antibody was used. For Illumina sequencing, ChIP-seq libraries were prepared using ~10 ng of immunoprecipitated DNA and were constructed by end repair and A-tailing using the NEBNext Ultra II End Repair Module (cat. # E7546S). Illumina adaptors were ligated to repaired DNA molecules using the NEBNext Ultra II Ligation Module (cat. # E7595S). Ligation products were amplified to generate dual-indexed libraries using NEBNext Ultra II Q5 Hot Start HiFi PCR Master Mix (cat. #

M0543S). Libraries were pooled and sequenced on a NextSeq500 instrument at the Georgia Genomics and Bioinformatics Core to generate single or paired-end 75-bp reads.

For ChIP-seq data analysis, short reads (<20-bp) and adaptor sequences were removed using TrimGalore (version 0.4.4) (56), cutadapt version 1.1 (57), and Python 2.7.8, with fastqc command (version 0.11.3). Trimmed Illumina reads were aligned to the current *Magnaporthe oryzae* 70-15 MG8 genome assembly (accession # GCA_000002495.2) using BWA (version 0.7.15) (58), mem algorithm, which randomly assign multi-mapped reads to a single location. Files were sorted and indexed using SAMtools (version 1.9) (58). To plot the relative distribution of mapped reads, read counts were determined for each 25-bp window across the genome using igvtools and data were displayed using the Integrated Genome Viewer (59). The Hypergeometric Optimization of Motif EnRichment (HOMER) software package (version 4.8) (60) was used to identify H3K27me3 peaks in wild type against input using "findPeaks.pl" with the following parameters: -style histone. Bedtools (version 2.27.1) "intersect" (version 2.26.0) (61) was used to determine the fraction/number of peaks that intersect with annotated genes. We defined genes enriched with H3K27me3 as genes that have 70% or more of an overlap with H3K27me3 regions using bedtools intersect (61) HOMER was also used to construct heatmaps using "annotatePeaks.pl" with the following options: -hist 10 -size 2000/6000 -ghist for heatmaps centered on TSS and centered on K27me3 peaks, respectively. Heatmaps for effector genes were constructed with R using the pheatmap package using a k-means of two. The 95th percentile value was set as the maximum value and pheatmap was used to generate heatmaps using HOMER -ghist matrix files as input. Karyotype plots of all assembled chromosomes with H3K27me3

enrichment and effector gene positions were generated using karyoploteR package (62) in R using the wild type H3K27me3 bam file. A custom cytoband file was created with the genomic coordinates of known effector genes from literature.

RNA isolation, quantitative RT-PCR, RNA-seq library construction and RNA-seq analysis

Total RNAs from mycelia and infected rice sheaths were extracted using a Trizol method (Invitrogen). Genomic DNA was removed by treatment with Turbo™ DNase (Ambion, Cat# AM1907) according to manufacturer's instructions. 2 µg of total RNA extracted from infected rice tissue or mycelia grown for 5 days in CM was used to synthesize cDNA with ImProm II Reverse Transcriptase system (Promega). qRT-PCR was performed with the MX3005P (Stratagene) and CFX96™ (Bio-Rad) systems using the PowerUp™ SYBR™ Green Master Mix (Thermo Fisher/applied biosystem). Thermocycler conditions were as follows: 2 min at 50°C, 10 min at 95°C, followed by 40 cycles of 95°C for 30 sec, 60°C for 30 sec, and 72°C for 30 sec. Primers used in qRT-PCR are listed in Table 2.8. A final dissociation cycle was incorporated to ensure the specificity of each primer pair. Each qRT-PCR mixture (final volume 14 µl) contained 7 µl of PowerUp™ SYBR™ Green Master Mix, 1.5 µl of forward and reverse primers (3.3 nM concentrations for each), 2 µl of cDNA template and 2 µl of distilled water. Primers used for qRT-PCR assays are listed in Table S12. The relative expression level of each gene was calculated by the $2^{-\Delta\Delta CT}$ method (63), with the *M. oryzae actin* gene (MGG_03982) as a control housekeeping gene. Briefly, the average threshold cycle (Ct) was normalized to that of *actin* gene for each of the treated samples as $2^{-\Delta Ct}$, where $\Delta Ct = (Ct_{\text{effector gene}} - Ct_{\text{actin}})$. The fold changes between the wild-type and mutants during

mycelia growth in liquid CM were calculated as $2^{-\Delta\Delta Ct}$, where $\Delta\Delta Ct = (Ct_{\text{effector gene}} - Ct_{\text{actin}})_{\text{mutant}} - (Ct_{\text{effector gene}} - Ct_{\text{actin}})_{\text{wild-type}}$. Two technical replicates for each of two or three biological replicates were performed. Mean and standard deviation were calculated from qRT-PCR results of three biological replicates.

RNA-seq libraries were prepared from 1µg of total RNA and constructed with the Illumina Truseq Stranded mRNA Library Prep kit (cat #20020594) according to manufacturer's directions. Libraries were pooled and sequenced on a NextSeq500 instrument at the Georgia Genomics and Bioinformatics Core to generate paired-end 75-bp reads.

For RNA-seq data analysis, Illumina paired-end reads were mapped to the current *Magnaporthe oryzae* 70-15 MG8 genome assembly (accession # GCA_000002495.2) using the Hierarchical Indexing for Spliced Alignment of Transcripts 2 (HISAT2: version 2.1.0) (64) with parameters –RNA-strandness RF then sorted and indexed using SAMtools (version 1.9) (65). FeatureCounts from Subread (version 1.6.2) (66) was used to generate gene level counts for all RNA bam files. Raw counts were imported into R and differential gene expression analysis was conducted using Bioconductor: DeSeq2 (67). Box plots generated in R using DeSeq2 and ggplot2 (68), respectively. Heatmaps were constructed using the pheatmap package with scale "row" for all sample heatmap, pairwise comparisons were generated the same with the addition of hierarchical clustering of rows for H3K27me3 enriched and unenriched groups separately.

Microarray analysis and comparison

Microarray data of *M. oryzae* O-137 infected rice sheath at 36 hpi were downloaded from NCBI Gene Expression Omnibus (GEO), accession number GSE8517

(www.ncbi.nlm.nih.gov/geo) (28). Probe IDs and corresponding sequences from previously published microarray dataset (GSE8517) were compiled into a FASTA file and their corresponding genes were identified using BLAST against a local database of the current *Magnaporthe oryzae* 70-15 MG8 genome assembly (accession# GCA_000002495.2). Only genes that had two-fold upregulated in comparison to wild type were used for further analysis and comparisons.

Effector prediction and data set

The *M. oryzae* secretome data was previously generated by Zhang et al. (2018), which included 1426 secretory proteins. This secretome was used to predict *M. oryzae* effector genes by EffectorP 2.0 (http://effectorp.csiro.au/)(69), a machine learning method trained with characterized fungal effectors to predict effector proteins from secretomes. A total of 449 effector genes were identified from EffectorP2.0. Then BLAST (e-value cut-off = 1e-5) search of 449 effector genes identified that 434 effectors had homologous sequences in the genome of *M. oryzae* isolate O-137. Among them, 36 known *M. oryzae* effector genes that were reported in literature were included, but another 14 known effector genes were not. Thus, we combined 434 predicted effector genes and 14 known effector genes as our dataset (a total of 448 known and predicted effector genes) for further analysis.

Pathogenicity assay

Rice (*Oryza sativa)* cultivar YT16 was planted as described (70). Long day conditions (14/10 h, day/night) in a growth chamber with daytime temperature of 28°C and nighttime temperature of 24 °C were applied for plant growth. Approximately 20-day-old rice was used for whole-plant infection with 3×10^4 spores/ml in distilled gelatin

solution (0.25%) to assess mutant phenotypes. Seven days after inoculation, symptoms on the inoculated plants were recorded (43) and evaluated (71) as described. Briefly, the youngest leaf that was expanded when being inoculated was examined and documented with EPSON perfection 4870 Photo with 24-bit color, 600 dpi resolution and same document size (8.5 inch of width and 11.7 inch of height). Then, ImageJ (72) was used to process and measure diseased leaf area. Pixels in images were converted to centimeters with the Set Scale command, and a single leaf was analyzed each time. To measure the whole and diseased leaf area, we used Color Threshold with HSB color space. Appropriate thresholds to whole leaf area and specifically diseased leaf area were selected and measured, respectively.

Confocal Microscopy

Confocal microscopy was performed on a Zeiss LSM 880 Confocal Microscope with an upright microscope stand. Excitation/emission wavelengths were 488 nm/496 to 544 nm for EGFP and 543 nm/565 to 617 nm for tdTomato. Images were processed using Zen Black software (version 10.0, Zeiss).

References

1. 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 66:513-45.10.1146/annurev-arplant-043014-114623

2. Sanchez-Vallet A, Fouche S, Fudal I, Hartmann FE, Soyer JL, Tellier A, Croll D. 2018. The Genome Biology of Effector Gene Evolution in Filamentous Plant Pathogens. Annual Review of Phytopathology doi:10.1146/annurev-phyto-080516-035303.10.1146/annurev-phyto-080516-035303

3. Kleemann J, Rincon-Rivera LJ, Takahara H, Neumann U, Ver Loren van Themaat E, van der Does HC, Hacquard S, Stuber K, Will I, Schmalenbach W, Schmelzer E, O'Connell RJ. 2012. Sequential delivery of host-induced virulence effectors by appressoria and intracellular hyphae of the phytopathogen

Colletotrichum higginsianum. PLoS Pathogen 8:e1002643.10.1371/journal.ppat.1002643

4. O'Connell RJ, Thon MR, Hacquard S, Amyotte SG, Kleemann J, Torres MF, Damm U, Buiate EA, Epstein L, Alkan N, Altmuller J, Alvarado-Balderrama L, Bauser CA, Becker C, Birren BW, Chen Z, Choi J, Crouch JA, Duvick JP, Farman MA, Gan P, Heiman D, Henrissat B, Howard RJ, Kabbage M, Koch C, Kracher B, Kubo Y, Law AD, Lebrun MH, Lee YH, Miyara I, Moore N, Neumann U, Nordstrom K, Panaccione DG, Panstruga R, Place M, Proctor RH, Prusky D, Rech G, Reinhardt R, Rollins JA, Rounsley S, Schardl CL, Schwartz DC, Shenoy N, Shirasu K, Sikhakolli UR, Stuber K, et al. 2012. Lifestyle transitions in plant pathogenic *Colletotrichum* fungi deciphered by genome and transcriptome analyses. Nature Genetics 44:1060-5.10.1038/ng.2372

5. Hacquard S, Kracher B, Maekawa T, Vernaldi S, Schulze-Lefert P, Ver Loren van Themaat E. 2013. Mosaic genome structure of the barley powdery mildew pathogen and conservation of transcriptional programs in divergent hosts. Proceedings of the National Academy of Sciences 110:E2219-28.10.1073/pnas.1306807110

6. Dong Y, Li Y, Zhao M, Jing M, Liu X, Liu M, Guo X, Zhang X, Chen Y, Liu Y, Liu Y, Ye W, Zhang H, Wang Y, Zheng X, Wang P, Zhang Z. 2015. Global genome and transcriptome analyses of *Magnaporthe oryzae* epidemic isolate 98-06 uncover novel effectors and pathogenicity-related genes, revealing gene gain and lose dynamics in genome evolution. PLoS Pathogen 11:e1004801.10.1371/journal.ppat.1004801

7. Gervais J, Plissonneau C, Linglin J, Meyer M, Labadie K, Cruaud C, Fudal I, Rouxel T, Balesdent MH. 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 18:1113-1126.10.1111/mpp.12464

8. Lanver D, Muller AN, Happel P, Schweizer G, Haas FB, Franitza M, Pellegrin C, Reissmann S, Altmuller J, Rensing SA, Kahmann R. 2018. The Biotrophic Development of Ustilago maydis Studied by RNA-Seq Analysis. Plant Cell 30:300-323.10.1105/tpc.17.00764

9. Berlowitz L. 1965. Correlation of Genetic Activity, Heterochromatization, and Rna Metabolism. Proc Natl Acad Sci U S A 53:68-73

10. Strahl BD, Allis CD. 2000. The language of covalent histone modifications. Nature 403:41-45

11. Saksouk N, Simboeck E, Dejardin J. 2015. Constitutive heterochromatin formation and transcription in mammals. Epigenetics Chromatin 8:3.10.1186/1756-8935-8-3

12. Wiles ET, Selker EU. 2016. H3K27 methylation: a promiscuous repressive chromatin mark. Curr Opin Genet Dev 43:31-37.10.1016/j.gde.2016.11.001

13. Soyer JL, El Ghalid M, Glaser N, Ollivier B, Linglin J, Grandaubert J, Balesdent MH, Connolly LR, Freitag M, Rouxel T, Fudal I. 2014. Epigenetic control of effector gene expression in the plant pathogenic fungus *Leptosphaeria maculans*. PLoS Genetics 10:e1004227.10.1371/journal.pgen.1004227

14. Connolly LR, Smith KM, Freitag M. 2013. The *Fusarium graminearum* histone H3K27 methyltransferase KMT6 regulates development and expression of secondary metabolite gene clusters. PLoS genetics 9:e1003916

15. Soyer JL, Grandaubert J, Haueisen J, Schotanus K, Holtgrewe Stukenbrock E. 2019. In planta chromatin immunoprecipitation in Zymoseptoria tritici reveals chromatin-based regulation of putative effector gene expression. bioRxiv doi:10.1101/544627.10.1101/544627

16. Lewis ZA. 2017. Polycomb group systems in fungi: new models for understanding polycomb repressive complex 2. Trends in Genetics 33:220-231.10.1016/j.tig.2017.01.006

17. Farman ML. 2007. Telomeres in the rice blast fungus *Magnaporthe oryzae*: the world of the end as we know it. FEMS Microbiology Letters 273:125-32.10.1111/j.1574-6968.2007.00812.x

18. Haas BJ, Kamoun S, Zody MC, Jiang RH, Handsaker RE, Cano LM, Grabherr M, Kodira CD, Raffaele S, Torto-Alalibo T, Bozkurt TO, Ah-Fong AM, Alvarado L, Anderson VL, Armstrong MR, Avrova A, Baxter L, Beynon J, Boevink PC, Bollmann SR, Bos JI, Bulone V, Cai G, Cakir C, Carrington JC, Chawner M, Conti L, Costanzo S, Ewan R, Fahlgren N, Fischbach MA, Fugelstad J, Gilroy EM, Gnerre S, Green PJ, Grenville-Briggs LJ, Griffith J, Grunwald NJ, Horn K, Horner NR, Hu CH, Huitema E, Jeong DH, Jones AM, Jones JD, Jones RW, Karlsson EK, Kunjeti SG, Lamour K, Liu Z, et al. 2009. Genome sequence and analysis of the Irish potato famine pathogen *Phytophthora infestans*. Nature 461:393-8.10.1038/nature08358

19. Peng Z, Oliveira-Garcia E, Lin G, Hu Y, Dalby M, Migeon P, Tang H, Farman M, Cook D, White FF. 2019. Effector gene reshuffling involves dispensable mini-chromosomes in the wheat blast fungus. PLoS Genetics 15:e1008272

20. Huang G, Wang H, Chou S, Nie X, Chen J, Liu H. 2006. Bistable expression of WOR1, a master regulator of white–opaque switching in *Candida albicans*. Proceedings of the National Academy of Sciences 103:12813-12818

21. Michielse CB, van Wijk R, Reijnen L, Manders EM, Boas S, Olivain C, Alabouvette C, Rep M. 2009. The nuclear protein Sge1 of *Fusarium oxysporum* is required for parasitic growth. PLoS Pathogen 5:e1000637.10.1371/journal.ppat.1000637

22. Santhanam P, Thomma BP. 2013. *Verticillium dahliae* Sge1 differentially regulates expression of candidate effector genes. Molecular Plant-Microbe Interactions 26:249-256

23. Brown DW, Busman M, Proctor RH. 2014. *Fusarium verticillioides* SGE1 is required for full virulence and regulates expression of protein effector and secondary metabolite biosynthetic genes. Molecular Plant-Microbe Interactions 27:809-23.10.1094/MPMI-09-13-0281-R

24. Mirzadi Gohari A, Mehrabi R, Robert O, Ince IA, Boeren S, Schuster M, Steinberg G, de Wit PJ, Kema GH. 2014. Molecular characterization and functional analyses of ZtWor1, a transcriptional regulator of the fungal wheat pathogen *Zymoseptoria tritici*. Molecular Plant Pathology 15:394-405.10.1111/mpp.12102

25. Okmen B, Collemare J, Griffiths S, van der Burgt A, Cox R, de Wit PJ. 2014. Functional analysis of the conserved transcriptional regulator CfWor1 in *Cladosporium fulvum* reveals diverse roles in the virulence of plant pathogenic fungi. Molecular Microbiology 92:10-27.10.1111/mmi.12535

26. Li Y, Wang G, Xu JR, Jiang C. 2016. Penetration peg formation and invasive hyphae development require stage-specific activation of *MoGTI1* in *Magnaporthe oryzae*. Molecular Plant-Microbe Interactions 29:36-45.10.1094/MPMI-06-15-0142-R

27. Tollot M, Assmann D, Becker C, Altmuller J, Dutheil JY, Wegner CE, 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 Pathogen 12:e1005697.10.1371/journal.ppat.1005697

28. Mosquera G, Giraldo MC, Khang CH, Coughlan S, Valent B. 2009. Interaction transcriptome analysis identifies *Magnaporthe oryzae* BAS1-4 as Biotrophy-associated secreted proteins in rice blast disease. Plant Cell 21:1273-90.10.1105/tpc.107.055228

29. Saitoh H, Fujisawa S, Mitsuoka C, Ito A, Hirabuchi A, Ikeda K, Irieda H, Yoshino K, Yoshida K, Matsumura H, Tosa Y, Win J, Kamoun S, Takano Y,

Terauchi R. 2012. Large-scale gene disruption in *Magnaporthe oryzae* identifies MC69, a secreted protein required for infection by monocot and dicot fungal pathogens. PLoS Pathogen 8:e1002711.10.1371/journal.ppat.1002711

30. Pham KT, Inoue Y, Vu BV, Nguyen HH, Nakayashiki T, Ikeda K, Nakayashiki H. 2015. MoSET1 (histone H3K4 methyltransferase in *Magnaporthe oryzae*) regulates global gene expression during infection-related morphogenesis. PLoS Genetics 11:e1005385.10.1371/journal.pgen.1005385

31. Fischer J, Muller SY, Netzker T, Jager N, Gacek-Matthews A, Scherlach K, Stroe MC, Garcia-Altares M, Pezzini F, Schoeler H, Reichelt M, Gershenzon J, Krespach MK, Shelest E, Schroeckh V, Valiante V, Heinzel T, Hertweck C, Strauss J, Brakhage AA. 2018. Chromatin mapping identifies BasR, a key regulator of bacteria-triggered production of fungal secondary metabolites. Elife 7.10.7554/eLife.40969

32. Mayran A, Khetchoumian K, Hariri F, Pastinen T, Gauthier Y, Balsalobre A, Drouin J. 2018. Pioneer factor Pax7 deploys a stable enhancer repertoire for specification of cell fate. Nature Genetics 50:259-269.10.1038/s41588-017-0035-2

33. Lohse MB, Zordan RE, Cain CW, Johnson AD. 2010. Distinct class of DNA-binding domains is exemplified by a master regulator of phenotypic switching in Candida albicans. Proceedings of the National Academy of Sciences 107:14105-14110

34. Jamieson K, Rountree MR, Lewis ZA, Stajich JE, Selker EU. 2013. Regional control of histone H3 lysine 27 methylation in *Neurospora*. Proceedings of the National Academy of Sciences 110:6027-32.10.1073/pnas.1303750110

35. Basenko EY, Sasaki T, Ji LX, Prybol CJ, Burckhardt RM, Schmitz RJ, Lewis ZA. 2015. Genome-wide redistribution of H3K27me3 is linked to genotoxic stress and defective growth. Proceedings of the National Academy of Sciences of the United States of America 112:E6339-E6348.10.1073/pnas.1511377112

36. Chujo T, Scott B. 2014. Histone H3K9 and H3K27 methylation regulates fungal alkaloid biosynthesis in a fungal endophyte-plant symbiosis. Molecular Microbiology 92:413-34.10.1111/mmi.12567

37. Wu J, Kou Y, Bao J, Li Y, Tang M, Zhu X, Ponaya A, Xiao G, Li J, Li C, Song MY, Cumagun CJ, Deng Q, Lu G, Jeon JS, Naqvi NI, Zhou B. 2015. Comparative genomics identifies the *Magnaporthe oryzae* avirulence effector AvrPi9 that triggers Pi9-mediated blast resistance in rice. New Phytologist 206:1463-75.10.1111/nph.13310

38. Tan KC, Oliver RP. 2017. Regulation of proteinaceous effector expression in phytopathogenic fungi. PLoS Pathogen 13:e1006241.10.1371/journal.ppat.1006241

39. Bendavit G, Aboulkassim T, Hilmi K, Shah S, Batist G. 2016. Nrf2 transcription factor can directly regulate mTOR: linking cytoprotective gene expression to a major metabolic regulation that generates redox activity. Journal of Biological Chemistry 291:25476-25488.10.1074/jbc.M116.760249

40. Koche RP, Smith ZD, Adli M, Gu H, Ku M, Gnirke A, Bernstein BE, Meissner A. 2011. Reprogramming factor expression initiates widespread targeted chromatin remodeling. Cell Stem Cell 8:96-105.10.1016/j.stem.2010.12.001

41. Yang Y, Cheng X, Tian W, Zhou B, Wu X, Xu H, Fang F, Fang M, Xu Y. 2014. MRTF-A steers an epigenetic complex to activate endothelin-induced pro-inflammatory transcription in vascular smooth muscle cells. Nucleic Acids Research 42:10460-72.10.1093/nar/gku776

42. Petruk S, Cai J, Sussman R, Sun G, Kovermann SK, Mariani SA, Calabretta B, McMahon SB, Brock HW, Iacovitti L. 2017. Delayed accumulation of H3K27me3 on nascent DNA is essential for recruitment of transcription factors at early stages of stem cell differentiation. Molecular Cell 66:247-257. e5

43. Valent B, Farrall L, Chumley FG. 1991. *Magnaporthe grisea* genes for pathogenicity and virulence identified through a series of backcrosses. Genetics 127:87-101

44. Khang CH, Park SY, Lee YH, Kang S. 2005. A dual selection based, targeted gene replacement tool for *Magnaporthe grisea* and *Fusarium oxysporum*. Fungal Genetics and Biology 42:483-92.10.1016/j.fgb.2005.03.004

45. Kim H-S, Park S-Y, Lee S, Adams EL, Czymmek K, Kang S. 2011. Loss of cAMP-dependent protein kinase A affects multiple traits important for root pathogenesis by *Fusarium oxysporum*. Molecular Plant-Microbe Interactions 24:719-732

46. van der Does HC, Duyvesteijn RG, Goltstein PM, van Schie CC, Manders EM, Cornelissen BJ, Rep M. 2008. Expression of effector gene *SIX1* of *Fusarium oxysporum* requires living plant cells. Fungal Genetics and Biology 45:1257-64.10.1016/j.fgb.2008.06.002

47. Wilson RA, Gibson RP, Quispe CF, Littlechild JA, Talbot NJ. 2010. An NADPH-dependent genetic switch regulates plant infection by the rice blast fungus. Proceedings of the National Academy of Sciences 107:21902-21907

48. Fernandez J, Wright JD, Hartline D, Quispe CF, Madayiputhiya N, Wilson RA. 2012. Principles of carbon catabolite repression in the rice blast fungus: Tps1, Nmr1-3, and a MATE-family pump regulate glucose metabolism during infection. PLoS genetics 8:e1002673.10.1371/journal.pgen.1002673

49. Khang CH, Berruyer R, Giraldo MC, Kankanala P, Park SY, Czymmek K, Kang S, Valent B. 2010. Translocation of *Magnaporthe oryzae* effectors into rice cells and their subsequent cell-to-cell movement. Plant Cell 22:1388-403.10.1105/tpc.109.069666

50. Chen S, Songkumarn P, Liu J, Wang GL. 2009. A versatile zero background T-vector system for gene cloning and functional genomics. Plant Physiology 150:1111-21.10.1104/pp.109.137125

51. Li X, Zhao X, Fang Y, Jiang X, Duong T, Fan C, Huang C-C, Kain SR. 1998. Generation of destabilized green fluorescent protein as a transcription reporter. Journal of Biological Chemistry 273:34970-34975

52. Mullins ED, Chen X, Romaine P, Raina R, Geiser DM, Kang S. 2001. *Agrobacterium*-mediated transformation of *Fusarium oxysporum*: an efficient tool for insertional mutagenesis and gene transfer. Phytopathology 91:173-180

53. Clarke JD. 2009. Cetyltrimethyl ammonium bromide (CTAB) DNA miniprep for plant DNA isolation. Cold Spring Harbor Protocol 2009:pdb prot5177.10.1101/pdb.prot5177

54. Kankanala P, Czymmek K, Valent B. 2007. Roles for rice membrane dynamics and plasmodesmata during biotrophic invasion by the blast fungus. Plant Cell 19:706-24.10.1105/tpc.106.046300

55. Ferraro AR, Lewis ZA. 2018. ChIP-Seq Analysis in *Neurospora crassa*, p 241-250, Fungal Genomics. Springer.

56. Krueger F. 2015. Trim galore. A wrapper tool around Cutadapt and FastQC to consistently apply quality and adapter trimming to FastQ files 516:517

57. Martin M. 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet journal 17:10-12

58. Li H. 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv preprint arXiv:13033997

59. Robinson JT, Thorvaldsdóttir H, Winckler W, Guttman M, Lander ES, Getz G, Mesirov JP. 2011. Integrative genomics viewer. Nature biotechnology 29:24-26

60. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Molecular cell 38:576-589

61. Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841-842

62. Gel B, Serra E. 2017. karyoploteR: an R/Bioconductor package to plot customizable genomes displaying arbitrary data. Bioinformatics 33:3088-3090

63. Livak KJ, Schmittgen TD. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2−ΔΔCT method. Methods 25:402-408

64. Kim D, Langmead B, Salzberg SL. 2015. HISAT: a fast spliced aligner with low memory requirements. Nature methods 12:357-360

65. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R. 2009. The sequence alignment/map format and SAMtools. Bioinformatics 25:2078-2079

66. Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923-930

67. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome biology 15:550

68. Wickham H. 2016. ggplot2: elegant graphics for data analysis. Springer.

69. Sperschneider J, Dodds PN, Gardiner DM, Singh KB, Taylor JM. 2018. Improved prediction of fungal effector proteins from secretomes with EffectorP 2.0. Molecular Plant Pathology 19:2094-2110.10.1111/mpp.12682

70. Jones K, Kim DW, Park JS, Khang CH. 2016. Live-cell fluorescence imaging to investigate the dynamics of plant cell death during infection by the rice blast fungus *Magnaporthe oryzae*. BMC Plant Biology 16:69

71. Matsunaga TM, Ogawa D, Taguchi-Shiobara F, Ishimoto M, Matsunaga S, Habu Y. 2017. Direct quantitative evaluation of disease symptoms on living plant leaves growing under natural light. Breed Science 67:316-319.10.1270/jsbbs.16169

72. Schneider CA, Rasband WS, Eliceiri KW. 2012. NIH Image to ImageJ: 25 years of image analysis. Nature Methods 9:671

Figures

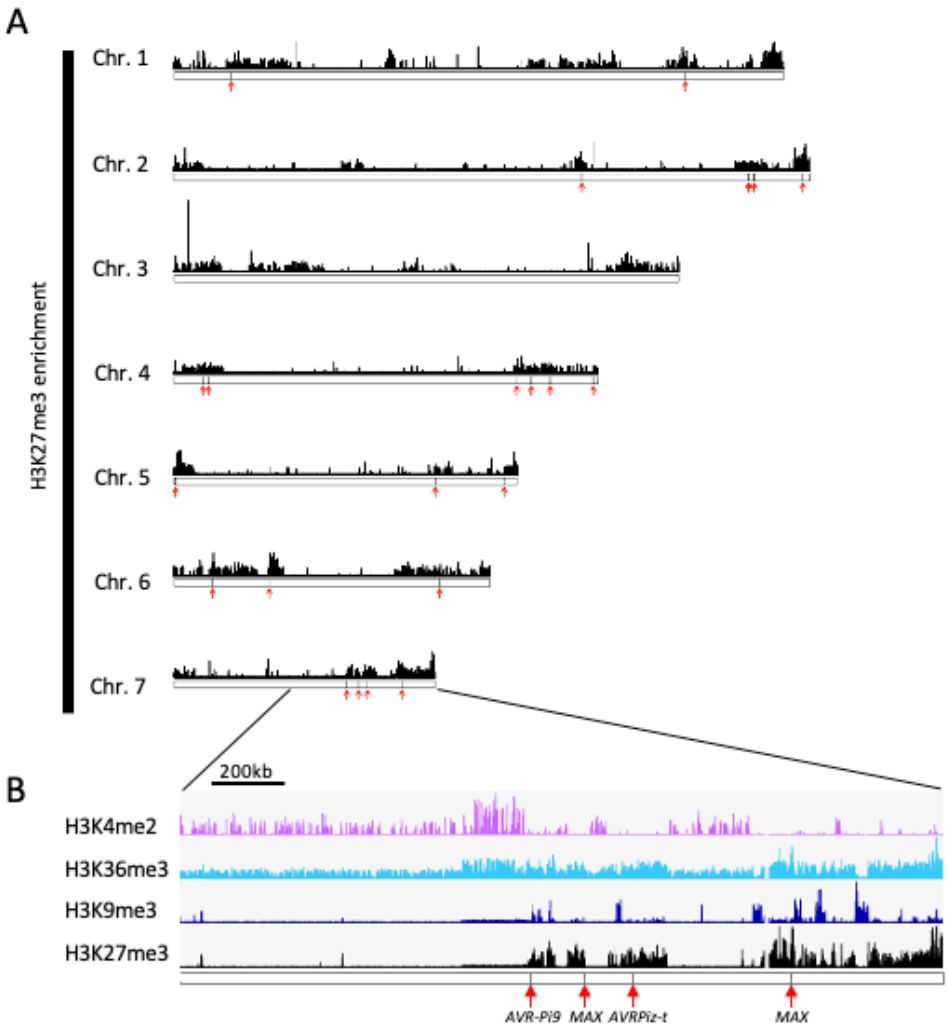


Figure 2.1: 24 known effector genes are found in regions enriched for H3K27me3

(A) Karyotype plot for all chromosomes, showing genome wide location of H3K27me3 ChIP enrichment. Red arrows indicate genomic location of 23 known effector genes. The location of one effector gene (MGG_10234) is unplaced in seven chromosomes but still found in an H3K27me3-enriched region. (B) IGV genome browser screenshot of ChIP

for H3K4me2, H3K9me3, H3K36me3, and H3K27me3 enrichment for partial of Chromosome VII. Gray bars and small red arrows indicate known effector genes.

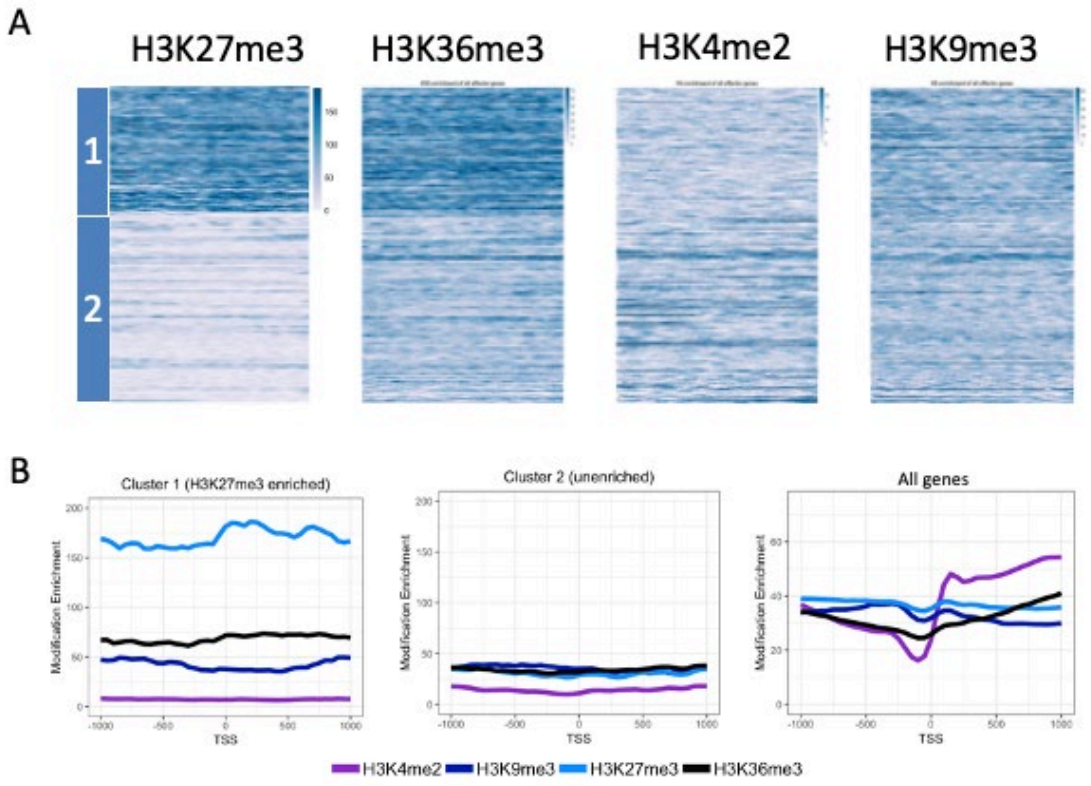


Figure 2.2: Effector genes cluster into two classes based on H3K27me3 enrichment status

(A) Heatmaps for all effector genes centered on the transcription start site (TSS) +/- 1000bp for a total window size of 2000bp. H3K27me3 heatmap was clustered using k-means of two, to divide effector genes into enriched (cluster 1: top), vs unenriched (cluster 2: bottom). All other modifications are plotted in the same order as the H3K27me3 heatmap. (B) Metaplot of H3K27me3, H3K36me3, H3K4me2, H3K9me3 at all effector genes centered on TSS and +/-1000bp for a total window size of 2000bp, for cluster 1, cluster 2 (left/middle: on same scale) and all genes (right: on different scale).

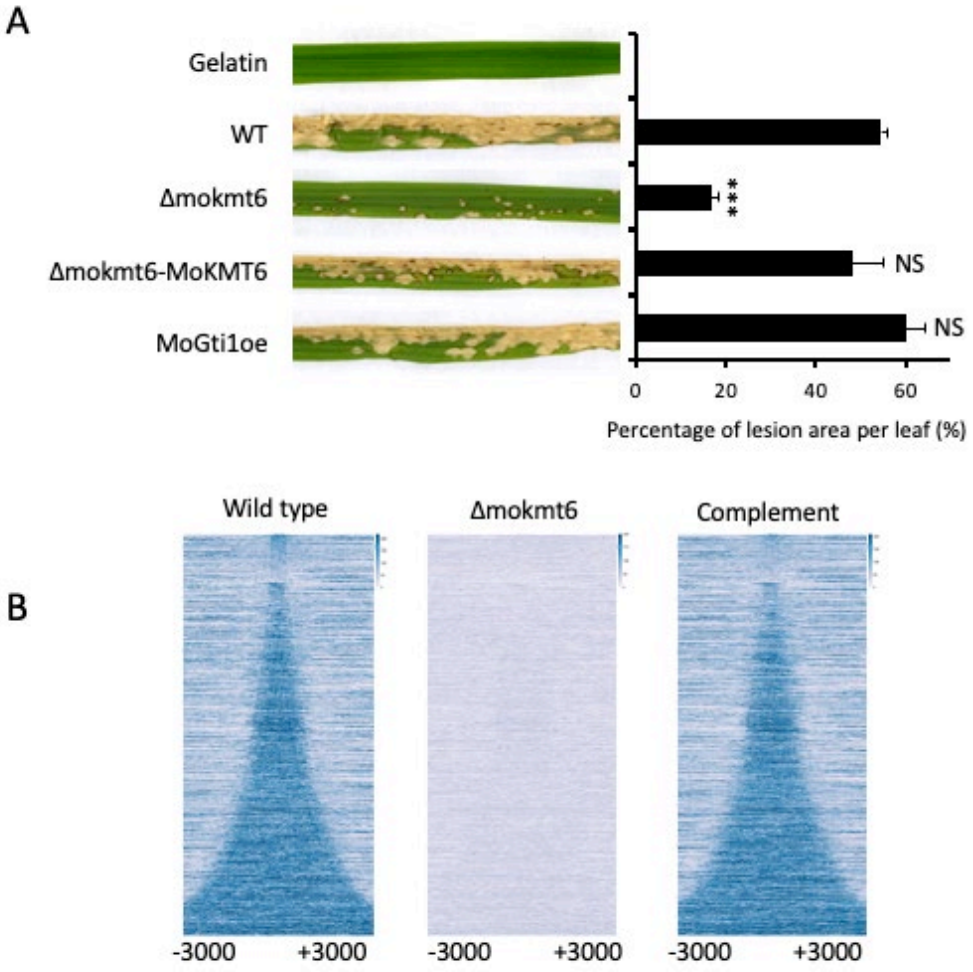


Figure 2.3: *MoKMT6* is required for the full virulence of *M. oryzae* and deposition of H3K27me3

(A) Targeted deletion of *MoKMT6* gene resulted in reduced pathogenicity on a fully susceptible rice cultivar YT16 in whole plant spray inoculation assay. Overexpression of *MoGti1* gene did not affect pathogenicity. Inoculation with 0.25% of gelatin was used as the negative control. Bar chart showed the percentage of lesion area per marked leaf after infection with fungal spores from different strains. Error bar equals standard deviation of the mean. *** $p<0.001$ (Two-tailed student t-test). NS means no significant difference.

(B) Heatmaps for all ~800 H3K27me3 peaks in genome ordered from smallest to largest.

Heatmaps are centered on peak center and +/-3000bp for a total window size of 6000bp. Deletion of *MoKMT6* abolishes H3K27me3 peaks (middle), and reintroduction of wild-type *MoKMT6* restores all H3K27me3 peaks (right).

