

CONTROL AND FUNCTION OF HISTONE H3 LYSINE 27

METHYLATION IN *NEUROSPORA CRASSA*

by

KEVIN JOSEPH MCNAUGHT

A DISSERTATION

Presented to the Department of Biology  
and the Graduate School of the University of Oregon  
in partial fulfillment of the requirements  
for the degree of  
Doctor of Philosophy

December 2019

DISSERTATION APPROVAL PAGE

Student: Kevin Joseph McNaught

Title: Control and Function of Histone H3 Lysine 27 Methylation in *Neurospora crassa*

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Biology by:

Dr. Jeffrey McKnight	Chairperson
Dr. Eric Selker	Advisor
Dr. Alice Barkan	Core Member
Dr. Kryn Stankunas	Core Member
Dr. Michael Harms	Institutional Representative

and

Kate Mondloch	Interim Vice Provost and Dean of the Graduate School
---------------	--

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded December 2019

© 2019 Kevin Joseph McNaught

## DISSERTATION ABSTRACT

Kevin Joseph McNaught

Doctor of Philosophy

Department of Biology

December 2019

Title: Control and Function of Histone H3 Lysine 27 Methylation in *Neurospora crassa*

In higher eukaryotes, proper development and maintenance of cellular identity requires the activity of Polycomb repressive complexes (PRCs). PRC2 catalyzes methylation of histone H3 lysine 27 (H3K27), which serves as a repressive chromatin mark. PRC2 is conserved in the filamentous fungus, *Neurospora crassa*, and H3K27 methylation covers around 7% of the genome, repressing scores of genes. My research focused on two basic questions: 1. What controls the distribution of H3K27 methylation? and 2. How does H3K27 methylation function as a repressive mark?

As one approach to identify factors that influence the deposition of H3K27 methylation, I screened gene knock-outs of *nst* (*neurospora sir two*) homologues for defects in H3K27 methylation. I discovered that loss of NST-3, a putative H3K56 deacetylase, results in accumulation of H3K27 methylation at centromeres, and that this is suppressed by strains lacking *rtt109* or bearing histones with the H3K56R substitution.

In collaboration with Kirsty Jamieson, we investigated what role chromosome position had on the deposition of H3K27 methylation by examining chromosomal rearrangement strains. We identified both position-dependent and position-independent H3K27 methylation. We further found that proximity to chromosome ends is necessary to maintain, and sufficient to induce, transcriptionally repressive H3K27 methylation, and



that this effect is mediated by telomere repeats, (TTAGGG)<sub>n</sub>.

As a part of a forward-genetics selection for factors necessary for Polycomb silencing, spearheaded by Elizabeth Wiles, I identified two alleles of a gene, *NCU04278*, that is required for the telomere-dependent H3K27 methylation uncovered in our previous work. Immunoprecipitation followed by mass spectrometry analyses of *NCU04278* and known PRC2 members established that *NCU04278* is a PRC2 accessory subunit and was therefore designated PAS.

The aforementioned forward genetic selection identified another gene, *epr-1* (effector of polycomb repression 1; *NCU07505*). EPR-1 associates with H3K27 methylation *in vivo* and *in vitro*, and loss of EPR-1 de-represses H3K27-methylated genes without loss of H3K27 methylation. EPR-1 is not fungal-specific; orthologs of EPR-1 are present in a diverse array of eukaryotic lineages, suggesting an ancestral EPR-1 was a component of a primitive Polycomb repression pathway.

This dissertation includes both previously published and unpublished co-authored material.

## CURRICULUM VITAE

NAME OF AUTHOR: Kevin Joseph McNaught

### GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene

### DEGREES AWARDED:

Doctor of Philosophy, Biology, 2019, University of Oregon

Bachelor of Science, Biology, 2012, University of Oregon

### AREAS OF SPECIAL INTEREST:

Epigenetics  
Chromatin  
Genomics

### PROFESSIONAL EXPERIENCE:

Graduate Student, Eric Selker laboratory, University of Oregon, 2013-2019

Teaching Fellow, Department of Biology, University of Oregon, 2013-2014

### GRANTS, AWARDS, AND HONORS:

Developmental Biology Training Grant, National Institutes of Health, 2015-2018

### PUBLICATIONS:

Jamieson, K.\*, **McNaught, K.J.\***, Ormsby, T., Leggett, N.A., Honda, S., and Selker, E.U. (2018). Telomere repeats induce domains of H3K27 methylation in *Neurospora*. *eLife* 7, 948.

Jamieson, K., Wiles, E.T., **McNaught, K.J.**, Sidoli, S., Leggett, N., Shao, Y., Garcia, B.A., and Selker, E.U. (2016). Loss of HP1 causes depletion of H3K27me3 from facultative heterochromatin and gain of H3K27me2 at constitutive heterochromatin. *Genome Research* 26, 97–107.

\*These authors contributed equally to this work.

## ACKNOWLEDGMENTS

I'd like to thank Eric Selker for his continued support and guidance throughout my graduate career. My lab mates, Kirsty Jamieson and Elizabeth Wiles, were incomparable collaborators and provided much valuable feedback. In addition, I owe a debt to all previous and current Selker lab members, who paved the way for my success. I also want to thank the members of my dissertation committee for their encouragement and advice, and the Fungal Genetics Stock Center for providing valuable *Neurospora crassa* strains. This investigation was supported in part by a National Institutes of Health Developmental Biology Training Grant.

I dedicate this dissertation to my parents, Joe and Marybeth,  
for their unwavering love and support.

## TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION .....	1
II. IMPLICATION OF ABERRANT H3K56 ACETYLATION IN ECTOPIC H3K27 METHYLATION .....	4
Introduction.....	4
Results.....	5
Deletion of <i>nst-3</i> results in centromeric H3K27 methylation.....	5
Centromeric H3K27 methylation suppressed in $\Delta nst-3$ and $\Delta hpo$ strains lacking H3K56 acetylation .....	6
H3K56 acetylation appears to accumulate at centromeres in $\Delta nst-3$ strains ...	7
Discussion.....	8
Materials and Methods.....	9
Strains, media and growth conditions.....	9
Chromatin immunoprecipitation (ChIP) and ChIP-sequencing.....	9
Replacement of endogenous histone H3 with H3K56R/Q mutants.....	9
Bridge to Chapter III.....	10
III. TELOMERE REPEATS INDUCE DOMAINS OF H3K27 METHYLATION IN NEUROSPORA.....	11
Introduction.....	11
Results.....	13
Analyses of classical chromosome rearrangements suggest that proximity to a chromosome end is key to subtelomeric H3K27 methylation .....	13

Chapter	Page
Identification of position-independent H3K27-methylated domains .....	15
<i>Cis-trans</i> test on H3K27 methylation in a segmental duplication strain reveals position effect.....	16
Altered gene expression in regions with changed H3K27me2/3.....	17
Loss of telomerase abolishes subtelomeric H3K27me2/3 .....	18
Internal telomere repeats are sufficient to induce <i>de novo</i> H3K27me2/3.....	19
Discussion.....	20
Materials and Methods.....	26
Strain, media and growth conditions .....	26
Deleting and targeting segments of LG VII.....	26
Deletion of <i>tert</i> .....	27
Targeting telomere repeats to <i>csr-1</i> and <i>his-3</i> .....	27
ChIP and preparation of libraries.....	28
RNA isolation and preparation of libraries.....	29
Translocation breakpoint analyses.....	29
Southern analysis .....	29
SNP mapping .....	30
Bridge to Chapter IV.....	30
IV. IDENTIFICATION OF A PRC2 ACCESSORY SUBUNIT REQUIRED FOR SUBTELOMERIC H3K27 METHYLATION IN NEUROSPORA .....	32
Introduction.....	32
Results.....	33

Chapter	Page
Isolation, mapping and identification of <i>NCU04278</i> as a potential PcG gene .....	33
NCU04278 is necessary for silencing subtelomeric H3K27-methylated genes .....	34
NCU04278 is necessary for maintenance of subtelomeric H3K27 methylation .....	36
NCU04278 is an accessory subunit of PRC2 .....	37
Discussion .....	38
Materials and Methods.....	40
Strains, media and growth conditions.....	40
Selection for mutants defective in H3K27 methylation-mediated silencing .....	40
Whole genome sequencing, mapping and identification of <i>NCU04278</i> alleles .....	40
Chromatin immunoprecipitation (ChIP), ChIP-qPCR and ChIP-seq .....	41
RNA isolation, RT-qPCR and mRNA-seq .....	41
Immunoprecipitation followed by mass spectrometry (IP-MS) .....	41
Replacement of <i>NCU04278</i> with <i>trpC::nat-1</i> .....	42
Generation of a <i>N. crassa</i> strain expressing <i>NCU04278::3xFLAG</i> .....	43
Bridge to Chapter V .....	43

Chapter	Page
V. EVOLUTIONARILY ANCIENT BAH-PHD PROTEIN MEDIATES POLYCOMB SILENCING .....	45
Introduction.....	45
Results.....	47
Genetic selection for factors necessary for H3K27 methylation-mediated repression .....	47
Mapping and identification of <i>epr-1</i> as <i>NCU07505</i> .....	48
EPR-1 and SET-7 repress an overlapping set of H3K27-methylated genes....	49
$\Delta epr-1$ and $\Delta set-7$ strains share a defect in sexual development .....	50
H3K27 methylation is essentially normal in $\Delta epr-1$ .....	50
EPR-1 forms telomere-associated foci dependent on EED .....	52
An intact BAH domain is required for normal nuclear distribution of EPR-1 .....	53
EPR-1 localizes to H3K27 methylation genome-wide .....	53
Both the BAH domain and PHD finger of EPR-1 are necessary for gene repression .....	54
BAH domain of EPR-1 binds to H3K27me3 <i>in vitro</i> .....	55
EPR-1 is a homolog of plant EBS/SHL and widely distributed across eukaryotes .....	56
Discussion.....	57
Materials and Methods.....	59
Strains, media and growth conditions.....	59