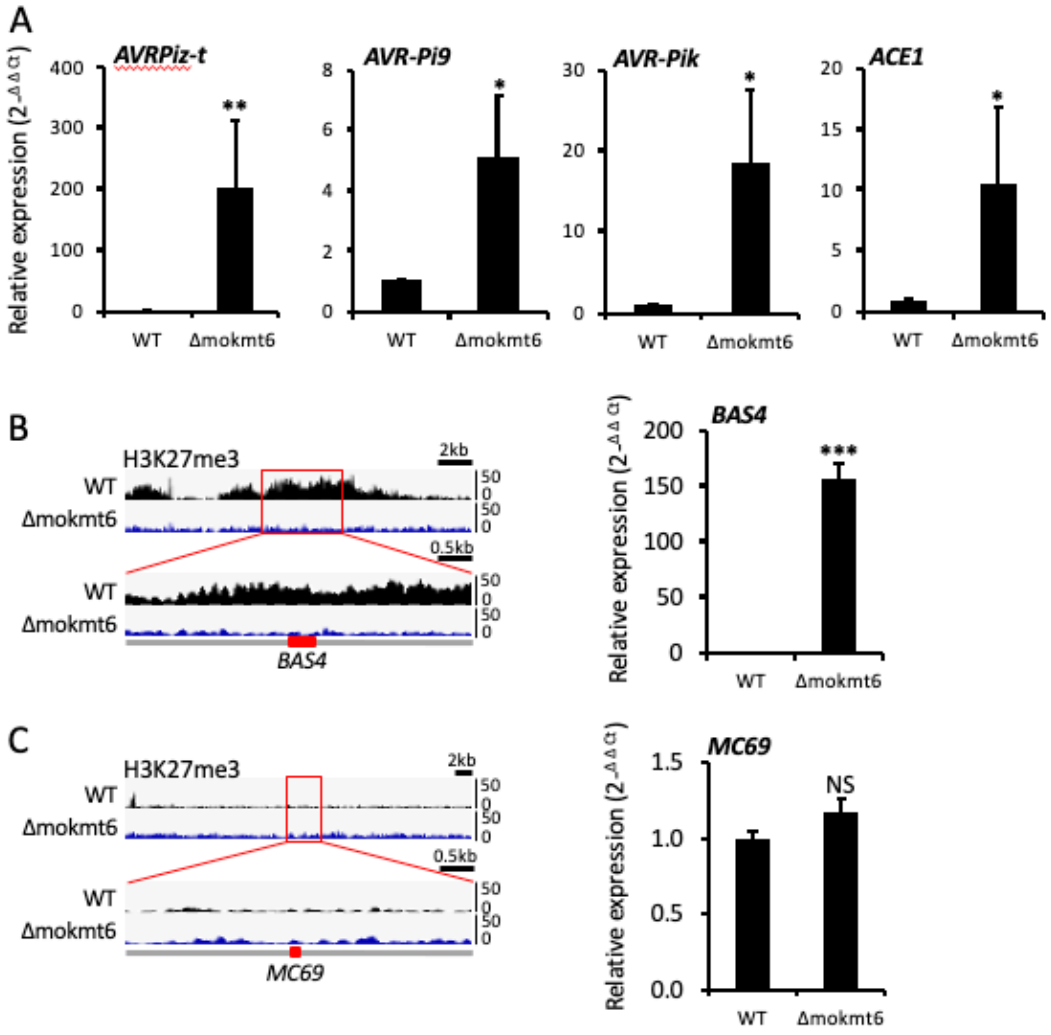


Figure 2.4: H3K27me3 is involved in repressing effector gene expression during mycelial growth of *M. oryzae*

(A) The expression of effector genes enriched with H3K27me3 were analyzed in strains of wild-type O-137 and *Δmokmt6* that were grown in CM. Quantitative RT-PCR analysis of expression was used to measure relative transcripts of avirulence genes to *M. oryzae*

actin gene. (B) *BAS4* expression was derepressed in the absence of H3K27me3 during mycelial growth. Left panel shows reduced H3K27me3 enrichment at *BAS4* and nearby loci in *Δmokmt6* compared to a wild-type. Right panel shows derepressed *BAS4* expression in mycelia of *Δmokmt6*. (C) *MC69* expression was remained in the absence of H3K27me3 during mycelial growth. No H3K27me3 enrichment was observed at *MC69* and nearby loci (left panel) and no expression change was detected in *Δmokmt6* compared to a wild-type (right panel). The abundance of effector gene transcripts in mutant is expressed relative to a value of 1 in the wild-type O-137. Mean values and standard deviation were calculated from three biological replicates. Two-tailed student t-test was performed to determine statistical difference. * indicates $p<0.05$, ** indicates $p<0.01$, *** indicates $p<0.001$ and NS means no significant difference was detected. See more examples in supplemental figure 5.

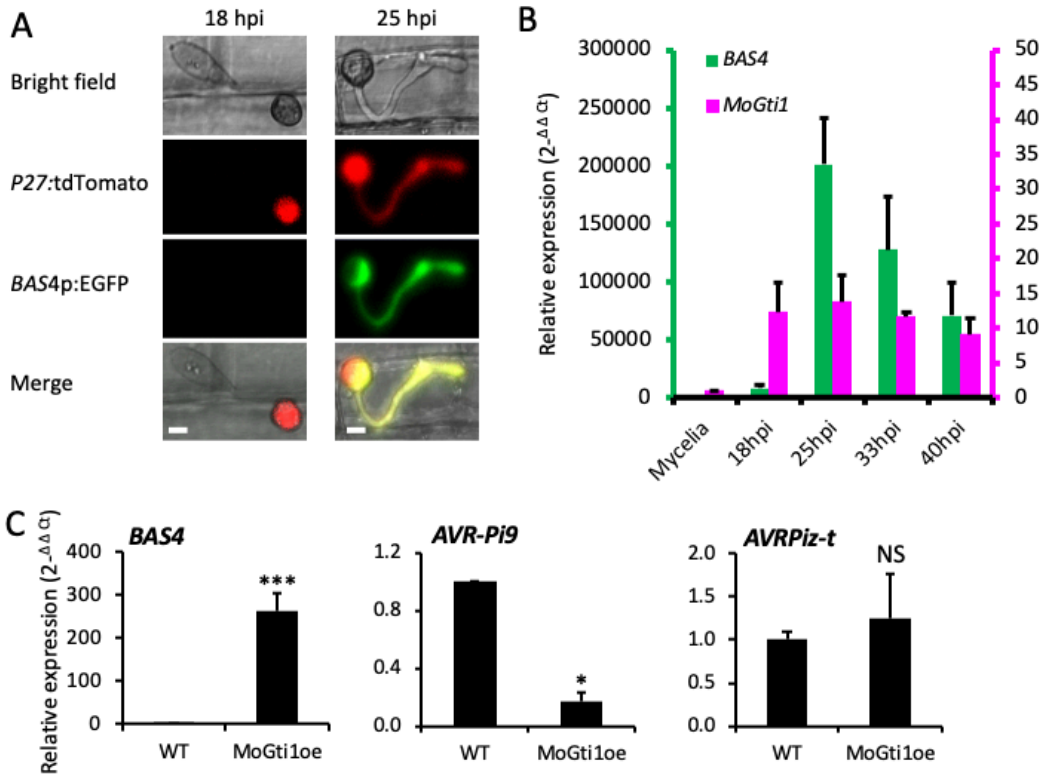


Figure 2.5: The transcription factor MoGti1 differentially regulates a subset of effector genes

Expression of *BAS4* during rice sheath infection was monitored by a fluorescent reporter strain (A) and qRT-PCR analysis (B). (A) At 18 hours post inoculation (hpi), only tdTomato control fluorescence was observed in an appressorium. But both EGFP fluorescence indicative of *BAS4* expression and tdTomato fluorescence were observed at 25hpi in appressorium and invasive hyphae. All images are projections of multiple z-stacks. Bars: 10 µm. (B) Time-course qRT-PCR analysis at different time points (18, 25, 33, 40 hpi) confirmed the *BAS4* expression pattern and also indicated a coincide expression of *BAS4* and *MoGti1* during plant infection. (C) MoGti1 overexpression upregulated *BAS4* expression during mycelial growth (left panel), downregulated *AVR*-

Pi9 expression (middle panel), and did not alter *AVRPiz-t* expression (right panel) when compared to that in a wild-type. See more examples in supplemental figure 8. Mean values and standard deviation were calculated from three biological replicates. * indicates p<0.05, *** indicates p<0.001 and NS means no significant difference was detected.

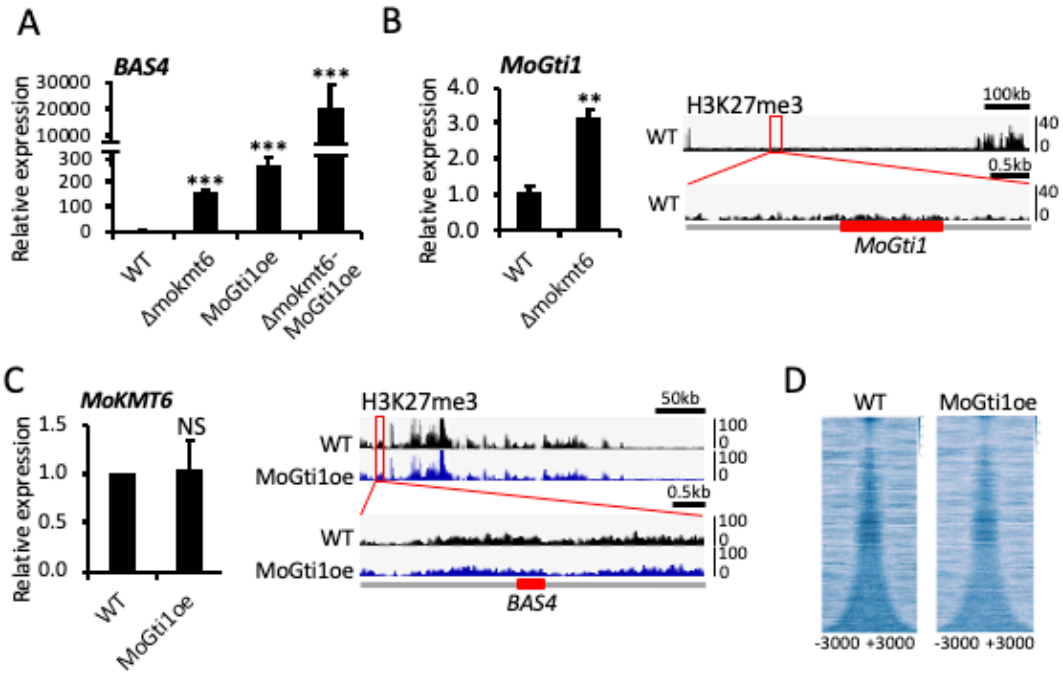


Figure 2.6: Double control of effector gene expression by H3K27me3 and MoGti1

(A) *BAS4* expression was synergistically upregulated during mycelial growth of Δ*mokmt6*-MoGti1oe strain compared to that of Δ*mokmt6* and MoGti1oe strain respectively by a qRT-PCR analysis. *BAS4* expression in Δ*mokmt6* and MoGti1oe strains were calculated from the same dataset to the Figure 4B and 5C. All four samples were prepared and run at same time. Mean values and standard deviation were calculated from three biological replicates. *** indicates p<0.001. (B) *MoGti1* expression was upregulated during mycelial growth of Δ*mokmt6* compared to wild-type (p<0.01, left

panel), though no H3K27me3 enrichment was observed at *MoGti1* and nearby loci (right panel). (C) Neither *MoKMT6* expression level (left panel) nor H3K27me3 pattern at *BAS4* and nearby loci (right panel) was affected by MoGti1 overexpression during mycelial growth. (D) Overexpression of MoGti1 does not change H3K27me3 in *M. oryzae.* Heatmaps for all ~800 H3K27me3 peaks in genome ordered from smallest to largest. Heatmaps are centered on peak center and +/-3000bp for a total window size of 6000bp. Wild type (left), and overexpression MoGti1 peaks (right).

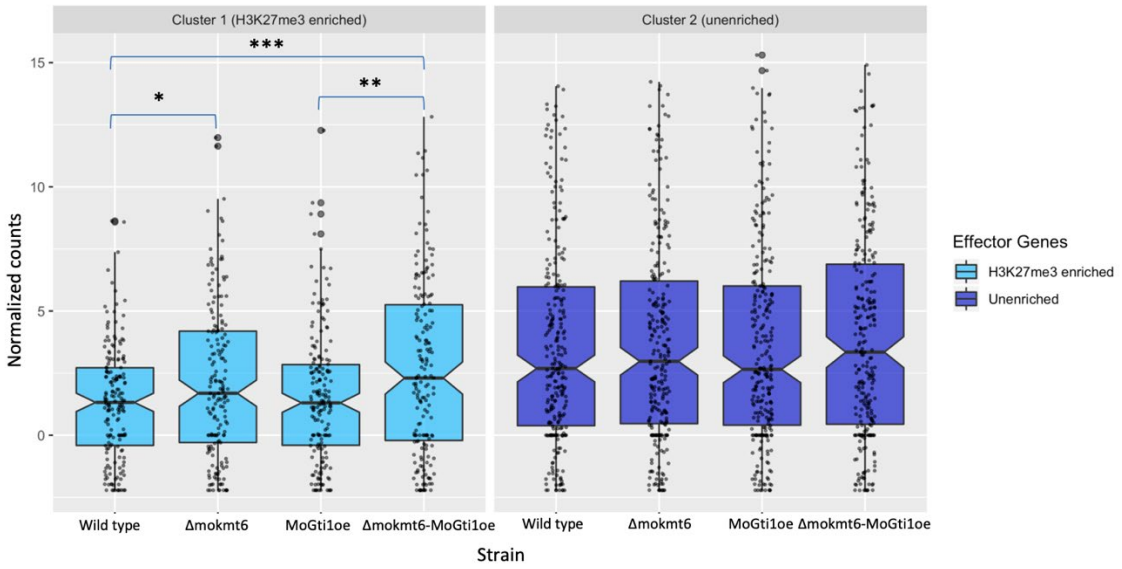


Figure 2.7: Effector genes show distinct expression changes according to H3K27me3 status

Boxplots of normalized counts for effector genes enriched in H3K27me3 (left) and unenriched (right). Statistical significance indicated by *, using the Kruskal-Wallis test.

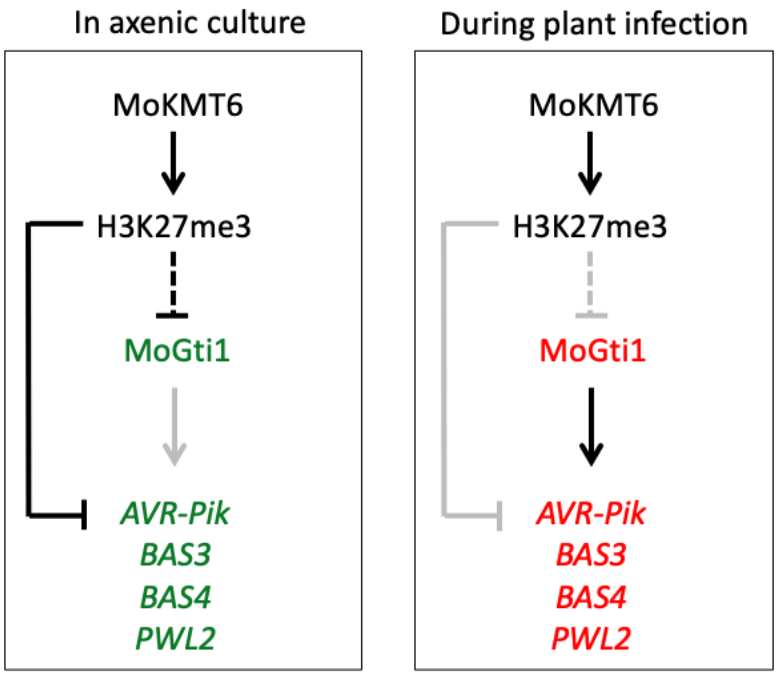


Figure 2.8. Expression of effector genes is coordinately controlled by histone modification H3K27me3 and transcription factor MoGti1

During growth in axenic culture, H3K27me3, deposited by MoKMT6, is highly enriched at effector gene loci and thereby involves repressing effector gene expression. Meanwhile, MoGti1 expression is repressed by H3K27me3, although no H3K27me3 enrichment is observed at MoGti1 locus. During plant infection, H3K27me3 enrichment is redistributed at effector gene loci and MoGti1 expression is upregulated. This allows MoGti1 or unknow transcription factors to efficiently access effector gene loci to regulate expression of effector genes. Green and red colors indicate repressed and upregulated expression respectively. Black and grey lines indicate enriched and removed H3K27me3 at effector gene loci.

Tables

Table 2.1: A total of 24 *M. oryzae* effector genes that are repressed during mycelial growth in culture but induced in planta

Gene Name	MGG_#	Chromosome	H3K27me3 enrichment[a]	Repressed in mycelia[b]	Induced in planta[c]	Reference
ACE1	MGG_12447	2	High	+	+, 16hpi	Bohnert et al., 2004; Fudal et al., 2007
AVR-Pi9	MGG_12655	7	Low	+	+, 12hpi	Wu et al., 2015
AVR-Pik	MGG_15972	2	High	+	+, 24hpi	de Guillen et al, 2015; Dong et al., 2015
AVR-Pita1	MGG_15370	6	Low	+	+	Mosquera et al., 2009;
AVRPiz-t	MGG_18041	7	High	+	+, 24 hpi	Li et al., 2009; Park et al., 2012; Dong et al., 2015
BAS1	MGG_04795	1	High	+	+	Mosquera et al., 2009
BAS2	MGG_09693	4	Low	+	+	Mosquera et al., 2009
BAS3	MGG_11610	5	High	+	+, 16hpi, 24hpi,	Mosquera et al., 2009; de Guillen et al, 2015; Mogga et al., 2016
BAS4	MGG_10914	5	High	+	+, 24hpi	Mosquera et al., 2009; Mogga et al., 2016
BAS107	MGG_10020	4	High	+	+	Mosquera et al., 2009; Dong et al., 2015
MAX	MGG_02546	7	High	+	+, 24, 48hpi	Dong et al., 2015; de Guillen et al, 2015
MAX	MGG_08414	2	High	+	+, 24hpi	de Guillen et al, 2015
MAX	MGG_08482	4	High	+	+, 24hpi	de Guillen et al, 2015
MAX	MGG_09675	7	High	+	+, 24hpi	de Guillen et al, 2015; Dong et al., 2015
MoCDIP3	MGG_07986	2	High	+	+, App, 96hpi	Chen et al., 2013
MoCDIP4	MGG_08409	2	Low	+	+, App, and 72hpi	Chen et al., 2013
MoCDIP5	MGG_10234	Unplaced	Low	+	+, App, 96hpi	Chen et al., 2013
MoHEG6	MGG_08506	4	High	+	+, 24hpi	Dong et al., 2015; Mogga et al., 2016
MoHEG9	MGG_00043	5	High	+	+, 24hpi	Mosquera et al., 2009; Dong et al., 2015; Mogga et al., 2016
MoHEG12	MGG_06224	4	High	+	+	Mosquera et al., 2009; Dong et al., 2015; Mogga et al., 2016
PWL2	MGG_04301	6	High	+	+	Sweigard et al., 1995; Mosquera et al., 2009
SLP1	MGG_10097	4	High	+	+	Mentlak et al., 2012; Dong et al., 2015
SPD5	MGG_02154	1	High	+	+	Dong et al., 2015; Sharpee et al., 2017
SPD8	MGG_09379	6	High	+	+	Mosquera et al., 2009; Dong et al., 2015; Sharpee et al., 2017

[a]H3K27me3 enrichment is profiled by ChIP-seq of *M. oryzae* mycelia grown in axenic culture.

[b]Undetectable expression of effector genes either by RT-PCR, qRT-PCR, fluorescent reporter or RNA-seq during fungal growth in axenic culture. "+" means yes.

[c]Induced expression during plant infection process including appressorium stage and invasive growth stage. "+" means yes.

Table 2.2: Phenotypic characterization of *MoKMT6* knock-out and *MoGti1* overexpression strains

Strain	Mycelial growth (mm)[a]	Conidiation ($x10^5$/mL)[b]	Conidia germination (%)[c]	Appressorium formation(%)[d]
Wild-type	69.0 ± 2.2	12.8 ± 4.1	99.0 ± 2.73	99.2 ± 1.27
Δmokmt6	66.9 ± 0.50	0.15 ± 0.05**	98.4 ± 2.97	93.7 ± 11.2*
Δmokmt6-MoKMT6	67.5 ± 1.5	3.3 ± 1.1*	98.5 ± 3.5	97.5 ± 3.54
MoGti1oe	59 ± 1.2**	14.3 ± 4.8	98.1 ± 2.87	96.4 ± 3.43*

[a]Vegetative growth of fungal mycelia was measured at 14 days post-inoculation on complete medium plates.
[b]Conidia were counted after 12 days of growth on 5cm of OMA media incubated at 24°C under continues light. Conidia were harvested by suspending them with 5 ml of distilled water per plate.
[c]Germination rate was measured 21-25 hours after inoculating on hydrophobic glass coverslips with conidia from 12-day-old OMA cultures. Germination rate was measured as the percentage ratio of germinated conidia to all counted conidia.
[d]Appressorium formation rates were calculated 21-25 hours after inoculating on glass coverslips from 12-day-old OMA cultures under a confocal microscope. Appressorium formation rate was measured as the percentage ratio of mature appressoria to germinated conidia.
All data are presented as means ± SD from four biological replicates. Data were analyzed with two-tailed student t-test. *, $p<0.05$; **, $p<0.01$.

Table 2.3: Summary of qRT-PCR tested effector genes in different backgrounds

Effector	MGG#	H3K27me enrichment	Induced expression in *Δmokmt6* (Y:Yes, N:No)	Regulated by MoGti1	Relative expression (qRT-PCR, Fold change)			Relative expression (RNA-seq,Log_2FC)		
					Δmokmt6	MoGti1oe	*Δmokmt6*-MoGti1oe	*Δmokmt6*	MoGti1oe	*Δmokmt6*-MoGti1oe
AVR-Pik	MGG_15972	High	Y	Positive	18.5	1.8	219.8	1.4[a,b]	-0.9[a,b]	4.9
BAS4	MGG_10914	High	Y	Positive	156.0	263.0	20090.0	3.9	5.2	10.6
BAS3	MGG_11610	High	Y	Positive	4.2	4.4	71.2	2.0	2.2	6.1
PWL2	MGG_04301	High	Y	Positive	3.5	7.2	235.6	0[a,b]	0[a,b]	1.0[a,b]
AVRPiz-t	MGG_18041	High	Y	No influence	198.9	1.2[b]	243.2	6.8	-0.9[b]	6.8
ACE1	MGG_12447	High	Y	No influence	10.5	0.8[b]	6.0	5.3	0.3[b]	4.8
Slp1	MGG_10097	High	Y	No influence	21.3	0.7[b]	38.9	4.9	0.8[b]	7.0
MC69	MGG_02848	Low	N	No influence	1.2[b]	0.9[b]	1.3[b]	-0.06[b]	-0.2[b]	-0.05[b]
AVR-Pi9	MGG_12655	Low	Y	Negative	5.1	0.2	1.1[b]	2.4	-5.1	0.02[b]

[a]Quantification of expression data in RNA-seq is impacted by lack of reads in wildtype. (For *AVR-Pik*, the wild-type, *Δmokmt6*, MoGti1oe samples all have very few or no reads, the *Δmokmt6*-MoGti1oe has significantly more. For *PWL2*, wild-type has no reads, *Δmokmt6* has very few, and the *Δmokmt6*-MoGti1oe has significantly more.)

[b]There is no statistical significant compared to wild-type ($p>0.05$).

Effector	MGG#	H3K27me enrichment	Induced in *Δmokmt6* (+:Yes, -:No)	Regulated by MoGti1	Relative expression (qRT-PCR, Log_2Fold change)			Relative expression (RNA-seq, Log_2FC)		
					Δmokmt6	MoGti1oe	*Δmokmt6*-MoGti1oe	*Δmokmt6*	MoGti1oe	*Δmokmt6*-MoGti1oe
AVR-Pik	MGG_15972	High	+	Positive	4.2	0.8	7.8	1.4[a,b]	-0.9[a,b]	4.9
BAS4	MGG_10914	High	+	Positive	7.3	8.0	14.3	3.9	5.2	10.6
BAS3	MGG_11610	High	+	Positive	2.1	2.1	6.2	2.0	2.2	6.1
PWL2	MGG_04301	High	+	Positive	1.8	2.8	7.9	0[a,b]	0[a,b]	1.0[a,b]
AVRPiz-t	MGG_18041	High	+	No influence	7.6	0.3[b]	7.9	6.8	-0.9[b]	6.8
ACE1	MGG_12447	High	+	No influence	3.4	-0.3[b]	2.6	5.3	0.3[b]	4.8
Slp1	MGG_10097	High	+	No influence	4.4	-0.5[b]	5.3	4.9	0.8[b]	7.0
MC69	MGG_02848	Low	-	No influence	0.3[b]	-0.2[b]	0.4[b]	-0.06[b]	-0.2[b]	-0.05[b]
AVR-Pi9	MGG_12655	Low	+	Negative	2.4	-2.3	0.1[b]	2.4	-5.1	0.02[b]

[a]Quantification of expression data in RNA-seq is impacted by lack of reads in wildtype. (For *AVR-Pik*, the wild-type, *Δmokmt6*, MoGti1oe samples all have very few or no reads, the *Δmokmt6*-MoGti1oe has significantly more. For *PWL2*, wild-type has no reads, *Δmokmt6* has very few, and the *Δmokmt6*-MoGti1oe has significantly more.)

[b]There is no statistical significant compared to wild-type ($p>0.05$).

Effector	MGG#	H3K27me enrichment	Induced in *Δmokmt6* (+:Yes, -:No)	Regulated by MoGti1	Relative expression (qRT-PCR, Log_2Fold change)			Relative expression (RNA-seq, Log_2FC)		
					Δmokmt6	MoGti1oe	*Δmokmt6*-MoGti1oe	*Δmokmt6*	MoGti1oe	*Δmokmt6*-MoGti1oe
AVR-Pik	MGG_15972	High	+	Positive	4.2	0.8	7.8	NA	NA	4.9
BAS4	MGG_10914	High	+	Positive	7.3	8.0	14.3	3.9	5.2	10.6
BAS3	MGG_11610	High	+	Positive	2.1	2.1	6.2	2.0	2.2	6.1
PWL2	MGG_04301	High	+	Positive	1.8	2.8	7.9	NA	NA	NA
AVRPiz-t	MGG_18041	High	+	No influence	7.6	NC	7.9	6.8	NC	6.8
ACE1	MGG_12447	High	+	No influence	3.4	NC	2.6	5.3	NC	4.8
Slp1	MGG_10097	High	+	No influence	4.4	NC	5.3	4.9	NC	7.0
MC69	MGG_02848	Low	-	No influence	NC	NC	NC	NC	NC	NC
AVR-Pi9	MGG_12655	Low	+	Negative	2.4	-2.3	NC	2.4	-5.1	NC

NA means Log_2Fold change is not available. Quantification of expression data in RNA-seq is impacted by lack of reads in wild-type.

NC means there is no change for expressions compared to wild-type ($p>0.05$).

Table 2.4: Percentage of up-/down-regulated genes and effector genes during mycelial growth of different fungal strains from transcriptomic analysis

		Up-regulated[d]			Down-regulated[d]		
		Δ*mokmt6* vs. wt	**MoGti1oe vs. wt**	**Δ*mokmt6*-MoGti1oe vs. wt**	**Δ*mokmt6* vs. wt**	**MoGti1oe vs. wt**	**Δ*mokmt6*-MoGti1oe vs. wt**
Whole-genome (13144)[a]		1858 (14.1%)	1524 (11.6%)	2979 (22.7%)	935 (7.1%)	1228 (9.3%)	2098 (16.0%)
Effector genes (448)	Predicted effectors[b]	110 (24.6%)	60 (13.4%)	167 (37.3%)	19 (4.2%)	25 (5.6%)	31 (6.9%)
	Known effectors[c]	15 (3.3%)	3 (0.7%)	16 (3.6%)	0	1 (0.2%)	0

[a]Total number of genes with MGG_#: 13144

[b]Total number of effector genes that are predicted by EffectorP: 424

[c]Total number of known effector genes of which expression are repressed in mycelia: 24

[d]Up regulated genes: Log_2foldchange>0; Down regulated genes: Log_2foldchange<0.

padjusted <0.05

Table 2.5: High percentage of H3K27me3-enriched effector genes are upregulated in Δ*mokmt6* mycelia and at 36hpi of WT

	Up-regulated in mycelia ($Log_2FC>2$)			**Up-regulated in planta (36hpi)[d] ($Log_2FC>2$)**
	Δ*mokmt6* vs. wt	**MoGti1oe vs. wt**	**Δ*mokmt6*-MoGti1oe vs. wt**	**WT**
Whole-genome (13144)[a]	1053 (8.0%)	462 (3.5%)	1470 (11.2%)	1351 (10.3%)
H3K27me3-enriched effector genes (178)[b]	75 (42.1%)	23 (12.9%)	97 (54.5%)	86 (48.3%)
H3K27me3-unenriched effector genes (268)[c]	31 (11.6%)	21 (7.8%)	58 (21.6%)	65 (24.3%)

[a]Total number of genes with MGG_#: 13144

[b]Total number of effector genes which show high enrichment of H3K27me3 during mycelial growth: 178

[c]Total number of effector genes which show low enrichment of H3K27me3 during mycelial growth: 268

[d]Data is from Mosquera et al., 2009.

padjusted <0.05

Table 2.6: Majority of effector genes upregulated both in *Δmokmt6* mycelia and at 36hpi of WT show high enrichment of H3K27me3 during mycelial growth.

#	MGG_#	H3K27me3 enrichment[a]	Note
1	MGG_01956	High	
2	MGG_02154	High	*SPD5*
3	MGG_05424	High	
4	MGG_06224	High	*MoHEG12*
5	MGG_07357	High	
6	MGG_07880	High	
7	MGG_07919	High	
8	MGG_08355	High	
9	MGG_08399	High	
10	MGG_08414	High	*MAX*
11	MGG_08480	High	
12	MGG_08482	High	*MAX*
13	MGG_08506	High	*MoHEG6*
14	MGG_08610	High	
15	MGG_08817	High	
16	MGG_09378	High	
17	MGG_10318	High	
18	MGG_10455	High	
19	MGG_10477	High	
20	MGG_10914	High	*BAS4*
21	MGG_11072	High	
22	MGG_15046	High	
23	MGG_15443	High	
24	MGG_15620	High	
25	MGG_15924	High	
26	MGG_16041	High	
27	MGG_16058	High	
28	MGG_16585	High	
29	MGG_16619	High	
30	MGG_16693	High	
31	MGG_16698	High	
32	MGG_17244	High	
33	MGG_17556	High	
34	MGG_17567	High	
35	MGG_17582	High	
36	MGG_18035	High	
37	MGG_18041	High	*AVRPiz-t*
38	MGG_02239	Low	
39	MGG_05403	Low	
40	MGG_07556	Low	
41	MGG_08407	Low	
42	MGG_08428	Low	
43	MGG_08435	Low	
44	MGG_09377	Low	
45	MGG_09693	Low	*BAS2*
46	MGG_12655	Low	*AVR-Pi9*

[a]H3K27me3 enrichment is profiled by ChIP-seq of *M. oryzae* mycelia grown in axenic culture.

Table 2.7: Majority of synergistically upregulated effector genes by H3K27me3 loss and MoGti1 overexpression are also upregulated at 36hpi of WT

#	MGG_#	Fold change			Synergistically up-regulated[a] (+:Yes; -:No)	Induced at 36 hpi[b] (+:Yes; -:No)	H3K27me3 enrichment[c]
		Δ*mokmt6*	MoGti1oe	Δ*mokmt6*-MoGti1oe			
1	MGG_01956	5.8	4.4	48.4	+	+	High
2	MGG_08355	44.9	21.5	309.3	+	+	High
3	MGG_08399	388.2	71.4	1651.4	+	+	High
4	MGG_08480	167.9	231.7	725.5	+	+	High
5	MGG_08610	913.7	187.0	2591.7	+	+	High
6	MGG_10455	38.3	100.3	1783.1	+	+	High
7	MGG_10914(*BAS4*)	14.8	37.6	1531.8	+	+	High
8	MGG_11072	386.8	712.5	1114.4	+	+	High
9	MGG_11610(*BAS3*)	4.0	4.7	69.2	+	+	High
10	MGG_15046	70.8	19.2	400.4	+	+	High
11	MGG_15620	19.4	11.0	2494.7	+	+	High
12	MGG_16057	3.7	3.6	10.4	+	+	High
13	MGG_16058	108.6	8.7	194.0	+	+	High
14	MGG_16585	17.2	37.0	1374.2	+	+	High
15	MGG_17556	5.1	6.0	19.0	+	+	High
16	MGG_18035	21.0	20.4	54.7	+	+	High
17	MGG_08428	33.0	61.6	763.6	+	+	Low
18	MGG_08376	23.0	7.6	89.6	+	-	Low
19	MGG_08941	4.4	7.0	12.4	+	-	Low
20	MGG_10456	7.2	2.1	11.2	+	-	Low
21	MGG_17580	32.1	5.6	340.0	+	-	High
22	MGG_02154(*SPD5*)	9.1	60.4	58.3	-	+	High
23	MGG_00614	2.1	5.0	3.9	-	+	Low
24	MGG_01974	1.8	2.1	3.6	-	+	Low
25	MGG_05531	2.2	2.5	1.3[d]	-	+	Low

26	MGG_07556	37.3	3.4	33.2	-	+	Low
27	MGG_07558	9.9	4.7	1.3[d]	-	-	High
28	MGG_16188	2.2	4.2	5.4	-	-	High
29	MGG_16553	192.3	19.4	52.2	-	-	High
30	MGG_05406	1.8	1.6	1.7	-	-	Low
31	MGG_05982	25.5	3.3	10.2	-	-	Low
32	MGG_08543	8.6	2.9	4.5	-	-	Low
33	MGG_09842	7.7	102.6	1.8[d]	-	-	Low
34	MGG_10531	41.9	56.2	50.4	-	-	Low
35	MGG_14195	44.5	15.9	36.3	-	-	Low
36	MGG_16869	2.2	7.5	7.5	-	-	Low

[a]Synergistically upregulated effector genes are defined that expression levels of effector genes in *Δmokmt6*-MoGti1oe are higher than the sum of individual expression levels in *Δmokmt6* mutant and MoGti1oe, respectively.

[b]Data is from Mosquera et al., 2009

[c]H3K27me3 enrichment is profiled by ChIP-seq of *M. oryzae* mycelia grown in axenic culture.

[d]Not statistically significant (padjusted>0.05).

Table 2.8. PCR primers used in this study

Name	Sequence [a] (5'-3')	Applications
CKP412	*GC*GTCGACTAAGTCCGGTAGTGTGAAGAGG	*MoKMT6* upstream for knockout
CKP413	*CG*GAATTCCTTGTCACTTTTGCTCCCTCC	*MoKMT6* upstream for knockout
CKP421	GGATCCGAAGTTCAACCGGCTCGC	*MoKMT6* downstream for knockout
CKP428	TCTAGAAAGTCCTACGGGCACAAGGC	*MoKMT6* downstream for knockout
CKP416	*CG*GAATTCTCGACAGAAGATGATATTG	*NTPII* sequence
CKP420	TCTAGATTAGAAGAACTCGTCAAG	*NTPII* sequence/ *MoKMT6* knockout screening
CKP443	AGGCTCGGTCGAGAATTGAC	*MoKMT6* knockout screening
CKP459	GGCAGGAGGGAGCAAAAGTGAC	*MoKMT6* knockout/ complementation screening
CKP460	ATTCCGCCTTGTGCCCGTAG	*MoKMT6* knockout/ complementation screening
CKP719J	GGATCCATGACGGGCAAGCTCGG	*MoKMT6* sequence for complementation
CKP720J	TCTAGAGAGCAGCACAGGCCAAGG	*MoKMT6* sequence for complementation
CKP565	GAATTCGCTGGGTAACTGATATTGAAGG	*Nat1* sequence
CKP566	CTCGAGTCAGGGGCAGGGCATG	*Nat1* sequence
CKP591	GGATCCATGTCGACTACGGGACAAGG	*MoGti1* sequence for overexpression
CKP592	AAGCTTTCGCTTGAGAAATAATATGTATTCAGGC	*MoGti1* sequence for overexpression
CKP110	GAATTCGGTAGCTTCTACGGATGC	*BAS4* 1-kb upstream for expression reporter
CKP234	GGATCCCAT TGTGAA AAGATTCGTTGTGG	*BAS4* 1-kb upstream for expression reporter
CKP548	GCGGCCGCGAGGGTTCTTTCACCTCG	*BAS4* 0.5-kb downstream for expression reporter
CKP549	AAGCTTCGGGGCTTTTGACAGTACCC	*BAS4* 0.5-kb downstream for expression reporter
CKP348	ACGGTGACCATATCGAGTGC	*ACE1* qRT-PCR
CKP349	CGCGCTTATACGTCTCCTGG	*ACE1* qRT-PCR
CKP795J	AGGTGACGCCAAGATTTCCG	*AVR-Pi9* qRT-PCR
CKP796J	ACCAGTGCGTCTTTTCGACT	*AVR-Pi9* qRT-PCR
CKP679J	CACTTTGGGAACTGTCGCTG	*AVR-Pik* qRT-PCR
CKP680J	TCGGGTACAGGAATACCAGGG	*AVR-Pik* qRT-PCR
CKP793J	CGATAAGGAAGAAGGCGGGT	*AVR-Pizt* qRT-PCR
CKP794J	TGTACGGGTGACGCGTTTTT	*AVR-Pizt* qRT-PCR
CKP611	TTGAGGAATTGTGCCCCGAC	*BAS3* qRT-PCR
CKP612	CGCAGTCGATGACGCAGAT	*BAS3* qRT-PCR
CKP329	TGCGACGACTGCACTATCTG	*BAS4* qRT-PCR
CKP330	CGCCAAGGTTAGGGCATTTC	*BAS4* qRT-PCR
CKP546	GTCATCACCCCCATCACCAAG	*MC69* qRT-PCR
CKP547	TTTGGCAGGTCCGCGAAG	*MC69* qRT-PCR
CKP333	CGACGTCCGAAAGGATCTGT	*Moactin* qRT-PCR
CKP334	TGCATACGGTCCGAAAGACC	*Moactin* qRT-PCR
CKP615	GCTCAGGTTACTTGTGGGCT	*MoGti1* qRT-PCR
CKP616	AGGTGATGGCTGACACCTTG	*MoGti1* qRT-PCR
CKP789J	TGTAATGGACACACGCACGA	*MoKMT6* qRT-PCR
CKP790J	GGAGGTGTAACTGGTCAGCC	*MoKMT6* qRT-PCR
CKP327	GGCGGGTGGACTAACAAACA	*PWL2* qRT-PCR
CKP328	TACCATCCTATCGGGCCCTC	*PWL2* qRT-PCR
CKP544	GTTCGCTACCATCACCACCC	*Slp1* qRT-PCR
CKP545	GTTGCCCTGCACCGTGTAG	*Slp1* qRT-PCR

[a]Underlined sequences correspond to restriction enzyme sites used for cloning:

*BamH*I (GGATCC), *EcoR*I (GAATTC), *Hind*III (AAGCTT), *Not*I (CGGCCG), *Sal*I (GTCGAC), *Xba*I (TCTAGA) and *Xho*I (CTCGAG).

CHAPTER 3

SHANNON ENTROPY AS A METRIC FOR CONDITIONAL GENE EXPRESSION IN *NEUROSPORA CRASSA*[3]

[3] Abigail J. Courtney and Zachary A. Lewis. To be submitted to *G3: Genes, Genomes, Genetics.*

Abstract

Neurospora crassa has been an important model organism for molecular biology and genetics for over 60 years. *N. crassa* has a complex life cycle, with over 28 distinct cell types and is capable of transcriptional responses to many environmental conditions including nutrient availability, temperature, and light. To quantify variation in *N. crassa* gene expression, we analyzed public expression data from 97 conditions and used Shannon entropy to calculate entropy values for *Neurospora's* approximately 11,000 genes. Entropy values can be used to estimate the variability in expression for a single gene over a range of conditions and to classify individual genes as constitutive or condition-specific. Shannon entropy has previously been used measure the degree of tissue specificity of multicellular plant or animal genes. We use this metric here to measure variable gene expression in a microbe and provide this information as a resource for the *N. crassa* research community.

Introduction

Across conditions, individual genes can display expression patterns that range from constitutive to highly condition-specific. When performing Quantitative Reverse Transcription PCR (qRT-PCR) it is crucial to identify constitutively expressed genes for experimental normalization. Conversely, highly regulated, condition-specific gene promoters are often used in molecular biology to drive conditional expression of a gene under investigation (e.g., an essential gene) or to control expression of reporter genes in certain cell types or environmental conditions (e.g., a gene encoding a fluorescent protein). Such genes or promoters are often identified by examining gene expression across just a handful of experimental conditions; however, with the increase in publicly

available transcriptomics data it is possible to quantify variation in gene expression across many conditions for a given organism.

In 1963, Claude Shannon laid the basis for information theory, and described the unit known as Shannon entropy (1). A simplistic definition of Shannon entropy is that it describes the amount of information a variable can hold (2). In our case, a variable is a gene, and the information is the collection of expression values from different conditions. If a gene is classified as having low entropy, then the expression values would be generally consistent across different conditions or possess a low amount of information. Instead, if a gene is classified as having high entropy, then the expression of this gene would be highly variable across different conditions and contain a high amount of information.

Since entropy describes information contained in a variable, there are a number of uses for such a metric. Previous studies have used entropy to investigate cell and tissue specific expression of genes (3), identify potential therapeutic targets (4), characterize periodicity in gene expression (5), identify cancerous tissue samples (6), and make genomic comparisons (7). Studies using entropy have been carried out in human cell lines (8), mouse (3), plants (9), yeast (10), bacteria, phage, and metagenomes (11) but not yet in filamentous fungi.

Neurospora crassa has a 43Mb genome encoding approximately ~11,000 genes (12). There is a whole genome knock out collection, and genetic, genomic, and epigenetic studies have been carried out with this organism for more than 60 years (13). Indeed, *N. crassa* has been used as a model organism for epigenetics, testing fungal enzymes for biomass degradation, and circadian clock studies (14-16). As a resource for *N. crassa*

researchers, we generated an entropy value for most genes in the *N. crassa* genome using publicly available RNA-seq data, and we validated this approach using previously published lists of housekeeping or inducible genes. This resource has a number of useful applications for the *N. crassa* community.

Methods

Public data collection:

Entropy calculations were made for all genes in the *N. crassa* genome using public RNA-seq data sets (97 conditions from a total of 173 separate sets, including replicates). Table S1 contains SRA accession numbers, short descriptions, total reads, and mapped reads for each public data set used.

Mapping, TPM and entropy calculations:

HiSat2 (version 2.1.0) (17) was used to map all of the SRA accessions to the NC12 genome (NCBI assembly: GCA_000182925.2) using appropriate parameters specific for paired or single end sequence reads (with parameters –RNA-strandness RF or R) to produce bam files which were then sorted and indexed using SAMtools (version 1.3) (18). If experiments contain replicates, the replicate bam files were merged together before obtaining counts with featureCounts from Subread (version 1.6.2) (19). FeatureCounts was used with parameters -T exon to generate all counts at the gene level. Counts were imported into R where we obtained TPM using the function calculateTPM from the R package scater (20). This package takes in feature-level (in our case, gene-level) counts and gene lengths and outputs the TPM values for each gene. TPM values were then used to calculate the Shannon entropy using the R package BioQC (21). The function entropySpecificy was used to calculate the entropy values for all genes in the

genome. To examine specific genes sets, we converted from NCU accession numbers to gene identifiers from NCBI Genome Assembly NC12 (GCA_000182925.2) and plotted the kernel density estimation with rug plots.

Results and Discussion

Shannon entropy values are useful in measuring the amount of variation in expression levels across different tissues or growth conditions**.** In order to calculate Shannon entropy values for all *Neurospora crassa* genes, we first compiled a list of available RNA-seq data sets present in the NCBI sequence read archive (SRA). To calculate accurate entropy values, we needed to gather many observations of gene expression across different conditions. We searched the SRA database (22) for *N. crassa* RNA-sequencing entries that were processed at different developmental stages or grown under different conditions. In total we gathered 173 accessions, which represent 97 developmental or growth conditions. We then developed a pipeline to generate entropy values for each gene (Figure 3.1A). We first mapped to the NC12 *N. crassa* genome using HiSat2 (17) to generate bam files. The bam files were then used to generate read counts for each gene in each condition using featureCounts (19), which assigns reads to genomic features. Once the count file was created, we calculated normalized expression values using the Transcripts per Million (TPM) normalization method to create a matrix of normalized expression values for all genes in all conditions. We then used this expression matrix to calculate the Shannon entropy value for each gene (21). This generated entropy values for 10,300 out of 10,398 genes. The remaining 98 genes had 0 read counts in all conditions, so we were unable to calculate entropy. Our final entropy values range from 0.0506 to 6.599. 70% of the genes in the genome possess low entropy

values between 0.05 and 1 (7,180/10,300) (Figure 3.1B). These values include the constitutively expressed genes in the genome. Entropy values above one represent only 30% of the genome (3,120/10,300), corresponding to genes with more condition-specific expression patterns.

Validation of entropy as a measure of gene expression variation in *N. crassa.*

In order to determine if entropy values are a reliable predictor of expression variability in a microbe, we examined the entropy values generated here for published gene sets expected to be enriched for constitutively expressed genes, or conversely, for genes with highly condition-specific expression. If entropy value is a reliable measure of gene expression variation across conditions, housekeeping genes should be enriched for genes with low entropy values, whereas sets of conditionally-induced genes are expected to be enriched for high entropy values. Two previous studies identified genes useful for RT-qPCR controls in *N. crassa*. One of which published a list of 38 genes classified as "housekeeping" genes based on previously generated microarray and RNA-seq datasets under three different conditions (quinic acid (QA) induction, circadian gene expression profiling, and light response) (23), and the other study identified four genes by using previous transcriptomic studies and genes used in related organisms to generate candidates that were validated by quantitative PCR under different conditions (24). To visualize the distribution of entropy values in this set of 42 "housekeeping" genes, we plotted a kernel density estimation (KDE) of entropy values (Figure 3.2A). The KDE is a smoothed version of a histogram estimated from the underlying data. As expected, the highest density of data points in the housekeeping data set is around 0.25 (low entropy) and the density falls sharply around 0.75 (Figure 3.2A). Two genes in this set possess

entropy values above 1.6 and they encode an exo-beta-1,3-glucanase and a UDP-glucose dehydrogenase. We plotted a heatmap depicting TPM values for each gene in each condition with genes ranked by entropy values from low to high (top to bottom) (Figure 3.2B). Genes with higher entropy values showed significant induction of gene expression under certain conditions, whereas genes with low entropy values displayed consistent expression values across all conditions. Thus, these data highlight the need for a comprehensive analysis of conditional gene expression when selecting constitutive control genes.

We further validated the use of entropy as a measure for constitutive gene expression by using the same approach with a published list of 2,624 genes involved in transcription and translation, reasoning that genes involved in these essential processes would be expressed at similar levels in all conditions. (25). The distribution of entropy values for transcription and translation genes resembles the distribution of entropy values for housekeeping genes, where the highest density is concentrated at the low end of entropy values (Figure 3.2C). Many of the genes that possess entropy values above 1.6 are either hypothetical proteins or genes associated with cellular transport or metabolism. We again examined the TPM values for each gene in this set in a heatmap ranked by entropy from low to high and again find mostly steady expression across conditions (Figure 3.2D).

We next asked if higher entropy values were associated with inducible genes. The highest entropy values imply that a gene must only be expressed under specific conditions and may only show expression in one or a few of the conditions in the entire RNA-seq dataset. Using genes induced in different conditions, we repeated the same

analysis to examine the distribution of entropy values across multiple gene sets. To determine if higher entropy values were indeed associated with tissue specific gene expression, we created KDE plots for 513 genes induced by light (Figure 3.3A) and 3,259 genes that have expression changes during sexual development (Figure 3.3C) (26) (27). In both cases, there is a shift in distribution of entropy values toward higher entropy values compared to "housekeeping" or "transcription and translation" genes. We examined TPM values for each gene in each condition using a heatmap ranked by entropy values from low to high (top to bottom) and find that a majority of genes in each gene set show variable expression across conditions, as expected (Figure 3.3B, 3.3D). Genes that have regulation changes during perithecial (sexual) development also show a shift to the right, but with retention of more low entropy genes than in the light induced gene set (Figure 3.3C). Plotting the TPM values in an entropy ranked heatmap shows that approximately half of these genes are constitutively expressed across conditions and half are variably expressed, which we see in the lower entropy values in the density plot (Figure 3.3D). This implies that half of these genes are not specific to sexual or vegetative cell types even though they show transcriptional changes throughout development (27).

As a final confirmation that entropy can be used as a reliable metric to assess the variation or lack of variation in gene expression levels across many conditions, we plotted the expression levels of 100 genes with the highest entropy values and 100 genes with the lowest entropy values. We took the $\log_2$ TPM values for all conditions (columns) and plotted them for each gene (row) in a heatmap for both the top and bottom 100 genes. As expected, genes in the high entropy group displayed highly variable and condition-

specific expression (Figure 3.4A), whereas genes with the lowest entropy values show mostly uniform expression across all conditions (Figure 3.4B). Together, these data demonstrate that entropy is an effective tool for measuring variation in gene expression levels.

We wanted to examine correlations between repressive chromatin modifications and our set of RNA-seq experiments. We plotted the TPM expression values for sets of genes that are found in repressive domains throughout the genome. Interestingly, we found that a large group of H3K27me2/3 marked genes were upregulated during sexual development (Figure 3.5). A particularly important open question is when establishment of H3K27 methylation takes place in *N. crassa*. This is the first report of H3K27me2/3-target gene expression changes during sexual development in *N. crassa*, which could be an indication of establishment taking place during sexual development. Not only is this finding important in understanding the mechanisms of establishment, it also demonstrates that cell type transitions in *N. crassa* may be governed by similar mechanisms as in higher eukaryotes.

The information and code generated in the course of this study could prove useful in a number of ways. First, identifying genes that are induced in a certain condition and display a high entropy value will help identify genes that are condition-specific. In addition, examining entropy values for individual genes can be a useful approach for finding new inducible promoters to use for genetic studies. Condition-specific expressed genes are good starting targets to test for this purpose. The entropy metric determined here can also be used to confirm constitutive expression of genes chosen as controls for RT-PCR. In examining the housekeeping genes from previously published studies it is

clear that not all will function as good controls under all conditions, a limitation recognized by the authors. We combined all of their housekeeping genes together, whereas they had them divided into housekeeping genes usable for different conditions in qRT-PCR (QA induction, light response studies, and circadian experiments). Here we can choose genes that will work across all conditions (provided the conditions were represented in the initial dataset). Our metric allows for a more quantitative way to identify condition-specific genes, as opposed to investigating individual datasets or using controls from previous studies which may not perform as expected. In addition, this methodology is scalable; the initial inclusion of more conditions will only increase the robustness of the metric produced. As more data are published, more datasets can be incorporated. This approach can be used across other fungi in addition to *N. crassa*, provided there are sufficient RNA-seq data publicly available.

References

1. Shannon CE. 1997. The mathematical theory of communication. 1963. MD Comput 14:306-17

2. Vajapeyam S. 2014. Understanding Shannon's Entropy metric for Information. arXiv preprint doi:arXiv:1405.2061 [cs.IT].arXiv:1405.2061 [cs.IT]

3. Schug J, Schuller WP, Kappen C, Salbaum JM, Bucan M, Stoeckert CJ, Jr. 2005. Promoter features related to tissue specificity as measured by Shannon entropy. Genome Biol 6:R33.10.1186/gb-2005-6-4-r33

4. Fuhrman S, Cunningham MJ, Wen X, Zweiger G, Seilhamer JJ, Somogyi R. 2000. The application of shannon entropy in the identification of putative drug targets. Biosystems 55:5-14.10.1016/s0303-2647(99)00077-5

5. Langmead CJ, McClung CR, Donald BR. 2002. A maximum entropy algorithm for rhythmic analysis of genome-wide expression patterns. Proc IEEE Comput Soc Bioinform Conf 1:237-45

6. van Wieringen WN, van der Vaart AW. 2011. Statistical analysis of the cancer cell's molecular entropy using high-throughput data. Bioinformatics 27:556-63.10.1093/bioinformatics/btq704

7. Machado JAT. 2012. Shannon Entropy Analysis of the Genome Code. Mathematical Problems in Engineering 2012.10.1155/2012/132625

8. Nathaniel D. Heintzman, Gary C. Hon, R. David Hawkins, Pouya Kheradpour, Alexander Stark, Lindsey F. Harp, Zhen Ye, Leonard K. Lee, Rhona K. Stuart, Christina W. Ching, Keith A. Ching, Jessica E. Antosiewicz-Bourget, Hui Liu, Xinmin Zhang, Roland D. Green, Victor V. Lobanenkov, Ron Stewart, James A. Thomson, Gregory E. Crawford, Manolis Kellis, Ren B. 2009. Histone modifications at human enhancers reflect global cell-type-specific gene expression. Nature 459:108-112.https://doi.org/10.1038/nature07829

9. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE, Ecker JR. 2006. Genome-wide high-resolution mapping and functional analysis of DNA methylation in arabidopsis. Cell 126:1189-201.10.1016/j.cell.2006.08.003

10. Timothy R. Lezon, Jayanth R. Banavar, Marek Cieplak, Amos Maritan, Fedoroff NV. 2006. Using the principle of entropy maximization to infer genetic interaction networks from gene expression patterns. PNAS 103:19033-19038.https://doi.org/10.1073/pnas.0609152103

11. Akhter S, Bailey BA, Salamon P, Aziz RK, Edwards RA. 2013. Applying Shannon's information theory to bacterial and phage genomes and metagenomes. Sci Rep 3:1033.10.1038/srep01033

12. Borkovich KA, Alex LA, Yarden O, Freitag M, Turner GE, Read ND, Seiler S, Bell-Pedersen D, Paietta J, Plesofsky N, Plamann M, Goodrich-Tanrikulu M, Schulte U, Mannhaupt G, Nargang FE, Radford A, Selitrennikoff C, Galagan JE, Dunlap JC, Loros JJ, Catcheside D, Inoue H, Aramayo R, Polymenis M, Selker EU, Sachs MS, Marzluf GA, Paulsen I, Davis R, Ebbole DJ, Zelter A, Kalkman ER, O'Rourke R, Bowring F, Yeadon J, Ishii C, Suzuki K, Sakai W, Pratt R. 2004. Lessons from the genome sequence of Neurospora crassa: tracing the path from genomic blueprint to multicellular organism. Microbiol Mol Biol Rev 68:1-108

13. Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC. 2006. A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. Proc Natl Acad Sci U S A 103:10352-10357.10.1073/pnas.0601456103

14. Tian C, Beeson WT, Iavarone AT, Sun J, Marletta MA, Cate JH, Glass NL. 2009. Systems analysis of plant cell wall degradation by the model filamentous fungus

Neurospora crassa. Proc Natl Acad Sci U S A 106:22157-62.10.1073/pnas.0906810106

15. Aramayo R, Selker EU. 2013. Neurospora crassa, a model system for epigenetics research. Cold Spring Harb Perspect Biol 5:a017921.10.1101/cshperspect.a017921

16. Dunlap JC, Loros JJ, Colot HV, Mehra A, Belden WJ, Shi M, Hong CI, Larrondo LF, Baker CL, Chen CH, Schwerdtfeger C, Collopy PD, Gamsby JJ, Lambreghts R. 2007. A circadian clock in Neurospora: how genes and proteins cooperate to produce a sustained, entrainable, and compensated biological oscillator with a period of about a day. Cold Spring Harb Symp Quant Biol 72:57-68.10.1101/sqb.2007.72.072

17. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37:907-915.10.1038/s41587-019-0201-4

18. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Proc GPD. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-2079.10.1093/bioinformatics/btp352

19. Liao Y, Smyth GK, Shi W. 2014. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics 30:923-930.10.1093/bioinformatics/btt656

20. McCarthy DJ, Campbell KR, Lun ATL, Wills QF. 2017. Scater: pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R. Bioinformatics 33:1179-1186.10.1093/bioinformatics/btw777

21. Zhang JD, Hatje K, Sturm G, Broger C, Ebeling M, Burtin M, Terzi F, Pomposiello SI, Badi L. 2017. Detect tissue heterogeneity in gene expression data with BioQC. BMC Genomics 18:277.10.1186/s12864-017-3661-2

22. Leinonen R, Sugawara H, Shumway M, International Nucleotide Sequence Database C. 2011. The sequence read archive. Nucleic Acids Res 39:D19-21.10.1093/nar/gkq1019

23. Hurley JH, Dasgupta A, Andrews P, Crowell AM, Ringelberg C, Loros JJ, Dunlap JC. 2015. A Tool Set for the Genome-Wide Analysis of Neurospora crassa by RT-PCR. G3 (Bethesda) 5:2043-9.10.1534/g3.115.019141

24. Cusick KD, Fitzgerald LA, Pirlo RK, Cockrell AL, Petersen ER, Biffinger JC. 2014. Selection and evaluation of reference genes for expression studies with quantitative PCR in the model fungus Neurospora crassa under different

environmental conditions in continuous culture. PLoS One 9:e112706.10.1371/journal.pone.0112706

25. Benz JP, Chau BH, Zheng D, Bauer S, Glass NL, Somerville CR. 2014. A comparative systems analysis of polysaccharide-elicited responses in Neurospora crassa reveals carbon source-specific cellular adaptations. Mol Microbiol 91:275-99.10.1111/mmi.12459

26. Wu C, Yang F, Smith KM, Peterson M, Dekhang R, Zhang Y, Zucker J, Bredeweg EL, Mallappa C, Zhou X, Lyubetskaya A, Townsend JP, Galagan JE, Freitag M, Dunlap JC, Bell-Pedersen D, Sachs MS. 2014. Genome-wide characterization of light-regulated genes in Neurospora crassa. G3 (Bethesda) 4:1731-45.10.1534/g3.114.012617

27. Wang Z, Lopez-Giraldez F, Lehr N, Farre M, Common R, Trail F, Townsend JP. 2014. Global gene expression and focused knockout analysis reveals genes associated with fungal fruiting body development in Neurospora crassa. Eukaryot Cell 13:154-69.10.1128/EC.00248-13

Figures

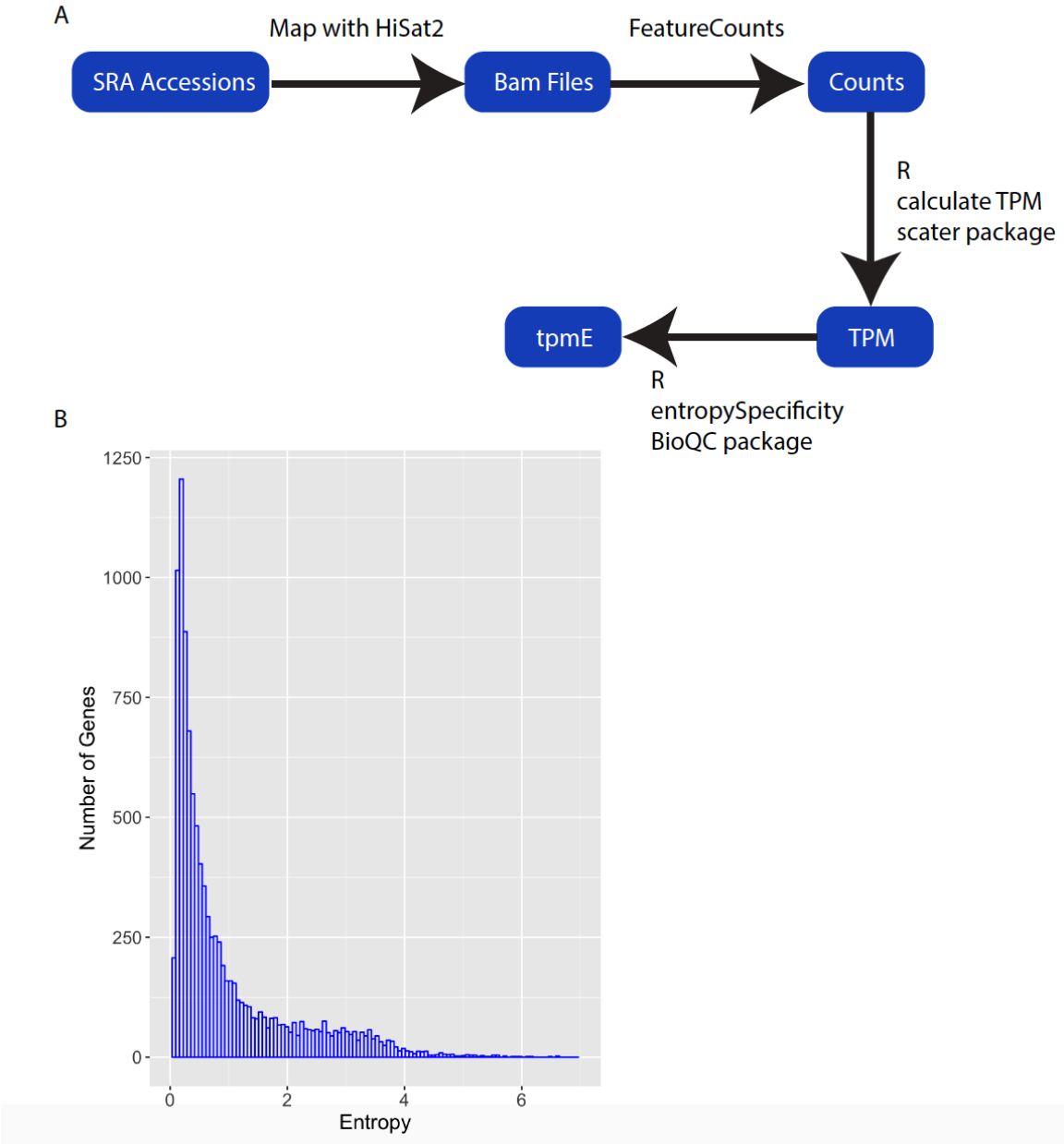


Figure 3.1: Pipeline schematic and histogram of entropy values for all genes

A) Schematic of pipeline starting at SRA accessions and endings at entropy values for each gene.

B) Histogram of entropy values for all genes. The y-axis is the number of genes found in each bin. The x-axis shows the binned entropy values.

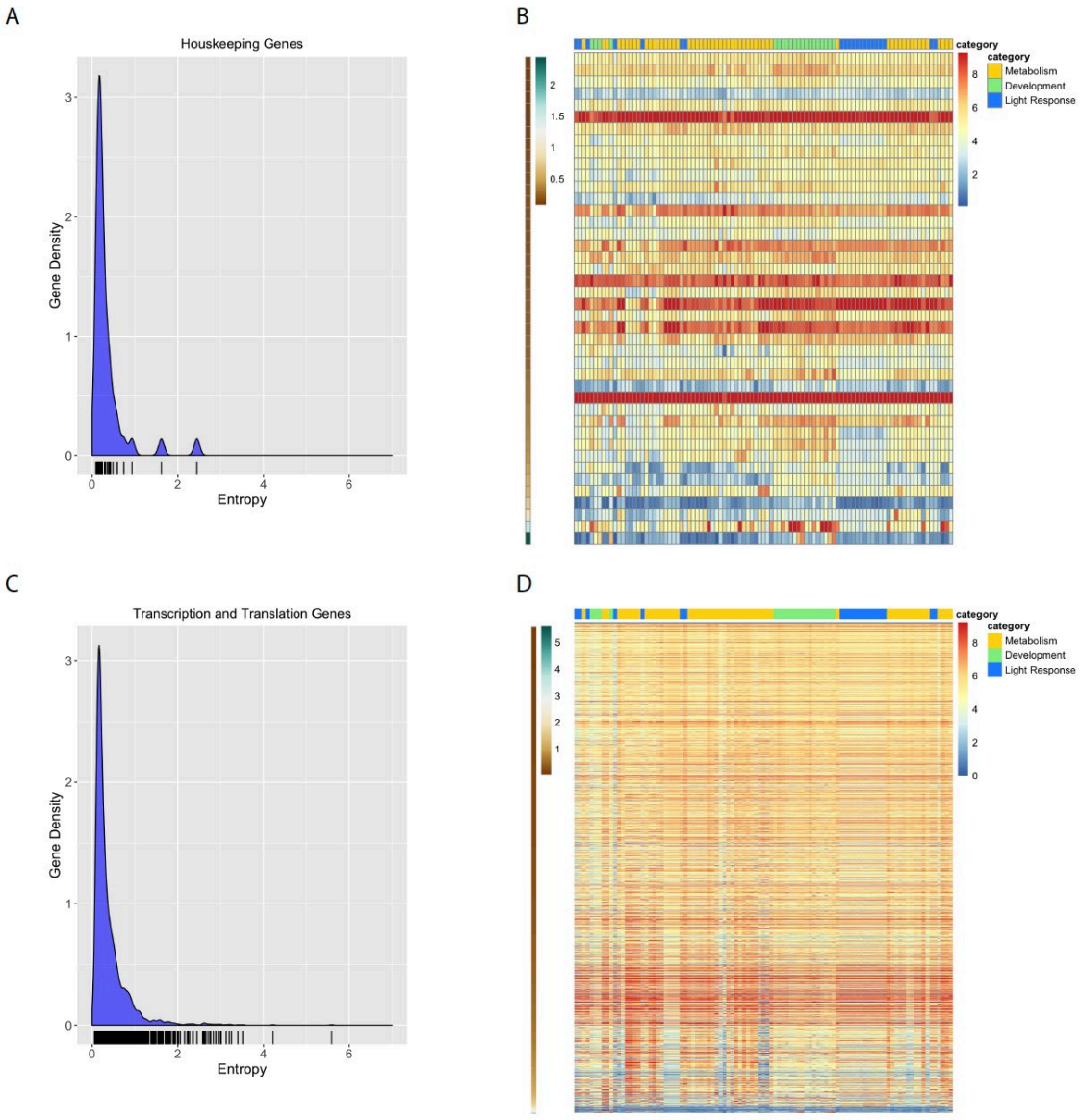


Figure 3.2: Validating entropy values with previously published housekeeping genes and genes related to transcription and translation

A) Kernel density estimation (KDE) plot of housekeeping genes. The rug plot, black lines on the bottom in the KDE plot represents the individual data points that create the estimation. The y-axis is the probability density, which is the probability for each unit (gene) on the x-axis. The total area below the KDE curve integrates to one. In short, the peaks of the curve are analogous to the values that appear most frequently in the data set.

B) Heatmap of $\log_2$ transformed TPM values for all housekeeping genes (rows) ranked by entropy (low to high). Entropy values are depicted by the brown to green heatmap on the left side where brown is low (top) and green is high (bottom). Each condition (column) has been assigned a category: Metabolism (gold), Development (green), or Light Response (blue). The categories are represented at the top of the heatmap in the three different colors.

C) Kernel density estimation plot of genes related to transcription and translation with rug plot. The rug plot, black lines on the bottom in the KDE plot represents the individual data points that create the estimation. The y-axis is the probability density, which is the probability for each unit (gene) on the x-axis. The total area below the KDE curve integrates to one. In short, the peaks of the curve are analogous to the values that appear most frequently in the data set.

D) Heatmap of $\log_2$ transformed TPM values from all transcription and translation related genes (rows) ranked by entropy (low to high). Entropy values are depicted by the brown to green heatmap on the left, where brown is low (top) and green is high (bottom). Each condition (column) has been assigned a category: Metabolism (gold), Development (green), or Light Response (blue). The categories are represented at the top of the heatmap in the three different colors.

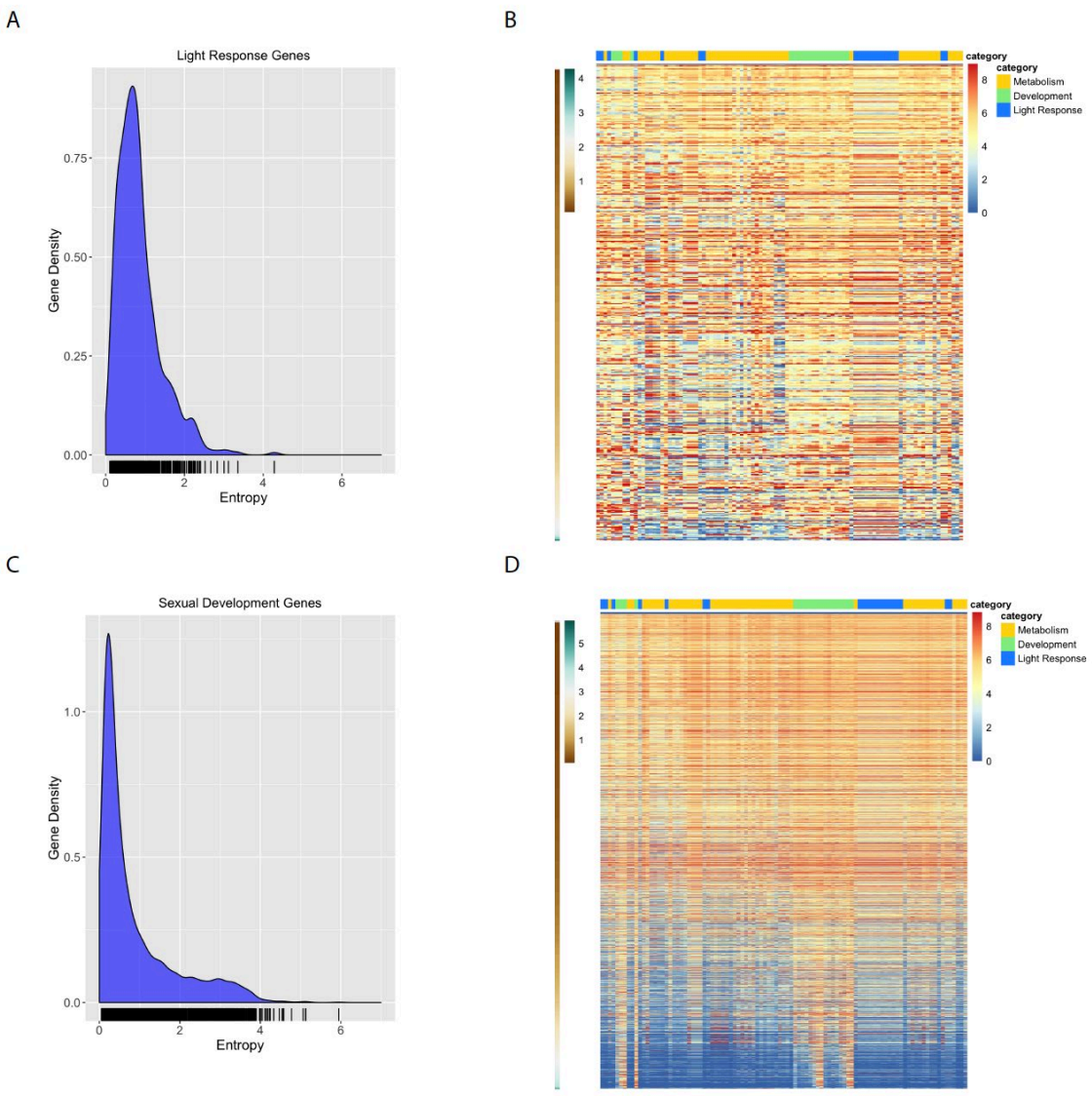


Figure 3.3: Validating entropy values with previously published light induced genes and genes induced during sexual development

A) Kernel density estimation plot of light induced genes with rug plot. The rug plot, black lines on the bottom in the KDE plot represents the individual data points that create the estimation. The y-axis is the probability density, which is the probability for each unit (gene) on the x-axis. The total area below the KDE curve integrates to one. In short, the peaks of the curve are analogous to the values that appear most frequently in the data set.

B) Heatmap of $\log_2$ TPM values for all light induced genes (rows) ranked by entropy (low to high). Entropy values are depicted by the brown to green heatmap on the left side where brown is low (top) and green is high (bottom). Each condition (column) has been assigned a category: Metabolism (gold), Development (green), or Light Response (blue). The categories are represented at the top of the heatmap in the three different colors.
C) Kernel density estimation plot of sexual development genes with rug plot. The rug plot, black lines on the bottom in the KDE plot represents the individual data points that create the estimation. The y-axis is the probability density, which is the probability for each unit (gene) on the x-axis. The total area below the KDE curve integrates to one. In short, the peaks of the curve are analogous to the values that appear most frequently in the data set.
D) Heatmap of $\log_2$ TPM values for all developmental genes (rows) ranked by entropy (low to high). Entropy values are depicted by the brown to green heatmap on the left side where brown is low (top) and green is high (bottom). Each condition (column) has been assigned a category: Metabolism (gold), Development (green), or Light Response (blue). The categories are represented at the top of the heatmap in the three different colors.

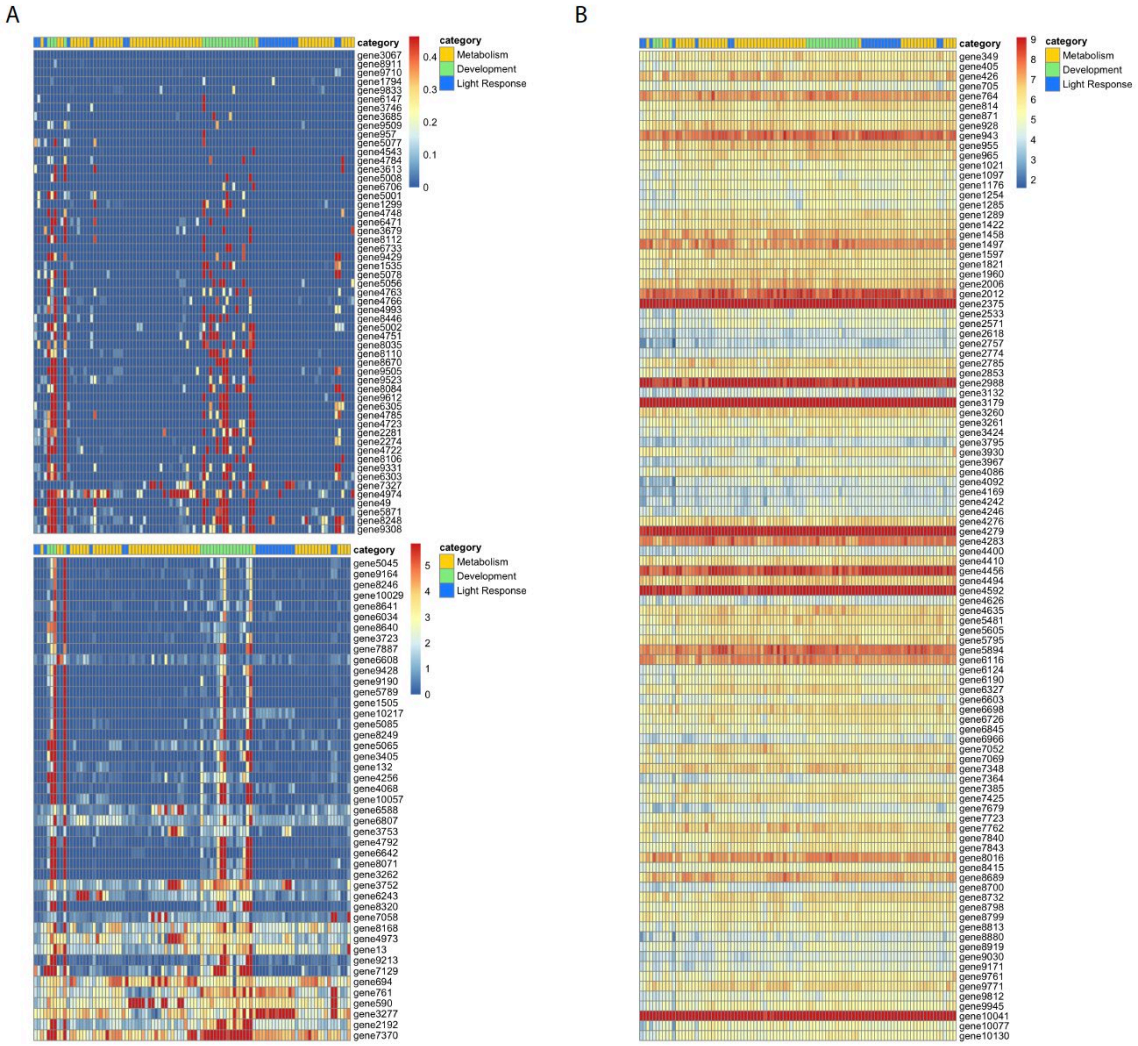


Figure 3.4: Log$_2$ TPM values for highest and lowest ranked genes

A) Log$_2$ TPM values for the 100 highest ranked genes (entropy) split by TPM values above 1 (top) and below 1 (bottom). Each row represents a genes and the gene names are listed on the right side of the heatmap. Each condition (column) has been assigned a category: Metabolism (gold), Development (green), or Light Response (blue). The categories are represented at the top of the heatmap in the three different colors.

B) Log2 TPM values for the 100 genes lowest ranked genes (entropy). Each row represents a genes and the gene names are listed on the right side of the heatmap. Each

condition (column) has been assigned a category: Metabolism (gold), Development (green), or Light Response (blue). The categories are represented at the top of the heatmap in the three different colors.

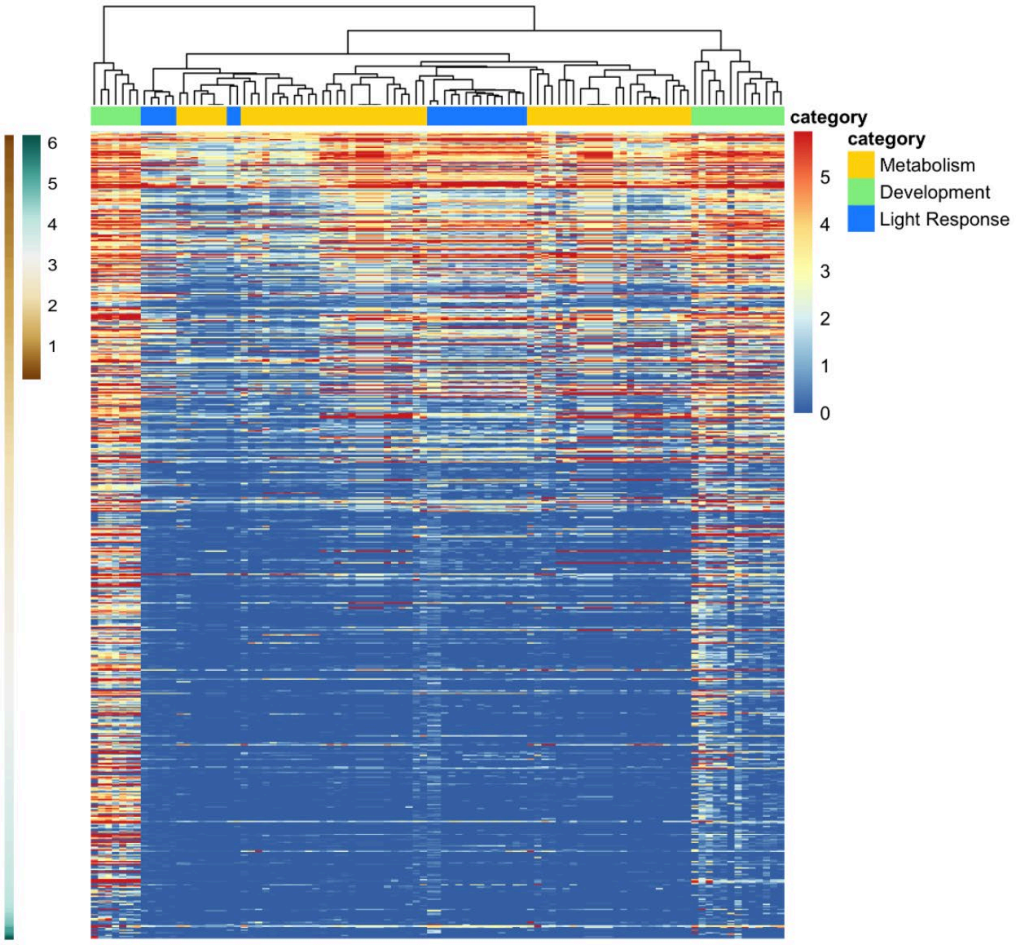


Figure 3.5: Log$_2$ TPM values for H3K27me2/3-target genes clustered by condition

Heatmap of log2 TPM values for 624 H3K27me2/3-target genes (rows) ranked by entropy (low to high) and clustered by condition (column). Entropy values are depicted by the brown to green heatmap on the left side where brown is low (top) and green is high (bottom). Each condition (column) has been assigned a category: Metabolism

(gold), Development (green), or Light Response (blue). The categories are represented at the top of the heatmap in the three different colors.

CHAPTER 4

NORMAL PATTERNS OF HISTONE H3K27 METHYLATION REQUIRE THE HISTONE VARIANT H2A.Z IN *NEUROSPORA CRASSA*[4]

[4] Abigail J. Courtney, Masayuki Kamei, Aileen R. Ferraro, Kexin Gai, Qun He, Shinji Honda, Zachary A. Lewis. 2020. *Genetics*. doi: 10.1534/genetics.120.303442. Reprinted here with permission from the publisher.

Abstract

Neurospora crassa contains a minimal Polycomb repression system, which provides rich opportunities to explore Polycomb-mediated repression across eukaryotes and enables genetic studies that can be difficult in plant and animal systems. Polycomb Repressive Complex 2 is a multi-subunit complex that deposits mono-, di-, and tri-methyl groups on lysine 27 of histone H3, and tri-methyl H3K27 is a molecular marker of transcriptionally repressed facultative heterochromatin. In mouse embryonic stem cells and multiple plant species, H2A.Z has been found to be co-localized with H3K27 methylation. H2A.Z is required for normal H3K27 methylation in these experimental systems, though the regulatory mechanisms are not well understood. We report here that *Neurospora crassa* mutants lacking H2A.Z or SWR-1, the ATP-dependent histone variant exchanger, exhibit a striking reduction in levels of H3K27 methylation. RNA-sequencing revealed downregulation of *eed,* encoding a subunit of PRC2, in an *hH2Az* mutant compared to wild type and overexpression of EED in a Δ*hH2Az*;Δ*eed* background restored most H3K27 methylation. Reduced *eed* expression leads to region-specific losses of H3K27 methylation suggesting that differential dependence on EED concentration is critical for normal H3K27 methylation at certain regions in the genome.