Chapter	Page
Selection for mutants defective in H3K27 methylation-mediated silencing .....	60
Whole genome sequencing, mapping and identification of <i>epr-1</i> alleles .....	60
Western blotting.....	61
Chromatin immunoprecipitation.....	62
ChIP-seq mapping and analysis .....	63
RNA isolation, mRNA-seq library prep, and RT-qPCR.....	63
mRNA-seq mapping and analysis.....	64
DamID Southern hybridizations and sequencing .....	64
False perithecia assay and image analysis .....	65
Microscopy image acquisition and analysis .....	65
Cloning, expression and purification of BAH domain from EPR-1 .....	66
NMR spectroscopy.....	67
Bioinformatic identification of EPR-1 and SET-7 homologs.....	68
Replacement of <i>NCU05173</i> and <i>NCU07152</i> ORFs with <i>hph</i> and <i>nat-1</i> .....	68
Creation and targeting of <i>his-3<sup>+</sup>::pCCG::N-GFP::EPR-1</i> plasmids .....	69
Replacement of <i>epr-1</i> with <i>trpC::nat-1</i> .....	70
Endogenous C-terminal-tagging of EPR-1 with 10xGly::Dam .....	70
VI. CONCLUDING SUMMARY .....	72
APPENDICES .....	76
A. FIGURES.....	76
Chapter II .....	76

Chapter	Page
Chapter III.....	77
Chapter IV.....	90
Chapter V .....	97
<b>B. STRAIN TABLES .....</b>	<b>111</b>
Chapter II .....	111
Chapter III.....	111
Chapter IV.....	112
Chapter V .....	113
<b>C. PRIMER TABLES.....</b>	<b>115</b>
Chapter II .....	115
Chapter III.....	115
Chapter IV.....	121
Chapter V .....	123
<b>D. PLASMID TABLES .....</b>	<b>127</b>
Chapter II .....	127
Chapter III.....	127
Chapter IV.....	127
Chapter V .....	127
<b>REFERENCES CITED.....</b>	<b>129</b>

## LIST OF FIGURES

Figure	Page
1. Aberrant H3K56ac implicated in ectopic H3K27me2/3.....	76
2. Chromosomal rearrangements are associated with altered H3K27me2/3 .....	77
3. Three kilobase segments from a natural H3K27me2/3 domain are insufficient to trigger <i>de novo</i> H3K27me2/3 at ectopic loci.....	78
4. H3K27me2/3 profiles of additional strains containing chromosomal rearrangements.....	79
5. Whole genome view of H3K27me2/3 ChIP-seq in wild-type and chromosomal translocation strains .....	80
6. qPCR validation of H3K27me2/3 ChIP-seq data from translocation strains .....	81
7. A segmental duplication confirms H3K27me2/3 position effect .....	82
8. Altered gene expression reflects changes in H3K27me2/3 .....	83
9. qPCR validation of RNA-seq expression changes in translocation OY350.....	84
10. Loss of <i>tert</i> disrupts chromosome ends and abolishes subtelomeric H3K27me2/3.....	85
11. Whole genome view of H3K27me2/3 ChIP-seq in wild type and <i>tert</i> .....	86
12. Telomere repeats targeted to the <i>csr-1</i> locus induce an approximately 225 kb H3K27me2/3 domain.....	87
13. Telomere repeats targeted to the <i>his-3</i> locus induce local H3K27me2/3 domain.....	88
14. Whole genome view of H3K27me2/3 ChIP-seq in wild type and $\Delta csr-1::(\text{TTAGGG})_{17}$ .....	89

Figure	Page
15. Identification of NCU04278 as a potential component of the Polycomb silencing pathway.....	90
16. Mapping and characterization of an additional <i>NCU04278</i> allele.....	91
17. Loss of NCU04278 upregulates a subset of SET-7 repressed genes near chromosome ends.....	92
18. Genome-wide distribution and level of H3K27 methylation in wild-type, $\Delta$ <i>NCU04278</i> , and $\Delta$ <i>npf</i> strains .....	93
19. The effect of loss of NCU04278 on H3K27 methylation and growth rate.....	94
20. NCU04278 associates with known <i>N. crassa</i> PRC2 members.....	95
21. PAS homologs appear restricted to fungal lineages .....	96
22. Forward genetics identifies a novel gene, <i>epr-1</i> , required for H3K27 methylation-mediated repression .....	98
23. Three additional alleles of <i>epr-1</i> were identified during selection for mutants defective in H3K27 methylation-mediated repression.....	99
24. $\Delta$ <i>epr-1</i> and $\Delta$ <i>set-7</i> strains share defects in transcriptional silencing and sexual development.....	100
25. Gene expression and phenotypic analysis of $\Delta$ <i>epr-1</i> and $\Delta$ <i>set-7</i> strains.....	102
26. <i>epr-1</i> is not required for H3K27 methylation .....	103
27. H3K27 methylation levels in $\Delta$ <i>epr-1</i> are comparable to wild type .....	104
28. EPR-1 forms telomere-associated foci that are dependent on EED and the BAH domain of EPR-1 .....	105

Figure	Page
29. EPR-1 <sup>WT</sup> and EPR-1 <sup>PHD</sup> exhibit no difference in distance to telomeres or EPR-1 foci number .....	106
30. EPR-1 directly interacts with H3K27 methylated chromatin through its BAH domain .....	107
31. EPR-1 associates with H3K27 methylation <i>in vivo</i> and <i>in vitro</i> .....	109
32. EPR-1 homologs are present in species beyond fungi .....	110

# CHAPTER I

## INTRODUCTION

*There is grandeur in this view of life, with its several powers, having been originally breathed into a few forms or into one; and that, whilst this planet has gone cycling on according to the fixed law of gravity, from so simple a beginning endless forms most beautiful and most wonderful have been, and are being, evolved.*

- Charles Darwin, *On the Origin of Species*, 1859

*Seen in the light of evolution, biology is, perhaps, intellectually the most satisfying and inspiring science. Without that light it becomes a pile of sundry facts—some of them interesting or curious but making no meaningful picture as a whole.*

- Theodosius Dobzhansky, *The American Biology Teacher*, 1973

The ability of an organism to properly regulate its gene expression is critical to its survival. For the single-celled baker's yeast, *Saccharomyces cerevisiae*, the functional output of such regulation is limited to concerns of proliferating and mating<sup>1</sup>. For multicellular organisms, like humans, precise gene regulation also ensures the proper development and maintenance of countless cell types<sup>2</sup>. If one such human cell should misbehave, *e.g.* express an improper amount of a gene product, a disease state, like cancer, may result<sup>3</sup>. Understanding fundamental principles of gene regulation, therefore, may lead to the development of therapies to alleviate conditions associated with aberrant gene expression.

The genetic material, DNA, from which genes are transcribed, is located in the nucleus of eukaryotic cells. DNA, along with its associated RNA and protein, constitute a nuclear complex called chromatin. The major constituent of the protein fraction in chromatin consists of histones, of which there are five classes: histone H1, H2A, H2B, H3 and H4. Histones are assembled into nucleosomes that consist of a histone octamer (H2A)<sub>2</sub>(H2B)<sub>2</sub>(H3)<sub>2</sub>(H4)<sub>2</sub>, which is encircled by ~146 base pairs of DNA. Histones can be post-translationally modified, and it is established that at least some of these modifications have roles in transcriptional regulation<sup>4</sup>. However, the molecular mechanisms by which these modifications are established, maintained and impact gene expression are largely unresolved.

Methylation of histone H3 lysine 27 (H3K27), catalyzed by Polycomb Repressive Complex 2 (PRC2), is known to be associated with maintenance of transcriptional silencing<sup>5</sup>. Components of PRC2 were originally genetically characterized in the fruit fly, *Drosophila melanogaster*, as necessary for repression of Hox genes during embryonic development<sup>6</sup>. H3K27 methylation has been more recently implicated in stem cell identity, cancer, imprinting, and X-chromosome inactivation<sup>7</sup>.

This dissertation investigates the control and function of H3K27 methylation utilizing the filamentous fungus, *Neurospora crassa*<sup>8</sup>. *N. crassa* is an attractive model organism to study H3K27 methylation, in part because more traditional fungal models, such as *Schizosaccharomyces pombe* and *S. cerevisiae*, lack this histone mark<sup>9</sup>. In addition, *N. crassa* possesses homologues of all core components of mammalian PRC2, and unlike higher eukaryotes these proteins are non-essential for normal development<sup>9</sup>.

The work described in this dissertation contains both published and unpublished co-authored material. Chapter II is an unpublished collaboration with Vincent Bicocca and Eric Selker that examines the role of H3K56 acetylation in ectopic H3K27 methylation. Chapter III was published in volume 7 of the journal *eLife* in January 2018 with co-authors Kirsty Jamieson, Tereza Ormsby, Neena Leggett, Shinji Honda and Eric Selker. It describes the identification of telomere repeats as a *cis*-acting signal for H3K27 methylation in *N. crassa*. Chapter IV is an unpublished collaboration with Elizabeth Wiles and Eric Selker that identifies and characterizes the phenotype of a *N. crassa* mutant lacking a PRC2 accessory subunit. Chapter V is an unpublished collaboration with Elizabeth Wiles, Saumya De Silva, Gurmeet Kaur, Jeanne Selker, L. Aravind, Catherine Musselman and Eric Selker. It describes the identification and characterization of the first, to our knowledge, ‘reader’ and effector of H3K27 methylation in fungi.



**CHAPTER II**  
**IMPLICATION OF ABERRANT H3K56 ACETYLATION**  
**IN ECTOPIC H3K27 METHYLATION**

This work was a collaboration with Vincent Bicocca. I initiated the investigation, created all relevant *Neurospora crassa* strains, performed all H3K27me<sub>2/3</sub> ChIP-seq experiments, and the summary of results presented was initially drafted by me. Vince Bicocca performed the H3K56ac ChIP-seq experiments and designed the wild-type, codon-optimized histone H3 plasmid. Eric Selker was the principal investigator for this work.

**Introduction**

Despite a great deal of effort, a complete understanding of the molecular mechanisms by which PRC2 is recruited to chromatin and subsequently catalyzes H3K27 methylation across the genome in wild-type cells is still unresolved<sup>10,11</sup>. For this reason, it is of interest to perturb chromatin generally, either genetically or chemically, in an effort to uncover principles that promote or prevent accumulation of H3K27 methylation genome-wide. Our research group previously demonstrated that perturbing constitutive heterochromatin machinery in *N. crassa* led to redistribution of H3K27 methylation from native sites to constitutive heterochromatin<sup>12</sup>. This phenomenon appears to be conserved in both plants<sup>13,14</sup> and animals<sup>15-17</sup>. Therefore, determining the features that make

constitutive heterochromatin an attractive target for PRC2 in *N. crassa* may yield insights into PRC2 function in a wide variety of species.

Here, in a limited genetic screen, we found that the deletion of *neurospora sir two-3* (*nst-3*) recapitulated the gain of H3K27 methylation at centromeres observed previously in *heterochromatin protein one* (*hpo*) mutants<sup>12</sup>. By introducing an unmodifiable histone H3 (*H3K56R*) into *N. crassa* and by removing the sole H3K56 histone acetyltransferase<sup>18</sup>, we found evidence that this effect is mediated by H3K56 acetylation, in both *nst-3* and *hpo* mutants. Finally, we observed that the gains of H3K27me<sub>2/3</sub> at the centromeres were coincident with increases of H3K56 acetylation. Thus, it is possible that H3K56 acetylation plays an important role in the ectopic H3K27 methylation observed in not only *N. crassa*, but also in plants and animals.

## Results

### Deletion of *nst-3* results in centromeric H3K27 methylation

A previous investigation<sup>19</sup> from our research group implicated SIR2 homologues of *Neurospora crassa* in subtelomeric gene silencing and revealed that H3K27 methylation appeared reduced at some chromosome ends in a mutant defective in three SIR2 homologs. To further investigate the relationship between SIR2 homologues and the control of H3K27 methylation, I performed H3K27me<sub>2/3</sub> chromatin immunoprecipitation followed by sequencing (ChIP-seq) on strains bearing deletions of *nst-1*, *nst-3*, *nst-4*, and *nst-5* (Fig. 1A). I found that while  $\Delta nst-1$ ,  $\Delta nst-4$ , and  $\Delta nst-5$  strains appeared to have a wild-type-like distribution of H3K27me<sub>2/3</sub>,  $\Delta nst-3$  strains

exhibited ectopic gains of H3K27me<sub>2/3</sub> at most centromeres, albeit to differing levels, as well as at some regions of interspersed constitutive heterochromatin (Fig. 1A). This phenotype was comparable to, though less pronounced than, that observed in strains defective in constitutive heterochromatin machinery<sup>12</sup>.

### **Centromeric H3K27 methylation suppressed in $\Delta nst-3$ and $\Delta hpo$ strains lacking H3K56 acetylation**

NST-3 is a homolog of *Saccharomyces cerevisiae* Hst4p, which acts on acetylated H3K56 (H3K56ac)<sup>20</sup>. To determine if the K56 residue on histone H3 was necessary for the ectopic H3K27 methylation observed in  $\Delta nst-3$  strains, I replaced the endogenous H3 gene with a codon-optimized version of H3 that substituted lysine 56 for arginine (*H3K56R*). Introduction of *H3K56R* into  $\Delta nst-3$  strains completely suppressed the accumulation of H3K27me<sub>2/3</sub> at constitutive heterochromatin (Fig. 1B). It is notable, however, that the *H3K56R* mutant alone, and in combination with  $\Delta nst-3$ , did display some ectopic H3K27me<sub>2/3</sub>, although not at constitutive heterochromatin (Fig. 1B).

To further investigate the potential role of H3K56ac in the ectopic H3K27 methylation in  $\Delta nst-3$  strains, I created a  $\Delta nst-3$  strain bearing an additional deletion of the sole H3K56 acetyltransferase, RTT109<sup>18</sup>. Similar to  $\Delta nst-3$ ;*H3K56R* double mutants,  $\Delta nst-3$ ;  $\Delta rtt109$  strains did not accumulate H3K27me<sub>2/3</sub> at centromeres or interspersed constitutive heterochromatin (Fig. 1B). Taken together, these findings suggest that the aberrant H3K27 methylation observed in  $\Delta nst-3$  strains requires H3K56 acetylation.

To determine if the previously observed redistribution of H3K27 methylation in strains lacking Heterochromatin Protein 1 (HP1)<sup>12</sup> is solely mediated by H3K56ac, I

created a  $\Delta hpo;H3K56R$  double mutant and examined the distribution of H3K27me<sub>2/3</sub>. While the introduction of *H3K56R* did partially suppress the redistribution of H3K27me<sub>2/3</sub> seen in  $\Delta hpo$ , it did not fully abolish it (Fig. 1B). This suggests that while H3K56ac may exacerbate the H3K27 methylation phenotype of  $\Delta hpo$ , other signaling pathways may be involved.

### **H3K56 acetylation appears to accumulate at constitutive heterochromatin in $\Delta nst-3$ strains**

To determine if there is an association between accumulation of H3K56ac and H3K27me<sub>2/3</sub> at constitutive heterochromatin in  $\Delta nst-3$  strains, Vince Bicocca performed H3K56ac ChIP-seq on wild-type and  $\Delta nst-3$  strains (Fig. 1C). The H3K56ac ChIP-seq showed enrichment at constitutive heterochromatin in the  $\Delta nst-3$  strain compared to the wild-type strain, even at regions that do not show ectopic H3K27me<sub>2/3</sub>. However, *H3K56R* and  $\Delta nst-3;H3K56R$  strains subjected to H3K56ac ChIP-seq as negative controls demonstrated that the H3K56ac antibody used is not specific to H3K56ac (Fig. 1C). Therefore, while the most parsimonious explanation is that the ectopic H3K56ac signal detected at constitutive heterochromatin in  $\Delta nst-3$  strains is indeed H3K56ac, this is not certain. Furthermore, the fact that H3K56ac appeared to accumulate at all constitutive heterochromatin suggests that while it may be necessary for ectopic H3K27me<sub>2/3</sub>, it is not sufficient.

## Discussion

The ectopic H3K27 methylation observed at constitutive heterochromatin, especially at centromeres, in  $\Delta nst-3$  strains, and subsequent epistasis experiments implicate H3K56ac as a mediator of this phenotype. It remains to be determined if H3K56ac *per se* is a signal, or it triggers a ‘downstream’ signal for PRC2 activity. Lysine 56 on histone H3 is close to the globular domain of the nucleosome particle, and acetylation of this residue has been implicated in increased nucleosome turnover<sup>21-23</sup>. Altered nucleosome dynamics at constitutive heterochromatin in  $\Delta nst-3$  strains could be responsible for the accumulation of H3K27 methylation. Indeed, nucleosome turnover dynamics can play a critical role in shaping histone methylation patterns genome-wide<sup>24</sup>.

In addition to promoting histone turnover, H3K56ac has been implicated in DNA damage-induced signaling upon replicative stress<sup>25</sup>. During S-phase, nearly all newly incorporated histones are acetylated at H3K56 in yeast<sup>26</sup>. Barring DNA damage, this nascent H3K56ac is subsequently removed by histone deacetylases<sup>20</sup>. Persistently elevated H3K56ac is not only an indicator of DNA damage, but necessary for the DNA damage response in yeast<sup>25</sup>. However, spontaneous suppressors of a hyperactive DNA damage response in cells with elevated H3K56ac do not necessarily abrogate H3K56ac levels<sup>27</sup>. Further epistasis experiments, perturbing known components of the DNA damage response pathway in  $\Delta nst-3$  strains will elucidate if the accumulation of H3K27 methylation at centromeres is a consequence of the DNA damage response. This is not wild speculation, as Polycomb repressors have known roles in the response to DNA damage<sup>28</sup>.

## Materials and Methods

### Strains, media and growth conditions

All *N. crassa* strains used in this study are listed in Appendix B. Liquid cultures were grown with shaking at 32 °C in Vogel's minimal medium with 1.5% sucrose<sup>29</sup>. Crosses were performed at 25 °C on modified Vogel's with 0.1% sucrose<sup>30</sup>.

### Chromatin immunoprecipitation (ChIP) and ChIP-sequencing

ChIP and ChIP-sequencing was performed as described in Chapter V Materials and Methods. Active Motif antibodies for H3K27me2/3 (#39536) and H3K56ac (#39281) were used.

### Replacement of endogenous histone H3 with H3K56R/Q mutants

Plasmid 3167, containing a codon-optimized version of *N. crassa hH3*, was subjected to site-directed mutagenesis using primers 5303 and 5304, or 5305 and 5306 for the introduction of H3K56R or H3K56Q mutations respectively, yielding plasmids 3168 and 3169. The 5' UTR of *hH3* was amplified from genomic DNA with primers 3498 and 5307, the 3' UTR of *hH3* was amplified from genomic DNA with primers 5333 and 5334, and the downstream 3' flank of *hH3* was amplified from genomic DNA with primers 5335 and 5336. The 5' UTR was PCR stitched to plasmid 3168 or 3169 with primers 3498 and 5337 (Fragment 1). The 3' UTR was PCR stitched to the *nat-1*-containing plasmid (3130) with primers 5333 and 4883 (Fragment 2). The 3' flank was PCR stitched to plasmid 3130 with primers 4884 and 5336 (Fragment 3). Fragment 1 and

2 were PCR stitched with primers 3498 and 4883 (Fragment 4). Fragment 3 and 4 were co-transformed into *N. crassa* strain N2931 and transformants were selected on Nourseothricin-containing media.

### **Bridge to Chapter III**

The observation that H3K27 methylation accumulates at centromeres in  $\Delta nst-3$  strains highlights the fact that PRC2 targeting is not strictly dependent on the underlying DNA sequences. While the AT-rich nature of centromeres may be necessary for this ectopic H3K27 methylation in  $\Delta nst-3$  strains, it is clearly not sufficient in wild type. In the next chapter, we further explore the role of underlying DNA sequence and chromosome position on PRC2 activity in wild type strains bearing chromosomal rearrangements. We uncover two fundamentally different classes of H3K27 methylation in *N. crassa*: telomere-dependent and telomere-independent.

**CHAPTER III**  
**TELOMERE REPEATS INDUCE DOMAINS OF**  
**H3K27 METHYLATION IN NEUROSPORA**

This work was published in volume 7 of the journal *eLife* in January 2018. I helped target genomic fragments to the *csr-1* locus, verified chromosomal translocation breakpoints, created and tested the segmental duplication strain for H3K27 methylation position effects, characterized the *tert* knock-out strain, targeted telomere repeats to the *csr-1* and *his-3* locus and examined induced H3K27 methylation, as well as wrote most of the manuscript. Kirsty Jamieson targeted genomic fragments to the *his-3* locus and tested for induced H3K27 methylation and characterized H3K27 methylation profile of the chromosomal translocation strains. Tereza Ormsby investigated the gene expression of chromosomal translocation strains. Neena Leggett identified candidate translocation breakpoints and provided bioinformatic support. Shinji Honda created and provided the *tert* knock-out strain to me. Eric Selker was the principal investigator for this work.

**Introduction**

Methylation of lysine 27 on histone H3 (H3K27me) has emerged as an important repressive mark of the Polycomb group (PcG) system, which is critical for development in higher organisms. PcG proteins were initially discovered in *Drosophila melanogaster* as repressors of homeotic (*HOX*) genes during early embryogenesis<sup>31</sup> and play integral roles in the maintenance of cellular identity and differentiation in a variety of eukaryotes.