Introduction

In eukaryotes, DNA-dependent processes in the nucleus are regulated by chromatin-based mechanisms (1). One heavily studied group of proteins that are particularly important for maintaining stable gene repression are the Polycomb Group (PcG) proteins. In plants and animal cells, PcG proteins assemble into Polycomb Repressive Complexes 1 and 2 (PRC1 and PRC2), which play key roles in repression of

developmental genes (as reviewed in 2, 3-6). PRC2 is a multi-subunit complex that deposits mono-, di-, and tri-methyl groups on lysine 27 of histone H3, and tri-methyl H3K27 is a molecular marker of transcriptionally repressed facultative heterochromatin (7-10). PcG proteins are absent from the model yeasts, *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe,* but core PRC2 components have been identified and characterized in several fungi, including *Neurospora crassa*, *Fusarium graminearum, Cryptococcus neoformans*, *Epichloë festucae*, and *Fusarium fujikuroi* (11-17). In these fungi, PRC2 is required for repression of key fungal genes suggesting that this enzyme complex is functionally conserved between fungi, plants, and animals (13, 14, 18).

In *N. crassa,* the catalytic subunit of PRC2 is SET-7, a protein with homology to EZH1/EZH2 in humans and curly leaf (CLF), medea (MEA), or swinger (SWN) in *Arabidopsis* (9, 10, 19-25). *Neurospora* EED is essential for catalysis and is a homolog of mammalian Enhanced Ectoderm Development (EED), *Drosophila* Extra Sex Combs (Esc) and *Arabidopsis* Fertilization Independent Endosperm (FIE) (24, 26, 27). SUZ-12 is the third essential component of PRC2 in *Neurospora* and is named SUZ12 in humans, su(z)12 in *Drosophila* and embryonic flower 2 (EMF2), vernalization 2 (VER2), or fertilization independent seed 2 (FIS2) in *Arabidopsis* (24, 28). *N. crassa* CAC-3/NPF is an accessory subunit homologous to mammalian retinoblastoma binding protein 46/48 (RBAP46/68) in humans, and multicopy suppressor of IRA1-5 (MSI1-5) in *Arabidopsis* (29-31). In contrast to PRC2, PRC1 components appear to be absent from the fungal kingdom (14, 32).

The presence of a minimal Polycomb repressive system in well studied fungi such as *N. crassa* provides an opportunity to explore the diversity of Polycomb-mediated

repression across eukaryotes and enables genetic studies that can be difficult in plant and animal systems. Indeed, genetic studies have provided insights into PRC2 control in *Neurospora*. Deletion of *cac-3/npf* causes region-specific losses of H3K27me3 at telomere-proximal domains, and telomere repeat sequences are sufficient to nucleate a new domain of H3K27me3-enriched chromatin (14, 33). In constitutive heterochromatin domains, heterochromatin protein-1 (HP1) prevents accumulaton of H3K27me3 (34, 35). Thus, regulation of H3K27 methylation occurs at multiple levels. Despite recent advances, the mechanisms that regulate PRC2 in fungal systems and eukaryotes in general is poorly understood.

In addition to the core histones (H2A, H2B, H3, and H4), eukaryotes also encode non-allelic histone variants. One of the most conserved and extensively studied histone variants is H2A.Z, which is enriched proximal to transcription start sites (TSS) and in vertebrate enhancers (36-43). Functional studies of H2A.Z have linked presence of this variant in nucleosomes to gene activation, gene repression, maintaining chromatin accessibility, and a multitude of other functions (38, 44-51). Notably, H2A.Z has been implicated in the direct regulation of H3K27 methylation in mouse Embryonic Stem Cells (mESCs) and in plants (52-54). In mESCs, there is a strong correlation between the activity of PRC2, enrichment of H3K27me3, and the presence of H2A.Z (55). Co-localization of SUZ12, a subunit of PRC2, and H2A.Z has been found in mESCs at developmentally important genes, such as HOX clusters (40). In addition, H2A.Z is differentially modified its N- and C- terminal tails at bivalent domains that are “poised” for activation or repression upon differentiation (54, 56). N-terminal acetylation (acH2A.Z) or C-terminal ubiquitylation (H2A.Zub) repress or stimulate the action of

PRC2 through interactions with the transcriptional activator BRD2 or the PcG protein complex PRC1 (54). It is important to note that functional studies of H2A.Z are challenging because this histone variant is essential for viability in most organisms, including *Drosophila*, *Tetrahymena*, mouse, and *Xenopus* (57-62).

In *Arabidopsis thaliana*, a genetic interaction between PICKLE (PKL), a chromatin remodeler which promotes H3K27me3, and PIE-1 (homolog to SWR-1), the remodeler which deposits H2A.Z, was recently reported (53). PKL has been found by ChIP-seq at loci enriched for H3K27me3 and is proposed to determine levels of H3K27me3 at repressed genes in *Arabidopsis* (63). In rice callus and seedlings, H2A.Z is found at the 5' and 3' ends of genes that are highly expressed. In repressed genes, H2A.Z is found along the gene body, and this pattern closely mimics the presence of H3K27me3 (52). This is a notable difference between plants and other eukaryotes.

We investigated the relationship between H2A.Z and PRC2 in the filamentous ascomycete *Neurospora crassa* and report that H2A.Z is required for normal enrichment of H3K27me2/3 across the genome. Our findings show that loss of H2A.Z leads to region-specific losses of H3K27me2/3 in *N. crassa*. Expression levels of *eed,* encoding a PRC2 subunit, are reduced in the absence of H2A.Z and ectopic expression of *eed* can restore H3K27me2/3 in an H2A.Z-deficient strain. Together, these data suggest that H2A.Z regulates facultative heterochromatin through transcriptional regulation of the PRC2 component EED and points to differential requirements for EED at discrete PRC2-target domains.

Materials and Methods

Strains and growth media: Strains used in this study are listed in (Table S1). Strains were grown at 32˚C in Vogel's Minimal Medium (VMM) with 1.5% sucrose or glucose for DNA based protocols, and RNA based protocols, respectively (64). Liquid cultures were shaken at 180 rpm. Crosses were performed on Synthetic Crossing (SC) medium in the dark at room temperature (64). Ascospores were collected 14 days after fertilization. To isolate cross progeny, spores were spread on solid VMM plates containing FGS (1X Vogel's salts, 2% sorbose, 0.1% glucose, 0.1% fructose, and 1.5% agar) and incubated at 65˚C for 1 hour as previously described (64), after which spores were picked using a sterile inoculating needle and transferred to agar slants with appropriate medium (typically VMM). To test for sensitivity to DNA damaging agents, 5 µL of a conidial suspension was spotted on VMM containing FGS (1X Vogel's salts, 2% sorbose, 0.1% glucose, 0.1% fructose, and 1.5% agar) plates containing concentrations of methyl methanesulfonate (Sigma Aldrich cat. # 129925-5g) between 0.010% and 0.03% (w/v).

To construct the N-terminal FLAG-tagged *eed* allele, we amplified the *eed* region with primers, MK #51: GGCGGAGGCGGCGCGATGCAAATTTGTCGGGACCG and MK #52: TTAATTAATGGCGCGTTACTTCCCCCACCGCTGAA (Table S7), from wild type genomic DNA (FGSC 4200). The amplified fragment was cloned into the *Asc*I site of pBM61::CCGp-N-3xFLAG (65) by InFusion cloning (Takara, cat. # 639648). The new plasmid was then digested with *Dra*I and transformed into a *his-3;mus-52::bar* strain. Primary transformants were selected on VMM plates, and then back-crossed to wild type to isolate homokaryons (*his-3::Pccg-1-3xflag-eed*). We next crossed the homokaryon (*his-3::Pccg-1-3xflag-eed)* to Δ*eed::hph* (FGSC 14852) to obtain Δ*eed;his-*

3::Pccg-1-3xflag-eed. 3xFLAG-EED expression and deletion of *eed* deletion were confirmed by western blots probed with anti-FLAG antibody (Sigma Aldrich, cat. # F1804) and genotyped by PCR with primers, LL #155: TCGCCTCGCTCCAGTCAATGACC and LL #466: TGTGGGCGATTTGAGCGTGC, respectively. The Δ*eed;his-3::Pccg-1-3xflag-eed* strain was then crossed to the Δ*hH2Az::hph* (FGSC 12088) strain to obtain Δ*hH2Az;*Δ*eed;his-3::Pccg-1-3xflag-eed*. 3xFLAG-EED expression and deletion of *eed* were confirmed by western blots with anti-FLAG antibody (Sigma Aldrich, cat. # F1804) and genotyping with *eed* deletion primers (see above). Deletion of *hH2Az* was confirmed by PCR with primers AC #24: GAACAAGCCGATTGCTGTCC and AC #23: TGTATAGAACGCTGCCAAGGA.

For the H2AZ-GFP gene replacement construct, a 1-kb segment including the end of the *hH2Az* coding region was amplified by PCR with primers #1577: CGGAAAGGGCAAGTCGTCTG and #1578: CCTCCGCCTCCGCCTCCGCCGCCTCCGCCAGCCTCCTGAGCCTTGGCCT and a 500-bp segment of the 3' flanking region was amplified with primers #1579: TGCTATACGAAGTTATGGATCCGAGCTCGCTGCACCGAAAAACTCGACG and #1580: GTGACGAGGGGAGATTGCTC. The cassettes containing the GFP segment and the *hph* gene were amplified using M13 forward and reverse primers from *pGFP::hph::loxP* (65). The three fragments were mixed and then assembled by overlapping PCR with primers #1577 and #1580 above. The cassette was transformed into the Δ*mus-52* strain (FGSC 15968) by electroporation.

Transformation and complementation assays: Transformations were performed as previously described (66). To carry out ectopic complementation of the Δ*hH2Az::hph*

strain, two linear gene fragments were electroporated into the mutant strain. Specifically, the *bar* (confers Basta resistance) was amplified with primers LL #148 CCGTCGACAGAAGATGATATTGAAGGAGC and LL #149 AATTAACCCTCACTAAAGGGAACAAAAGC (67) and the wild-type *hH2Az* gene fragment including its native promoter (genomic coordinate 1390154-1393398 of GCA_000182925.2 assembly accession) was amplified with primers AC #27 CCCAATCCTAGAATCCCGTCG and AC #21 TAAAAGAGCTGCTGTCGCACG, and fragments were co-integrated into the Δ*hH2Az::hph* strain, followed by selection of transformants on Basta-containing plates (VMM with 2% sorbose, 0.1% glucose, 0.1% fructose, 1.5% agar, and 200 ug/mL Basta). Transformants were transferred to agar slants and then screened by PCR, and Southern blots with the North2South Biotin Random Prime Labeling and Detection Kit (ThermoFisher cat. #17175) and the wild-type *hH2Az* gene fragment generated from primers AC#27: CCCAATCCTAGAATCCCGTCG and AC#21: TAAAAGAGCTGCTGTCGCACG was used as a probe. Genomic DNA was extracted from wild type, Δ*hH2Az*, and the two complemented strains (ACt9-3, ACt12-1). 500ng of DNA from all strains were subjected to double restriction digests with Eco-RI-HF (NEB cat # R3101S) and Eco-RV-HF (NEB cat # R3195S) in CutSmart Buffer. Digests were incubated at 37˚C overnight and then subjected to heat inactivation for 20 minutes at 65˚C. Digests were run on a 0.7% agarose gel and transferred overnight to a nylon membrane which was then UV crosslinked. Hybridization and detection were followed from manufacturer's instructions.

Race tube assay: Race tubes were prepared with 15 mL of VMM plus 1.5% sucrose and 1.5% agar. Strains were grown on VMM plates with 1.5% sucrose and 1.5% agar for 16

hours before using a 6mm cork borer to extract mycelial agar plugs from the edge of growing hyphae. This plug was used for inoculating each tube at one end. Strains were inoculated in triplicate. Measurements were taken at 9, 23, 47 and 60 hours to determine linear growth rates.

Protein extraction and western blotting: Strains were grown at 32˚C shaken in 18x150mm glass test tubes at 180rpm in 5 mL VMM with 1.5% sucrose. After 16 hours, tissue was harvested using filtration, washed once in phosphate buffered saline (PBS), and suspended in 1 mL of ice-cold protein extraction buffer (50mM HEPES pH 7.5, 150mM NaCl, 0.02% NP-40, 1mM EDTA, 1mM phenylmethylsulfonyl fluoride [PMSF; Sigma, P7626], one tablet Roche cOmplete mini EDTA-free Protease Inhibitor Cocktail [Roche, cat. # 11836170001]). Tissue was subjected to sonication by Diagenode Bioruptor UCD-200 to deliver 22.5 30 second pulses at 4˚C. After two rounds of centrifugation at 13,200 rpm for 10 minutes, supernatant was mixed with 2x Laemmli buffer and boiled for 5 minutes. Samples for FLAG and H2A.Z western blots were separated by SDS-polyacrylamide gel electrophoresis (SDS-PAGE) and transferred to polyvinylidene difluoride (PVDF) membranes in Tris-Glycine transfer buffer (25mM Tris, 200mM glycine) or CAPS transfer buffer (10mM, pH 11) containing 20% methanol at constant 100V for 1 hour at 4˚C respectively. Membranes were blocked with Tris-buffered saline (TBS; 10mM Tris, pH 7.5, 150mM NaCl) including 3% milk powder or Phosphate-buffered saline including Tween 20 (PBS-T; 137mM NaCl, 2.7mM KCl, 10mM Na_2HPO_4, KH_2PO_4, 0.01% Tween-20) includeing 30% milk powder for 1 hour and incubated overnight with anti-FLAG antibody (Sigma Aldrich, cat. # F1804) in TBS plus 3% milk or H2A.Z antibody in PBS-T plus 3% milk. Detection was performed with

horseradish peroxidase-conjugated secondary antibodies and SuperSignal West Femto chemiluminescent substrate (ThermoFisher, cat. # 34094).

Chromatin immunoprecipitation (ChIP): To carry out ChIP, conidia were inoculated in 5 mL of liquid VMM plus 1.5% sucrose and grown for 18 hours for wild type and other strains with typical growth rates. Slow growing Δ*hH2Az::hph* strains were grown for 24 hours to isolate cultures at a similar developmental stage. ChIP was performed as described previously (68-70). In brief, mycelia were harvested using filtration and were washed once in PBS prior to cross-linking for 10 minutes in PBS containing 1% formaldehyde on a rotating platform at room temperature. After 10 minutes, the reaction was quenched using 125mM glycine and placed back on the rotating platform for five minutes. Mycelia were harvested again using filtration, washed once with PBS, then resuspended in 600 µl of ChIP lysis buffer (50mM HEPES, pH 7.5, 140mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, one tablet Roche cOmplete mini EDTA-free Protease Inhibitor Cocktail (Roche, cat. # 11836170001) in 15 mL conical tubes. Chromatin was sheared by sonication after lysing cell walls with the QSONICA Misonix S-4000 ultrasonic processor (amplitude 10, 30 second processing, one second on, one second off), using the Diagenode Bioruptor UCD-200 (Intensity level: Medium, three rounds of 15 minutes (30 seconds on, 30 seconds off) to deliver 22.5 30 second pulses at 4˚C. Water temperature was kept at a constant 4˚C by using a Biorad cooling module (cat. # 170-3654) with variable speed pump to circulate 4˚C water while processing samples. Lysates were centrifuged at 13,000 rpm in an Eppendorf 5415D microcentrifuge for five minutes at 4˚C. For ChIP reactions with antibodies against *N. crassa* H2A.Z, 1 µl, 2.5 µl, or 5 µl of antibody was used (antibody supplied by Dr. Qun

He, China Agricultural University (71, 72)). For detection of H3K27 di- and tri-methylation (H3K27me2me3; Active Motif 39535), and GFP-tagged H2A.Z (GFP; Rockland 600-301-215) 1 µl of the relevant antibody was used. Protein A/G beads (20 µl) (Santa Cruz, cat. # sc-2003) were added to each sample. Following overnight incubation, beads were washed twice with 1 mL lysis buffer without protease inhibitors, once with lysis buffer containing 500mM NaCl, once with lysis buffer containing 50mM LiCl, and finally with TE (10mM Tris-HCl, 1mM EDTA). Bound chromatin was eluted in TES (50mM Tris pH 8.0, 10mM EDTA, 10% SDS) at 65˚C for 10 minutes. Chromatin was de-crosslinked overnight at 65˚C. The DNA was treated with RNase A for two hours at 50˚C, then with proteinase K for two hours at 50˚C and extracted using phenol-chloroform-isoamyl alcohol (25:24:1) followed by chloroform extraction. After final chloroform extraction, DNA was precipitated using ethanol precipitation with two volumes of ethanol, 1/10 volume 3M NaOAc, pH 5.2, and 0.025mg/mL glycogen overnight at -20˚C. DNA pellets were washed with 70% ethanol and resuspended in TE buffer. Samples were then prepared for Illumina sequencing.

RNA extraction: Conidia were inoculated into 100 x 15mm plates containing 25 mL of VMM + 1.5% glucose and grown for 36-48 hours to generate mycelial mats. Using a 9mm cork borer, 5-7 disks were cut out of the mycelial mat and transferred to 125 mL flasks with 50 mL of VMM + 1.5% glucose and allowed to grow for 12 hours at 29˚C in constant light while agitating at ~90-100 rpm. Disks were harvested using filtration and flash frozen with liquid nitrogen. Frozen tissue was transferred to 1.5 mL RNase-free tubes with 100 µl sterile RNase-free glass beads and vortexed to lyse tissue in phenol:chloroform (5:1) pH 4.5. Three sequential acid phenol:chloroform extractions

were performed followed by ethanol precipitation using two volumes of ethanol and 1/10 volume of 3M NaOAc pH 5.2, incubated overnight at -20˚C. Samples were centrifuged at 13,2000 rpm in 4˚C for 30 minutes and pellets were then washed in RNase-free 70% ethanol, and resuspended in RNase-free water. Samples were quantified using the Invitrogen Qubit 2.0 fluorometer (cat. # Q32866) and RNA quality was checked on a denaturing agarose gel. After quality was verified 10 µg of RNA for each sample was subjected to Turbo DNase treatment (Invitrogen, cat. # AM2238) at 37˚C for 30 minutes and then another acid phenol:chloroform extraction was performed to inactivate enzyme and purify the RNA. Samples were subjected to another ethanol precipitation as described above, this time with the addition of 1 µL of RNase-free glycogen (5 mg/mL). Samples were centrifuged at 13,200 rpm in 4˚C for 30 minutes and the pellets were washed with RNase-free 70% ethanol, then resuspended in RNase-free water. Quality and quantity were again checked with denaturing gel and with the Invitrogen Qubit 2.0 fluorometer. Samples were then prepared for Illumina sequencing.

ChIP library preparation: Libraries were constructed as described (68-70). In brief, the NEBNext Ultra II End Repair/dA-tailing Module (cat. # E7546S), NEBNext Ultra II Ligation Module (cat. # E7546) were used to clean and A-tail DNA after which Illumina adapters were ligated. The ligation products were amplified to generate dual-indexed libraries using NEBNext Ultra II Q5 Hot Start HiFi PCR Master Mix (cat. # M0543S). Size selection with magnetic beads was performed after the adapter ligation and PCR steps with Sera-Mag SpeedBeads (cat. # 65152105050250) suspended in a solution of (20mM PEG 8000, 1mM NaCl, 10mM Tris-HCl, 1mM EDTA) (73).

RNA library preparation: Libraries were prepared according to the Illumina TruSeq mRNA stranded Library Kit (cat. # RS-122-2101). In brief, mRNA selection via polyA tails was performed using RNA purification beads and washed with bead washing buffer. Fragmentation and cleanup were performed enzymatically using the Fragment, Prime, Finish Mix and incubated at 94˚C for eight minutes. First strand synthesis using the SuperScript II RT enzyme and First Strand Synthesis Act D Mix was incubated as described and second strand synthesis used the Second Strand Marking Mix with resuspension buffer was incubated for one hour to generate cDNA. The final steps in the library preparation are the same as the above ChIP-seq library preparation with exception of two extra bead cleanup steps: one prior to A-tailing and adapter ligation, two after adapter ligation.

Libraries were pooled and sequenced on a NextSeq500 instrument at the Georgia Genomics and Bioinformatics Core to generate single or paired-end reads.

Data Analysis

For ChIP-seq data, short reads (<20 bp) and adaptor sequences were removed using TrimGalore (version 0.4.4), cutadapt version 1.14 (74), and Python 2.7.8, with fastqc command (version 0.11.3). Trimmed Illumina reads were aligned to the current *N. crassa* NC12 genome assembly available from NCBI (accession # GCA_000182925.2) using the BWA (version 0.7.15) (75), mem algorithm, which randomly assign multi-mapped reads to a single location. Files were sorted and indexed using SAMtools (version 1.9) (76). To plot the relative distribution of mapped reads, read counts were determined for each 50 bp window across the genome using DeepTools to generate bigwigs (version 3.3.1) (77) with the parameters –normalizeUsing CPM (counts per million) and data were displayed

using the Integrated Genome Viewer (78). The Hypergeometric Optimization of Motif EnRichment (HOMER) software package (version 4.8) (79) was used to identify H3K27me3 peaks in wild type and Δ*hH2Az* against input using "findPeaks.pl" with the following parameters: -style histone. Bedtools (version 2.27.1) "intersect" (version 2.26.0) was used to determine the number of peaks that intersect with other peak or gff files, for determining PRC2-target genes the parameter of -f 0.70 was used. Heatmaps, Spearman correlation matrix (Figure S5) and line plots were constructed with DeepTools (version 3.3.1) (77).

For RNA-seq data, short reads (<20 bp) and adaptor sequences were removed using TrimGalore (version 0.4.4), cutadapt version 1.14 (74), and Python 2.7.8, with fastqc command (version 0.11.3). Trimmed Illumina single-end reads were mapped to the current *N. crassa* NC12 genome assembly using the Hierarchical Indexing for Spliced Alignment of Transcripts 2 (HISAT2: version 2.1.0) (80) with parameters –RNA-strandness R then sorted and indexed using SAMtools (version 1.9) (76). FeatureCounts from Subread (version 1.6.2) (81) was used to generate gene level counts for all RNA bam files. Raw counts were imported into R and differential gene expression analysis was conducted using Bioconductor: DeSeq2 (82). Volcano plot and box plots were generated in R using DeSeq2 and ggplot2 (83).

Data Deposition: Raw sequence data associated with this paper are available through the NCBI GEO database (accession # GSE146611). Supplementary data have been uploaded to figshare.

Results

Normal patterns of H3K27me2/3 enrichment require the presence of H2A.Z or SWR-1

Normal H3K27me2/3 patterns in plants and in mESCs depend on the histone variant H2A.Z (40, 53), but the underlying mechanism is ill-understood. To determine if H2A.Z also plays a role in Polycomb Group repression in *N. crassa,* we performed ChIP-seq to examine H3K27me2/3 enrichment in an H2A.Z deletion strain (Δ*hH2Az::hph, hereafter* Δ*hH2Az)* and compared this to wild type and Δ*set-7*. Inspection of the data in the IGV genome browser (78) revealed that the Δ*hH2Az* mutant displayed a significant reduction in H3K27me2/3 (Figure 4.1A). This reduction was apparent on all chromosomes (Figure S1) but H3K27me2/3 was not completely abolished, as observed in the Δ*set-7* strain, which lacks the catalytic subunit of PRC2 (Figure 4.1A). To quantify the change in H3K27me2/3 patterns, we called peaks of H3K27me2/3 enrichment using Hypergeometric Optimization of Motif EnRichment (HOMER; version 4.8) (79). We identified 325 peaks of H3K27me2/3 in wild type, hereafter referred to as PRC2-target domains (Table S2). Consistent with previous studies, these peaks comprised ~6% of the *N. crassa* genome (14, 34). These regions are typically larger than single genes, ranging in size from 500 bp to 108 kb, with an average size of 7.7 kb. We next plotted H3K27me2/3 levels across the 5' end of all 325 domains for wild type and Δ*hH2Az* (Figure 4.1B). Inspection of heatmaps and the genome browser revealed that H3K27me2/3 levels were reduced in many, but not all PRC2-target domains in Δ*hH2Az*. Using HOMER software to identify PRC2-target domains in Δ*hH2Az* revealed 239 peaks (Table S3). These were slightly smaller, with an average size of 5.5 kb, and comprised

only 3% of the *N. crassa* genome. To determine if the peaks observed in the Δ*hH2Az* strain are in wild type locations we only compared peaks from assembled contigs. Using bedtools intersect we found that all peaks in Δ*hH2Az* overlap with wild type peaks, indicating that Δ*hH2Az* exhibits significant loss of H3K27me2/3 from normal domains but does not gain H3K27me2/3 in new locations (Table S4).

Since H2A.Z is required for maintaining genome stability in yeast and animals, our findings raised the possibility that a second site mutation could be responsible for the observed phenotype (47, 48, 84, 85). To confirm that loss of H3K27me2/3 was due to the absence of H2A.Z, we first backcrossed the original deletion strain (FGSC 12088) to wild type (86). Four independent Δ*hH2Az* progeny all displayed similar reduction in H3K27me2/3 levels (Figure S2). In addition, the backcrossed Δ*hH2Az* strain displayed slow and variable growth (Figure S3) and was hypersensitive to the DNA damaging agent MMS. This is consistent with previous studies that have demonstrated poor growth of Δ*hH2Az* in *Saccharomyces cerevisiae* and in *N. crassa* (71, 87).

We next introduced a wild type copy of the *hH2Az* gene with its native promoter into Δ*hH2Az* (Figure S4A, B). This complemented defects in growth and MMS-sensitivity, and fully restored H3K27 methylation, suggesting loss of H2A.Z was responsible for all observed phenotypes in the deletion mutant (Figure 4.1C and 4.1D, S3). Using HOMER software to identify PRC2-target domains in the complemented strains (n=2) revealed an average of 449 peaks with an average size of 6.6 kb comprising ~7% of the *N. crassa* genome. To determine if the peaks observed in the complemented strains are in wild type locations, we only compared peaks from assembled contigs. Using bedtools intersect we found that most peaks in overlap with wild type peaks. The

six peaks that were reported as not intersecting were viewable in the genome browser in the same location as wild type, but of a slightly different amplitude which affects peak calling and is likely due to ChIP efficiency from multiple different experiments. Because a specific chromatin remodeling complex, SWR1, exchanges H2A.Z for H2A in plants, yeast and animals (88-92), we next examined H3K27me2/3 in a deletion strain lacking the *N. crassa* homolog of the SWR1 ATPase (Δ*swr-1* also known as Δ*crf1-1)*(93). The *swr-1* mutant displayed a similar reduction in H3K27me2/3 (Figure 4.1C and 4.1D). Together, these data demonstrate that H2A.Z is required for normal H3K27me2/3 in *N. crassa*.

Deletion of *hH2Az* results in region-specific loss of H3K27me2/3

Visual inspection of the ChIP-seq data revealed losses of H3K27me2/3 from PRC2-target domains located at internal (i.e., non-subtelomeric regions >200kb from the telomere repeats) chromosome sites, but not at telomere-proximal sites (i.e., <200kb from the telomere repeats) (Figure 4.2A). To quantify this, we inspected ChIP-seq results for H3K27me2/3 for both classes and found retention of H3K27 methylation in telomere-proximal regions with progressive loss in domains farther from chromosome ends. Previously published work showed that a *cac-3/npf* deficient strain has H3K27me2/3 loss which was primarily observed in the telomere-proximal regions (14); *cac-3/npf* encodes an accessory subunit of PRC2 in *N. crassa* homologous to the conserved PRC2 components Msl1-5, NURF55, Rpbp46/48, found in plants, *Drosophila,* and humans, respectively.

To better visualize which regions of the genome in the Δ*cac-3/npf* or Δ*hH2Az* strains lose enrichment of H3K27me2/3, we again divided all 325 H3K27me2/3 peaks in

the wild-type strain into telomere-proximal sites (123 peaks, average size 8,261 bp) (Figure 4.2B, top) and internal sites (186 peaks, average size 7,509 bp) (Figure 4.2B, bottom). The loss was again most dramatic at the internal regions in the *hH2Az* deletion strain, where most PRC2-target domains showed significant reduction of H3K27me2/3 levels. Specifically, we found that telomere-proximal regions show normal levels of H3K27me2/3.

Previous work demonstrated that the placement of repetitive telomere repeat sequences (5'-TTAGGG-3') in a euchromatic locus can induce *de novo* H3K27 methylation across large regions (33). Together, these data demonstrate that the absence of H2A.Z is more detrimental for the establishment and/or maintenance of internal domains of H3K27me2/3 in *N. crassa*.

***Neurospora* H2A.Z localizes to promoter regions but not to PRC2-target domains**

We next asked if H2A.Z co-localizes with H3K27 methylation, as has been reported for plants and mESCs (40, 52, 53). We used a strain expressing a C-terminal H2A.Z-GFP fusion protein to perform ChIP-seq with antibodies against H3K27me2/3 and GFP. We confirmed the H2A.Z-GFP was functional by testing the growth rate with race tubes, as disruption of H2A.Z function leads to significantly impaired growth in *Neurospora* (Figures S3 and S5).Visual inspection of the enrichment profiles in a genome browser revealed a mostly mutually exclusive localization pattern (Figure 4.3A). There are some small H2A.Z peaks in PRC2-target domains, such as in Figure 4.3A; however, these were rare (Figure 4.3B). The genomic locations with the highest enrichment for H2A.Z-GFP are the regions immediately before and after the TSS of most genes, with low enrichment in gene bodies and 3' ends (Figure 4.3C). On average we find little

enrichment of H2A.Z-GFP in the promoters and gene bodies of H3K27me2/3 enriched genes or at the center of H3K27me2/3 peaks, confirming that H3K27me2/3 and H2A.Z are largely mutually exclusive (Figure 4.3D and 4.3E).

To validate H2A.Z enrichment, we also performed ChIP-seq on wild type using an antibody raised against the native *N. crassa* H2A.Z protein (71). We performed a western blot using the H2A.Z antibody on both the chromatin and soluble fractions of wild type, H2A.Z-GFP, and Δ*hH2Az* and identify a 15 kD band in the chromatin fraction only in wild type confirming our ability to use this antibody to recognize H2A.Z (Figure S6C). These H2A.Z ChIP-seq experiments show the same localization as the H2A.Z-GFP ChIP-seq experiments (Figure S6A and S6B). Using HOMER we called peaks and found that there were far fewer peaks identified, which is likely due to the higher signal-to-noise ratio when using the H2A.Z antibody. Overall, 695 peaks were identified as compared to 3,104 in the H2A.Z-GFP strain; however, when viewed in the genome browser the peaks clearly overlap (Figure S6A) and when we intersected the peaks to determine if they were in the same location we find that 80% of the peaks were in common with the GFP peaks (Figure S6D). The localization of H2A.Z-GFP at the TSS of 5,704 genes (over half of all genes) is similar to findings in multiple other organisms (36-43).

H2A.Z is crucial for proper regulation of 11% of the genes in *N. crassa*, including *eed*

Previous studies have implicated H2A.Z in multiple roles related to transcription including gene activation and repression (40, 48, 51, 54, 94, 95). We therefore asked if H2A.Z regulates H3K27me2/3 by regulating expression of one or more PRC2

components. We performed RNA sequencing of wild type, Δ*hH2Az*, Δ*set-7*, and the double mutant Δ*hH2Az;*Δ*set-7* to determine which genes exhibit differential expression in the absence of H2A.Z. Deletion of histone variant H2A.Z causes both positive and negative mis-regulation of a large number of genes (Figure 4.4A). After Benjamini-Hochberg correction (96), there are 1,066 genes with differential transcription (adjusted *p* value < 0.05 and fold change >= 1.5). Of these 1,066 genes, there are similar numbers of genes up- and downregulated in the absence of H2A.Z (559 genes upregulated and 507 downregulated) (Figure 4.4A, Table S6).

H2A.Z distribution flanks the TSS throughout much of the genome, and we wanted to know if this distribution was different for the transcripts that were mis-regulated in the Δ*hH2Az* strain. There is higher enrichment at the TSS for transcripts that get downregulated in the absence of H2A.Z than there is for transcripts that are upregulated in the absence of H2A.Z (Figure 4.4C). These data suggest that in *Neurospora* H2A.Z may be more important for activation of genes than for repression.

We looked more closely at the PRC2-target genes and their differential expression status. Using the previously generated domains of internal and telomere-proximal H3K27me2/3 domains, we found 506 internal and 262 telomere-proximal genes. Out of these, 78 PRC2-target genes are mis-regulated (FDR corrected *p* value < 0.05 and fold change >= 1.5) in the Δ*hH2Az* strain. 31 are located in internal domains and 41 are located in telomere-proximal domains. Due to the stringency of the intersections between domains and genes, there are six that are denoted as other (six genes are not covered by 70% or more H3K27me2/3). There are approximately even numbers of genes that are up (46) and downregulated (32); however, we find a large number of the internal genes are

upregulated (27) in comparison to the telomere-proximal genes that are almost evenly up and downregulated (Figure 4D).

We next examined expression levels of genes encoding individual PRC2 components (Figure 4.4B). We found that expression of *eed* is significantly reduced in Δ*hH2Az* by more than 9-fold (FDR-corrected *p* value = 8.96 x 10^{-07}), whereas *cac-3/npf*, *suz-12,* and *set-7* were expressed at similar levels in both wild type and Δ*hH2Az* (Figures 4.4A and 4.4B).

The *eed* gene showed the most dramatic change in expression compared to wild type in either Δ*hH2Az* or Δ*hH2Az;*Δ*set-7*, but is expressed normally in the single mutant Δ*set-7*. This indicated that deletion of H2A.Z is likely responsible for its downregulation. As an essential component of PRC2, EED is required for catalytic activity. EED is also important for recognition of the H3K27me2/3 mark and has been implicated in maintenance and/or spreading of H3K27me3 from nucleation sites (97, 98). Since H2A.Z is localized proximal to the promoters of a little over half the genes (5,704) in the *N. crassa* genome, we examined the H2A.Z localization at the *eed* gene. There is a large peak of H2A.Z enrichment at the promoter of *eed* (Figure 4.4B), which appears to be crucial for normal *eed* expression. Promoters of other PRC2 components are also enriched for H2A.Z, but apparently are not dependent on H2A.Z for their expression. Together, these data suggest that H2A.Z is required for the proper expression of *eed*.

Overexpression of EED rescues H3K27 methylation levels in the absence of H2A.Z

To determine if downregulation of *eed* is responsible for the depletion of H3K27me2/3 observed in Δ*hH2Az*, we constructed a strain which lacks both *eed* and *hH2Az,* and we introduced N-terminal tagged 3x*flag-eed* into the *his-3* locus driven by

the strong constitutive *clock controlled gene-1/glucose-repressible gene-1* (*ccg-1/grg-1*) promoter (*his-3::Pccg1-3xflag-eed*). We calculated expected expression levels of this construct using native *ccg-1* levels observed in our RNA-seq experiment, and we expect *eed* to be expressed at approximately 100 times the native level. To confirm this construct was being expressed at the same level in both the Δ*eed* and Δ*eed;*Δ*hH2Az* backgrounds, we performed an anti-FLAG western blot (Figure S4C). Our results confirm that the deletion of H2A.Z does not alter 3xFLAG-EED expression driven by the *ccg-1* promoter. After performing H3K27me2/3 ChIP-seq in this strain, we find that the majority of H3K27me2/3 peaks are recovered in the genome (Figure 4.5A), but the growth phenotype of the Δ*eed;*Δ*hH2Az* strain is only partially rescued (Figure S4D). Strains with deletions of either *set-7* or *eed*, which lack any H3K27me2/3, and exhibit a wild type like growth rate (34, 35). Therefore, we conclude that the defective growth of Δ*hH2Az* strains is likely due other functions of H2A.Z and not the loss of H3K27me2/3. There are some qualitative differences in peak shape and not all peaks are fully restored (Figure 4.5B), which could indicate that H2A.Z contributes to normal H3K27me2/3 via additional mechanisms. To quantify the restoration of peaks in the overexpression strain, we called peaks of H3K27me2/3 using HOMER. We identified a total of 445 peaks, which comprised ~6% of the genome, analogous to wild type. The restored domains tended to be smaller than wild type ranging from 500 bp to 59 kb, with an average size of 6.1 kb. To determine if the peaks observed in this strain are in wild type locations, we again compared only peaks from assembled contigs. Using bedtools intersect we found 301 peaks in common (out of 308 wild type peaks on assembled contigs) indicating that the peaks in the overexpression strain were smaller but located in the same regions as in wild

type. Nevertheless, the significant restoration of H3K27me2/3 suggests that reduced *eed* expression is the major contributor to the loss of H3K27me2/3 in the Δ*hH2Az* strain.

Discussion

H2A.Z is a highly conserved histone variant that has been linked to gene activation and repression, and control of H3K27 methylation. We report here that *N. crassa* H2A.Z is required for normal methylation of H3K27 in facultative heterochromatin domains. In contrast to the situation in plants and animals, we find that *N. crassa* H2A.Z does not co-localize with H3K27me2/3. In undifferentiated mammalian cells and in plant cells, H2A.Z co-localized with PRC2 components, H3K27me3, SUZ12 or both (40, 52-56). In mESCs, H2A.Z is found at developmentally important loci where SUZ12 is also enriched (40, 56). In addition, this histone variant is proposed to regulate lineage commitment by functioning as a "molecular rheostat" to drive either activation or repression of genes (52, 54, 99). This co-localization of PRC2 and H2A.Z is not seen in differentiated murine cells, and ubiquitylated residues on the C-terminal tail of H2A.Z have been hypothesized as integral for cells to maintain undifferentiated status (54, 56). In plants, H2A.Z displays significant co-localization with H3K27me3 in the gene bodies of PcG-repressed genes even in differentiated tissues (52, 53). Our work highlights an important structural difference between facultative heterochromatin in plants and filamentous fungi. In both plants and animals, PRC1 and PRC2 work together to maintain stable gene repression in a number of ways including, the classic hierarchical model (100), ubiquitylation of H2A proteins (101), and inhibition of machinery involved in transcription (102). PRC1 monoubiquitylates H2A.Z in these organisms, which is important for Polycomb binding and repression (43, 54). H2A.Z may not be required for

H3K27me2/3 in fungi due to the absence of PRC1. Although we did not observe co-localization of H2A.Z and H3K27me2/3 in *N. crassa*, it remains possible that these two chromatin features overlap in specific developmental cell types (e.g., during sexual development or meiosis). Future work is needed to test this possibility.

In *N. crassa* we generally find histone H2A.Z at the promoters of a large number of genes in the genome. When viewing the localization using a metaplot, which averages the enrichment of all H2A.Z marked nucleosomes, it appears that H2A.Z flanks the TSS. Genome-wide localization of H2A.Z has been performed in a variety of organisms including *Arabidopsis*, *C. elegans*, *S. cerevisiae*, mouse, and *Drosophila*. H2A.Z is generally found in the promoters of active and inactive genes, as well as at in vertebrate enhancers (36, 38-43). The +1 nucleosome, first nucleosome after the TSS, containing H2A.Z has been postulated as a lower energy barrier to transcription elongation in *Drosophila* and *Arabidopsis* (36, 37). Our data are consistent with an important promoter-specific role for *N. crassa* H2A.Z.

Indeed, in *N. crassa* we find that the *eed* gene contains a large peak of H2A.Z in the +1 nucleosome, and we find that H2A.Z is required for the proper expression of *eed*. To our knowledge this is the first report of H2A.Z specifically regulating the *eed* gene. Previous studies in mESCs demonstrate that appropriate binding of multiple factors to the *eed* promoter are required for the normal expression of *eed* (103, 104). It is also possible that there are transcription factors that can bind to the sequence associated with the H2A.Z-containing nucleosome. Nucleosomes that contain H2A.Z protect approximately 120 bp of DNA from MNase digestion as opposed to nucleosomes with canonical H2A

that protect 147 bp (105). This may leave more sequence available for transcription factor binding between H2A.Z-containing nucleosomes.