Moreover, dysfunction of the PcG system commonly leads to disease, including cancer<sup>32,33</sup>. Biochemical work demonstrated that PcG proteins form two distinct histone-modifying complexes known as Polycomb Repressive Complex 1 and 2 (PRC1 and PRC2)<sup>34,35</sup>. PRC1 mono-ubiquitinates lysine 119 on histone H2A (H2AK119ub1) with its E3-ubiquitin ligase subunit, Ring1, while PRC2 catalyzes mono-, di-, and trimethylation of histone H3 lysine 27 (H3K27me1/2/3) by its SET-domain component, EZH2<sup>34,35</sup>. PRC2, but not PRC1, is widely conserved in eukaryotes, including the filamentous fungus *Neurospora crassa*, but is absent in some simple eukaryotes such as the well-studied yeasts *Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*<sup>9</sup>. H3K27me2/3 covers approximately 7% of the *N. crassa* genome, including about 1000 fully-covered genes, all of which are transcriptionally quiescent<sup>9,36</sup>. The greater than 200 H3K27me2/3 domains, which range from 0.5-107 kb, are widely distributed throughout the genome but are enriched at subtelomeric regions<sup>9</sup>, as also reported for other fungi<sup>37-40</sup>.

In *D. melanogaster*, DNA regulatory regions known as Polycomb Response Elements (PREs) recruit PcG proteins to specific chromatin targets to maintain transcriptional silencing<sup>41</sup>. Recently, similar cis-acting elements have been detected in *Arabidopsis thaliana*<sup>42</sup>. In vertebrates and other organisms, however, the mechanism by which PcG proteins are directed to particular loci is elusive<sup>43</sup>. We show that the control of H3K27 methylation is fundamentally different from that of epigenetic marks in constitutive heterochromatin. Not only is the genomic distribution of H3K27 methylation much more plastic than that of H3K9 methylation<sup>12-17,44</sup>, but also, unlike DNA methylation and methylation of histone H3K9, which are faithfully methylated *de novo* when introduced at arbitrary ectopic genomic sites<sup>45,46</sup>, we show that H3K27 methylation

is often position-dependent. We further demonstrate that telomere repeats underpin the observed position effect on H3K27me by showing that loss of telomerase abolishes subtelomeric H3K27me<sub>2/3</sub> and that artificial introduction of telomere repeats at interstitial sites trigger deposition of H3K27me<sub>2/3</sub>. That is, telomere repeats themselves are both necessary for subtelomeric H3K27me<sub>2/3</sub> and sufficient to trigger ectopic H3K27 methylation at internal chromosomal sites.

## Results

### **Analyses of classical chromosome rearrangements suggest that proximity to a chromosome end is key to subtelomeric H3K27 methylation**

To search for *cis*-acting sequences that trigger facultative heterochromatin formation in *N. crassa*, perhaps analogous to PREs in *D. melanogaster*, we dissected a 47 kb H3K27me<sub>2/3</sub> domain on linkage group (LG) VIL. A series of eight, partially overlapping, three kb fragments from this domain were separately targeted to both the *histidine 3* (*his-3*) and the cyclosporin A resistance I (*csr-1*) euchromatic loci in a strain in which we had deleted the endogenous H3K27me<sub>2/3</sub> domain (Fig. 3A). H3K27me<sub>2/3</sub> chromatin immunoprecipitation (ChIP) followed by qPCR demonstrated the absence of *de novo* H3K27me<sub>2/3</sub> within all eight segments when transplanted to either *his-3* or *csr-1* (Fig. 3B and C).

To address the possibility that the failure to induce H3K27me<sub>2/3</sub> in the transplantation experiments was simply due to the size of the test fragments, we utilized strains with large chromosomal rearrangements<sup>47</sup>. We first examined translocations that

involved the H3K27me<sub>2/3</sub> domain on LG VII that we had dissected, starting with UK3-41, an insertional translocation strain that has an approximately 1.88 Mb segment of LG VR inserted into a distal position on LG VII (Fig. 2A). The translocation shifted most of LG VII to a more interior chromosomal position. Interestingly, H3K27me<sub>2/3</sub> ChIP-seq of UK3-41 showed an absence of H3K27me<sub>2/3</sub> in the region that was displaced from the chromosome end (lost H3K27me<sub>2/3</sub>; indicated in orange in Fig. 2A). This result is consistent with our finding that transplantation of segments from this region did not induce methylation at *his-3* or *csr-1* and suggests that the normal methylation of this subtelomeric domain is position-dependent.

These results raised the question of whether H3K27me<sub>2/3</sub> in this domain absolutely depends on its normal location, or whether the methylation would occur if the region were moved near another chromosome end. To address this possibility, we utilized OY350, a quasiterminal translocation that transfers a distal segment of LG VII to the end of LG IR (Fig. 2B). Analysis of OY350 by H3K27me<sub>2/3</sub> ChIP-seq showed that the translocated LG VII domain preserved its normal H3K27me<sub>2/3</sub> distribution when placed adjacent to the adoptive telomere. This finding supported the hypothesis that H3K27me<sub>2/3</sub> in this domain depends on its proximity to the chromosome end and suggested that this requirement is non-specific; the different chromosome end apparently substituted for the native one. A similar situation was observed on LG VIR in OY320 (Fig. 2F).

The results described above suggested that proximity to a chromosome end may be sufficient to induce H3K27me. Interestingly, in addition to consistent losses of previously subtelomeric H3K27me<sub>2/3</sub> in the examined genomic rearrangements, we

identified new H3K27me<sub>2/3</sub> domains at novel subtelomeric regions. Striking examples of *de novo* H3K27me<sub>2/3</sub> were independently observed in seven genome rearrangements at their new chromosome ends (Fig. 2B-E and Fig. 4A-C; marked in green). These novel domains of H3K27me<sub>2/3</sub> extended an average of approximately 180 kb into the new subtelomeric regions. The only apparent exception concerns translocation UK2-32 (Fig. 4A), which involved the left end of LG V with approximately one Mb of tandemly repeated rDNA that is not shown in the *N. crassa* genome assembly. The H3K27me<sub>2/3</sub> distribution on all chromosomes unrelated to the genome rearrangements was unaltered (Fig. 5). Altogether, our survey of chromosomal translocations strongly suggests that chromosome ends promote the deposition of H3K27me on neighboring chromatin.

### **Identification of position-independent H3K27-methylated domains**

While domains of H3K27me<sub>2/3</sub> are enriched near the ends of chromosomes in *N. crassa*, substantial domains are also found elsewhere. To examine the possibility of position-independent H3K27me<sub>2/3</sub> residing within LG VIR, we examined the distribution of H3K27me<sub>2/3</sub> in two translocations that effectively moved LG VIR to the middle of a chromosome. ALS159 is a reciprocal translocation that shifted wild-type LG VIR approximately 4.5 Mb away from the new telomere (Fig. 2C). Although the shifted segment lost the H3K27me<sub>2/3</sub> that was previously closest to the chromosome end, the more internal H3K27me<sub>2/3</sub> was unaffected by the translocation (Fig. 2C). Similarly, when LG VIR was inserted in the middle of LG III in OY329 (Fig. 2D), most of the H3K27me<sub>2/3</sub> on LG VIR was retained. Again, the H3K27me<sub>2/3</sub> that was originally

closest to the telomere was lost, however, consistent with the idea that subtelomeric H3K27me<sub>2/3</sub> domains depend on their proximity to the chromosome ends.

To determine if the retention of H3K27me<sub>2/3</sub> from LG VIR was unique, we examined the position-dependence of two other H3K27me<sub>2/3</sub> domains. In translocation NM149, approximately 600 kb of LG IIL was translocated onto the end of LG VR, effectively shifting the native right arm of LG V about 600 kb away from the chromosome end (Fig. 2E). Like the behavior of LG VIR in ALS159 and OY329, losses of H3K27me<sub>2/3</sub> occurred in the most distal section of LG VR, but most H3K27me<sub>2/3</sub> was retained, even when moved away from the chromosome end. In addition, a large H3K27me<sub>2/3</sub> domain on LG IIR was shifted approximately 450 kb away from the chromosome end in OY320, yet it did not incur losses (Fig. 2F). Taken together, these findings demonstrate that some H3K27me<sub>2/3</sub> domains can be maintained when translocated to internal chromosomal sites, *i.e.* some H3K27me domains appear to be position-independent.

### ***Cis-trans* test on H3K27 methylation in a segmental duplication strain reveals position effect**

The observation that some chromosomal segments that are normally marked with H3K27me<sub>2/3</sub> lose the mark when translocated to another genomic location suggested that H3K27me<sub>2/3</sub> can be position-dependent. It remained formally possible, however, that the loss of H3K27me<sub>2/3</sub> was an indirect effect of the translocation, perhaps due to an altered transcriptional program. To examine the possibility that one or more *trans*-acting factors could be responsible for the loss of H3K27me<sub>2/3</sub> in OY329, we crossed this strain, which

is in an Oak Ridge (OR) background, to a highly polymorphic wild-type strain, Mauriceville (MV), to obtain progeny with a segmental duplication, *i.e.* with the OR translocated LG VIR domain inserted into LG II in a strain that also contained the corresponding region on the native LG VI of MV (Fig. 7A). The high density of single nucleotide polymorphisms (SNPs) in the MV background allowed us to separately map the H3K27me<sub>2/3</sub> distribution in the translocated (OR) segment and the wild-type (MV) segment<sup>48</sup>. SNP-parsed H3K27me<sub>2/3</sub> ChIP-seq showed that the duplicated chromosomal segments have distinct H3K27me<sub>2/3</sub> profiles (Fig. 7B). The previously subtelomeric H3K27me<sub>2/3</sub> that was lost in OY329 did not return in the duplication strain, nor was the loss in this region recapitulated on the native LG VIR of MV. The independent behavior of these homologous segments suggests that the loss of H3K27me<sub>2/3</sub> in OY329 was not a result of *trans*-acting factors, but rather was a result of the translocation itself, *i.e.* it is a *bona fide* position effect.