We observed that reduced *eed* expression levels leads to region-specific losses of H3K27me2/3, rather than a more general, or global, reduction. In contrast to our work, reduced *Eed* is reported to cause a global decrease in H3K27me3 in mESCs. In these cells, reduced expression of *Eed* was observed following downregulation of *Oct3/4*, which in turn led to a global reduction of H3K27me3 (103, 104). In *N. crassa*, repetitive sequences (e.g., the canonical telomere repeats) are sufficient to induce an artificial H3K27me3 domain when inserted into a locus normally devoid of H3K27me3 (33). It is interesting that even though we also observe the loss of H3K27 methylation throughout much of the genome, regions proximal to the telomeres retained H3K27me2/3. This suggests that PRC2 is being recruited to the telomeric region and the downregulation of *eed* causes a defect in the propagation of the H3K27me2/3 modification into topologically associated, nearby regions, and we will address this possibility in future Hi-C studies. Another possibility is that the internal domains have a special requirement for EED in spreading, or for the maintenance of H3K27 methylation following DNA replication. Alternatively, EED may interact directly with transcription factors that control assembly of facultative heterochromatin at certain internal domains, while other PRC2-associated proteins may be more important for targeting PRC2 to telomeres. Future studies are needed to distinguish between these possible working models.

References

1. Luger K. 2003. Structure and dynamic behavior of nucleosomes. Current Opinion in Genetics & Development 13:127-135

2. Muller J. 1995. Transcriptional silencing by the Polycomb protein in Drosophila embryos. EMBO J 14:1209-20

3. Hennig L, Derkacheva M. 2009. Diversity of Polycomb group complexes in plants: same rules, different players? Trends Genet 25:414-23.10.1016/j.tig.2009.07.002

4. Simon JA, Kingston RE. 2009. Mechanisms of polycomb gene silencing: knowns and unknowns. Nat Rev Mol Cell Biol 10:697-708.10.1038/nrm2763

5. Schuettengruber B, Bourbon HM, Di Croce L, Cavalli G. 2017. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. Cell 171:34-57.10.1016/j.cell.2017.08.002

6. Kuroda MI, Kang H, De S, Kassis JA. 2020. Dynamic Competition of Polycomb and Trithorax in Transcriptional Programming. Annu Rev Biochem doi:10.1146/annurev-biochem-120219-103641.10.1146/annurev-biochem-120219-103641

7. Muller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, Miller EL, O'Connor MB, Kingston RE, Simon JA. 2002. Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell 111:197-208.10.1016/s0092-8674(02)00976-5

8. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. 2002. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 298:1039-43.10.1126/science.1076997

9. Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. 2002. Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal polycomb sites. Cell 111:185-196.Doi 10.1016/S0092-8674(02)00975-3

10. Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D. 2002. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev 16:2893-905.10.1101/gad.1035902

11. Veerappan CS, Avramova Z, Moriyama EN. 2008. Evolution of SET-domain protein families in the unicellular and multicellular Ascomycota fungi. BMC Evol Biol 8:190.10.1186/1471-2148-8-190

12. Aramayo R, Selker EU. 2013. Neurospora crassa, a model system for epigenetics research. Cold Spring Harb Perspect Biol 5:a017921.10.1101/cshperspect.a017921

13. Connolly LR, Smith KM, Freitag M. 2013. The Fusarium graminearum histone H3 K27 methyltransferase KMT6 regulates development and expression of secondary metabolite gene clusters. PLoS Genet 9:e1003916.10.1371/journal.pgen.1003916

14. Jamieson K, Rountree MR, Lewis ZA, Stajich JE, Selker EU. 2013. Regional control of histone H3 lysine 27 methylation in Neurospora. Proc Natl Acad Sci U S A 110:6027-32.10.1073/pnas.1303750110

15. Chujo T, Scott B. 2014. Histone H3K9 and H3K27 methylation regulates fungal alkaloid biosynthesis in a fungal endophyte-plant symbiosis. Mol Microbiol 92:413-34.10.1111/mmi.12567

16. Schotanus K, Soyer JL, Connolly LR, Grandaubert J, Happel P, Smith KM, Freitag M, Stukenbrock EH. 2015. Histone modifications rather than the novel regional centromeres of Zymoseptoria tritici distinguish core and accessory chromosomes. Epigenetics & Chromatin 8.ARTN 4110.1186/s13072-015-0033-5

17. Studt L, Rosler SM, Burkhardt I, Arndt B, Freitag M, Humpf HU, Dickschat JS, Tudzynski B. 2016. Knock-down of the methyltransferase Kmt6 relieves H3K27me3 and results in induction of cryptic and otherwise silent secondary metabolite gene clusters in Fusarium fujikuroi. Environmental Microbiology 18:4037-4054.10.1111/1462-2920.13427

18. Dumesic PA, Homer CM, Moresco JJ, Pack LR, Shanle EK, Coyle SM, Strahl BD, Fujimori DG, Yates JR, Madhani HD. 2015. Product Binding Enforces the Genomic Specificity of a Yeast Polycomb Repressive Complex. Cell 160:204-218.10.1016/j.cell.2014.11.039

19. Jones RS, Gelbart WM. 1993. The Drosophila Polycomb-group gene Enhancer of zeste contains a region with sequence similarity to trithorax. Mol Cell Biol 13:6357-66.10.1128/mcb.13.10.6357

20. Chen H, Rossier C, Antonarakis SE. 1996. Cloning of a human homolog of the Drosophila enhancer of zeste gene (EZH2) that maps to chromosome 21q22.2. Genomics 38:30-7.10.1006/geno.1996.0588

21. Abel KJ, Brody LC, Valdes JM, Erdos MR, McKinley DR, Castilla LH, Merajver SD, Couch FJ, Friedman LS, Ostermeyer EA, Lynch ED, King MC, Welcsh PL, Osborne-Lawrence S, Spillman M, Bowcock AM, Collins FS, Weber BL. 1996. Characterization of EZH1, a human homolog of Drosophila Enhancer of zeste near BRCA1. Genomics 37:161-71.10.1006/geno.1996.0537

22. Goodrich J, Puangsomlee P, Martin M, Long D, Meyerowitz EM, Coupland G. 1997. A Polycomb-group gene regulates homeotic gene expression in Arabidopsis. Nature 386:44-51.10.1038/386044a0

23. Grossniklaus U, Vielle-Calzada JP, Hoeppner MA, Gagliano WB. 1998. Maternal control of embryogenesis by MEDEA, a polycomb group gene in Arabidopsis. Science 280:446-50.10.1126/science.280.5362.446

24. Chanvivattana Y, Bishopp A, Schubert D, Stock C, Moon YH, Sung ZR, Goodrich J. 2004. Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. Development 131:5263-76.10.1242/dev.01400

25. Schwartz YB, Pirrotta V. 2007. Polycomb silencing mechanisms and the management of genomic programmes. Nat Rev Genet 8:9-22.10.1038/nrg1981

26. Schumacher A, Lichtarge O, Schwartz S, Magnuson T. 1998. The murine Polycomb-group gene eed and its human orthologue: functional implications of evolutionary conservation. Genomics 54:79-88.10.1006/geno.1998.5509

27. Ng J, Hart CM, Morgan K, Simon JA. 2000. A Drosophila ESC-E(Z) protein complex is distinct from other polycomb group complexes and contains covalently modified ESC. Mol Cell Biol 20:3069-78.10.1128/mcb.20.9.3069-3078.2000

28. Birve A, Sengupta AK, Beuchle D, Larsson J, Kennison JA, Rasmuson-Lestander A, Muller J. 2001. Su(z)12, a novel Drosophila Polycomb group gene that is conserved in vertebrates and plants. Development 128:3371-9

29. Derkacheva M, Steinbach Y, Wildhaber T, Mozgova I, Mahrez W, Nanni P, Bischof S, Gruissem W, Hennig L. 2013. Arabidopsis MSI1 connects LHP1 to PRC2 complexes. EMBO J 32:2073-85.10.1038/emboj.2013.145

30. Huang S, Lee WH, Lee EY. 1991. A cellular protein that competes with SV40 T antigen for binding to the retinoblastoma gene product. Nature 350:160-2.10.1038/350160a0

31. Qian YW, Wang YC, Hollingsworth RE, Jr., Jones D, Ling N, Lee EY. 1993. A retinoblastoma-binding protein related to a negative regulator of Ras in yeast. Nature 364:648-52.10.1038/364648a0

32. Lewis ZA. 2017. Polycomb Group Systems in Fungi: New Models for Understanding Polycomb Repressive Complex 2. Trends Genet 33:220-231

33. Jamieson K, McNaught KJ, Ormsby T, Leggett NA, Honda S, Selker EU. 2018. Telomere repeats induce domains of H3K27 methylation in Neurospora. Elife 7.10.7554/eLife.31216

34. Basenko EY, Sasaki T, Ji LX, Prybol CJ, Burckhardt RM, Schmitz RJ, Lewis ZA. 2015. Genome-wide redistribution of H3K27me3 is linked to genotoxic stress and

defective growth. Proceedings of the National Academy of Sciences of the United States of America 112:E6339-E6348.10.1073/pnas.1511377112

35. Jamieson K, Wiles ET, McNaught KJ, Sidoli S, Leggett N, Shao YC, Garcia BA, Selker EU. 2016. Loss of HP1 causes depletion of H3K27me3 from facultative heterochromatin and gain of H3K27me2 at constitutive heterochromatin. Genome Research 26:97-107.10.1101/gr.194555.115

36. Weber CM, Ramachandran S, Henikoff S. 2014. Nucleosomes are context-specific, H2A.Z-modulated barriers to RNA polymerase. Mol Cell 53:819-30.10.1016/j.molcel.2014.02.014

37. Dai X, Bai Y, Zhao L, Dou X, Liu Y, Wang L, Li Y, Li W, Hui Y, Huang X, Wang Z, Qin Y. 2017. H2A.Z Represses Gene Expression by Modulating Promoter Nucleosome Structure and Enhancer Histone Modifications in Arabidopsis. Mol Plant 10:1274-1292.10.1016/j.molp.2017.09.007

38. Guillemette B, Bataille AR, Gevry N, Adam M, Blanchette M, Robert F, Gaudreau L. 2005. Variant histone H2A.Z is globally localized to the promoters of inactive yeast genes and regulates nucleosome positioning. Plos Biology 3:2100-2110.ARTN e38410.1371/journal.pbio.0030384

39. Barski A, Cuddapah S, Cui K, Roh TY, Schones DE, Wang Z, Wei G, Chepelev I, Zhao K. 2007. High-resolution profiling of histone methylations in the human genome. Cell 129:823-37.10.1016/j.cell.2007.05.009

40. Creyghton MP, Markoulaki S, Levine SS, Hanna J, Lodato MA, Sha K, Young RA, Jaenisch R, Boyer LA. 2008. H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment. Cell 135:649-61.10.1016/j.cell.2008.09.056

41. Bargaje R, Alam MP, Patowary A, Sarkar M, Ali T, Gupta S, Garg M, Singh M, Purkanti R, Scaria V, Sivasubbu S, Brahmachari V, Pillai B. 2012. Proximity of H2A.Z containing nucleosome to the transcription start site influences gene expression levels in the mammalian liver and brain. Nucleic Acids Res 40:8965-78.10.1093/nar/gks665

42. Latorre I, Chesney MA, Garrigues JM, Stempor P, Appert A, Francesconi M, Strome S, Ahringer J. 2015. The DREAM complex promotes gene body H2A.Z for target repression. Genes Dev 29:495-500.10.1101/gad.255810.114

43. Gomez-Zambrano A, Merini W, Calonje M. 2019. The repressive role of Arabidopsis H2A.Z in transcriptional regulation depends on AtBMI1 activity. Nat Commun 10:2828.10.1038/s41467-019-10773-1

44. Bruce K, Myers FA, Mantouvalou E, Lefevre P, Greaves I, Bonifer C, Tremethick DJ, Thorne AW, Crane-Robinson C. 2005. The replacement histone H2A.Z in a hyperacetylated form is a feature of active genes in the chicken. Nucleic Acids Res 33:5633-9.10.1093/nar/gki874

45. Neves LT, Douglass S, Spreafico R, Venkataramanan S, Kress TL, Johnson TL. 2017. The histone variant H2A.Z promotes efficient cotranscriptional splicing in S. cerevisiae. Genes & Development 31:702-717

46. Xu Y, Ayrapetov MK, Xu C, Gursoy-Yuzugullu O, Hu Y, Price BD. 2012. Histone H2A.Z controls a critical chromatin remodeling step required for DNA double-strand break repair. Mol Cell 48:723-33.10.1016/j.molcel.2012.09.026

47. Rangasamy D, Greaves I, Tremethick DJ. 2004. RNA interference demonstrates a novel role for H2A.Z in chromosome segregation. Nat Struct Mol Biol 11:650-5.10.1038/nsmb786

48. Dhillon N, Oki M, Szyjka SJ, Aparicio OM, Kamakaka RT. 2006. H2A.Z functions to regulate progression through the cell cycle. Mol Cell Biol 26:489-501.10.1128/MCB.26.2.489-501.2006

49. Meneghini MD, Wu M, Madhani HD. 2003. Conserved histone variant H2A.Z protects euchromatin from the ectopic spread of silent heterochromatin. Cell 112:725-36.10.1016/s0092-8674(03)00123-5

50. Adam M, Robert F, Larochelle M, Gaudreau L. 2001. H2A.Z is required for global chromatin integrity and for recruitment of RNA polymerase II under specific conditions. Molecular and Cellular Biology 21:6270-6279

51. Hu G, Cui K, Northrup D, Liu C, Wang C, Tang Q, Ge K, Levens D, Crane-Robinson C, Zhao K. 2013. H2A.Z facilitates access of active and repressive complexes to chromatin in embryonic stem cell self-renewal and differentiation. Cell Stem Cell 12:180-92.10.1016/j.stem.2012.11.003

52. Zhang K, Xu W, Wang C, Yi X, Zhang W, Su Z. 2017. Differential deposition of H2A.Z in combination with histone modifications within related genes in Oryza sativa callus and seedling. Plant J 89:264-277.10.1111/tpj.13381

53. Carter B, Bishop B, Ho KK, Huang R, Jia W, Zhang H, Pascuzzi PE, Deal RB, Ogas J. 2018. The Chromatin Remodelers PKL and PIE1 Act in an Epigenetic Pathway That Determines H3K27me3 Homeostasis in Arabidopsis. Plant Cell 30:1337-1352.10.1105/tpc.17.00867

54. Surface LE, Fields PA, Subramanian V, Behmer R, Udeshi N, Peach SE, Carr SA, Jaffe JD, Boyer LA. 2016. H2A.Z.1 Monoubiquitylation Antagonizes BRD2 to

Maintain Poised Chromatin in ESCs. Cell Rep 14:1142-1155.10.1016/j.celrep.2015.12.100

55. Wang Y, Long H, Yu J, Dong L, Wassef M, Zhuo B, Li X, Zhao J, Wang M, Liu C, Wen Z, Chang L, Chen P, Wang QF, Xu X, Margueron R, Li G. 2018. Histone variants H2A.Z and H3.3 coordinately regulate PRC2-dependent H3K27me3 deposition and gene expression regulation in mES cells. BMC Biol 16:107.10.1186/s12915-018-0568-6

56. Ku M, Jaffe JD, Koche RP, Rheinbay E, Endoh M, Koseki H, Carr SA, Bernstein BE. 2012. H2A.Z landscapes and dual modifications in pluripotent and multipotent stem cells underlie complex genome regulatory functions. Genome Biol 13:R85.10.1186/gb-2012-13-10-r85

57. van Daal A, Elgin SC. 1992. A histone variant, H2AvD, is essential in Drosophila melanogaster. Mol Biol Cell 3:593-602.10.1091/mbc.3.6.593

58. Clarkson MJ, Wells JR, Gibson F, Saint R, Tremethick DJ. 1999. Regions of variant histone His2AvD required for Drosophila development. Nature 399:694-7.10.1038/21436

59. Liu X, Li B, GorovskyMa. 1996. Essential and nonessential histone H2A variants in Tetrahymena thermophila. Mol Cell Biol 16:4305-11.10.1128/mcb.16.8.4305

60. Faast R, Thonglairoam V, Schulz TC, Beall J, Wells JR, Taylor H, Matthaei K, Rathjen PD, Tremethick DJ, Lyons I. 2001. Histone variant H2A.Z is required for early mammalian development. Curr Biol 11:1183-7.10.1016/s0960-9822(01)00329-3

61. Iouzalen N, Moreau J, Mechali M. 1996. H2A.ZI, a new variant histone expressed during Xenopus early development exhibits several distinct features from the core histone H2A. Nucleic Acids Res 24:3947-52.10.1093/nar/24.20.3947

62. Ridgway P, Brown KD, Rangasamy D, Svensson U, Tremethick DJ. 2004. Unique residues on the H2A.Z containing nucleosome surface are important for Xenopus laevis development. J Biol Chem 279:43815-20.10.1074/jbc.M408409200

63. Zhang H, Bishop B, Ringenberg W, Muir WM, Ogas J. 2012. The CHD3 remodeler PICKLE associates with genes enriched for trimethylation of histone H3 lysine 27. Plant Physiol 159:418-32.10.1104/pp.112.194878

64. Davis R, de Serres F. 1970. Genetic and microbiological research techniques for Neurospora crassa. Methods in Enzymology 17:79-143.https://doi.org/10.1016/0076-6879(71)17168-6

65. Honda S, Selker EU. 2009. Tools for fungal proteomics: multifunctional neurospora vectors for gene replacement, protein expression and protein purification. Genetics 182:11-23.10.1534/genetics.108.098707

66. Margolin BS, Freitag, M., and Selker, E.U. 1997. Improved plasmids for gene targeting at the his-3 locus of Neurospora crassa by electroporation. Fungal Genetics Newsletter 44:2

67. Avalos J, Geever RF, Case ME. 1989. Bialaphos resistance as a dominant selectable marker in Neurospora crassa. Curr Genet 16:369-72.10.1007/bf00340716

68. Ferraro AR, Lewis ZA. 2018. ChIP-Seq Analysis in Neurospora crassa. Fungal Genomics: Methods and Protocols, Second Edition 1775:241-250.10.1007/978-1-4939-7804-5_19

69. Seymour M, Ji L, Santos AM, Kamei M, Sasaki T, Basenko EY, Schmitz RJ, Zhang X, Lewis ZA. 2016. Histone H1 Limits DNA Methylation in Neurospora crassa. G3 (Bethesda) 6:1879-89.10.1534/g3.116.028324

70. Sasaki T, Lynch KL, Mueller CV, Friedman S, Freitag M, Lewis ZA. 2014. Heterochromatin controls gammaH2A localization in Neurospora crassa. Eukaryot Cell 13:990-1000.10.1128/EC.00117-14

71. Liu X, Dang Y, Matsu-ura T, He Y, He Q, Hong CI, Liu Y. 2017. DNA Replication Is Required for Circadian Clock Function by Regulating Rhythmic Nucleosome Composition. Molecular Cell 67:203-213.e4

72. Dong Q, Wang Y, Qi S, Gai K, He Q, Wang Y. 2018. Histone variant H2A.Z antagonizes the positive effect of the transcriptional activator CPC1 to regulate catalase-3 expression under normal and oxidative stress conditions. Free Radic Biol Med 121:136-148.10.1016/j.freeradbiomed.2018.05.003

73. Rohland N, Reich D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Research 22:939-946.10.1101/gr.128124.111

74. Martin M. 2011. Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. EMBnet 17:2.https://doi.org/10.14806/ej.17.1.200

75. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754-60.10.1093/bioinformatics/btp324

76. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Proc GPD. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-2079.10.1093/bioinformatics/btp352

77. Ramirez F, Ryan DP, Gruning B, Bhardwaj V, Kilpert F, Richter AS, Heyne S, Dundar F, Manke T. 2016. deepTools2: a next generation web server for deep-sequencing data analysis. Nucleic Acids Res 44:W160-5.10.1093/nar/gkw257

78. Thorvaldsdottir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14:178-92.10.1093/bib/bbs017

79. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38:576-89.10.1016/j.molcel.2010.05.004

80. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. 2019. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol 37:907-915.10.1038/s41587-019-0201-4

81. Liao Y, Smyth GK, Shi W. 2013. The Subread aligner: fast, accurate and scalable read mapping by seed-and-vote. Nucleic Acids Res 41:e108.10.1093/nar/gkt214

82. Love MI, Huber W, Anders S. 2014. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biology 15.ARTN 55010.1186/s13059-014-0550-8

83. Wickham H. 2009. Ggplot2 : elegant graphics for data analysis. Springer, New York.

84. Krogan NJ, Baetz K, Keogh MC, Datta N, Sawa C, Kwok TCY, Thompson NJ, Davey MG, Pootoolal J, Hughes TR, Emili A, Buratowski S, Hieter P, Greenblatt JF. 2004. Regulation of chromosome stability by the histone H2A variant Htz1, the Swr1 chromatin remodeling complex, and the histone acetyltransferase NuA4. Proc Natl Acad Sci USA 101:13513-13518

85. Greaves IK, Rangasamy D, Ridgway P, Tremethick DJ. 2007. H2A.Z contributes to the unique 3D structure of the centromere. Proc Natl Acad Sci U S A 104:525-30.10.1073/pnas.0607870104

86. Colot HV, Park G, Turner GE, Ringelberg C, Crew CM, Litvinkova L, Weiss RL, Borkovich KA, Dunlap JC. 2006. A high-throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. Proc Natl Acad Sci U S A 103:10352-10357.10.1073/pnas.0601456103

87. Jackson JD, Gorovsky MA. 2000. Histone H2A.Z has a conserved function that is distinct from that of the major H2A sequence variants. Nucleic Acids Research 28:3811-3816

88. Kobor MS, Venkatasubrahmanyam S, Meneghini MD, Gin JW, Jennings JL, Link AJ, Madhani HD, Rine J. 2004. A protein complex containing the conserved Swi2/Snf2-related ATPase Swr1p deposits histone variant H2A.Z into euchromatin. PLoS Biol 2:E131.10.1371/journal.pbio.0020131

89. Mizuguchi G, Shen X, Landry J, Wu WH, Sen S, Wu C. 2004. ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. Science 303:343-8.10.1126/science.1090701

90. Luk E, Ranjan A, Fitzgerald PC, Mizuguchi G, Huang Y, Wei D, Wu C. 2010. Stepwise histone replacement by SWR1 requires dual activation with histone H2A.Z and canonical nucleosome. Cell 143:725-36.10.1016/j.cell.2010.10.019

91. Deal RB, Topp CN, McKinney EC, Meagher RB. 2007. Repression of flowering in Arabidopsis requires activation of FLOWERING LOCUS C expression by the histone variant H2A.Z. Plant Cell 19:74-83.10.1105/tpc.106.048447

92. Wong MM, Cox LK, Chrivia JC. 2007. The chromatin remodeling protein, SRCAP, is critical for deposition of the histone variant H2A.Z at promoters. J Biol Chem 282:26132-9.10.1074/jbc.M703418200

93. Borkovich KA, Alex LA, Yarden O, Freitag M, Turner GE, Read ND, Seiler S, Bell-Pedersen D, Paietta J, Plesofsky N, Plamann M, Goodrich-Tanrikulu M, Schulte U, Mannhaupt G, Nargang FE, Radford A, Selitrennikoff C, Galagan JE, Dunlap JC, Loros JJ, Catcheside D, Inoue H, Aramayo R, Polymenis M, Selker EU, Sachs MS, Marzluf GA, Paulsen I, Davis R, Ebbole DJ, Zelter A, Kalkman ER, O'Rourke R, Bowring F, Yeadon J, Ishii C, Suzuki K, Sakai W, Pratt R. 2004. Lessons from the genome sequence of Neurospora crassa: tracing the path from genomic blueprint to multicellular organism. Microbiol Mol Biol Rev 68:1-108

94. Kim K, Punj V, Choi J, Heo K, Kim JM, Laird PW, An W. 2013. Gene dysregulation by histone variant H2A.Z in bladder cancer. Epigenetics & Chromatin 6.Artn 3410.1186/1756-8935-6-34

95. Valdes-Mora F, Song JZ, Statham AL, Strbenac D, Robinson MD, Nair SS, Patterson KI, Tremethick DJ, Stirzaker C, Clark SJ. 2012. Acetylation of H2A.Z is a key epigenetic modification associated with gene deregulation and epigenetic remodeling in cancer. Genome Res 22:307-21.10.1101/gr.118919.110

96. Yoav Benjamini, Hochberg Y. 1995. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society: Series B (Methodological) 57:289-300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x

97. Xu C, Bian C, Yang W, Galka M, Ouyang H, Chen C, Qiu W, Liu H, Jones AE, MacKenzie F, Pan P, Li SS, Wang H, Min J. 2010. Binding of different histone marks differentially regulates the activity and specificity of polycomb repressive complex 2 (PRC2). Proc Natl Acad Sci U S A 107:19266-71.10.1073/pnas.1008937107

98. Hansen KH, Bracken AP, Pasini D, Dietrich N, Gehani SS, Monrad A, Rappsilber J, Lerdrup M, Helin K. 2008. A model for transmission of the H3K27me3 epigenetic mark. Nat Cell Biol 10:1291-300.10.1038/ncb1787

99. Subramanian V, Fields PA, Boyer LA. 2015. H2A.Z: a molecular rheostat for transcriptional control. F1000Prime Rep 7:01.10.12703/P7-01

100. Wang L, Brown JL, Cao R, Zhang Y, Kassis JA, Jones RS. 2004. Hierarchical recruitment of polycomb group silencing complexes. Mol Cell 14:637-46.10.1016/j.molcel.2004.05.009

101. Wang H, Wang L, Erdjument-Bromage H, Vidal M, Tempst P, Jones RS, Zhang Y. 2004. Role of histone H2A ubiquitination in Polycomb silencing. Nature 431:873-8.10.1038/nature02985

102. King IF, Francis NJ, Kingston RE. 2002. Native and recombinant polycomb group complexes establish a selective block to template accessibility to repress transcription in vitro. Mol Cell Biol 22:7919-28.10.1128/mcb.22.22.7919-7928.2002

103. Ura H, Usuda M, Kinoshita K, Sun C, Mori K, Akagi T, Matsuda T, Koide H, Yokota T. 2008. STAT3 and Oct-3/4 control histone modification through induction of Eed in embryonic stem cells. Journal of Biological Chemistry 283:9713-9723.10.1074/jbc.M707275200

104. Ura H, Murakami K, Akagi T, Kinoshita K, Yamaguchi S, Masui S, Niwa H, Koide H, Yokota T. 2011. Eed/Sox2 regulatory loop controls ES cell self-renewal through histone methylation and acetylation. Embo Journal 30:2190-2204.10.1038/emboj.2011.126

105. Tolstorukov MY, Kharchenko PV, Goldman JA, Kingston RE, Park PJ. 2009. Comparative analysis of H2A.Z nucleosome organization in the human and yeast genomes. Genome Res 19:967-977

Figures

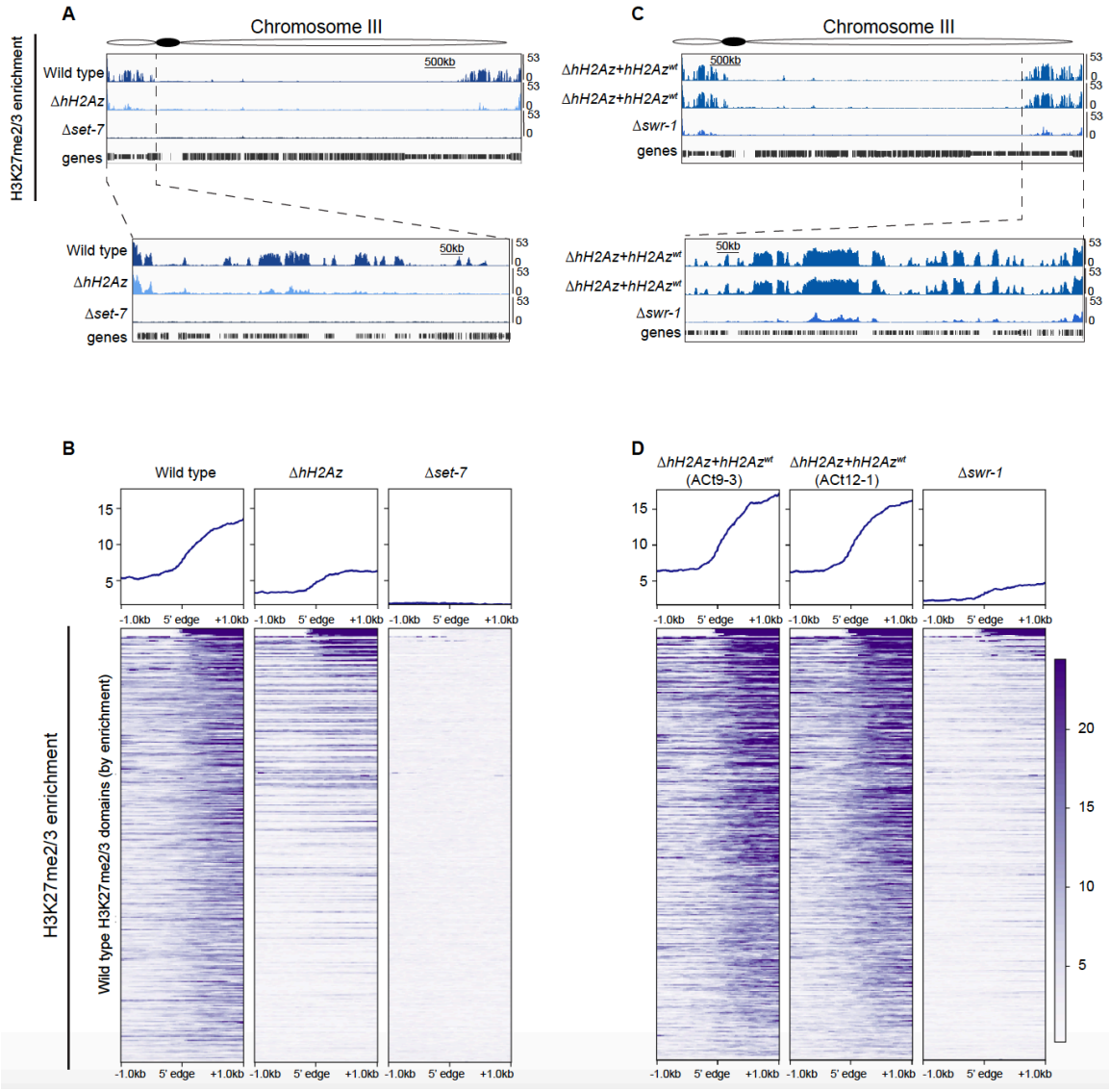


Figure 4.1: H2A.Z is required for normal patterns of H3K27 methylation

A) Genome browser images illustrate H3K27me2/3 enrichment on *N. crassa* Linkage Group ("chromosome") III for wild type, *ΔhH2Az*, and *Δset-7*. A segment of chromosome III is displayed at higher resolution to illustrate depletion of internal H3K27me2/3 domains.

B) H3K27me2/3 in the *ΔhH2Az* strain exhibits striking depletion of the most H3K27me2/3 domains, with overall lower enrichment of this modification. Heatmaps display 325 PRC2-target domains (rows) ordered by wild type enrichment for wild type, *ΔhH2Az*, and *Δset-7* strains centered on the 5' end of each domain + or – 1,000 bp for a total window size of 2,000 bp.

C) Genome browser images illustrate H3K27me2/3 enrichment on chromosome III for two ectopic complemented strains of *ΔhH2Az+hH2Az*^{wt} and the Δ*swr-1* strain. The segment of chromosome III is displayed at higher resolution to illustrate rescue by complementation and depletion of H3K27me2/3 in Δ*swr-1* background.

D) Heatmaps of H3K27me2/3 rescue in complemented strains (*ΔhH2Az+hH2Az*^{wt} [ACt9-3 and ACt12-1]) and depletion in the Δ*swr-1* strain. The heatmaps are ordered as in *B* and depict the domain boundary + or – 1,000 bp for a total window size of 2,000bp.

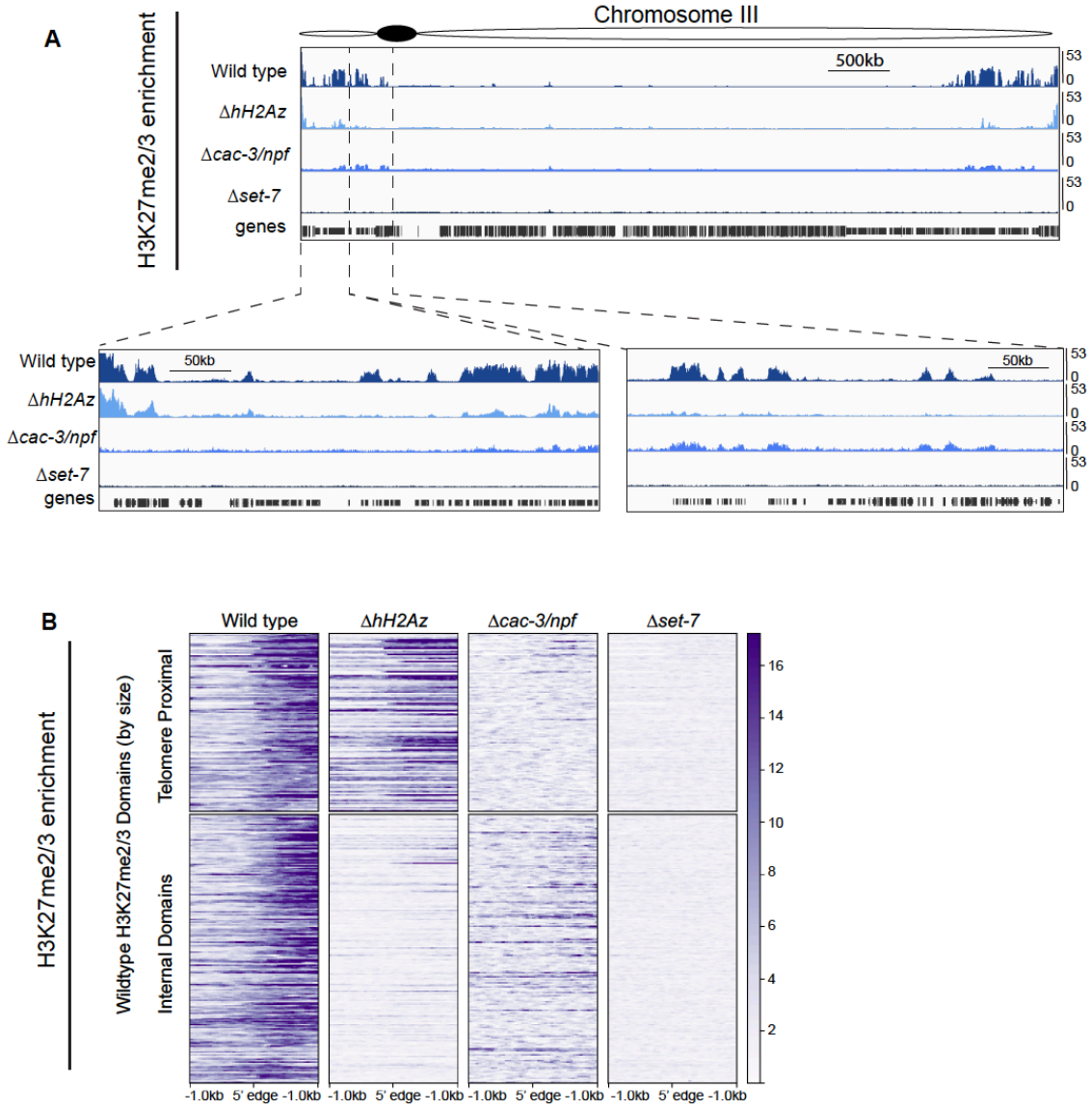


Figure 4.2: Deletion of *hH2Az* results in region-specific loss of H3K27me2/3

A) Genome browser images illustrate H3K27me2/3 enrichment on Linkage Group ("chromosome") III for wild type, *ΔhH2Az*, *Δcac-3/npf*, and *Δset-7*. The two segments of chromosome III are displayed at higher resolution to visualize region-specific loss. Left panel displays the end of the chromosome to ~300 kb (left panel) and from ~400-600 kb (right panel). The telomere-proximal H3K27me2/3 regions are only moderately affected

by the deletion of *hH2Az*, whereas internal domains show a more dramatic loss of H3K27me2/3.

B) Heatmaps of H3K27me2/3 enrichment for wild type, *ΔhH2Az*, *Δcac-3/npf*, and *Δset-7* across PRC2-target domains organized by their proximity to the telomere. The top section is restricted to domains that are <200 kb away from the chromosome ends ("telomere-proximal domains"), plotted from largest to smallest (123 domains). The bottom of the heatmaps contain the domains that are >200 kb away from chromosome ends ("internal domains"), also plotted from largest to smallest (186 domains). Heatmaps are centered on the 5' edge of PRC2-target domains + or –1,000 bp for a total window size of 2,000 bp. The *ΔhH2Az* strain retains most telomere-proximal H3K27me2/3, as opposed to the *Δcac-3/npf* strain where almost all H3K27me2/3 enrichment is lost from telomere-proximal regions.

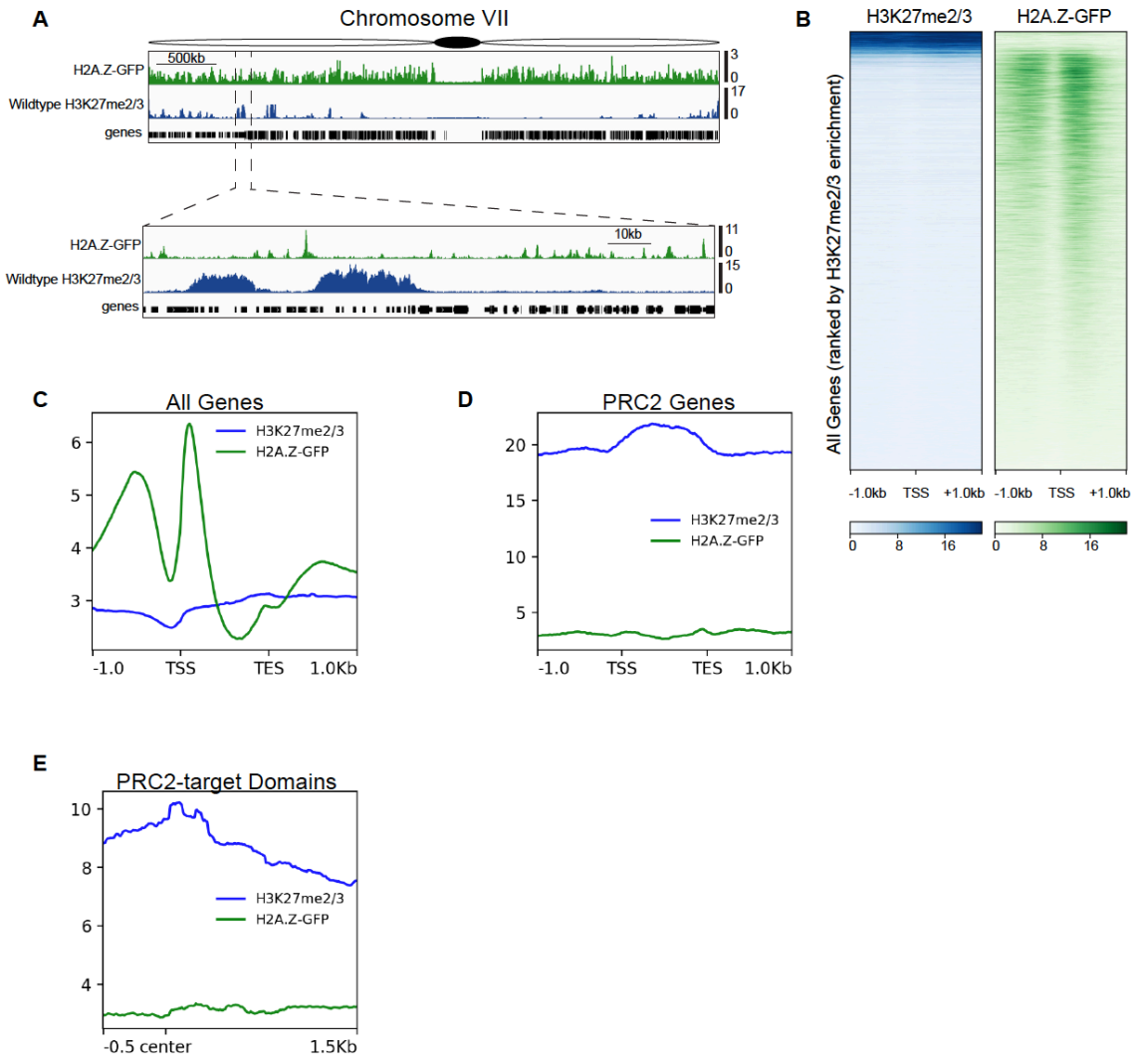


Figure 4.3: H3K27me2/3 and H2A.Z are not co-localized in *N. crassa* mycelium

A) Genome browser images of ChIP-seq for H2A.Z-GFP (green) and H3K27me2/3 (blue) enrichment across Linkage Group ("chromosome") VII. A segment of chromosome VII is displayed at higher resolution to visualize the distinct patterns of each modification. Distinct peaks are located at the start of many genes in the genome yet few H2A.Z peaks are present within transcriptionally silent PRC2-target domains.