### **Altered gene expression in regions with changed H3K27me<sub>2/3</sub>**

Considering that H3K27me<sub>2/3</sub> normally marks transcriptionally quiescent chromatin in *N. crassa*<sup>9</sup>, it was of obvious interest to ascertain whether ectopic H3K27me<sub>2/3</sub> resulting from chromosomal translocations could cause gene silencing. To determine if the loss or gain of H3K27me<sub>2/3</sub> in the translocation strains activated or repressed gene expression, respectively, we performed poly-A mRNA sequencing on wild-type and seven translocation strains (ALS159, AR16, OY329, OY337, OY350, UK2-32, and UK3-41). Fig. 8A and B show gene expression and associated H3K27me<sub>2/3</sub> levels in the wild-type and translocation OY350 strains at representative

regions where former subtelomeres were moved and new ones created. The gain of H3K27me<sub>2/3</sub> on LG VIL in OY350, due to formation of a novel subtelomere, coincided with reduced transcript levels relative to wild-type (Fig. 8A). Conversely, the loss of H3K27me<sub>2/3</sub> on LG IR in OY350, due to shifting the subtelomeric region, was associated with an increase in transcript abundance compared to wild-type (Fig. 8B). Results of qPCR analyses of select genes confirmed these findings (Fig. 9). Indeed, a comprehensive analysis of poly-A mRNA-seq data from all seven translocation strains showed increases in gene expression at previously subtelomeric regions that lost H3K27me<sub>2/3</sub> and decreases in gene expression at novel subtelomeres that gained H3K27me<sub>2/3</sub> (Fig. 8C). These findings demonstrate that chromosomal rearrangements can cause marked changes in both H3K27me<sub>2/3</sub> and gene expression of the affected regions.

### **Loss of telomerase abolishes subtelomeric H3K27me<sub>2/3</sub>**

Our analyses of translocation strains demonstrated that domains of H3K27me<sub>2/3</sub> can be position-dependent and suggested that a feature of chromosome ends directly or indirectly recruits H3K27me<sub>2/3</sub> to subtelomeric regions. To investigate the basis of this, we utilized a mutant lacking the single telomerase reverse transcriptase (*tert*) that is responsible for all (TTAGGG)<sub>n</sub> telomere repeats normally found on *N. crassa* chromosome ends<sup>49</sup>. Southern analysis of the *tert* strain using a (TTAGGG)<sub>n</sub> probe revealed major reductions in the hybridization signals and showed that the majority of telomere fragments present in wild-type were either undetectable or greatly reduced in the mutant (Fig. 10A). In *S. pombe*, the majority of *trt1*<sup>-</sup> (yeast homologue of *N. crassa*

*tert*) survivors have circularized their chromosomes<sup>50</sup>. To determine if *N. crassa tert* survivors also have circular chromosomes, we designed outwardly directed PCR primers to amplify sequences near the ends of each chromosome and tested whether fragments were generated by fusions of the right and left chromosome ends. Indeed, the *tert* genomic DNA, but not control DNA from a wild-type strain, supported amplification using the divergent primer pairs, confirming intra-chromosomal fusions (Fig. 10B).

We performed H3K27me2/3 ChIP-seq to determine whether the subtelomeric domains of H3K27me2/3 were retained or lost in the *tert* strain and found clear evidence of loss of H3K27 methylation from chromosome ends (Fig. 10C). With the exception of LG V, which is unique in that one of its ends is capped with approximately 150 copies of the approximately 9 kb rDNA repeat<sup>51</sup>, all chromosomes lost their subtelomeric H3K27me2/3 domains, which typically extend tens of thousands of bp from their ends. Internal H3K27me2/3 domains were not noticeably altered (Fig. 10C and Fig. 11). Sequencing of *tert* input DNA from the ChIP showed no reduction in coverage near the chromosome ends, confirming that the losses of H3K27me2/3 ChIP signal observed in *tert* are not caused by chromosome degradation. We conclude that either (TTAGGG)<sub>n</sub> telomere repeats or linear chromosome ends are required for position-dependent domains of H3K27me2/3 in *N. crassa*.

### **Internal telomere repeats are sufficient to induce *de novo* H3K27me2/3**

To determine if the presence of telomere repeats is sufficient to trigger the deposition of H3K27 methylation, telomere repeats were separately targeted to the *csr-1* and *his-3* loci. Due to the repetitive nature of these sequences and our cloning strategy,



strains containing a variable number of (TTAGGG) repeats could be obtained through homologous recombination. Sequencing of inserted DNA in eleven independent transformants revealed that the number of inserted telomere repeats ranged from 5 to 23. Native *N. crassa* telomeres have an average of 20 repeats<sup>49</sup>. H3K27me<sub>2/3</sub> ChIP-qPCR demonstrated that ectopic telomere repeats are sufficient to induce local H3K27me<sub>2/3</sub> at the *csr-1* and *his-3* loci (Fig. 12A and Fig. 13). Even insertion of (TTAGGG)<sub>8</sub> triggered some methylation and insertion of (TTAGGG)<sub>17</sub> led to a high level of H3K27me<sub>2/3</sub> at *csr-1*. ChIP-seq on the (TTAGGG)<sub>17</sub> strain revealed new peaks as far as 170 kb from the insertion site and the semi-continuous H3K27me<sub>2/3</sub> domain spanned approximately 225 kb including 30 genes (Fig. 12B and Fig. 14). This is comparable to the size of subtelomeric H3K27me<sub>2/3</sub> domains that were lost in the *tert* strain.

## Discussion

The proper distribution of the facultative heterochromatin mark, H3K27me<sub>2/3</sub>, deposited by PRC2, is necessary for appropriate gene expression in a variety of plants, animals and fungi<sup>11</sup>. Unfortunately, the control of H3K27 methylation remains largely unknown. In *D. melanogaster*, DNA elements known as PREs are important to define domains of H3K27me and associated silencing, but even in this organism, PREs are neither sufficient in all genomic contexts, nor fully penetrant<sup>52,53</sup>. In addition, it was recently found that deletion of a well-studied PRE in *D. melanogaster* did not significantly affect either gene silencing or H3K27 methylation at its native locus<sup>54</sup>. The

control of H3K27 methylation is less defined in mammals, in which only a few candidate PRE-like elements have been identified<sup>55-57</sup>.

We took advantage of a relatively simple eukaryote bearing H3K27me, the filamentous fungus *N. crassa*, to explore the control of this chromatin mark. Conveniently, H3K27me is non-essential in this organism, in spite of being responsible for silencing scores of genes<sup>9,58</sup>. Unlike the situation with constitutive heterochromatin in *N. crassa*<sup>45,46</sup>, we demonstrated that not all facultative heterochromatin is entirely controlled by underlying sequence elements; transplanted gene-sized segments of a H3K27me domain do not, in general, become faithfully H3K27-methylated. However, translocations of large chromosomal segments revealed two distinct classes of transcriptionally repressive H3K27me domains, position-dependent and position-independent. Mechanistic insights from subsequent experiments defined these classes as telomere repeat-dependent and telomere repeat-independent H3K27me. The identification of telomere repeats as potent signals for H3K27 methylation represents a major advance in our understanding of facultative heterochromatin formation in eukaryotes.

Our study took advantage of a collection of spontaneous and UV-induced chromosome rearrangement strains of *N. crassa* that were primarily collected and characterized by David Perkins<sup>47</sup>. We surveyed representative rearrangement strains in which chromosomal regions containing domains of H3K27me were translocated to novel genomic positions. Strikingly, ChIP-seq revealed multiple cases in which subtelomeric H3K27me<sub>2/3</sub> domains completely lost this modification when moved to an interior chromosomal location. Indeed, no exceptions to this rule were found. Generation and

analysis of a segmental duplication strain containing differentially marked translocated and normal segments confirmed that the changed distribution of H3K27me<sub>2/3</sub> was a *bona fide* position effect rather than an effect of a *trans*-acting factor (Figure 2). Moreover, new subtelomeric regions gained H3K27me<sub>2/3</sub> over sequences that were previously devoid of this mark (Fig. 2). We conclude that proximity to a chromosome end, *per se*, somehow induces deposition of H3K27me<sub>2/3</sub> in domains that can span hundreds of kilobases. Analyses of polyA<sup>+</sup> mRNA from the strains revealed losses of gene expression associated with the new H3K27me<sub>2/3</sub> and gains in gene expression in regions that lost this mark; *i.e.* the changes in H3K27me<sub>2/3</sub> were reflected in underlying gene expression levels (Fig. 8).

Although gene silencing near chromosome ends, sometimes dubbed “telomere position effect” (TPE) has been observed in fungi<sup>19,59-62</sup>, *D. melanogaster*<sup>63</sup>, *Mus musculus*<sup>64</sup>, and human cells<sup>65</sup>, previous studies did not implicate H3K27me. Interestingly, the extent of heterochromatin spreading from telomeres in *N. crassa* is substantially greater than previously reported. This could be due to the conspicuous absence of canonical subtelomere sequences in *N. crassa*<sup>49,66-68</sup>. Not all organisms that exhibit TPE have H3K27 methylation machinery, but in *D. melanogaster*, which does, systematic analysis of PcG mutants demonstrated they do not disrupt telomere silencing<sup>63</sup>. In human cells, TPE is thought to be mediated by histone deacetylation, H3K9 methylation, and heterochromatin protein HP1 $\alpha$ <sup>67,69</sup>. Although the phenomenon of TPE has been studied for decades, our findings may reflect the first documented case of H3K27me-mediated TPE.

The ability of chromosome ends to induce large domains of H3K27me<sub>2/3</sub> in chromosomal translocation strains motivated us to investigate the role of telomere sequences in the establishment of facultative heterochromatin. We found that deletion of *tert*, the sole telomerase reverse transcriptase gene, results in a dramatic loss of H3K27me<sub>2/3</sub> at subtelomeres, concomitant with loss of (TTAGGG)<sub>n</sub> repeats and chromosome circularization (Fig. 10). To directly test the possibility that telomere repeats can trigger domains of H3K27me<sub>2/3</sub>, we inserted an array of telomere repeats at two euchromatic, interstitial sites and used ChIP to check induction of H3K27me<sub>2/3</sub>. We found that, indeed, even a 48 bp array, (TTAGGG)<sub>8</sub>, could induce some local H3K27me<sub>2/3</sub> at the *csr-1* locus. The independent induction of local H3K27me<sub>2/3</sub> at the *his-3* locus suggests this effect is not position-specific (Fig. 13). Remarkably, a 102 bp array, (TTAGGG)<sub>17</sub>, induced an H3K27me<sub>2/3</sub> domain that covered approximately 225 kb, including 30 genes (Fig. 12). These results strongly suggest that wild-type subtelomeric H3K27me<sub>2/3</sub> is dependent on telomere repeats. The recruitment of PRC2 to telomere repeats may be a widespread phenomenon, considering that enrichment of H3K27 methylation at telomeres has been observed in fungi<sup>9,37-39</sup>, plants<sup>70,71</sup> and animals<sup>72</sup>. Indirect recruitment of PRC2 by telobox-binding transcription factors<sup>42</sup> and direct binding of PRC2 to G-quadruplex RNA resulting from transcription of telomere repeats<sup>73</sup> represent possible molecular mechanisms. Indeed, it was recently reported that telomeric repeat-containing RNAs (TERRA) directly interact with EZH2 and that H3K27me<sub>3</sub> ChIP-seq peaks are strongly correlated with TERRA-associated chromatin in mouse embryonic stem cells<sup>74</sup>.