B) Heatmaps of H3K27me2/3 (blue) and H2A.Z-GFP (green) enrichment ordered by H3K27me2/3 enrichment. Heatmaps are centered on the transcription start site (TSS) for all 10,397 genes, + or – 1,000 bp for a full window size of 2,000 bp.

C) Gene profile of H2A.Z-GFP (green line) and H3K27me2/3 (blue line) enrichment for all 10, 397 genes (fit to 1000 bp for gene body length) in the genome – 1,000 bp upstream of TSS and + 1,000 bp downstream of TES.

D) Gene profile of H2A.Z-GFP (green line) and H3K27me2/3 (blue line) enrichment for only H3K27me2/3 genes (589 genes) enriched in the genome – 1,000 bp upstream of TSS and + 1,000 bp downstream of TES.

E) Line plot centered on 325 PRC2-target domains displays very low enrichment for H2A.Z-GFP (green line) and high enrichment for H3K27me2/3 (blue line).

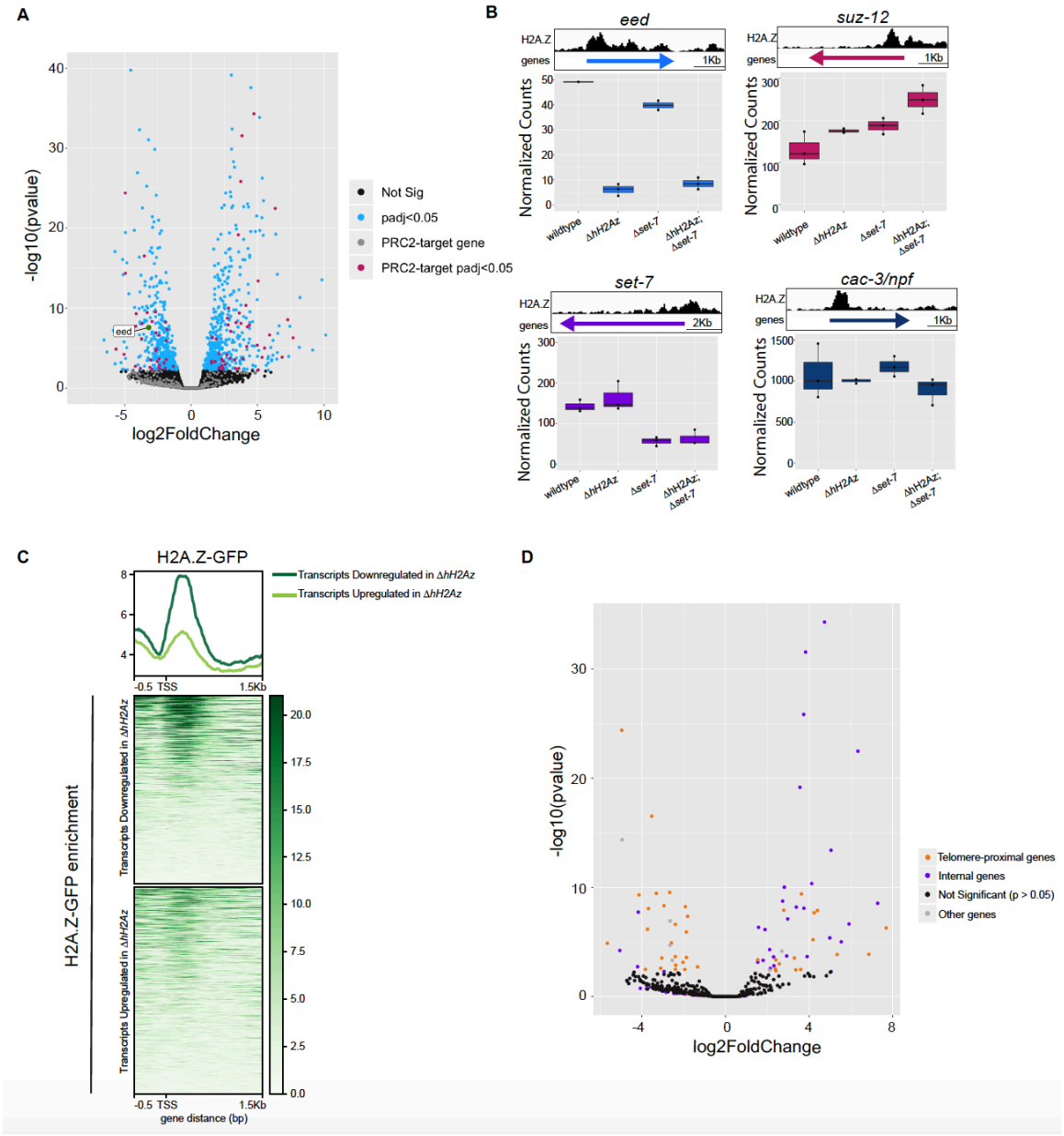


Figure 4.4: H2A.Z is important for the proper regulation of a large number of genes in *N. crassa*, including *eed*

A) Volcano plot of differentially expressed genes in Δ*hH2Az*. Deletion of *hH2Az* misregulates a large number of genes in both directions; however, there are slightly more genes that are upregulated upon deletion of *hH2Az*. It is likely that H2A.Z is necessary for the proper expression of a large percentage of genes in *N. crassa*. EED is labeled and

highlighted in green on the plot. Genes enriched for H3K27me2/3 are in grey (FDR corrected p-value > 0.05) or pink (FDR corrected p-value < 0.05) corresponding to their significance values. The *eed* gene is significantly downregulated in the deletion strain.
B) Genome browser images of each gene and its corresponding H2A.Z enrichment, for all PRC2 components, there is enrichment of H2A.Z near the TSS. Boxplots of normalized counts for all subunits of PRC2 (*eed* [light blue]*, set-7* [purple]*, suz-12* [pink] *cac-3/npf* [dark blue]*)* in wild type, Δ*hH2Az*, Δ*set-7*, and Δ*hH2Az;*Δ*set-7* backgrounds. Downregulation of *eed* is dependent on *hH2Az* deletion.
C) Heatmap showing the difference of H2A.Z-GFP enrichment for all transcripts that are misregulated in the Δ*hH2Az* strain. Heatmap is ordered by enrichment with the top panel transcripts that are upregulated in ΔhH2Az (559), and the bottom panel transcripts that are downregulated in Δ*hH2Az* (507).
D) Volcano plot of 548 differentially expressed PRC2-target genes divided by their proximity to the end of the chromosome. Genes that are not significant are in black, significant telomere-proximal genes are in orange, significant internal genes are in purple, and other genes are in gray (other are genes that do not overlap 70% or more with H3K27me2/3 domains). More internal genes show upregulation (27) than downregulation, and the telomere-proximal genes show a similar number of up and downregulated genes.

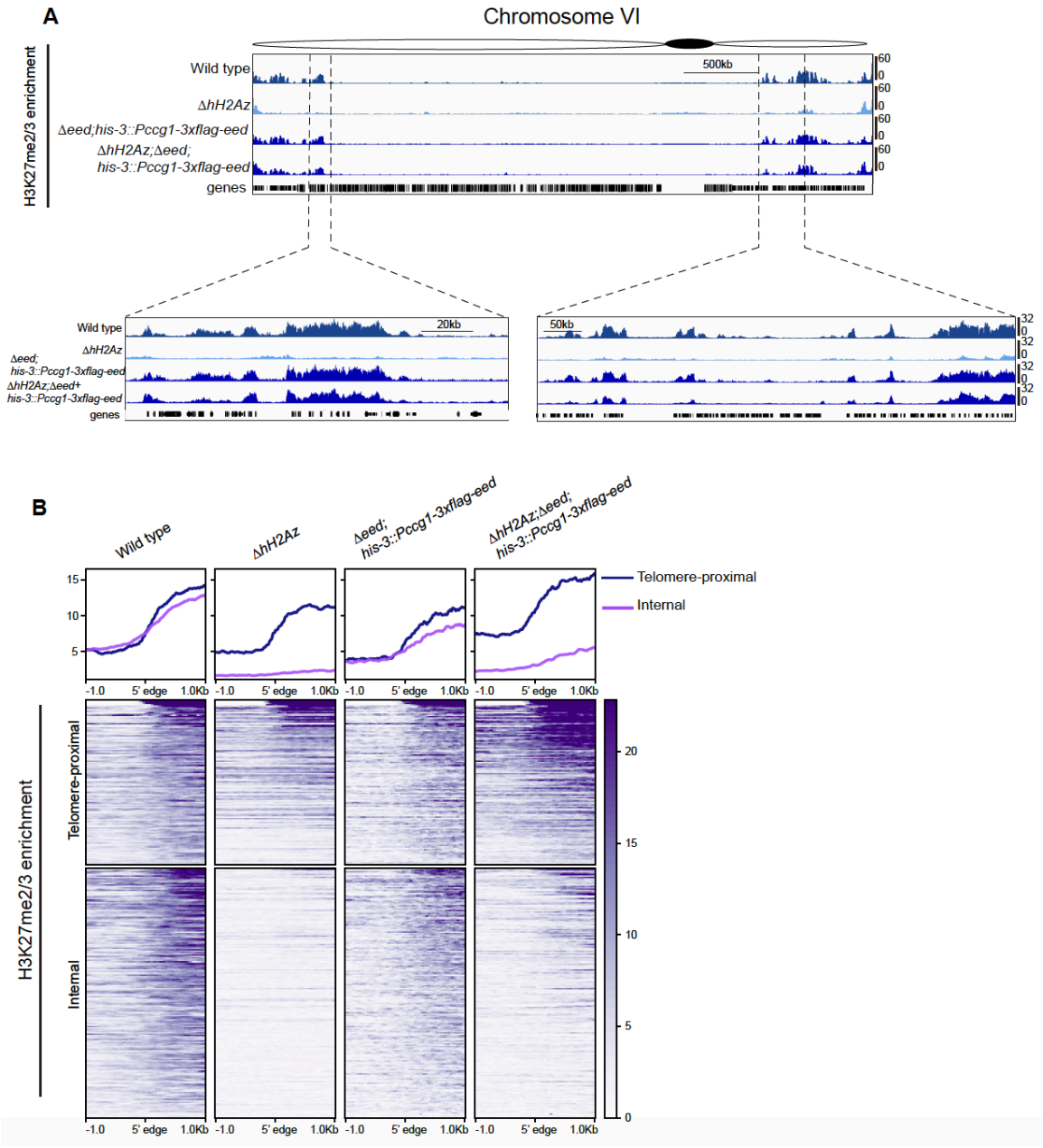


Figure 4.5: Overexpression of *eed* rescues H3K27 methylation levels in the absence of H2A.Z

A) Partial restoration of H3K27 methylation throughout the genome in a Δ*hH2Az*;Δ*eed* strain containing *his-3::Pccg-1-3xFlag-eed* and overexpressing *eed* at ~100x the native level. Most H3K27 methylation is restored, though there are some qualitative differences in the peak patterns between the overexpression strain and wild type.

B) Heatmaps of H3K27me3 enrichment across PRC2-target domains sorted by first by telomere-proximal (123 domains) or internal domains (186 domains) and then by enrichment centered on each domain + or – 1,000 bp for a full window size of 2,000 bp for wild type, Δ*hH2Az*, Δ*eed;his-3::Pccg-1-3xflag-eed,* and Δ*eed;*Δ*hH2Az;his-3::Pccg-1-3xflag-eed*. Not all domains are fully rescued to wild type levels.

CHAPTER 5

RTT109 IS REQUIRED FOR NORMAL PATTERNS OF H3K27ME2/3 IN *NEUROSPORA CRASSA*

Introduction

RTT109 (regulator of Ty1 transposition 109) is a fungal-specific acetyltransferase which can acetylate multiple lysine resides on the H3 histone protein (1, 2). Histone acetyltransferases are a diverse group of enzymes that can acetylate histone and non-histone proteins (3). Originally, *rtt109* was discovered through a mutant screen which sought to identify mutations that resulted in increased Ty1 mobility in the *Saccharomyces cerevisiae* genome (4). It was further characterized by two genetics screens in 2001 with the goal of identifying mutants required for resistance to genotoxic agents (5, 6). In 2007, its function as a histone H3 lysine 56 (H3K56) acetyltransferase was discovered in yeast and it was implicated in transcriptional regulation (7-9). Acetylation of histones is an important post-translational modification which is imperative in transcriptional activity as well as chromatin structure and assembly (3, 10).

Because K56 is located on the globular domain of the histone and not the unstructured tail, acetylation of this residue dramatically changes the contacts between the nucleosome and the DNA wrapped around in a way that opens up the DNA (11-13). Two chaperones are implicated in directing the specificity of RTT109, ASF-1 and VPS75 (1). H3K56 acetylation occurs on nascent H3-H4 dimers prior to nucleosome assembly generally leading to transcriptional activity where this modification is found (12).

However, in 2011, there was a report of *rtt109* acting as both an activator and repressor of gene expression depending on the interplay between ASF-1 and RTT109 (14).

One important group of proteins that are required for maintaining stable gene repression are the Polycomb Group (PcG) proteins. These proteins can assemble into multiple different complexes, namely Polycomb Repressive Complexes 1 and 2 (PRC1 and PRC2) in animals and plants and have important roles in regulation of developmental genes (15-19). PRC2 is responsible for depositing mono-, di-, and trimethylation of the histone H3 lysine 27 residue and is a molecular hallmark of facultative heterochromatin, which is conditionally repressed (20-23).

In human cells, EZH2, the catalytic subunit of the human PRC2, is acetylated at lysine 348 by P300/CBP-associated factor (PCAF) (24). RTT109 is a structural homolog to the metazoan p300/CBP coactivator, which contains an acetyltransferase domain (25). In order to determine if any lysine acetyltransferases (KATs) regulate PRC2 in *Neurospora crassa*, we performed a KAT RNA-seq screen to determine which, if any, were involved in the regulation of PRC2. Interestingly, we find that the deletion strain of *rtt109* has broad upregulation of a large subset of PRC2-target genes and loss of select H3K27me2/3 domains on certain chromosomes.

Results

Δ*rtt109* exhibits upregulation of a subset of PRC2-target genes

PRC2-target genes are known to be repressed by the H3K27me3 modification, at least partially (26). However, when the H3K27 methyltransferase itself is deleted in *N. crassa*, only a subset of genes are upregulated, suggesting these genes are controlled by a complex system of regulation (27, 28). PRC2 is known to be acetylated at lysine 348 in

human cells on EZH2 (24), the catalytic subunit, and to determine if any KATs are involved in H3K27me2/3 we performed an RNA-seq screen. This screen revealed that there are many PRC2-target genes upregulated in the Δ*rtt109* strain as compared to wild type (Fig. 5.1). In strains lacking *set-7*, the H3K27 methyltransferase, only a subset of genes are upregulated; however, in strains lacking *rtt109* we see broader upregulation of H3K27me2/3 marked genes.

***rtt109* deletion exhibits depletion of K27me3 only in specific regions**

Because there was a substantial impact on expression of PRC2-target genes we wanted to know if the H3K27me2/3 pattern was also altered in the Δ*rtt109* strain. We performed chromatin immunoprecipitation followed by high-throughput sequencing (ChIP-seq) on three replicates and found that some of the PRC2-target domains show losses, but not all (Fig. 5.2). Each ChIP-seq replicate shows conflicting results (Fig. 5.3A, 5.3B). We found almost no H3K27me2/3 at all in the first strain tested; however, this strain had a second mutation in the background, *mus-51::bar*. The *mus-51* deletion was part of the protocol used to generate the whole genome knock-out collection; this deletion impairs the non-homologous end joining machinery in *N. crassa* allowing for efficient homologous recombination.

To obtain a strain that only contained the *rtt109* deletion we backcrossed the strain to wild type and isolated progeny. The second ChIP-seq experiment was performed on the backcrossed strain ACx81-1 (progeny from the backcross to wild type). This experiment either shows that there is little to no H3K27me2/3 present, or that the experiment failed and that is what we are detecting. In the third replicate, which is an

aged version of the second replicate, we find almost normal H3K27me2/3 at many regions in the genome but still observe losses at other regions.

We used ChIP followed by qPCR to investigate specific regions of the genome to determine H3K27 methylation enrichment and found that subtelomeric regions tested retain their normal methylation (data not shown) as opposed to three separate internal regions (NCU04707, NCU04877, NCU09834) which appear to lose most of their enrichment (Fig. 5.4A). This region-specific loss is similar to recent findings in another strain, Δ*hH2Az*.

Since our previous results in a strain lacking H2A.Z showed region-specific loss due to mis-regulation of *eed* (29), a component of PRC2, we wanted to know if the expression of the PRC2 subunits were equivalent to wild type levels. We checked the TPMs for each component of PRC2 (and *hH2Az*) and find that they are expressed at levels similar to wild type in the Δ*rtt109* strain (Table 5.1). Together, this indicates some other mechanism is causing the depletion of H3K27me2/3 observed in the *rtt109* deletion strain.

H3K27 methylation molecular phenotype is rescued by complementation

To determine if the depletion of H3K27me2/3 is due to the absence of *rtt109* we performed an ectopic complementation using the *bar* gene fragment, which confers basta resistance, and a fragment of the *rtt109* gene including the native promoter. We find that H3K27me2/3 is restored in these strains (DDt1-7, DDt1-14) by using ChIP-qPCR (Fig. 5.4A). Strains lacking *rtt109* or K56 acetylation are known to be sensitive to MMS and other DNA damaging agents (30). We spot tested on MMS containing plates to confirm

that the complemented strain rescued the MMS sensitivity phenotype as well and indeed it did (Fig 5.4B).

Discussion

Deletion of *rtt109* changes the genome-wide patterns of H3K27 methylation. As an acetyltransferase with structural homology to P300, it is possible that RTT109 is acetylating PRC2, and the removal of this enzyme inhibits the function of PRC2. It is worth noting that p300 itself does not acetylate EZH2, but the PCAF complex (24, 25). It is also possible that RTT109 is regulating an unknown component of PRC2, or another gene that acts upstream of PRC2, similar to a function of the histone variant H2A.Z. Future studies are necessary to identify what role RTT109 plays in the regulation of H3K27me2/3.

These conflicting replicates of H3K27me2/3 ChIP experiments in Δ*rtt109* strains are not easily explained. These data may point to a connection between the *mus-51* deletion when in combination with other mutations that cause loss/depletion of H3K27me2/3. This has been observed in one other strain (unpublished data), *acf-1* when combined with *mus-51*. Further experiments are needed to validate this hypothesis.

Because K56 is the substrate for RTT109, we have obtained an engineered strain with arginine in place of the lysine at residue 56. This strain mimics unmodified K56. If lack of acetylation of K56 is causative for the H3K27me2/3 depletion observed, this engineered strain should phenocopy the Δ*rtt109* strain results. Other follow-up experiments include H3K27me2/3 ChIP-seq of the RTT109 chaperones Δ*asf-1* and Δ*vps-75*. These experiments are the first steps toward elucidating which pathway the depletion of H3K27me2/3 is connected to.

Materials and Methods

Strains and Media

Strains used in this study are listed in (Table 5.2). Strains were grown at 32˚C in Vogel's Minimal Medium (VMM) with 1.5% sucrose or 2% sucrose for DNA based protocols, and RNA based protocols, respectively (31). Liquid cultures were shaken at 180 rpm. Crosses were performed on Synthetic Crossing (SC) medium in the dark at room temperature (31). Ascospores were collected 14 days after fertilization. To isolate cross progeny, spores were spread on solid VMM plates containing FGS (1X Vogel's salts, 2% sorbose, 0.1% glucose, 0.1% fructose, and 1.5% agar) and incubated at 65˚C for 1 hour as previously described (31), after which spores were picked using a sterile inoculating needle and transferred to agar slants with appropriate medium (typically VMM). To test for sensitivity to DNA damaging agents, 5 µL of a conidial suspension was spotted on VMM containing FGS (1X Vogel's salts, 2% sorbose, 0.1% glucose, 0.1% fructose, and 1.5% agar) plates containing concentrations of methyl methanesulfonate (Sigma Aldrich cat. # 129925-5g) between 0.010% and 0.03% (w/v).

Transformation and complementation assays

Transformations were performed as previously described (32). To carry out ectopic complementation of the Δ*rtt109::hph* strain, two linear gene fragments were electroporated into the mutant strain. Specifically, the *bar* (confers Basta resistance) was amplified with primers LL #148 CCGTCGACAGAAGATGATATTGAAGGAGC and LL #149 AATTAACCCTCACTAAAGGGAACAAAAGC (33) and the wild type *rtt109* gene fragment including its native promoter (from assembly accession GCA_000182925.2) was amplified with primers NCU09825_3Fcomp:

ATAGGACGGCGGGTTGTATG and NCU09825_5Rcomp: ACTGGACTGGACTGGATGGA, and fragments were co-integrated into the Δ*rtt109::hph* strain, followed by selection of transformants on Basta-containing plates (VMM with 2% sorbose, 0.1% glucose, 0.1% fructose, 1.5% agar, and 200 ug/mL Basta). Transformants were transferred to agar slants and then screened by spot testing on Basta and MMS containing plates.

Chromatin immunoprecipitation (ChIP)

To carry out ChIP, conidia were inoculated in 5 mL of liquid VMM plus 1.5% sucrose and grown for 18 hours for wild type and other strains with typical growth rates. ChIP was performed as described previously (34-36). In brief, mycelia were harvested using filtration and were washed once in PBS prior to cross-linking for 10 minutes in PBS containing 1% formaldehyde on a rotating platform at room temperature. After 10 minutes, the reaction was quenched using 125mM glycine and placed back on the rotating platform for five minutes. Mycelia were harvested again using filtration, washed once with PBS, then resuspended in 600 µl of ChIP lysis buffer (50mM HEPES, pH 7.5, 140mM NaCl, 1mM EDTA, 1% Triton X-100, 0.1% sodium deoxycholate, one tablet Roche cOmplete mini EDTA-free Protease Inhibitor Cocktail (Roche, cat. # 11836170001) in 15 mL conical tubes. Chromatin was sheared by sonication after lysing cell walls with the QSONICA Misonix S-4000 ultrasonic processor (amplitude 10, 30 second processing, one second on, one second off), using the Diagenode Bioruptor UCD-200 (Intensity level: Medium, three rounds of 15 minutes (30 seconds on, 30 seconds off) to deliver 22.5 30 second pulses at 4˚C. Water temperature was kept at a constant 4˚C by using a Biorad cooling module (cat. # 170-3654) with variable speed pump to circulate

4˚C water while processing samples. Lysates were centrifuged at 13,000 rpm in an Eppendorf 5415D microcentrifuge for five minutes at 4˚C. For detection of H3K27 di- and tri-methylation (H3K27me2me3; Active Motif 39535) 1 µl of the relevant antibody was used. Protein A/G beads (20 µl) (Santa Cruz, cat. # sc-2003) were added to each sample. Following overnight incubation, beads were washed twice with 1 mL lysis buffer without protease inhibitors, once with lysis buffer containing 500mM NaCl, once with lysis buffer containing 50mM LiCl, and finally with TE (10mM Tris-HCl, 1mM EDTA). Bound chromatin was eluted in TES (50mM Tris pH 8.0, 10mM EDTA, 10% SDS) at 65˚C for 10 minutes. Chromatin was de-crosslinked overnight at 65˚C. The DNA was treated with RNase A for two hours at 50˚C, then with proteinase K for two hours at 50˚C and extracted using phenol-chloroform-isoamyl alcohol (25:24:1) followed by chloroform extraction. DNA pellets were washed with 70% ethanol and resuspended in TE buffer. Samples were then prepared for Illumina sequencing or ChIP-qPCR.

ChIP-qPCR

ChIP-qPCR was carried out with primers to three genic regions (NCU04707, NCU09834, NCU04877) that are normally marked by H3K27me2/3, and a control region of euchromatin LL#13, LL#14, primers are listed in Table 5.3. Primers were designed to generate ~100bp amplicons. Input samples were diluted 1:50 and IP samples were diluted 1:10 with sterile distilled water prior to qPCR. 10 µL reactions were assembled for each well were assembled as follows: 5ul Biorad SYBR Green mastermix, 0.5 µL Forward Primer, 0.5 µL Reverse Primer, 2 µL distilled sterile water and 2 µL of input or IP as template. qPCR was carried out on the Biorad C1000 thermocycler as previously described (36).

ChIP library preparation

Libraries were constructed as described (34-36). In brief, the NEBNext Ultra II End Repair/dA-tailing Module (cat. # E7546S), NEBNext Ultra II Ligation Module (cat. # E7546) were used to clean and A-tail DNA after which Illumina adapters were ligated. The ligation products were amplified to generate dual-indexed libraries using NEBNext Ultra II Q5 Hot Start HiFi PCR Master Mix (cat. # M0543S). Size selection with magnetic beads was performed after the adapter ligation and PCR steps with Sera-Mag SpeedBeads (cat. # 65152105050250) suspended in a solution of (20mM PEG 8000, 1mM NaCl, 10mM Tris-HCl, 1mM EDTA) (37).

Data Analysis

For ChIP-seq data, short reads (<20 bp) and adaptor sequences were removed using TrimGalore (version 0.4.4), cutadapt version 1.14 (38), and Python 2.7.8, with fastqc command (version 0.11.3). Trimmed Illumina reads were aligned to the current *N. crassa* NC12 genome assembly available from NCBI (accession # GCA_000182925.2) using the BWA (version 0.7.15) (39), mem algorithm, which randomly assign multi-mapped reads to a single location. Files were sorted and indexed using SAMtools (version 1.9) (40). To plot the relative distribution of mapped reads, read counts were determined for each 25-bp window across the genome using igvtools and data were displayed using the Integrated Genome Viewer (version 2.3.72) (41). The Hypergeometric Optimization of Motif EnRichment (HOMER) software package (version 4.8) (42) was used to identify H3K27me3 peaks in wild type and Δ*rtt109* against input using “findPeaks.pl” with the following parameters: -style histone.

References

1. D'Arcy S, Luger K. 2011. Understanding histone acetyltransferase Rtt109 structure and function: how many chaperones does it take? Curr Opin Struct Biol 21:728-34.10.1016/j.sbi.2011.09.005

2. Fillingham J, Recht J, Silva AC, Suter B, Emili A, Stagljar I, Krogan NJ, Allis CD, Keogh MC, Greenblatt JF. 2008. Chaperone control of the activity and specificity of the histone H3 acetyltransferase Rtt109. Mol Cell Biol 28:4342-53.10.1128/MCB.00182-08

3. Lee KK, Workman JL. 2007. Histone acetyltransferase complexes: one size doesn't fit all. Nat Rev Mol Cell Biol 8:284-95.10.1038/nrm2145

4. Scholes DT, Banerjee M, Bowen B, Curcio MJ. 2001. Multiple regulators of Ty1 transposition in Saccharomyces cerevisiae have conserved roles in genome maintenance. Genetics 159:1449-65

5. Chang M, Bellaoui M, Boone C, Brown GW. 2002. A genome-wide screen for methyl methanesulfonate-sensitive mutants reveals genes required for S phase progression in the presence of DNA damage. Proc Natl Acad Sci U S A 99:16934-9.10.1073/pnas.262669299

6. Bennett CB, Lewis LK, Karthikeyan G, Lobachev KS, Jin YH, Sterling JF, Snipe JR, Resnick MA. 2001. Genes required for ionizing radiation resistance in yeast. Nat Genet 29:426-34.10.1038/ng778

7. J S, P B, FC J, SR B, A. S. 2006. Rtt109 is required for proper H3K56 acetylation: a chromatin mark associated with the elongating RNA polymerase II. Journal of Biological Chemistry 281:37270-37274

8. Driscoll R, Hudson A, Jackson SP. 2007. Yeast Rtt109 promotes genome stability by acetylating histone H3 on lysine 56. Science 315:649-52.10.1126/science.1135862

9. Han J, Zhou H, Horazdovsky B, Zhang K, Xu RM, Zhang Z. 2007. Rtt109 acetylates histone H3 lysine 56 and functions in DNA replication. Science 315:653-5.10.1126/science.1133234

10. Verdone L, Caserta M, Di Mauro E. 2005. Role of histone acetylation in the control of gene expression. Biochem Cell Biol 83:344-53.10.1139/o05-041

11. Xu F, Zhang K, Grunstein M. 2005. Acetylation in histone H3 globular domain regulates gene expression in yeast. Cell 121:375-85.10.1016/j.cell.2005.03.011

12. Masumoto H, Hawke D, Kobayashi R, Verreault A. 2005. A role for cell-cycle-regulated histone H3 lysine 56 acetylation in the DNA damage response. Nature 436:294-8.10.1038/nature03714

13. Xu F, Zhang Q, Zhang K, Xie W, Grunstein M. 2007. Sir2 deacetylates histone H3 lysine 56 to regulate telomeric heterochromatin structure in yeast. Mol Cell 27:890-900.10.1016/j.molcel.2007.07.021

14. Lin LJ, Schultz MC. 2011. Promoter regulation by distinct mechanisms of functional interplay between lysine acetylase Rtt109 and histone chaperone Asf1. Proc Natl Acad Sci U S A 108:19599-604.10.1073/pnas.1111501108

15. Muller J. 1995. Transcriptional silencing by the Polycomb protein in Drosophila embryos. EMBO J 14:1209-20

16. Hennig L, Derkacheva M. 2009. Diversity of Polycomb group complexes in plants: same rules, different players? Trends Genet 25:414-23.10.1016/j.tig.2009.07.002

17. Simon JA, Kingston RE. 2009. Mechanisms of polycomb gene silencing: knowns and unknowns. Nat Rev Mol Cell Biol 10:697-708.10.1038/nrm2763

18. Schuettengruber B, Bourbon HM, Di Croce L, Cavalli G. 2017. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. Cell 171:34-57.10.1016/j.cell.2017.08.002

19. Kuroda MI, Kang H, De S, Kassis JA. 2020. Dynamic Competition of Polycomb and Trithorax in Transcriptional Programming. Annu Rev Biochem doi:10.1146/annurev-biochem-120219-103641.10.1146/annurev-biochem-120219-103641

20. Cao R, Wang L, Wang H, Xia L, Erdjument-Bromage H, Tempst P, Jones RS, Zhang Y. 2002. Role of histone H3 lysine 27 methylation in Polycomb-group silencing. Science 298:1039-43.10.1126/science.1076997

21. Czermin B, Melfi R, McCabe D, Seitz V, Imhof A, Pirrotta V. 2002. Drosophila enhancer of Zeste/ESC complexes have a histone H3 methyltransferase activity that marks chromosomal polycomb sites. Cell 111:185-196.Doi 10.1016/S0092-8674(02)00975-3

22. Kuzmichev A, Nishioka K, Erdjument-Bromage H, Tempst P, Reinberg D. 2002. Histone methyltransferase activity associated with a human multiprotein complex containing the Enhancer of Zeste protein. Genes Dev 16:2893-905.10.1101/gad.1035902

23. Muller J, Hart CM, Francis NJ, Vargas ML, Sengupta A, Wild B, Miller EL, O'Connor MB, Kingston RE, Simon JA. 2002. Histone methyltransferase activity of a Drosophila Polycomb group repressor complex. Cell 111:197-208.10.1016/s0092-8674(02)00976-5

24. Wan J, Zhan J, Li S, Ma J, Xu W, Liu C, Xue X, Xie Y, Fang W, Chin YE, Zhang H. 2015. PCAF-primed EZH2 acetylation regulates its stability and promotes lung adenocarcinoma progression. Nucleic Acids Res 43:3591-604.10.1093/nar/gkv238

25. Tang Y, Holbert MA, Wurtele H, Meeth K, Rocha W, Gharib M, Jiang E, Thibault P, Verreault A, Cole PA, Marmorstein R. 2008. Fungal Rtt109 histone acetyltransferase is an unexpected structural homolog of metazoan p300/CBP. Nat Struct Mol Biol 15:738-45.10.1038/nsmb.1448

26. Trojer P, Reinberg D. 2007. Facultative heterochromatin: Is there a distinctive molecular signature? Molecular Cell 28:1-13.10.1016/j.molcel.2007.09.011

27. Basenko EY, Sasaki T, Ji LX, Prybol CJ, Burckhardt RM, Schmitz RJ, Lewis ZA. 2015. Genome-wide redistribution of H3K27me3 is linked to genotoxic stress and defective growth. Proceedings of the National Academy of Sciences of the United States of America 112:E6339-E6348.10.1073/pnas.1511377112

28. Jamieson K, Rountree MR, Lewis ZA, Stajich JE, Selker EU. 2013. Regional control of histone H3 lysine 27 methylation in Neurospora. Proc Natl Acad Sci U S A 110:6027-32.10.1073/pnas.1303750110

29. Courtney AJ, Kamei M, Ferraro AR, Gai K, He Q, Honda S, Lewis ZA. 2020. Normal Patterns of Histone H3K27 Methylation Require the Histone Variant H2A.Z in Neurospora crassa. Genetics doi:10.1534/genetics.120.303442.10.1534/genetics.120.303442

30. Zhang Z, Yang Q, Sun G, Chen S, He Q, Li S, Liu Y. 2014. Histone H3K56 acetylation is required for quelling-induced small RNA production through its role in homologous recombination. J Biol Chem 289:9365-71.10.1074/jbc.M113.528521

31. Davis R, de Serres F. 1970. Genetic and microbiological research techniques for Neurospora crassa. Methods in Enzymology 17:79-143.https://doi.org/10.1016/0076-6879(71)17168-6

32. Margolin BS, Freitag, M., and Selker, E.U. 1997. Improved plasmids for gene targeting at the his-3 locus of Neurospora crassa by electroporation. Fungal Genetics Newsletter 44:2

33. Avalos J, Geever RF, Case ME. 1989. Bialaphos resistance as a dominant selectable marker in Neurospora crassa. Curr Genet 16:369-72.10.1007/bf00340716

34. Ferraro AR, Lewis ZA. 2018. ChIP-Seq Analysis in Neurospora crassa. Fungal Genomics: Methods and Protocols, Second Edition 1775:241-250.10.1007/978-1-4939-7804-5_19

35. Seymour M, Ji L, Santos AM, Kamei M, Sasaki T, Basenko EY, Schmitz RJ, Zhang X, Lewis ZA. 2016. Histone H1 Limits DNA Methylation in Neurospora crassa. G3 (Bethesda) 6:1879-89.10.1534/g3.116.028324

36. Sasaki T, Lynch KL, Mueller CV, Friedman S, Freitag M, Lewis ZA. 2014. Heterochromatin controls gammaH2A localization in Neurospora crassa. Eukaryot Cell 13:990-1000.10.1128/EC.00117-14

37. Rohland N, Reich D. 2012. Cost-effective, high-throughput DNA sequencing libraries for multiplexed target capture. Genome Research 22:939-946.10.1101/gr.128124.111

38. Martin M. 2011. Cutadapt Removes Adapter Sequences From High-Throughput Sequencing Reads. EMBnet 17:2.https://doi.org/10.14806/ej.17.1.200

39. Li H, Durbin R. 2009. Fast and accurate short read alignment with Burrows-Wheeler transform. Bioinformatics 25:1754-60.10.1093/bioinformatics/btp324

40. Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, Proc GPD. 2009. The Sequence Alignment/Map format and SAMtools. Bioinformatics 25:2078-2079.10.1093/bioinformatics/btp352

41. Thorvaldsdottir H, Robinson JT, Mesirov JP. 2013. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. Brief Bioinform 14:178-92.10.1093/bib/bbs017

42. Heinz S, Benner C, Spann N, Bertolino E, Lin YC, Laslo P, Cheng JX, Murre C, Singh H, Glass CK. 2010. Simple combinations of lineage-determining transcription factors prime cis-regulatory elements required for macrophage and B cell identities. Mol Cell 38:576-89.10.1016/j.molcel.2010.05.004

Figures

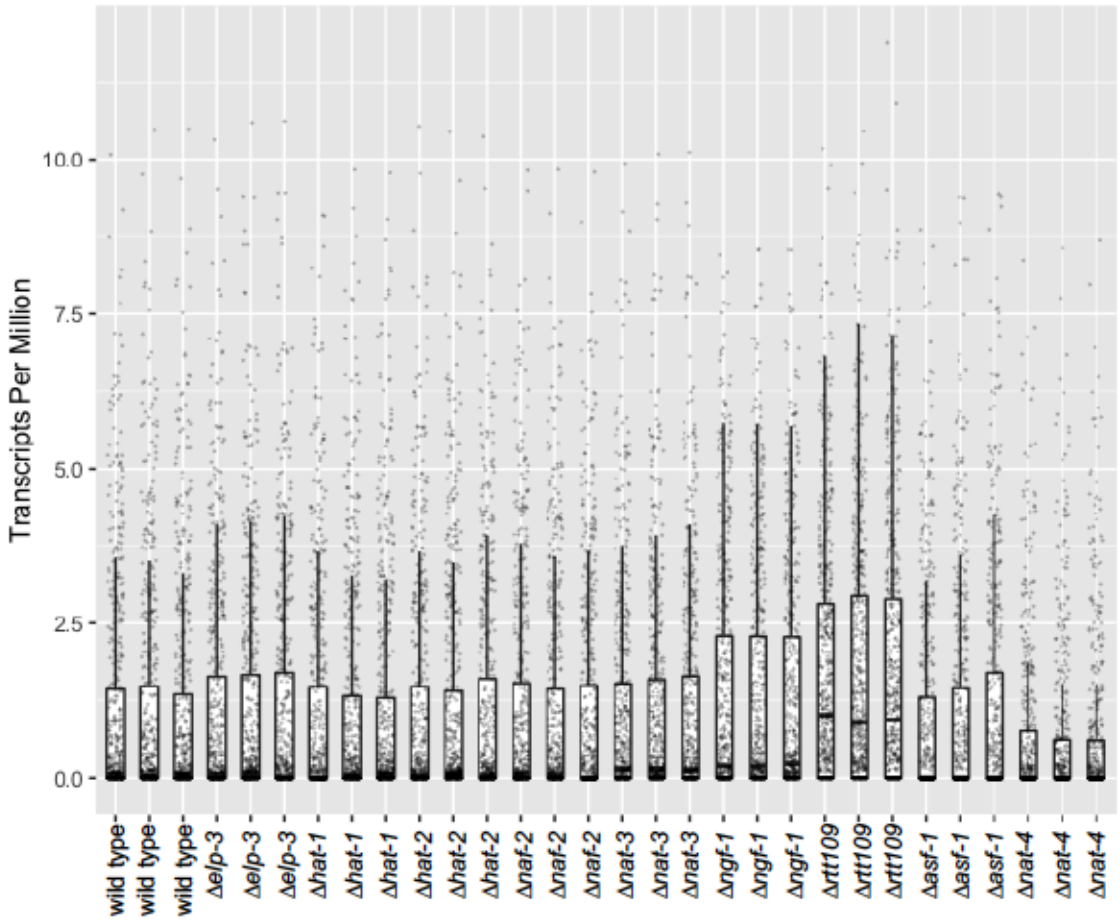


Figure 5.1: Lysine acetyltransferase expression of PRC2-target genes

Boxplot showing transcripts per million values for three replicates each of multiple putative and known lysine acetyltransferases reveals upregulation of PRC2-target genes in Δ*rtt109.* In wildtype, expression levels are near 0 and only Δ*rtt109* shows marked upregulation of these genes.