In addition to telomere repeats capping the ends of chromosomes, telomere repeat-like elements are scattered interstitially in the genomes of many organisms<sup>75</sup>. In *N. crassa*, interstitial telomere sequences are rare, limited to less than four tandem repeats, and are not preferentially associated with H3K27me<sub>2/3</sub> regions. While *S. cerevisiae* and *S. pombe* lack PRC2 components and H3K27me, their internal telomere repeat-like elements can promote heterochromatin formation<sup>76,77</sup>. The genome of *Nicotiana tabacum* also has heterochromatic internal telomere repeats, but they lack the H3K27 methylation present at genuine telomeres<sup>78</sup>. Curiously, insertion of about 130 telomere repeats in Chinese ovary cells failed to significantly alter local transcription<sup>79</sup>. It should be interesting to study the effects of terminal and interstitial telomere repeats on heterochromatin formation in a variety of organisms.

We inferred the existence of position-independent H3K27 methylation from our observation that some H3K27me<sub>2/3</sub> domains were unaffected when moved to ectopic chromosomal locations (*e.g.* see domains on LG IIIR in OY320, LG VR in NM149, and LG VIR in ALS159 and OY329; Fig. 2). The recapitulation of normal H3K27me<sub>2/3</sub> profiles at ectopic chromosomal sites is consistent with the possibility that these chromosomal regions contain *cis*-acting signals that trigger the deposition of H3K27 methylation, perhaps comparable to PREs in *D. melanogaster*. It will be of interest to define the presumptive elements responsible for such position-independent H3K27me. While there is no consensus sequence for PREs in *D. melanogaster*, they are composed of unique combinations of binding sites for a variety of factors<sup>80</sup>. One DNA-binding factor known to affect PcG silencing in *D. melanogaster*, GRH, has an obvious homolog in *N. crassa* and could be a candidate for helping establish or maintain position-independent

domains of H3K27me2/3 *via* sequence-specific interactions<sup>81-83</sup>. It remains possible that deposition of H3K27 methylation at position-independent domains is not directed by sequence-specific elements. In this context, it is worth noting that inhibition of transcription, caused by either exposure to RNA polymerase II inhibitors<sup>84</sup> or deletion of a transcriptional start site<sup>85</sup> was found to promote the deposition of H3K27me in mammalian cells. Conceivably, the position-independent domains of H3K27me2/3 in *N. crassa* could be directed by blocs of inherently low-expressing genes or by some other feature of the region, such as its location in the nucleus, which may be controlled by yet undefined factors. The fact that only about 10% of H3K27me2/3-marked genes are upregulated when the sole H3K27 methyltransferase is removed supports the notion that transcriptional shut-off may precede PRC2 recruitment<sup>9,58</sup>. Still, low transcriptional activity cannot entirely explain the targeting of PRC2 in *N. crassa* since many low-expressed genes are not H3K27 methylated.

Genome rearrangements are a common occurrence in malignant cells and can drive tumorigenesis through the creation of gene fusions, enhancer hijacking and oncogene amplification<sup>86</sup>. Our findings suggest that genome rearrangements associated with cancer may additionally impact gene expression through effects on facultative heterochromatin. Indeed, there are already some reports of human diseases being driven by position effects<sup>87,88</sup>. A great deal of research has focused on how characteristics of local chromatin influence translocation breakpoint frequencies<sup>89</sup>, but the effects of the resulting translocations on chromatin state are still ill-defined. Altogether, our work shows that changes in genome organization can have sweeping effects on both the

distribution of an epigenetic mark and gene expression. Thus, chromosome rearrangements may have unappreciated roles in evolution and cancer etiology.

## Materials and Methods

### Strains, media and growth conditions

*N. crassa* strains are listed in Appendix B and were grown and maintained according to standard procedures<sup>29</sup>. All genome rearrangement strains are available through the Fungal Genetics Stock Center ([www.fgsc.net](http://www.fgsc.net)).

### Deleting and targeting segments of LG VIL

A strain containing a 47.4 kb deletion of LG VIL (N4933) was constructed by homologous recombination using primers listed in Appendix C<sup>90</sup>. For *his-3* constructs, segments from LG VIL were PCR-amplified from wild-type (N3752) genomic DNA with primers containing restriction enzyme sites (Appendix C) and directly cloned into pCR (Life Technologies TA Cloning Kit), subcloned into the *his-3*-targeting vector pBM61 (Appendix D) and transformed into *N. crassa* strain N5739<sup>91</sup>. For *csr-1* replacements, segments from LG VIL were PCR-amplified from wild-type (N3752) genomic DNA with primers containing homology to the 5' and 3' flanks of the *csr-1* locus (Appendix C). Amplified segments were subsequently assembled by PCR ("PCR-stitched") to the 5' and 3' flanks of the *csr-1* locus respectively. Stitched PCR products were co-transformed into *N. crassa* strain N2931 and transformants were selected on cyclosporin A-containing medium<sup>92</sup>.

### **Deletion of *tert***

The *nat1* gene with the *trpC* promoter was amplified by PCR using the pAL12-Lifeact as the template with primers 3902 and 1369. The 5' and 3' flanking fragments of the *tert* gene were amplified by PCR with primers 3406-3409, which have specific 29-bp overhang sequence with the 5' and 3' *nat1* gene with the *trpC* promoter. The three PCR products were gel-purified, combined and PCR-stitched with primers 3406 and 3409 to construct the knockout cassette. The cassette was gel-purified and transformed into a  $\Delta$ *mus-51* strain (N2929) by electroporation and resulting strains were crossed with a *mus-51*<sup>+</sup> strain (N3752) to recover progeny with the wild-type allele of *mus-51*.

### **Targeting telomere repeats to *csr-1* and *his-3***

Concatamers of telomere repeats were amplified in a polymerase chain reaction lacking template as previously described<sup>49</sup>. Resulting PCR products of ~500 bp were gel-purified and cloned into the pCR4-TOPO TA vector (Life Technologies TOPO TA Cloning kit for Sequencing). Telomere repeats from the cloning vector were PCR-stitched separately to 5' and 3' flanks of the targeting locus. For *csr-1* targeting, stitched PCR products were co-transformed into strain N5739 and homokaryotic transformants were selected on cyclosporin A-containing medium<sup>92</sup>. For *his-3* targeting, stitched PCR products were co-transformed into strain N2834 and heterokaryotic transformants were selected on medium lacking histidine.



## ChIP and preparation of libraries

ChIP was performed as previously described<sup>12</sup>. An anti-H3K27me2/3 antibody (Active Motif, 39535) was used for all experiments. ChIPs for Fig. 2-5 and Fig. 8 were performed in biological triplicate. Real-time qPCR was performed as previously described<sup>9</sup>. ChIP-seq libraries were also prepared as previously described<sup>12</sup>, except N51, N6089, N6228 and N6383 libraries were subjected to eight cycles of amplification. Sequencing was performed using the Illumina NextSeq 500 using paired-end 100 nucleotide reads for input chromatin and some ChIP experiments. The remainder of the ChIP experiments were sequenced using the Illumina HiSeq 2000 using single-end 100 nucleotide reads or the NextSeq500 using single-end 75 nucleotide reads. All sequences were mapped to the corrected *N. crassa* OR74A (NC12) genome<sup>36</sup> using Bowtie2<sup>93</sup>. ChIP-seq read coverage was averaged over 100 bp sliding (50 bp increment) windows with BEDTools<sup>94</sup> and normalized to wild-type by coverage. ChIP-seq data were visualized with Gviz<sup>95</sup>. To adjust for different sequencing depths between wild-type and translocation samples, a scaling factor was determined for each sample as a ratio of total number of mapped reads in the given sample relative to wild-type. Each scaling factor was then used to normalize ChIP-seq samples. ChIP-seq data were visualized with either IGV<sup>96</sup> for Fig. 2-5, 7, 10-12 and 14 or Gviz for Fig. 8<sup>95</sup>. Normalized H3K27me2/3 ChIP-seq data were displayed using a 350 or 500 bp sliding window unless otherwise specified in the figure legend.

## **RNA isolation and preparation of libraries**

RNA isolation and poly(A)-RNA enrichment, RNA-seq library preparation, and subsequent differential expression analysis were performed with replicate samples as previously described<sup>58</sup>. RNA-seq data were visualized with Gviz<sup>95</sup>.

## **Translocation breakpoint analyses**

Incongruous-paired and split-read alignments for input DNA samples were found using bwa-mem. The discordant and split-read alignments were analyzed with LUMPY v0.2.9 to determine the location of chromosomal breakpoints<sup>97</sup>. A list of predicted breakpoints with a minimum call weight of five was generated and calls with an evidence set score <0.05. The chromosomal breakpoints were compared to mapping data determined by recombination mapping<sup>47</sup>, RFLP coverage and inverse PCR (<http://hdl.handle.net/10603/68>) to remove calls that were not associated with the translocation of interest. PCR analyses of genome rearrangements ALS159, NM149, OY329, OY320 and UK3-41 were consistent with predicted breakpoint patterns.

## **Southern analysis**

Southern hybridization analyses were performed as previously described<sup>98</sup>. *N. crassa* telomere restriction fragments were examined as described previously with a HindIII/NotI double digest<sup>49</sup>. Primers used to make probe are listed in Appendix C.

## SNP Mapping

ChIP-seq data for the duplication strain were converted into fasta format, trimmed to 70 nt, and filtered for 70 nt long reads using FASTX-Toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit](http://hannonlab.cshl.edu/fastx_toolkit)). Processed data were SNP-parsed with Hashmatch<sup>99</sup> using the Mauriceville-Oakridge SNPome (fasta file listing both versions of each SNP)<sup>48</sup>. The SNPome file was modified to align with read lengths of 70 nt. Only reads with 100% alignment to either genome were kept for further analysis and were allocated into two files according to SNP mapping. Each file was then converted back into fastq format using SeqTK 1.0 (<https://github.com/lh3/seqtk>) and remapped with Bowtie2 to the genome corresponding with its SNP alignment. To adjust for the variable mapping efficiency of each SNP-parsed file to either Mauriceville or Oakridge, samples were normalized using a scaling factor. The scaling factor was calculated as a ratio of the number of alignments for each SNP-parsed file relative to the file with the smallest number of alignments. H3K27me2/3 profiles were displayed in IGV<sup>96</sup> to examine the individual H3K27me2/3 profiles for wild-type and rearranged chromosome segments.

## Bridge to Chapter IV

Our examination of chromosomal rearrangement strains and engineered strains of *N. crassa* established that at least two classes of H3K27 methylation exist: telomere-dependent and telomere-independent. However, how these two kinds of H3K27 methylation are established and maintained was left unanswered. In the next chapter, using a forward-genetics approach we unexpectedly gain some insight into the apparent

division of labor by PRC2 suggested in our findings described in Chapter III. In particular, we identified an accessory subunit of PRC2 that is responsible for subtelomeric H3K27 methylation.

**CHAPTER IV**

**IDENTIFICATION OF A PRC2 ACCESSORY SUBUNIT REQUIRED FOR  
SUBTELOMERIC H3K27 METHYLATION IN NEUROSPORA**

This work was a collaborative effort. I mapped and complemented one of the two identified *NCU04278* mutant alleles, analyzed linear growth rates of  $\Delta$ *NCU04278* strains, performed and verified H3K27me<sub>2/3</sub> ChIP-seq on  $\Delta$ *NCU04278* strains, verified gene expression changes in  $\Delta$ *NCU04278* strains by RT-qPCR, performed all immunoprecipitations for mass spectrometry, and wrote first draft of this chapter. Elizabeth Wiles created the forward-genetic selection strain and generated the *NCU04278* UV alleles, as well as performed the mRNA-seq experiments. Eric Selker was the principal investigator for this work.

**Introduction**

Specification and maintenance of cellular fate in higher eukaryotes requires the concerted effort of transcription factor networks and chromatin-modifying complexes. Polycomb repressive complex 2 (PRC2), which catalyzes methylation of lysine 27 on histone H3 (H3K27), is one such chromatin-modifying complex essential for normal development in many organisms<sup>100,101</sup>. The three core components of PRC2, which are conserved in lineages of plants, animals, and fungi<sup>7,102</sup>, are EED, SUZ12, and the methyltransferase EZH2. In addition to associating with RBB4/7, human PRC2 interacts with accessory proteins, defining two main subtypes designated PRC2.1 and PRC2.2<sup>103</sup>.

Distinct subcomplexes of PRC2 have also been described in *Drosophila melanogaster*<sup>104-106</sup>, suggesting that this may be a common feature.

Here, using a forward genetic selection for factors defective in Polycomb silencing (Chapter V), we isolated two alleles of a gene, *NCU04278*, necessary for subtelomeric H3K27 methylation in the filamentous fungus, *Neurospora crassa*. As expected, the loss of H3K27 methylation at the subtelomeres led to concomitant upregulation of subtelomeric genes. Immunoprecipitation followed by mass spectrometry of *NCU04278*-interacting protein demonstrated that *NCU04278* is a PRC2 accessory subunit, and was therefore deemed PAS. Although PAS homologs appear restricted to the fungal lineage, these findings reiterate the crucial role PRC2 accessory subunits can have in target site selection.

## Results

### Isolation, mapping and identification of *NCU04278* as potential PcG gene

Using a forward genetic selection for factors necessary for Polycomb silencing (Chapter V), we isolated a Hygromycin B- and Nourseothricin-resistant (*hnr*) mutant (Fig. 15A). We mapped the causative mutation in the *hnr* strain (Oak Ridge genetic background) by crossing it to the highly polymorphic “Mauriceville” strain<sup>107</sup>, pooling genomic DNA from progeny with an antibiotic-resistant phenotype, and scoring the percentage of Oak Ridge single nucleotide polymorphisms (SNPs) from whole-genome sequencing data<sup>48</sup>. We observed an enrichment of Oak Ridge SNPs on the right arm of

linkage group (LG) V that included a frameshift mutation (ATG -> TTTG) in *NCU04278* (Fig. 15B).

To verify that the observed mutation in *NCU04278* is the critical mutation in *hnr*, we targeted a wild-type copy of *NCU04278* to the *his-3* locus in the *hnr* mutant background. This ectopic copy of *NCU04278* complemented the *hnr* mutant, *i.e.*, restored Hygromycin B sensitivity (Fig. 15C). In addition, deletion of *NCU04278* was sufficient to confer Hygromycin B resistance (Fig. 15C). We subsequently isolated and mapped a second allele of *NCU04278* (Fig. 16A), further confirming the involvement of *NCU04278* in the antibiotic drug-resistant phenotype.

### **NCU04278 is necessary for silencing subtelomeric H3K27-methylated genes**

Although our forward genetic selection was designed to identify novel components of the Polycomb repression pathway, mutations in *NCU04278* could possibly confer resistance to Hygromycin B and Nourseothricin not through the derepression of the H3K27-methylated antibiotic-resistance genes, but rather through an alternative means, such as more efficient drug efflux<sup>108</sup>. Alternatively, our selection could have isolated a mutant that had global transcriptional defects, which led to the derepression of *hph* and *nat-1*, but was not defective in a specific component of the Polycomb repression mechanism.

To determine if loss of *NCU04278* has specific defects in Polycomb silencing, we performed mRNA-seq on  $\Delta$ *NCU04278* and wild-type strains in biological replicate and compared the gene expression profiles with previously generated wild-type and  $\Delta$ *set-7* data sets<sup>58</sup>. We found that 33 genes were upregulated, and 12 downregulated, greater

than two-fold in  $\Delta NCU04278$  strains compared to wild-type strains (Fig. 17A,B). Although less than 9% of all genes are H3K27-methylated in a wild-type strain, 64% of the upregulated genes in  $\Delta NCU04278$  strains were in this select group. Moreover, there was significant overlap between the upregulated genes in  $\Delta NCU04278$  and  $\Delta set-7$  strains ( $P = 6.095 \times 10^{-29}$ ), although loss of SET-7 appeared to derepress more H3K27-methylated genes (Fig. 17A), *i.e.*,  $\Delta NCU04278$  strains derepress a subset of SET-7 targets.

We have previously demonstrated that there are at least two types of H3K27 methylation in *N. crassa*, telomere-dependent and telomere-independent<sup>109</sup>. We reasoned that since loss of NCU04278 only affected a subset of SET-7 targets, NCU04278 may be responsible for specifically silencing either telomere-proximal or telomere-distal H3K27-methylated genes. We therefore analyzed the proximity of H3K27-methylated genes upregulated in  $\Delta NCU04278$  and  $\Delta set-7$  strains, respectively, and found that the genes derepressed by loss of NCU04278 are closer to the chromosome ends compared to the genes derepressed by loss of SET-7 (Fig. 17C). This suggested that NCU04278 is responsible for silencing telomere-proximal H3K27-methylated genes.

We verified the results of our mRNA-seq experiments by performing reverse transcription followed by quantitative polymerase chain reaction (RT-qPCR) on RNA isolated from biological triplicates of wild-type,  $\Delta set-7$ , and  $\Delta NCU04278$  strains. In addition to confirming the derepression of the native genes (*NCU07152* and *NCU05173*) replaced by the antibiotic-resistance genes (*nat-1* and *hph* respectively) in the selection strain, we further confirmed a significant increase in gene expression for five out of six H3K27-methylated genes that appeared upregulated by mRNA-seq in both  $\Delta NCU04278$



and  $\Delta set-7$  strains (Fig. 17D). We conclude that NCU04278 represses a subset of SET-7 targets that are telomere-proximal.

### **NCU04278 is necessary for maintenance of subtelomeric H3K27 methylation**

Considering that the loss of NCU04278 led to the derepression of H3K27-methylated genes near the chromosome ends, we wondered if there was also a concomitant loss of H3K27 methylation at subtelomeric regions. To address this question, we performed H3K27me2/3 chromatin immunoprecipitation followed by sequencing (ChIP-seq) on two  $\Delta NCU04278$  siblings and compared it to the wild-type distribution of H3K27me2/3<sup>110</sup> (Fig. 18A). We found that the majority of H3K27-methylated genes did not exhibit changes in H3K27me2/3 level in  $\Delta NCU04278$  strains compared to wild type (Fig. 18A,B). However, we did detect 284 genes in  $\Delta NCU04278$  strains that had at least two-fold reductions in H3K27me2/3 compared to wild-type strains (blue dots; Fig. 18B) and 34 genes that showed at least two-fold greater H3K27me2/3 than wild-type strains (red dots; Fig. 18B). The losses of H3K27me2/3 were concentrated at chromosome ends (Fig. 18A) and were closer to the telomere than the gains (Fig. 19A). To confirm the H3K27me2/3 ChIP-seq results, we performed H3K27me2/3 ChIP followed by qPCR (ChIP-qPCR) on wild-type,  $\Delta set-7$ , and  $\Delta NCU04278$  strains in biological triplicate. We confirmed three regions expected to have no change in H3K27me2/3 (*NCU05086*, *NCU08085*, and *NCU08251*), three regions with expected losses in H3K27me2/3 (*NCU05173*, *NCU07152*, and Telomere IIIR), and one region with an expected gain in H3K27me2/3 (*NCU07801*) (Fig. 18C). Thus NCU04278 is critical for maintaining subtelomeric H3K27 methylation.

To determine if there was a correlation between the significant gene expression changes in  $\Delta NCU04278$  strains and the losses or gains of H3K27 methylation in  $\Delta NCU04278$ , we examined the intersection of these two data sets. We found that all but one (*NCU08790*) of the 21 H3K27-methylated genes upregulated in  $\Delta NCU04278$  strains were on the list of genes that lost H3K27 methylation in  $\Delta NCU04278$  strains. In contrast, there was no overlap between the genes that gained H3K27 methylation in  $\Delta NCU04278$  strains and the genes that were downregulated  $\Delta NCU04278$  strains. Analysis of the genes that gained H3K27me<sub>2/3</sub> in  $\Delta NCU04278$  strains revealed that the majority (88 %) fall into the lowest quartile of gene expression in wild-type strains, suggesting that they were already effectively ‘off’ (Fig. 19B). In summary, the losses of H3K27 methylation in  $\Delta NCU04278$  strains are associated with, but not sufficient for, increased gene transcription, and the gains of H3K27 methylation in  $\Delta NCU04278$  strains are mostly on genes that are normally poorly transcribed.