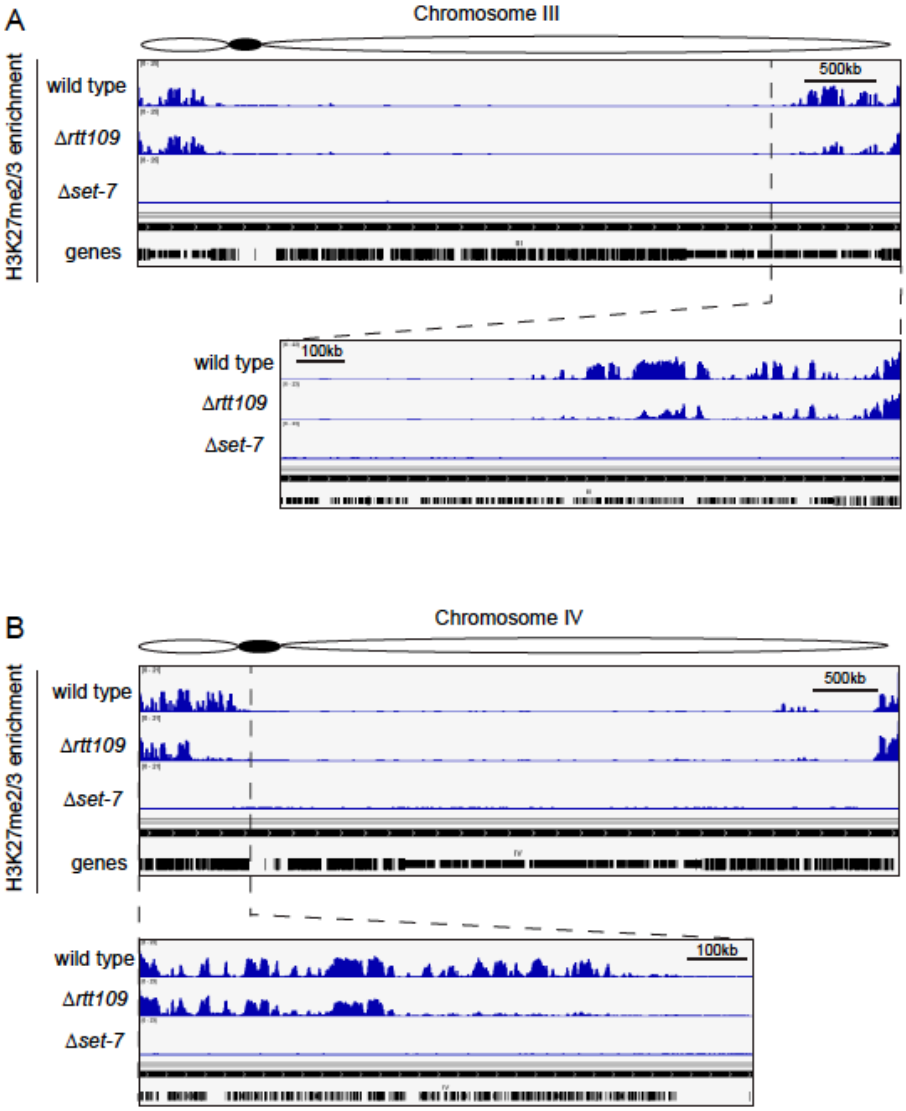


Figure 5.2: Selective H3K27me2/3 loss in Δ*rtt109* strain

A) IGV genome browser screenshot showing ChIP-seq for H3K27me2/3 for the entire length of chromosome III. The bottom panel shows a higher resolution view of the selective losses of H3K27me2/3 domains on the long arm of the chromosome.

B) IGV genome browser screenshot showing ChIP-seq for H3K27me2/3 for the entire length of chromosome IV. The bottom panel shows a higher resolution view of the selective losses of H3K27me2/3 domains on the short arm of the chromosome.

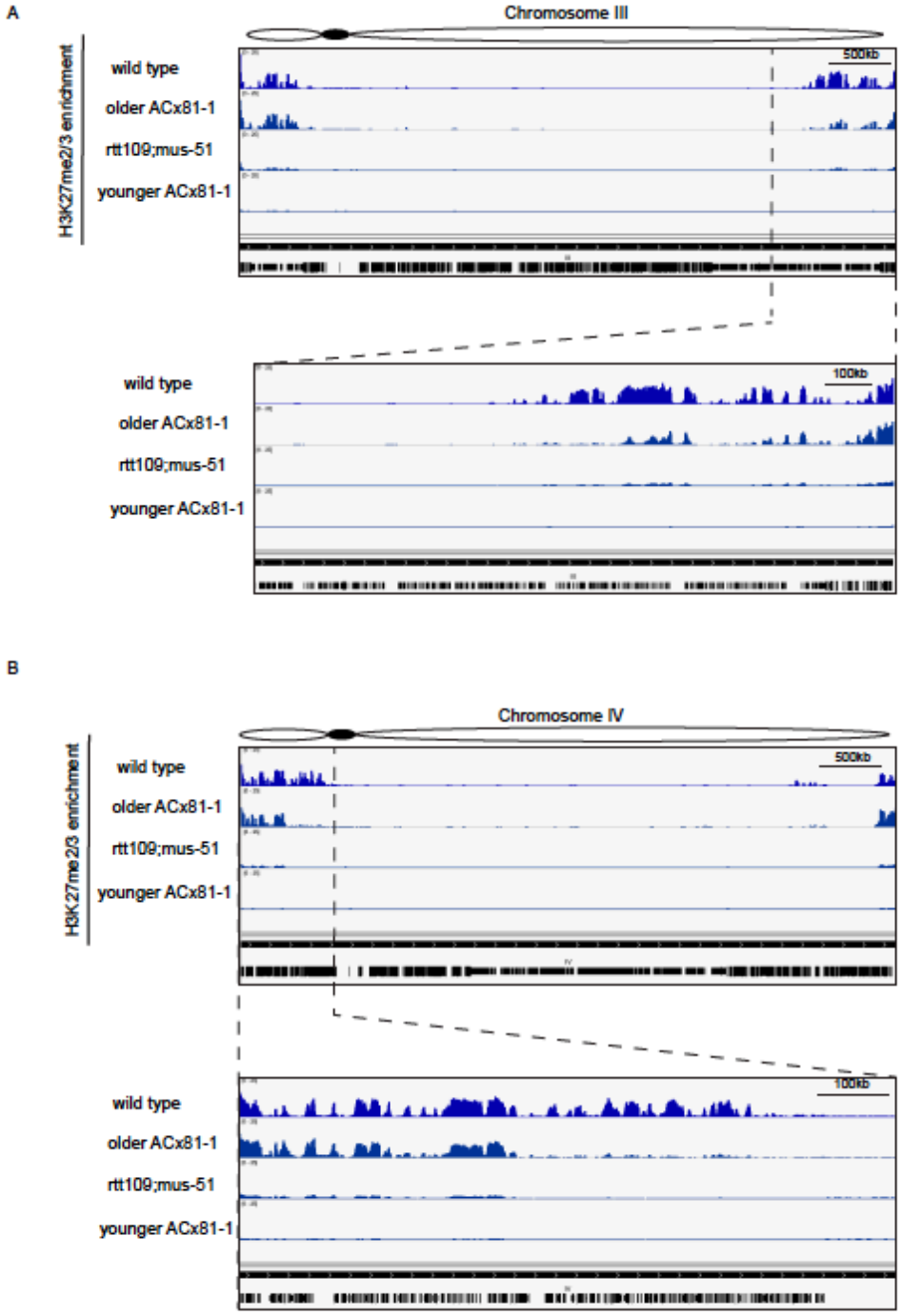
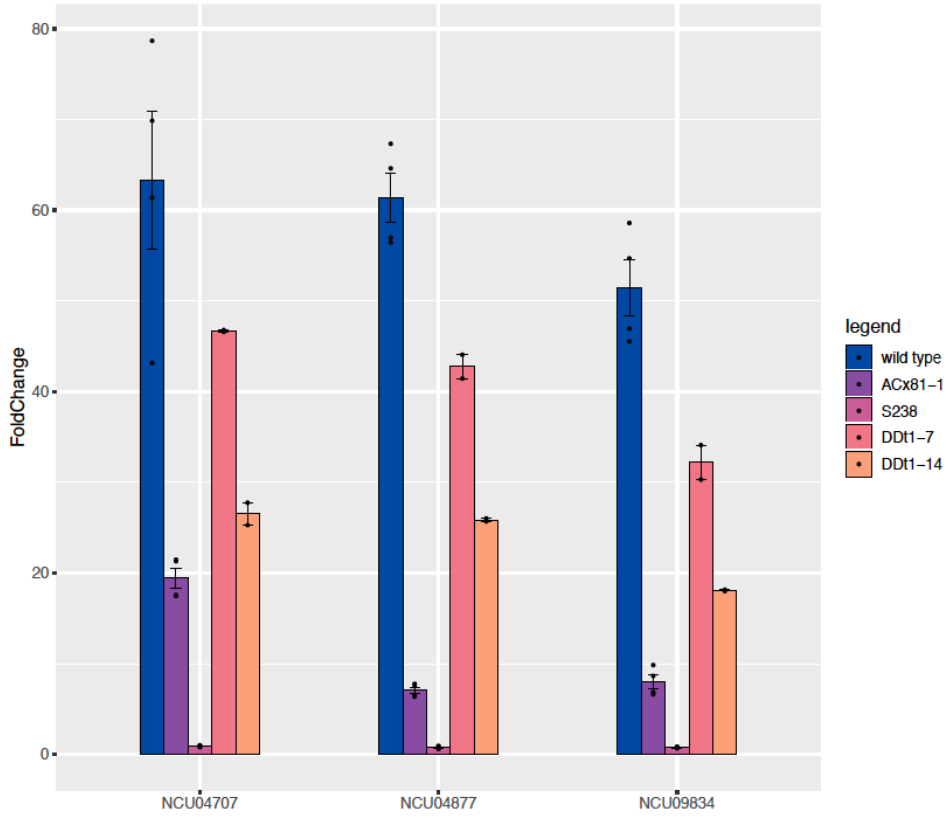

A
Chromosome III
H3K27me2/3 enrichment
wild type
older ACx81-1
rtt109;mus-51
younger ACx81-1
500kb
100kb
B
Chromosome IV
H3K27me2/3 enrichment
wild type
older ACx81-1
rtt109;mus-51
younger ACx81-1
500kb
100kb

Figure 5.3: Replicates Δ*rtt109*;Δ*mus51* and Δ*rtt109* strains show variable H3K27me2/3 depletion

A) IGV genome browser screenshot showing ChIP-seq for H3K27me2/3 for the entire length of chromosome III. The bottom panel shows a higher resolution view of the variable losses of H3K27me2/3 domains on the long arm of the chromosome.

B) IGV genome browser screenshot showing ChIP-seq for H3K27me2/3 for the entire length of chromosome IV. The bottom panel shows a higher resolution view of the variable losses of H3K27me2/3 domains on the short arm of the chromosome.

A


80
60
40
20
0
FoldChange
NCU04707
NCU04877
NCU09834
legend
wild type
ACx81−1
S238
DDt1−7
DDt1−14

B

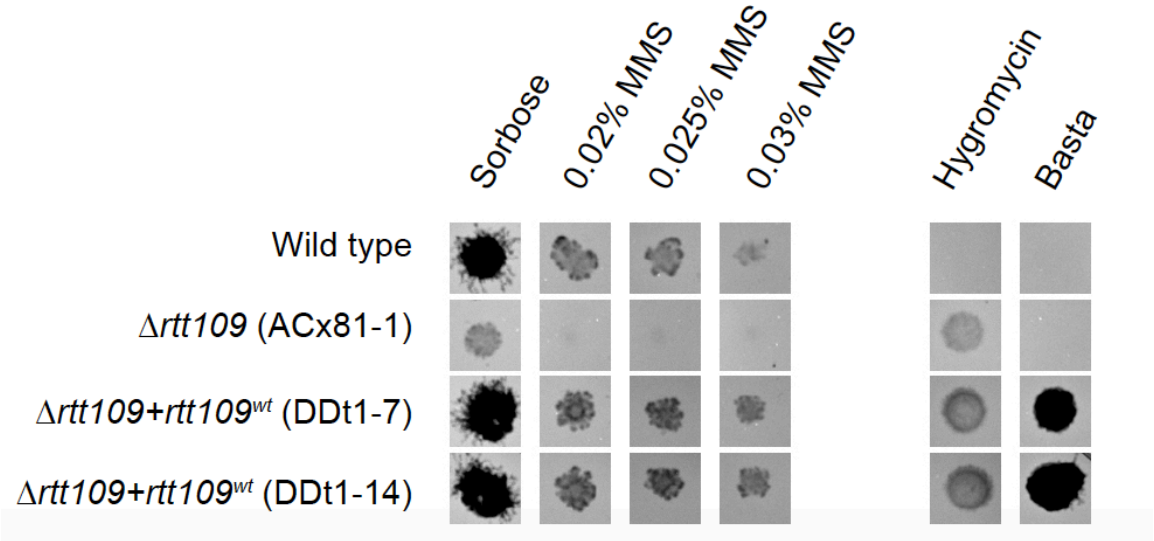

Sorbose
0.02% MMS
0.025% MMS
0.03% MMS
Hygromycin
Basta
Wild type
Δrtt109 (ACx81-1)
Δrtt109+rtt109wt (DDt1-7)
Δrtt109+rtt109wt (DDt1-14)

Figure 5.4: Complementation of Δ*rtt109* with ectopic *rtt109* fragment mostly restores H3K27me2/3 and rescues MMS sensitivity

A) ChIP-qPCR results for three replicates of wild type, two replicates of Δ*rtt109* (ACx81-1 older strain), Δset-7 (S238), and two replicates of two complemented primary transformants Δ*rtt109*+*rtt109*^{wt} (DDt1-7, DDt1-14). Δrtt109 has a dramatic loss at the three loci tested and complemented strains show higher enrichment of H3K27me2/3 than the deletion strain, although they are not completely restored to wild type levels.

B) Spot tests on plates containing different concentrations of MMS demonstrate complementation of rtt109 rescues the MMS sensitivity of the deletion strain.

Tables

Table 5.1: Comparison of average TPM values from three replicates for PRC2 components

Gene	Average TPM in Δ*rtt109*	Average TPM in wild type
eed	3.820	3.348
set-7	3.085	3.631
cac-3	44.224	53.134
suz12	7.548	8.253
hh2az	75.663	68.689

Table 5.2: Strains used in this study

Genotype	Strain #	Additional information	Reference
Δ*rtt109::hph;*Δ*mus-51::bar*	NCU09825	FGSC12340	(1)
Δ*rtt109::hph*	ACx81-1	progeny from cross of NCU09825 x S1	This study
Δ*rtt109::hph;rtt109+*	DDt1-7	Primary transformant *rtt109* complementation	This study
Δ*rtt109::hph;rtt109+*	DDt1-14	Primary transformant *rtt109* complementation	This study

Wild type	S2	FGSC2489	(1)
Wild type	S1	FGSC4200	(1)
Δ*set-7::hph*	S238	NCU07496 (FGSC11182)	(1)

1. Colot, H. V., G. Park, G. E. Turner, C. Ringelberg, C. M. Crew *et al.*, 2006 A high throughput gene knockout procedure for Neurospora reveals functions for multiple transcription factors. Proc Natl Acad Sci U S A 103**:** 10352-10357.

Table 5.3: Primers used in this study

Name	**Sequence**
NCU04707rev	TGACGAATCAGGGTTTG
NCU04707fwd	GAGCTCGAGAACACACAGATAC
NCU09834rev	TGGTACAGATGTCATTGTCTGG
NCU09834fwd	GAATTCGGACGAGGAAGAAGAG
NCU04877rev	GTTGAGGAACAAATGGCAAGAG
NCU04877fwd	GACGCCGACACAGTTATTCT
LL 13	CACACTTCACTTGGTCGCA
LL 14	ATGACCCCTTGATGTGATGTA
LL 148	CCGTCGACAGAAGATGATATTGAAGGAGC
LL 149	AATTAACCCTCACTAAAGGGAACAAAAGC
NCU09825_3FComp	ATAGGACGGCGGGTTGTATG
NCU09825_5Rcomp	ACTGGACTGGACTGGATGGA

CHAPTER 6

DISCUSSION

The work presented in this dissertation represents meaningful contributions to understanding the function and control of facultative heterochromatin in filamentous fungi. First, detailed analyses showed how facultative heterochromatin regulates effector gene expression in *Magnaporthe oryzae,* an economically important pathogen of rice. Additionally, we identified similar expression patterns between the overexpression of transcription factor *MoGti1* to post-infection microarray data and found a synergistic effect on effector gene expression when both H3K27me3 was removed and *MoGti1* was overexpressed. This is the first investigation to show combinatorial regulation of effector genes by the Polycomb Repressive Complex 2 and a transcription factor in a filamentous fungus. Next, my work has generated a metric to determine the extent of conditional gene expression based on entropy. This metric can be used to identify and/or validate constitutive and inducible genes using a quantitative approach. In the process of creating this metric we have demonstrated for the first time that expression of PRC2-target genes are induced during specific stages of sexual development in the fungus *N. crassa.* We next investigated the regulation of facultative heterochromatin using *N. crassa* as a model organism. We found that deletion of H2A.Z leads to region-specific depletion of H3K27me3 in *N. crassa,* and we determined that the histone variant H2A.Z regulates facultative heterochromatin indirectly by controlling *eed* expression. Finally, we have

begun the characterization of an acetyltransferase mutant, *rtt109*, which causes broad upregulation of PRC2-target genes with variable losses in H3K27me2/3.

The histone modification H3K27me3 and transcription factor MoGti1 coordinately control expression of effector genes in rice blast fungus

Understanding the biological impacts of facultative heterochromatin on *Magnaporthe oryzae* has far reaching implications in potential plant pathogen control. The destruction of cereal crops by *M. oryzae* affects more than 60 million people annually and has been a historically difficult disease to control (1). Elucidation of basic mechanisms governing how effector genes, the genes that are responsible for hijacking the plant's immune response, will prove to be valuable in designing crop protection strategies (2) (3). Effector genes are in a repressed state during mycelial growth on plates and have rapid transcriptional changes upon plant infection (4-9). We wanted to know if previous reports of effector gene regulation by H3K27me3 (10), was also true in *M. oryzae*. Collaboratively, we determined that most previously published and experimentally validated known *M. oryzae* effector genes were located in regions of the genome enriched for H3K27me3 and removal of this modification altered their expression patterns to mimic post-infection expression patterns. Not only did removal of H3K27me3 upregulate effector genes that are normally silent in axenic culture, but it also attenuated the virulence of this pathogen when tested on rice leaves. In addition, we found the transcription factor *MoGti1* has transcriptional consequences to effector genes when overexpressed. Some effector genes are upregulated, and some are downregulated. The mis-regulation of these genes is amplified when the overexpression of MoGti1 and the deletion of the methyltransferase for H3K27me3 are combined in one strain. This

demonstrates that the mechanisms governing expression of effector genes are not found in a single pathway. We discovered that there are at least two layers of regulation, one epigenetic modification and one transcription factor, that activate effectors, which suggests the proper transcriptional regulation of these genes is remarkably important in pathogenic fungi. Additionally, these data demonstrate the complexities of Polycomb repression in this filamentous fungus and how Polycomb regulated genes contribute to the infection life cycle of this devastating plant pathogen.

Shannon entropy as a metric for conditional gene expression in *Neurospora crassa*

In order to understand more about the expression changes of these conditionally repressed genes, we created a metric by which all genes in the genome were classified by their expression variation across many conditions. By gathering public datasets of RNA-seq data performed under numerous conditions we generated an entropy value for most genes in the *Neurospora* genome. This is the first quantitative analysis of expression variation for single genes in filamentous fungi and these data can be used to validate housekeeping genes for normalization in qRT-PCR as well as in selecting a constitutive or inducible promoters for synthetic biology and molecular genetics approaches. Thus, this project provides a community resource of entropy values for *Neurospora* researchers to use. In addition, including the pipeline used to create the entropy values allows others the ability to create their own values for another organism, or to produce new values for *Neurospora* by adding their own datasets to the list we provided.

An obvious question arose while working on this project: what expression patterns do the conditionally repressed genes, PRC2-target genes, have across all the conditions collected? I plotted TPM values for all PRC2-target genes for all conditions

and found a group of genes that was upregulated during specific stages of sexual development. A particularly important open question is when establishment of H3K27 methylation takes place in *N. crassa*. This is the first report of PRC2-target gene expression changes during sexual development in *N. crassa*, which could be an indication of this establishment taking place during sexual development.

Loss of *set-7* has been shown by our lab to promote sexual development in a Δ*dim-5* background. In contrast, homozygous crosses of Δ*dim-5* produce barren perithecia. (11). In a recent study, the loss of *epr-1,* effector of Polycomb repression 1*,* was also shown to promote premature sexual development on unfertilized plates (12). *Epr-1* is a homolog to ebs and shl in plants and these genes have been linked to gene repression and to binding H3K27 methylation (12). Deletion of *set-7* results in loss of H3K27me2/3 throughout the genome (13) and deletion of *epr-1* results in upregulation of PRC2-target genes without the loss of H3K27 methylation (12). However, both deletions cause premature sexual development (12). My work ties these findings together showing that PRC2-target genes exhibit de-repression during specific stages of sexual development, which suggests that cell type transitions, even in *Neurospora*, are governed by Polycomb group proteins. In *Arabidopsis thaliana,* in the fertilization-independent seed (FIS) mutants developing seeds can be found even when the flowers do not possess pollen, suggesting premature sexual development (14). The FIS proteins are Polycomb group proteins and possess homology to SUZ12 (15). Polycomb group proteins are also critically important during development in metazoans (16). This demonstrates a conserved mechanism for Polycomb group proteins in determining cell type/fate that would have existed in the ancestor to plants, fungi, and animals.

Normal patterns of histone H3K27 methylation require the histone variant H2A.Z

There are additional important open questions about the control of H3K27 methylation in *Neurospora*. H2A.Z, the most well studied histone variant, has been proposed to be important for Polycomb repression in other organisms (17-19). I investigated the histone variant H2A.Z and its role in Polycomb regulated repression in *Neurospora*. My work uncovered that H2A.Z is critical for the proper expression of EED, a core subunit of PRC2. This regulation of expression ensures that the proper patterns of H3K27me2/3 are formed in specific regions of the genome, namely internal regions, or regions 200kb away from the chromosome ends. In addition, the genes found in the internal regions exhibit upregulation, which was not fully expected since the removal of H3K27me2/3 alone is not sufficient to upregulate all PRC2-target genes (13). This adds to the collective knowledge of PRC2, by revealing that specific regions of the genome are more dependent on EED related mechanisms than other regions of the genome.

RTT109 is required for normal patterns of H3K27 methylation in *Neurospora crassa*

Characterization of a second mutant strain, *rtt109*, shows upregulation of PRC2-target genes and related depletion signatures to the H2A.Z mutant strain with internal depletion and retention at telomere-proximal sites. Conflicting H3K27me2/3 ChIP-seq replicates have exposed many unanswered questions. The depletion observed in the replicates could be due to an unknown factor RTT109 is interacting with under certain conditions. The absence of *rtt109* is causative for upregulation of PRC2-target genes, and whether that is direct or indirect is an open question. Recently, *epr-1*, was found to upregulate H3K27me2/3 genes without the loss of H3K27me2/3 (12). All PRC2 components have normal expression in this deletion strain and the loss of H3K27me2/3

in certain regions of the genome presents a more puzzling question and is currently more difficult to draw conclusions from. It is quite possible that this situation is similar to the H2A.Z story where the presence of RTT109 regulates an upstream regulator of PRC2, or even acetylates PRC2 itself. Future work is needed to elucidate the role of RTT109 in controlling facultative heterochromatin.

To determine if acetylation of lysine 56, a known substrate of RTT109, is causative for the loss of H3K27 methylation, I will perform ChIP-seq for H3K27me2/3 in a strain that has an arginine in place of the lysine 56 residue. This strain mimics an unmodified lysine residue and cannot be acetylated. If loss of this critical modification is important for the facultative heterochromatin pathway, this ChIP-seq experiment will phenocopy the losses exhibited in the Δ*rtt109* strain. If this experiment looks like wild type instead of phenocopying the deletion strain it would imply that RTT109 likely has another substrate that is of importance to this pathway. Other follow-up experiments include H3K27me2/3 ChIP-seq of the RTT109 chaperones Δ*asf-1* and Δ*vps-75* as well as acetylation assays to determine if RTT109 is in fact acetylating PRC2. These experiments are the first steps toward elucidating which pathway the depletion of H3K27me2/3 is connected to.

References

1. Fernandez J, Orth K. 2018. Rise of a Cereal Killer: The Biology of Magnaporthe oryzae Biotrophic Growth. Trends Microbiol 26:582-597.10.1016/j.tim.2017.12.007

2. Sanchez-Vallet A, Fouche S, Fudal I, Hartmann FE, Soyer JL, Tellier A, Croll D. 2018. The Genome Biology of Effector Gene Evolution in Filamentous Plant Pathogens. Annu Rev Phytopathol 56:21-40.10.1146/annurev-phyto-080516-035303

3. 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. Annu Rev Plant Biol 66:513-45.10.1146/annurev-arplant-043014-114623

4. Lanver D, Muller AN, Happel P, Schweizer G, Haas FB, Franitza M, Pellegrin C, Reissmann S, Altmuller J, Rensing SA, Kahmann R. 2018. The Biotrophic Development of Ustilago maydis Studied by RNA-Seq Analysis. Plant Cell 30:300-323.10.1105/tpc.17.00764

5. Gervais J, Plissonneau C, Linglin J, Meyer M, Labadie K, Cruaud C, Fudal I, Rouxel T, Balesdent MH. 2017. Different waves of effector genes with contrasted genomic location are expressed by Leptosphaeria maculans during cotyledon and stem colonization of oilseed rape. Mol Plant Pathol 18:1113-1126.10.1111/mpp.12464

6. Dong Y, Li Y, Zhao M, Jing M, Liu X, Liu M, Guo X, Zhang X, Chen Y, Liu Y, Liu Y, Ye W, Zhang H, Wang Y, Zheng X, Wang P, Zhang Z. 2015. Global genome and transcriptome analyses of Magnaporthe oryzae epidemic isolate 98-06 uncover novel effectors and pathogenicity-related genes, revealing gene gain and lose dynamics in genome evolution. PLoS Pathog 11:e1004801.10.1371/journal.ppat.1004801

7. Hacquard S, Kracher B, Maekawa T, Vernaldi S, Schulze-Lefert P, Ver Loren van Themaat E. 2013. Mosaic genome structure of the barley powdery mildew pathogen and conservation of transcriptional programs in divergent hosts. Proc Natl Acad Sci U S A 110:E2219-28.10.1073/pnas.1306807110

8. O'Connell RJ, Thon MR, Hacquard S, Amyotte SG, Kleemann J, Torres MF, Damm U, Buiate EA, Epstein L, Alkan N, Altmuller J, Alvarado-Balderrama L, Bauser CA, Becker C, Birren BW, Chen Z, Choi J, Crouch JA, Duvick JP, Farman MA, Gan P, Heiman D, Henrissat B, Howard RJ, Kabbage M, Koch C, Kracher B, Kubo Y, Law AD, Lebrun MH, Lee YH, Miyara I, Moore N, Neumann U, Nordstrom K, Panaccione DG, Panstruga R, Place M, Proctor RH, Prusky D, Rech G, Reinhardt R, Rollins JA, Rounsley S, Schardl CL, Schwartz DC, Shenoy N, Shirasu K, Sikhakolli UR, Stuber K, et al. 2012. Lifestyle transitions in plant pathogenic Colletotrichum fungi deciphered by genome and transcriptome analyses. Nat Genet 44:1060-5.10.1038/ng.2372

9. Kleemann J, Rincon-Rivera LJ, Takahara H, Neumann U, Ver Loren van Themaat E, van der Does HC, Hacquard S, Stuber K, Will I, Schmalenbach W, Schmelzer E, O'Connell RJ. 2012. Sequential delivery of host-induced virulence effectors by appressoria and intracellular hyphae of the phytopathogen Colletotrichum higginsianum. PLoS Pathog 8:e1002643.10.1371/journal.ppat.1002643

10. Connolly LR, Smith KM, Freitag M. 2013. The Fusarium graminearum histone H3 K27 methyltransferase KMT6 regulates development and expression of secondary metabolite gene clusters. PLoS Genet 9:e1003916.10.1371/journal.pgen.1003916

11. Basenko EY, Sasaki T, Ji LX, Prybol CJ, Burckhardt RM, Schmitz RJ, Lewis ZA. 2015. Genome-wide redistribution of H3K27me3 is linked to genotoxic stress and defective growth. Proceedings of the National Academy of Sciences of the United States of America 112:E6339-E6348.10.1073/pnas.1511377112

12. Wiles ET, McNaught KJ, Kaur G, Selker JML, Ormsby T, Aravind L, Selker EU. 2020. Evolutionarily ancient BAH-PHD protein mediates Polycomb silencing. Proc Natl Acad Sci U S A 117:11614-11623.10.1073/pnas.1918776117

13. Jamieson K, Rountree MR, Lewis ZA, Stajich JE, Selker EU. 2013. Regional control of histone H3 lysine 27 methylation in Neurospora. Proc Natl Acad Sci U S A 110:6027-32.10.1073/pnas.1303750110

14. Chaudhury AM, Ming L, Miller C, Craig S, Dennis ES, Peacock WJ. 1997. Fertilization-independent seed development in Arabidopsis thaliana. Proc Natl Acad Sci U S A 94:4223-8.10.1073/pnas.94.8.4223

15. Chanvivattana Y, Bishopp A, Schubert D, Stock C, Moon YH, Sung ZR, Goodrich J. 2004. Interaction of Polycomb-group proteins controlling flowering in Arabidopsis. Development 131:5263-76.10.1242/dev.01400

16. Boyer LA, Plath K, Zeitlinger J, Brambrink T, Medeiros LA, Lee TI, Levine SS, Wernig M, Tajonar A, Ray MK, Bell GW, Otte AP, Vidal M, Gifford DK, Young RA, Jaenisch R. 2006. Polycomb complexes repress developmental regulators in murine embryonic stem cells. Nature 441:349-53.10.1038/nature04733

17. Carter B, Bishop B, Ho KK, Huang R, Jia W, Zhang H, Pascuzzi PE, Deal RB, Ogas J. 2018. The Chromatin Remodelers PKL and PIE1 Act in an Epigenetic Pathway That Determines H3K27me3 Homeostasis in Arabidopsis. Plant Cell 30:1337-1352.10.1105/tpc.17.00867

18. Creyghton MP, Markoulaki S, Levine SS, Hanna J, Lodato MA, Sha K, Young RA, Jaenisch R, Boyer LA. 2008. H2AZ is enriched at polycomb complex target genes in ES cells and is necessary for lineage commitment. Cell 135:649-61.10.1016/j.cell.2008.09.056

19. Surface LE, Fields PA, Subramanian V, Behmer R, Udeshi N, Peach SE, Carr SA, Jaffe JD, Boyer LA. 2016. H2A.Z.1 Monoubiquitylation Antagonizes BRD2 to Maintain Poised Chromatin in ESCs. Cell Rep 14:1142-1155.10.1016/j.celrep.2015.12.100

APPENDIX A

THE MATERNAL TO ZYGOTIC TRANSITION REGULATES GENOME-WIDE HETEROCHROMATIN ESTABLISHMENT IN THE ZEBRAFISH EMBRYO[5]

[5] Kathrin Laue, Srivarsha Rajshekar, Abigail J. Courtney, Zachary A. Lewis & Mary G. Goll. 2019. *Nature Communications.* doi: 10.1038/s41467-019-09582-3. Springer Nature. Reprinted here with the permission of the publisher.

Abstract

The segregation of eukaryotic genomes into euchromatin and heterochromatin represents a fundamental and poorly understood process. Here, we demonstrate that genome-wide establishment of heterochromatin is triggered by the maternal to zygotic transition (MZT) during zebrafish embryogenesis. We find that prior to MZT, zebrafish lack hallmarks of heterochromatin including histone H3 lysine 9 trimethylation (H3K9me3) and condensed chromatin ultrastructure. Global establishment of heterochromatic features occurs following MZT and requires both activation of the zygotic genome and degradation of maternally deposited RNA. Mechanistically, we demonstrate that zygotic transcription of the micro RNA miR-430 promotes degradation of maternal RNA encoding the chromatin remodeling protein Smarca2, and that clearance of Smarca2 is required for global heterochromatin establishment in the early embryo. Our results identify MZT as a key developmental regulator of heterochromatin establishment during vertebrate embryogenesis and uncover a novel role for Smarca2 in protecting the embryonic genome against heterochromatinization.

Introduction

The segregation of eukaryotic genomes into regions of euchromatin and heterochromatin is fundamental to genome organization. At the molecular level, these domains are distinguished by different levels of chromatin compaction and unique sets of histone modifications. Highly condensed, constitutive heterochromatin is marked by trimethylation of histone H3 lysine 9 (H3K9me3) and is found predominately at repetitive sequences across the genome. Heterochromatin formation at these sequences promotes transcriptional repression, as well as genome stability, and depletion of

H3K9me3 marked heterochromatin severely impairs viability in mice, flies and zebrafish [1-3].

Although the timing and extent varies between species, developmental reprogramming of H3K9me3 marked heterochromatin has been noted in diverse metazoa including mammals, flies and C. elegans [4-9]. Heterochromatin establishment coincides with diminishing cellular plasticity in these species, and heterochromatin formation also represents a major barrier to cellular reprogramming [10, 11]. The inverse relationship between chromatin compaction and developmental potential suggests that the timing of heterochromatin formation must be tightly regulated during development [11, 12]. However, the mechanisms that control global establishment of heterochromatic states remain largely unknown, especially in the context of vertebrate embryogenesis.

In this study, we set out to characterize heterochromatin regulation in the context of early zebrafish development. We find that the zebrafish embryo undergoes at least 10-rounds of cell division in the absence of condensed chromatin ultrastructure. Global establishment of H3K9me3 and chromatin compaction is first noted following the maternal to zygotic transition (MZT) and the establishment of heterochromatin is dependent on this transition. At the molecular level, we show that zygotic transcription of the microRNA miR-430 promotes degradation of maternal RNA encoding the ATP dependent chromatin remodeler Smarca2, and that clearance of Smarca2 is required for H3K9me3 establishment and chromatin compaction following MZT. Our study identifies MZT as a key regulator of de novo heterochromatin establishment in the early zebrafish embryo and reveals a novel function for Smarca2 in antagonizing the de novo establishment of heterochromatin during early vertebrate embryogenesis.

Results

Chromatin of the early zebrafish embryo is globally deficient in H3K9me3

In order to characterize the heterochromatic compartment of the early zebrafish genome, we collected embryos during the period between early blastula (512-cell, 2.7 hpf), and mid gastrula (shield, 6 hpf) stages and assessed embryonic nuclei for H3K9me3 rich foci (**Fig. 1a**). These H3K9me3 marked foci, termed chromocenters serve as a classical cytological marker of constitutive heterochromatin, and are readily detected in mouse embryos as early as the two-cell stage [3]. We observed that while chromocenters were obvious in nuclei of shield stage zebrafish embryos (6 hpf), they were undetectable in nuclei at the 512-cell stage (2.7 hpf) (**Fig. 1b**). Faint, diffuse nuclear H3K9me3 labeling was first observed in embryos at oblong stage (3.7 hpf), and emergent foci were noted by dome stage (4.5 hpf) (**Fig. 1b**). H3K9me3 was also undetectable by western blot in bulk histones isolated at 4.0 hpf and earlier time points, but was abundant in dome (4.5 hpf) and shield stage (6 hpf) embryos **(Fig. 1c, Extended data 1a)**. Chromatin immunoprecipitation with sequencing (ChIP-Seq) supported these findings. At 6 hpf, regions enriched for H3K9me3 were detected across the genome, with ~96% of called H3K9me3 peaks overlapping annotated repeats (**Fig. 1d-e, Extended data 1b-d**). At 4.5 hpf, low amplitude peaks were observed within a subset of regions showing enrichment at 6 hpf. However, little signal was detected within these regions at 2.7 hpf. Enrichment of H3K9me3 at 4.5 and 6 hpf was observed at DNA, LINE, LTR and SINE transposons, with the most peaks detected in LTR transposons **(Extended data 1e).** The small fraction of peaks not corresponding to repetitive elements were found mainly in intergenic regions. Similar temporal enrichment of H3K9me3 was observed when H3K9me3

enrichment was directly examined at pericentromeric Satellite-1 (Sat1) repeats **(Fig. 1f)**. These findings suggest that the early zebrafish embryo undergoes many rounds of cell division in an environment that is globally depleted for H3K9me3 marked heterochromatin, and that a major wave of H3K9me3 establishment initiates around 3.7 hpf.

The early zebrafish embryo lacks condensed chromatin ultrastructure

To characterize embryonic heterochromatin at the ultrastructure level, we next turned to transmission electron microscopy (TEM) (**Fig. 2a-h and Extended data 2a**). As expected, electron dense aggregates indicative of condensed chromatin ultrastructure were clearly visible within nuclei from shield stage embryos (6 hpf) (**Fig. 2d, h**). However, at the 512-cell stage (2.7 hpf), these aggregates were undetectable (**Fig. 2a, e**). Aggregates were first noted in some embryonic nuclei at oblong stage (3.7 hpf), and appeared more common by dome stage (4.5 hpf) **(Fig. 2b-c, f-g)**. To quantify these observations, maximum entropy thresholding was used to define and count the number of electron dense particles relative to nuclear area in individual cells from three embryos per time point (**Fig. 2i)**[13]. At the 512-cell stage (2.7 hpf), electron dense aggregates exceeding a particle size of 0.03 μm^2 were not detected in embryonic nuclei, suggesting a lack of condensed ultrastructure. Significant increases in the number of nuclear aggregates per μm^2 and the percent nuclear area covered by aggregates were first noted at dome stage (4.5 hpf), and increased 7- and 9-fold by shield stage (6 hpf) **(Fig. 2j, k**). Consistent with these increases in chromatin compaction, we observed decreased expression of transcripts derived from repetitive elements between 4 and 6 hpf **(Extended data 2b-e).** These data indicate that the genome of the early zebrafish is packaged in an atypically decondensed

chromatin state, and that embryonic chromatin undergoes a profound reorganization involving the establishment of condensed chromatin ultrastructure between 3.7 and 6 hpf.

Blocking zygotic transcription impairs heterochromatin establishment

We noted that the timing of H3K9me3 establishment and chromatin compaction in the zebrafish embryo roughly coincided with the maternal to zygotic transition (MZT), which occurs during the 10th cell division in zebrafish (3 hpf). This critical transition in embryonic development marks the stage at which the zygotic genome is activated for the first time, and maternally deposited RNAs are degraded. To test whether blocking zygotic transcription would impact heterochromatin establishment during zebrafish embryogenesis, we injected embryos with the RNA polymerase II inhibitor α-amanatin at the 1-cell stage and assessed H3K9me3 levels at 4.5 hpf. At this time point, α-amanatin injected embryos exhibited strong reductions in total H3K9me3 levels and reduced H3K9me3 enrichment at pericentromeric repeats relative to controls (**Fig. 3a, b, Extended data 3a**). Similar results were obtained using triptolide, an unrelated inhibitor of RNA polymerase II dependent transcription (**Extended data 3b,c**). In the absence of zygotic transcription, electron dense aggregates indicative of condensed chromatin ultrastructure were also essentially undetectable in TEM images of nuclei from embryos at 5 hpf, whereas these aggregates were readily detected in mock-injected controls (**Fig. 3c-e**). These findings demonstrate that blocking zygotic transcription leads to impaired H3K9me3 establishment and chromatin compaction during early embryogenesis. This effect cannot be explained by regulation of H3K9 methyltransferase transcript levels, as RNAs encoding known H3K9 methyltransferase enzymes are present at comparable or slightly higher levels prior to zygotic genome activation (ZGA) (**Extended data 3d**).

The microRNA miR-430 is required for heterochromatin establishment

Next, we investigated whether degradation of maternal RNA was required for heterochromatin establishment during MZT. The microRNA miR-430 is required for the degradation of maternal transcripts during zebrafish MZT, and transcription of miR-430 is dependent on ZGA [14]. Antisense morpholino targeting miR-430 was injected into embryos at the 1-cell stage, and the impact on H3K9me3 establishment was then assessed. At 5hpf, embryos injected with miR-430 morpholino exhibited sustained expression of genes that are normally down regulated at MZT, suggesting the effectiveness of the morpholino (**Extended data 4a**). Injected embryos also exhibited lower global levels of H3K9me3 at 4.5 hpf when compared to siblings, as well as a derepression of transcripts from pericentromeric repeats (**Fig. 4a, c, Extended data 4b**). Similar results were obtained when we depleted the exoribonuclease Dicer, which is required for miR-430 processing **(Fig. 4b,c, Extended data 4c).** These findings suggest that miR-430 is important for the establishment of H3K9me3 at MZT and implicate the degradation of maternal mRNA in the developmental regulation of heterochromatin establishment.

miR-430 mediates clearance of maternal RNA encoding the chromatin remodeling protein Smarca2

To identify potentially relevant miR-430 targets, we turned to previous work by Giraldez et al., which identified the subset of maternal transcripts that are degraded by miR-430 [14]. Interrogation of this dataset revealed that *smarca2*, an ATP dependent chromatin remodeler, was among the five genes that showed the most significant elevation in transcript levels after miR-430 processing was blocked. We confirmed that

smarca2 3- UTR contained multiple miR-430 binding sites (**Extended data 4d**). Consistent with these findings, we found that *smarca2* transcripts were downregulated in wildtype embryos during MZT and that that following MZT, embryos injected with miR-430 or dicer morpholino had elevated levels of *smarca2* compared to mock-injected controls (**Fig 5a, b**). Downregulation of Smarca2 at MZT is also observed in published RNA-seq data (**Extended data Fig 5a)** [15].

Smarca2 degradation is required for H3K9me3 establishment during MZT

To determine whether prolonged expression of *smarca2* in the developing embryo would impact heterochromatin establishment, we injected 1-cell stage embryos with *Smarca2* mRNA lacking miR-430 recognition sites, and assayed H3K9me3 levels at dome stage (4.5 hpf). Consistent with a role for Smarca2 in inhibiting heterochromatin establishment, we found that, at dome stage, embryos that had been injected with miR-430 resistant RNA encoding *Smarca2* had reduced H3K9me3 levels compared to mock injected controls (**Fig 5c, Extended data 5b**). Conversely, knockdown of *smarca2* using two independently designed morpholinos was sufficient to accelerate the onset of H3K9me3 incorporation (**Fig 5d, Extended data 5c, d**). We observed strong H3K9me3 signal by western blot in morpholino injected embryos at 3.5 hpf, while H3K9me3 signal was not observed in wildtype embryos at this stage. Precocious H3K9me3 incorporation was associated with abnormal development at 24 hpf, using either of the two smarca2 morpholinos (**Extended data 6a-c**). All abnormal embryos were dead by 7 dpf.

Precocious accumulation of H3K9me3 was noted at pericentromeric Sat1 repeats by ChIP, suggesting that Smarca2 depletion allowed for early incorporation of H3K9me3 at sequences that were destined for later enrichment in wildtype embryos (**Fig 5e**).

Consistent with this model, genome wide ChIP-seq analysis of *smarca2* morpholino-injected embryos at 3.5 hpf identified 4,692 H3K9me3 peaks that were not present in control embryos, with roughly 90% of these precocious peaks overlapping annotated repeats **(Fig. 5f, g, Extended data 7a)**. ChIP-seq profiles from embryos injected with Smarca2 clustered with profiles from 4.5 and 6 hpf embryos, wherease controls were clustered with profiles from 2.7 hpf embryos (**Extended data 7b**).

To confirm these results using a morpholino independent approach, we inhibited Smarca2 function in the early embryo using the Bromo domain binding small molecule PFI-3 which can displace the Smarca2 Bromo domain from chromatin [16]. We found that PFI-3 injection led to precocious establishment of H3K9me3 and anormal development similar to that observed in morpholino experiments (**Fig 6a, Extended data 6a-c**). The effect on H3K9me3 was dose dependent (**Fig 6b and Extended data 5e**)

Smarca2 inhibition is sufficient to accelerate the establishment of condensed chromatin ultrastructure

Finally, we asked whether early establishment of H3K9me3 following inhibition of Smarca2 would be sufficient to promote the precocious establishment of condensed chromatin ultrastructure in the developing embryo. To this end, embryos were injected with PFI-3 at the 1-cell stage and embryonic nuclei were examined by transmission electron microscopy at 4 hpf. At this stage, we found that embryos that had been injected with PFI-3 had significantly more heterochromatic aggregates than mock injected controls, and that these aggregates covered a larger percent of nuclear area in PFI-3 injected embryos compared to controls (**Fig. 6c, d**). These observations combined with the molecular studies described above suggest that Smarca2 serves as a maternally

supplied inhibitor of heterochromatin establishment during early embryogenesis. Collectively, our findings identify MZT and the miR-430/Smarca2 axis as an important mechanism controlling the de novo establishment of vertebrate heterochromatin in the early embryo (**Fig 7)**.

Discussion

Reprogramming of H3K9me3 marked constitutive heterochromatin has been widely noted in metazoa. However, very little is known regarding the mechanisms that govern this process, especially in vertebrate systems [5-7]. In this study, we demonstrate that global establishment of H3K9me3 marked heterochromatin is controlled by the maternal to zygotic transition in the zebrafish embryo and we identify Smarca2 as an essential gatekeeper of H3K9me3 establishment and global chromatin compaction during zebrafish embryogenesis. In addition to identifying MZT as a key regulator of heterochromatin establishment, these studies identify maternal inhibition, rather than targeting of the de novo methyltransferase machinery as a primary mechanism controlling de novo establishment of H3K9me3 in the embryo. Our results also demonstrate that precocious establishment of H3K9me3 is sufficient to accelerate the timeline of chromatin compaction at the ultrastructure level, implicating H3K9me3 deposition as a primary barrier to heterochromatin establishment in the early embryo.

The extent and timing of heterochromatin reprogramming varies significantly between species. For example, heterochromatin depletion in the early mouse embryo preferentially affects the paternal genome, with major gains in H3K9me3 noted by the 2-cell stage [4, 5]. In contrast, immunohistochemistry suggests that the early C. elegans embryo is broadly depleted for H3K9me3, with significant gains in heterochromatic

markers noted in 5-20 cell stage embryos [7]. Our studies reveal that the zebrafish embryo is globally deficient in H3K9me3 and condensed chromatin ultrastructure until at least the 1000- cell stage. This observation indicates that zebrafish undergo roughly 10 rounds of embryonic cell division in the absence of condensed chromatin ultrastructure. This extensive and prolonged heterochromatin depletion raises questions regarding how genome integrity is sustained during these early rounds of rapid cell division.

The coupling of heterochromatin establishment and the maternal zygotic transition in zebrafish assures a globally decondensed genome at the onset of ZGA. This open structure has the potential to facilitate initiation of ZGA and to reduce global barriers to transcription factor binding during the critical window immediately following ZGA. It is possible, that the benefits of increased chromatin plasticity during this period outweigh the potential costs of increased genome instability associated with heterochromatin depletion. Indeed, we find that precocious accumulation of H3K9me3 is associated with abnormal development in embryos injected with Smarca2 morpholino or PFI-3. The timing of heterochromatin establishment roughly correlates with MZT in other species as well, raising the possibility that MZT may be a conserved regulator of heterochromatin establishment. Supporting this hypothesis, blocking zygotic transcription disrupts replication timing in the Drosophila embryo and spatial reorganization of pericentromeric domains in the mouse embryo, two features that are associated with heterochromatin [17, 18].

In this study, we identify zygotic transcription of miR-430 and subsequent clearance of Smarca2 mRNA as a key mechanism tying MZT to heterochromatin establishment. There have been few previous reports identifying potential mechanisms

regulating the establishment of H3K9me3 marked heterochromatin during development in any species. In mouse, a single siRNA based study has implicated the chromatin assembly factor 1a (Chaf1a) in the regulation of H3K9me3 establishment and silencing at a subset of LTR transposons [5]. However, global heterochromatin formation was unaffected in these embryos. Mutation of Lsd1/Su(VAR)3-3, was shown to reduce H3K9me3 establishment at pericentromeric sequences in Drosophila, but similar effects were not reported in mice [6, 19]. More recently, nuclear accumulation of the histone methyltransferase Setdb1/MET-2 was implicated in regulating the timing of heterochromatin establishment in C. elegans [7]. While it remains to be seen whether this pathway is also conserved in vertebrate systems, there is potential for the miR-430/Smarca2 regulatory axis to act upstream of Met-2 and other as yet unidentified targeting pathways to tightly control the precise window of heterochromatin establishment.

Smarca2 and the related Smarca4 protein function as alternative catalytic subunits for the BRG1/BRM associated factor (BAF) complex, which is broadly involved in nucleosome remodeling. BAF complexes containing Smarca4 as the catalytic subunit have been shown to inhibit establishment of H3K27me3, a known marker of facultative heterochromatin [20]. To our knowledge, Smarca2 has not been previously implicated in vertebrate heterochromatin regulation. Our findings raise the possibility that BAF may have broad functions in antagonizing the establishment of repressive chromatin states, with complex specificity designated by subunit composition. Although Smarca2 is downregulated at MZT, it is expressed at low levels in many adult tissues [21], raising the

possibility that Smarca2 may also antagonize H3K9me3 in other somatic contexts or disease states.

Data availability

The ChIP-seq data sets generated and analyzed during the current study are available in the NCBI GEO database (accession # GSE113086). All other datasets generated during and/or analyzed during the current study are available from the corresponding author on reasonable request.

Author Contributions

KL and MGG conceived the study. KL designed and performed most of the experiments. SR contributed ChIP experiments. SR and AJC performed ChIP-seq experiments. All data was analyzed by KL and MGG with the exception of the ChIP-seq analysis performed by AJC and ZAL. KL and MGG prepared the manuscript with input from SR, AJC and ZAL.

References

1. Peters, A.H. *et al.* Loss of the Suv39h histone methyltransferases impairs mammalian heterochromatin and genome stability. *Cell* **107**, 323-337 (2001).

2. Penke, T.J., McKay, D.J., Strahl, B.D., Matera, A.G. & Duronio, R.J. Direct interrogation of the role of H3K9 in metazoan heterochromatin function. *Genes Dev* **30**, 1866-1880 (2016).

3. Allshire, R.C. & Madhani, H.D. Ten principles of heterochromatin formation and function. *Nat Rev Mol Cell Biol* (2017).

4. Santos, F., Peters, A.H., Otte, A.P., Reik, W. & Dean, W. Dynamic chromatin modifications characterise the first cell cycle in mouse embryos. *Dev Biol* **280**, 225-236 (2005).

5. Wang *et al.* Reprogramming of H3K9me3-dependent heterochromatin during mammalian embryo development. *Nature CellBiology* **epublished** (2018).

6. Rudolph, T. *et al.* Heterochromatin formation in Drosophila is initiated through active removal of H3K4 methylation by the LSD1 homolog SU(VAR)3-3. *Mol Cell* **26**, 103-115 (2007).

7. Mutlu, B. *et al.* Regulated nuclear accumulation of a histone methyltransferase times the onset of heterochromatin formation in C. elegans embryos. *Sci Adv* **4**, eaat6224 (2018).

8. Ahmed, K. *et al.* Global chromatin architecture reflects pluripotency and lineage commitment in the early mouse embryo. *PLoS One* **5**, e10531 (2010).

9. van der Heijden, G.W. *et al.* Parental origin of chromatin in human monopronuclear zygotes revealed by asymmetric histone methylation patterns, differs between IVF and ICSI. *Mol Reprod Dev* **76**, 101-108 (2009).

10. Boskovic, A. *et al.* Higher chromatin mobility supports totipotency and precedes pluripotency in vivo. *Genes Dev* **28**, 1042-1047 (2014).

11. Becker, J.S., Nicetto, D. & Zaret, K.S. H3K9me3-Dependent Heterochromatin: Barrier to Cell Fate Changes. *Trends Genet* **32**, 29-41 (2016).

12. Fadloun, A., Eid, A. & Torres-Padilla, M.E. Mechanisms and dynamics of heterochromatin formation during mammalian development: closed paths and open questions. *Curr Top Dev Biol* **104**, 1-45 (2013).

13. Sahoo, P.K., Soltani, S., Wong, K.C. & Y.C., C. A survey of Thresholding Techiques. *Computer Vision, Graphics and Image Processing,* **41**, 233-260

(1988).

14. Giraldez, A.J. *et al.* Zebrafish MiR-430 promotes deadenylation and clearance of maternal mRNAs. *Science* **312**, 75-79 (2006).

15. White, R.J. *et al.* A high-resolution mRNA expression time course of embryonic development in zebrafish. *Elife* **6** (2017).

16. Fedorov, O. *et al.* Selective targeting of the BRG/PB1 bromodomains impairs embryonic and trophoblast stem cell maintenance. *Sci Adv* **1**, e1500723 (2015).

17. Shermoen, A.W., McCleland, M.L. & O'Farrell, P.H. Developmental control of late replication and S phase length. *Curr Biol* **20**, 2067-2077 (2010).

18. Casanova, M. *et al.* Heterochromatin reorganization during early mouse development requires a single-stranded noncoding transcript. *Cell Rep* **4**, 1156-1167 (2013).

19. Ancelin, K. *et al.* Maternal LSD1/KDM1A is an essential regulator of chromatin and transcription landscapes during zygotic genome activation. *Elife* **5** (2016).

20. Stanton, B.Z. *et al.* Smarca4 ATPase mutations disrupt direct eviction of PRC1 from chromatin. *Nat Genet* **49**, 282-288 (2017).

21. Reisman, D.N., Sciarrotta, J., Bouldin, T.W., Weissman, B.E. & Funkhouser, W.K. The expression of the SWI/SNF ATPase subunits BRG1 and BRM in normal human tissues. *Appl Immunohistochem Mol Morphol* **13**, 66-74 (2005).

22. Kimmel, C.B., Ballard, W.W., Kimmel, S.R., Ullmann, B. & Schilling, T.F. Stages of embryonic development of the zebrafish. *Dev Dyn* **203**, 253-310 (1995).

23. Langenau, D.M. & Zon, L.I. The zebrafish: a new model of T-cell and thymic development. *Nat Rev Immunol* **5**, 307-317 (2005).

Figures

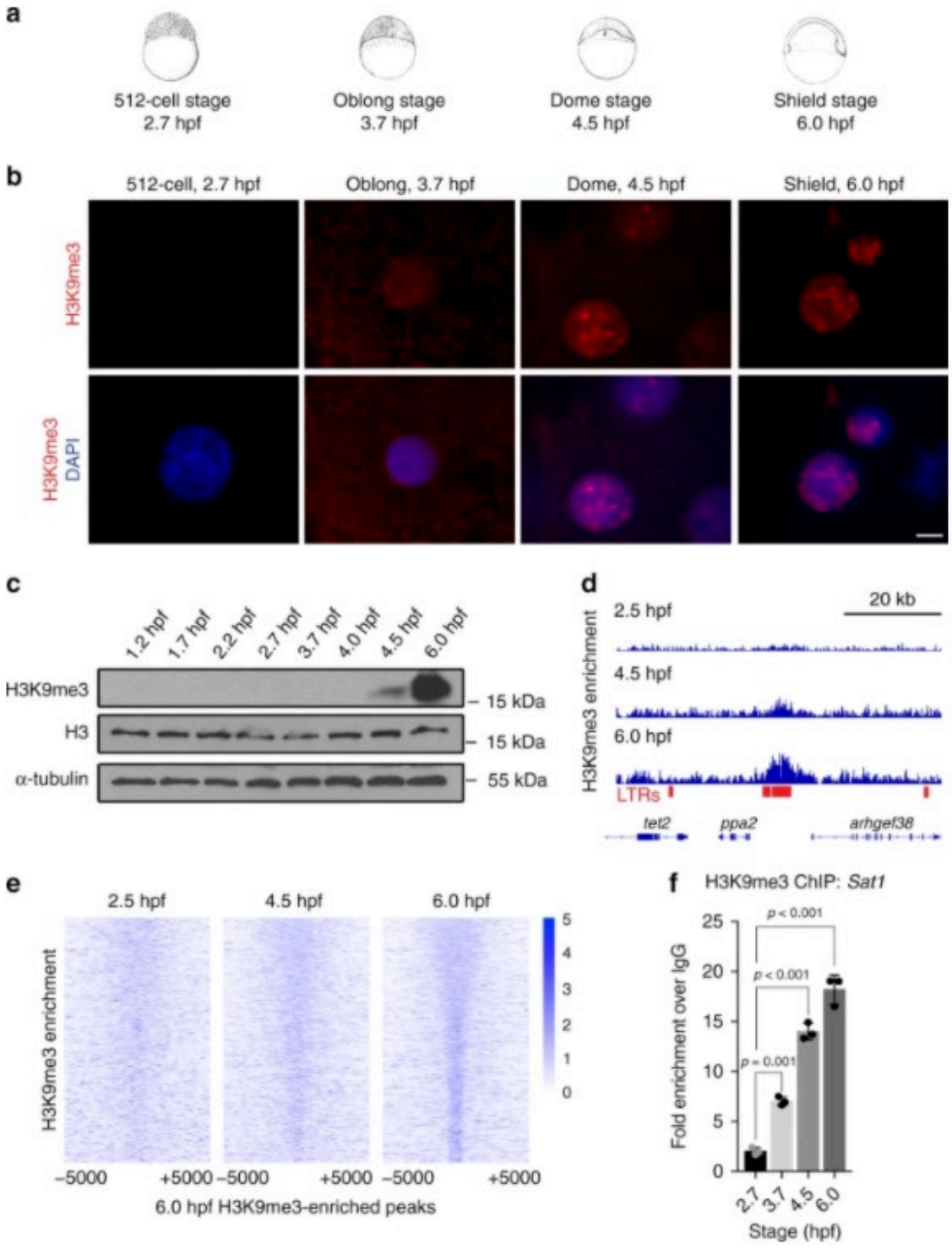

a
512-cell stage
2.7 hpf
Oblong stage
3.7 hpf
Dome stage
4.5 hpf
Shield stage
6.0 hpf
b
512-cell, 2.7 hpf
Oblong, 3.7 hpf
Dome, 4.5 hpf
Shield, 6.0 hpf
H3K9me3
H3K9me3
DAPI
c
1.2 hpf
1.7 hpf
2.2 hpf
2.7 hpf
3.7 hpf
4.0 hpf
4.5 hpf
6.0 hpf
H3K9me3
H3
α-tubulin
15 kDa
15 kDa
55 kDa
d
20 kb
2.5 hpf
4.5 hpf
6.0 hpf
H3K9me3 enrichment
LTRs
tet2
ppa2
arhgef38
e
2.5 hpf
4.5 hpf
6.0 hpf
H3K9me3 enrichment
5
4
3
2
1
0
−5000
+5000
−5000
+5000
−5000
+5000
6.0 hpf H3K9me3-enriched peaks
f
H3K9me3 ChIP: Sat1
p < 0.001
p < 0.001
p = 0.001
Fold enrichment over IgG
25
20
15
10
5
0
2.7
3.7
4.5
6.0
Stage (hpf)

Fig 1: Chromatin of the early zebrafish embryo is globally deficient in H3K9me3.

(a) Developmental stages used for heterochromatin analysis. Images are derived from Kimmel et al 1995 [22]. (b) Representative images showing H3K9me3 immunofluorescent antibody labeling of nuclei from fixed embryos at 2.7, 3.7, 4.5 and 6 hpf. Images are representative of three independent experiments with at least 15 embryos observed at each time point in each experiment. Top panels show H3K9me3 labeling (red). Bottom panels show H3K9me3 (red) overlaid with DAPI (blue). All images were taken under identical conditions, the scale bar represents 1uM. (c) Western blots for H3K9me3, histone H3 and α-tubulin. For each time point, total protein extracts were isolated from 20 embryos, and one third of the protein extract for each sample was loaded for western blot. (d-e) H3K9me3-enrichment measured by ChIP-seq at 2.5 hpf, 4.5 hpf, and 6 hpf. (d) Screen shot of H3K9me3 enrichment at an LTR transposon in a representative genomic region. (e) Heatmap depicting H3K9me3-enrichment across all 17621 H3K9me3 peaks identified in 6 hpf embryos. Signals are centered on peak centers and include +/- 5000 base pairs. (f) H3K9me3 enrichment at Sat1 pericentromeric repeats at 2.7, 3.7, 4.5 and 6 hpf, measured by ChIP and presented as fold enrichment over an IgG only control. p-values were calculated by Ordinary one-way ANOVA, with corrections for multiple testing. Error bars indicate the standard error of the mean (SEM).

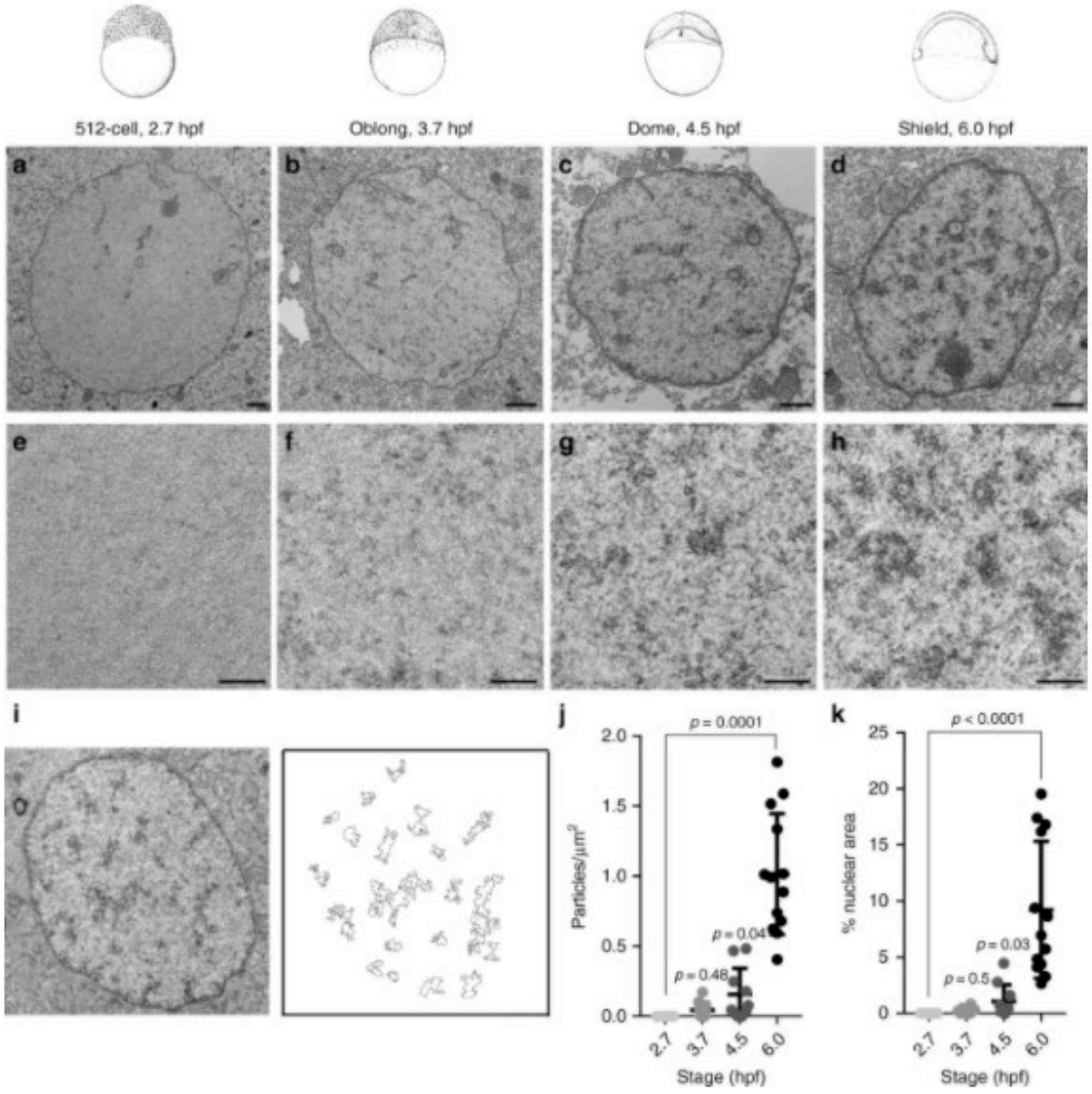


Fig 2: The early zebrafish embryo lacks condensed chromatin ultrastructure. (a-h) Transmission electron micrographs of representative nuclei from embryos at 2.7, 3.7, 4.5 and 6 hpf. Images of representative nuclei at specified time points. (e-h) Higher magnification images (20000x) of nuclear interior at specified time points. All scale bars (a-h) indicate 1 μm (i) Representative image illustrating particle selection. (j) Quantification of the number of particles per nuclear μm^2 in TEM images at 2.7, 3.7, 4.5 and 6 hpf. (k) Quantification of the percent nuclear area covered by particles in TEM

images at 2.7, 3.7, 4.5 and 6 hpf. Nuclei from three embryos were assessed per time point, and electron dense aggregates were quantified in 4-6 representative nuclei from each embryo. Each point represents data one nucleus. p-values were calculated using the Kruskal-Wallis test and corrected for multiple comparisons, error bars indicate standard deviation (SD).

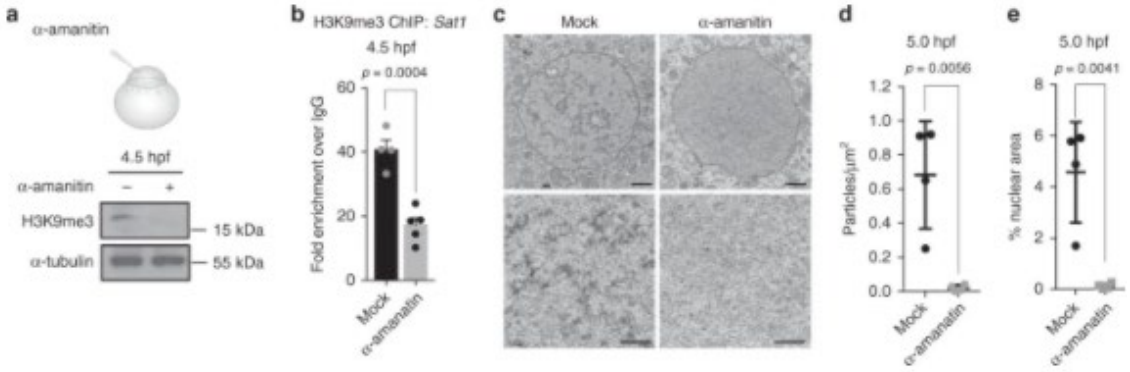


Fig 3: Blocking zygotic transcription impairs heterochromatin establishment.

(a) Western blot for H3K9me3 (top) and α-tubulin (bottom) using protein extracted from embryos that were either mock injected (–) or injected with 0.2ng of α-amanitin (+) at the 1-cell stage. Protein was collected for analysis at 4.5 hpf. Injection scheme derived from Langenau and Zon 2005[23]. (b) ChIP for H3K9me3 enrichment at pericentromeric Sat1 repeats in wildtype and α-amanatin injected embryos. Embryos were collected for analysis at 4.5 hpf. p-values were calculated using the students t test, error bars indicate SEM. (c) TEM images demonstrating a lack of condensed chromatin ultrastructure in 4.5 hpf embryos that were injected with α-amanatin at the 1-cell stage. Bottom panels represent higher magnification images (20000x) of nuclear interior in mock and α-amanatin injected embryos at specified time points. Scale bars indicate 1 μm. (d-e) Quantification of the number of electron dense aggregates per nuclear μm^2 (d) and

percent nuclear area covered by aggregates (e) in mock and α-amanitin injected embryos at 4.5 hpf. Particles/um^2 and percent nuclear area were measured in nuclei from each of four α-amanitin and four mock injected embryos. Each graphed data point represents data from one embryo, values for each embryo are the average of 6-10 representative nuclei. Error bars indicate SD.

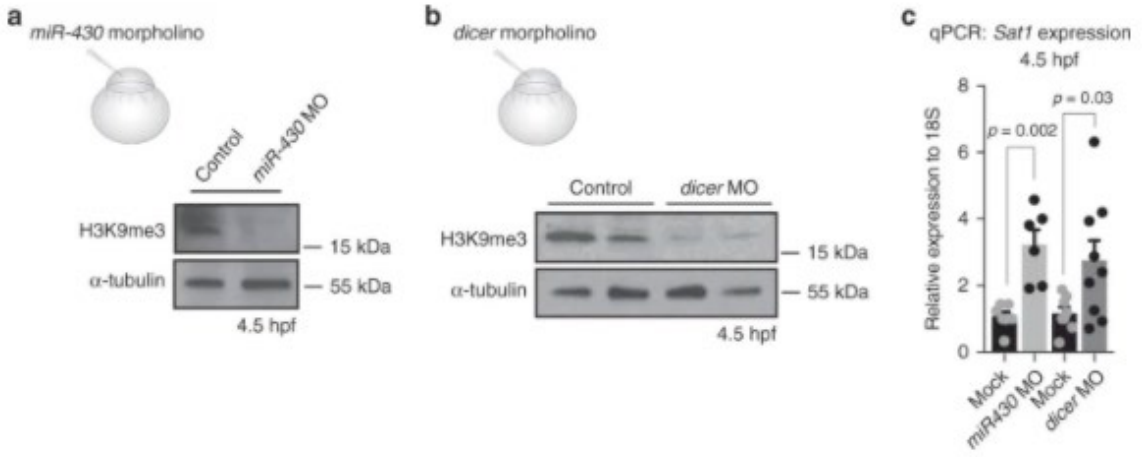


Fig 4: The microRNA miR-430 is required for heterochromatin establishment.

(a) Western blot for H3K9me3 (top) and α-tubulin (bottom) in embryos that were either mock injected or injected with 2 ng of miR-430 morpholino at the 1-cell stage. Protein was collected for analysis at 4.5 hpf. (b) Western blot for H3K9me3 (top) and α-tubulin (bottom) using protein extracted from embryos that were either injected with 2ng of control morpholino, or injected with dicer morpholino at the 1-cell stage. Protein was collected for analysis at 4.5 hpf. (c) Sat1 pericentromeric transcript levels in 4.5 hpf embryos that were either injected with control morpholino or injected with miR-430 or dicer morpholino at the 1-cell stage. p-values were calculated using the students t test, error bars indicate SEM.

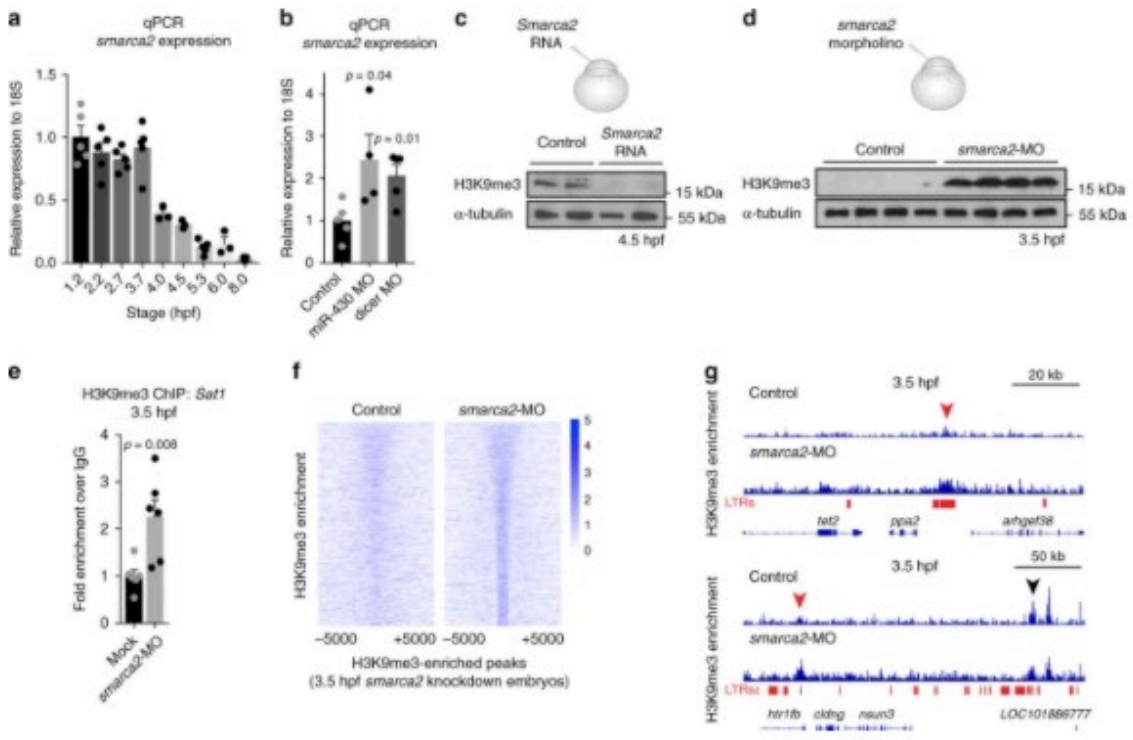


Fig 5: Smarca2 degradation is required for H3K9me3 establishment during MZT.

(a) Quantitative RT-PCR showing *smarca2* RNA decreases during MZT. Error bars indicate SEM. (b) Quantitative RT-PCR confirming elevated levels of *smarca2* transcripts in 5 hpf embryos that were injected with 2ng of miR-430 or dicer morpholino at the 1-cell stage. p-values were calculated using the students-t test, error bars indicate SEM. (c) Western blot for H3K9me3 (top) and α-tubulin (bottom) in embryos that were either injected with 200 pg of *in vitro* transcribed antisense RNA or *Smarca2* mRNA at the 1-cell stage. Protein was collected for analysis at 4.5 hpf. (d) Western blot for H3K9me3 (top) and α-tubulin (bottom) in embryos that were either mock-injected or injected with 2 ng of *smarca2* morpholino at the 1-cell stage. Protein was collected for analysis at 3.5 hpf. (e) ChIP for H3K9me3 enrichment at Sat1 sequences in embryos injected with a GFP control morpholino or with smarca2 morpholino. Analysis was performed at 3.5 hpf. p-values were calculated using the students-t test, error bars

indicate SEM. (f-g) ChIP for H3K9me3-enrichment in embryos injected with a scrambled control morpholino or morpholinos targeting *smarca2.* (f) Heatmap depicting H3K9me3-enrichment in control embryos (left) across all H3K9me3 peak centers (+/- 5000 base pairs) identified in *smarca2* morpholino injected embryos (right). (g) Screen shots of two representative genomic regions. Red arrows indicate regions that gain H3K9me3 in *smarca2* morpholino injected samples. The black arrow indicates a region that is enriched for H3K9me3 in both control and morpholino injected embryos.

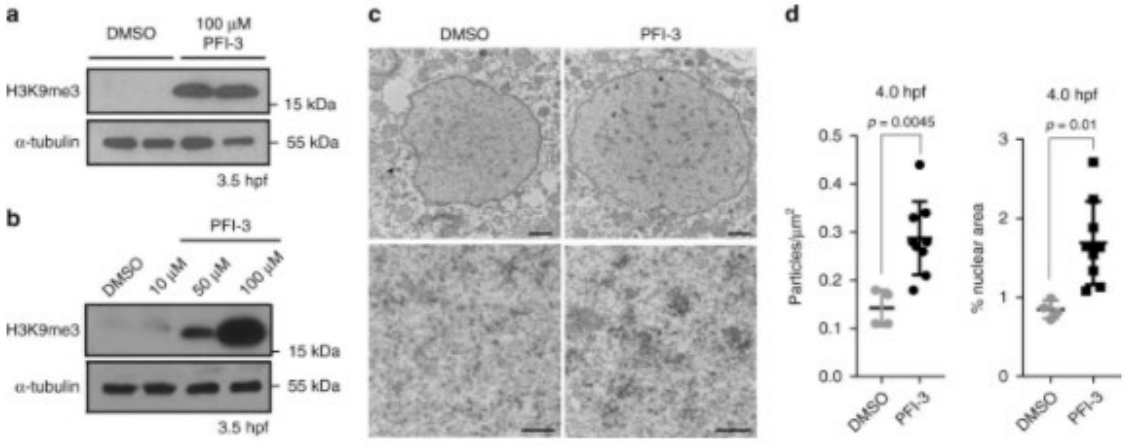


Fig 6: Smarca2 inhibition is sufficient to accelerate the timeline of chromatin compaction. (a) Western blot for H3K9me3 (top), and α-tubulin (bottom) in embryos that were either mock-injected or injected with 100 µM of the Smarca2 inhibitor PFI-3 at the 1-cell stage. Protein was collected for analysis at 3.5 hpf. (b) Western blot using embryos injected with increasing concentrations of PFI-3. Protein was collected for analysis at 3.5 hpf. (c) Electron micrographs demonstrating increased levels of chromatin compaction in 4.5 hpf embryos that were injected with DMSO or 100 µM PFI-3 at the 1-cell stage. Bottom panels represent higher magnification images (20000x) of nuclear interior in mock and PFI-3 injected embryos at specified time points. All scale bars

indicate 1 μm. (d) Quantification of the number of particles per nuclear μm^2 and percent nuclear area. Each data point indicates an individual embryo, for each embryo (four DMSO and nine PFI-3 injected embryos) values for particles per μm^2 and percent nuclear area were averaged from 5-10 representative nuclei. p-values were calculated using the students t-test, error bars indicate SD.

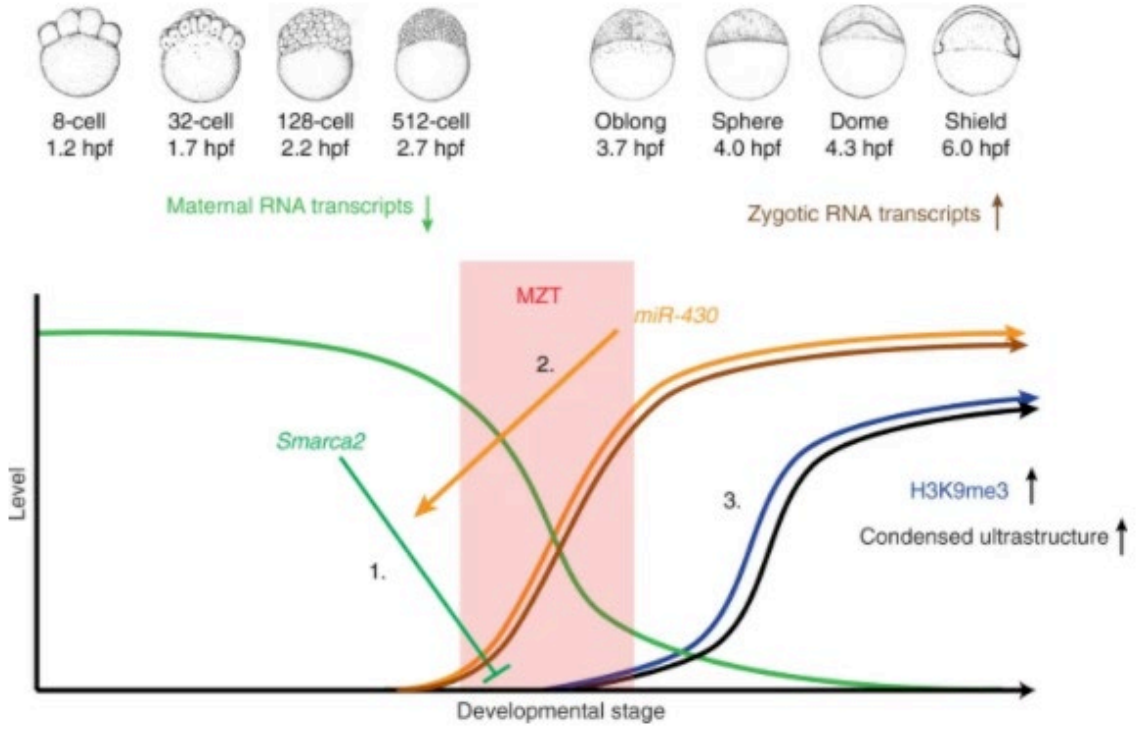


Fig 7: Model for Smarca2-miR430 dependent heterochromatin formation at the maternal to zygotic transition. The early embryo relies exclusively on maternally deposited RNA transcripts and protein, and only starts transcribing the zygotic genome at the maternal to zygotic transition. This model shows that maternal *smarca2,* a catalytic subunit of the BAF complex, inhibits H3K9me3 incorporation and chromatin condensation prior to the maternal zygotic transition (1). At the onset of zygotic transcription, the microRNA miR-430 is transcribed and targets maternal *smarca2* for degradation (2). Decreasing levels of *smarca2* are sufficient to allow for the

incorporation of H3K9me3 and formation of condensed chromatin ultrastructure post-MZT (3).