### **NCU04278 is an accessory subunit of PRC2**

To determine if NCU04278 acts in a complex, we immunopurified NCU04278-3xFLAG from *N. crassa* lysate and identified its copurifying proteins by mass spectrometry. We identified NCU04278 itself, as well as the four known components of PRC2: SET-7, SUZ12, EED, and NPF (Fig. 20A)<sup>9</sup>. We also immunopurified SUZ12-3xFLAG and a negative control (3xFLAG-control), and analyzed their copurifying proteins by mass spectrometry. In addition, we examined previously collected mass spectrometry data from a 3xFLAG-EED purification<sup>9</sup>. The purifications of the known PRC2 components, SUZ12 and EED, yielded NCU04278 as well as the rest of the known

members of the complex (Fig. 20A). The 3xFLAG-control purification did identify peptides from NPF, but did not detect any peptides from SET-7, SUZ12, EED, or NCU04278 (Fig. 20A). Besides the proteins listed in Fig. 20A, the only other proteins with peptides detected in all the NCU04278, SUZ12, and EED purifications, but absent in the control, were NCU01249 (importin  $\alpha$ ), NCU02407 (dihydrolipoyl dehydrogenase), and NCU06482 (pyruvate dehydrogenase). Thus, it appears that PRC2 in *N. crassa* can form a five-member complex and that NCU04278 is an accessory component. For this reason, we have deemed NCU04278 as PAS (PRC2 accessory subunit) (Fig. 20B).

## Discussion

Ever since the identification of Polycomb response elements (PREs) in *Drosophila melanogaster*, sequences that reliably induce H3K27 methylation<sup>41</sup>, there has been an intense search for other *cis*- and *trans*-acting factors that direct PRC2 activity in a wide variety of organisms. In plants, this line of inquiry has been fruitful, with researchers demonstrating the direct interaction of PRC2 with transcription factors that bind specific DNA sequences<sup>42,111,112</sup>. In mammals, unmethylated CpG islands can apparently recruit PRC2<sup>113-115</sup>, but the mechanism of this recruitment is unclear. Non-core subunits of mammalian PRC2.1 and PRC2.2 are collectively required to target PRC2<sup>116,117</sup>, but how these structurally distinct components can independently target the same genomic sites is unsettled. Defining the full complement of *cis*- and *trans*-acting factors and how they specifically recruit PRC2 remains an open challenge.

Utilizing a relatively simple eukaryote, the fungus *N. crassa*, we performed a forward genetic selection to identify novel factors implicated in Polycomb repression. This yielded an unknown, fungal-specific, accessory component of PRC2, PAS (NCU04278), that is necessary for subtelomeric H3K27 methylation and associated gene silencing. Interestingly, we had already observed an H3K27 methylation defect similar to that observed in  $\Delta pas$  strains in strains lacking the PRC2 component, NPF<sup>9</sup>. Indeed, H3K27me<sub>2/3</sub> ChIP-seq profiles from  $\Delta pas$  and  $\Delta npf$  strains bear a striking resemblance (Fig. 18A), although they are clearly different. It was not certain that the lack of NPF in PRC2 was the direct cause of the observed defects in  $\Delta npf$  strains, as this protein also associates with other complexes<sup>38</sup>,  $\Delta npf$  strains exhibit greater changes in gene expression than  $\Delta set-7$  strains<sup>58</sup> and  $\Delta npf$  strains display growth defects not seen in other PRC2 member knockout strains<sup>9</sup>. It is notable that  $\Delta pas$  strains, although quite similar to  $\Delta npf$  strains in H3K27 methylation defects, have more limited changes in gene expression (Fig. 17A,B) and no detectable growth defect (Fig. 19C). Despite difficulties in establishing direct or indirect effects of NPF loss on PRC2 recruitment *in vivo* in *N. crassa*, such experiments are not feasible in mammalian cell lines as RBBP4/7 are essential for proliferation<sup>10</sup>.

Our research group has previously identified telomere repeats, (TTAGGG)<sub>n</sub>, as effective inducers of H3K27 methylation at chromosome ends in *N. crassa*<sup>109</sup>. This subtelomeric H3K27 methylation is largely lost in  $\Delta pas$ ,  $\Delta npf$  and  $\Delta tert$  (telomerase) strains (Fig. 18A)<sup>109</sup>. Although the exact link between the *cis*-acting telomere repeats and induced H3K27 methylation remains unclear, the identification of PAS should help bridge this gap in our knowledge.

## **Materials and Methods**

### **Strains, media and growth conditions**

All *N. crassa* strains used in this study are listed in Appendix B. Preparation of media and growth conditions for experiments were carried out as previously described (Chapter V).

### **Selection for mutants defective in H3K27 methylation-mediated silencing**

Conidia from strain N6279 were mutagenized with ultraviolet light and challenged with Hygromycin B and Nourseothricin to select for drug-resistant colonies as previously described (Chapter V). Primary mutants were rendered homokaryotic by crossing to strain N3756.

### **Whole genome sequencing, mapping and identification of *NCU04278* alleles**

This was performed as previously described (Chapter V). Briefly, drug-resistant, homokaryotic mutants were crossed to a “Mauriceville” strain<sup>107</sup> and drug-resistant progeny were pooled for whole genome sequencing (SRA accession: PRJNA559544). Mapping of the causative mutations was performed as previously described<sup>48</sup>. FreeBayes and VCFtools were used to identify novel genetic variants present in our pooled mutant genomic DNA<sup>118,119</sup>.

### **Chromatin immunoprecipitation (ChIP), ChIP-qPCR, and ChIP-seq**

H3K27me<sub>2/3</sub> ChIP using anti-H3K27me<sub>2/3</sub> antibody (Active Motif, 39536) was performed as previously described (Chapter V) and the isolated DNA used for qPCR or prepared for sequencing (Chapter V). Mapping, visualization, and analysis of H3K27me<sub>2/3</sub> ChIP-sequencing reads was performed as previously described (Chapter V).

### **RNA isolation, RT-qPCR, and mRNA-seq**

Extraction of total RNA from germinated conidia was performed as previously described (Chapter V) and either used for cDNA synthesis and subsequent qPCR (Chapter V) or mRNA-seq library preparation<sup>58</sup>. Mapping and analysis of gene expression levels was performed as previously described (Chapter V).

### **Immunoprecipitation followed by mass spectrometry (IP-MS)**

Approximately  $1 \times 10^9$  conidia collected from 7-10 days of growth of strain N7807 (over-expressing *NCU04278::C-Gly::3xFLAG*) or strain N4666 (endogenous *suz12::C-Gly::3xFLAG*) in a 250 mL flask containing 50 mL of Vogel's minimal medium containing 1.5% sucrose and 1.5% agar was filtered through sterile cheesecloth. Filtered conidia were used to inoculate 1 L of Vogel's minimal medium containing 1.5% sucrose and then grown for 16 hours shaking (150 RPM) at 30 °C. Tissue was collected by filtration using a Buchner funnel and washed with water. Dry tissue (~10 g) was ground using a 6870 Freezer/Mill Cryogenic Grinder (SPEXSamplePrep). Ground tissue was added to 40 mL Extraction Buffer (EB) (50 mM HEPES [pH=7.5], 150 mM NaCl, 10

mM EDTA, 10% glycerol, 0.02% NP-40, 1X Halt Protease Inhibitor Cocktail (Thermo Scientific, 78438)) to achieve a final volume of 50 mL, and then suspended by rotation at 4 °C for 1.5 hours. Insoluble material was pelleted with one 10-minute centrifugation step at 2,000 RPM and two consecutive 10-minute centrifugation steps at 8,000 RPM. Soluble material was pre-cleared with 250 µL of Protein A agarose (Invitrogen, 15918014) equilibrated in EB with rotation for 1 hour at 4 °C. The Protein A agarose was pelleted by centrifugation at 2,000 RPM and the supernatant collected, which was then incubated with 400 µL of ANTI-FLAG M2 affinity gel (Sigma-Aldrich, A2220) equilibrated with EB overnight, rotating at 4 °C. Resin was pelleted by centrifugation at 1,000 RPM and supernatant removed. Resin was washed with EB, rotated at 4 °C for 10 minutes, and pelleted with a 1,000 RPM spin five consecutive times. All liquid was removed after the final spin. Protein was eluted twice from the resin by incubating with 300 µL of 500 µg/mL 3X Flag Peptide (APExBIO, A6001) rotating at 4 °C for 20 minutes, and one final wash with 300 µL of EB. Eluate was precipitated with trichloroacetic acid (10% final concentration) on ice for 1 hour, pelleted by centrifugation at 14,000 RPM, and washed three times with ice-cold acetone. Pellet was air-dried by placing in heat block at 100 °C for 30 seconds. Samples were sent to and processed by the UC Davis Proteomics Core Facility for mass spectrometry and subsequent analysis.

### **Replacement of *NCU04278* with *trpC::nat-1***

The 5' and 3' flanks of *NCU04278* were PCR-amplified from wild-type genomic DNA with primer pairs 6405 and 6406 (5') and 6409 and 6410 (3'). The 5' and 3' flanks were separately PCR-stitched to plasmid 3237 (source of *trpC::nat-1*) using primer pairs

6401 and 4883, and 4882 and 6351, respectively. These two ‘split-marker’ PCR products were co-transformed into strain N2930 and *NCU04278* replacements were selected on Nourseothricin-containing medium.

### **Generation of a *N. crassa* strain expressing *NCU04278::3xFLAG***

*NCU04278* was PCR-amplified from wild-type genomic DNA with primers 6397 and 6398, and cloned into plasmid 2401<sup>120</sup> using XbaI and PacI restriction sites to create plasmid 3340. Plasmid 3340 was subjected to Sanger sequencing using primers 6397, 6407, 6426, 6427, and 6715 to verify that the sequence matched that of wild-type. Plasmid 3340 was linearized with NdeI and targeted to *his-3* in N6762, as previously described<sup>91</sup>. A *his-3*<sup>+</sup> primary transformant was then crossed to N7742 to generate N7807.

## **Bridge to Chapter V**

In this chapter, we utilized a powerful forward genetic selection, developed by Elizabeth Wiles, and identified an unknown accessory subunit of PRC2 that is necessary for subtelomeric H3K27 methylation. H3K27-methylated genes near the chromosome ends lose their methylation and a subset of these are derepressed. While this result does give us some insight into the workings of PRC2 targeting/activity, the mechanism of how H3K27 methylation mediates repression is unresolved. In the next chapter, we address this issue when we identify a ‘reader’ and effector of H3K27 methylation using



the aforementioned genetic selection. This protein, which we deemed EPR-1, is not limited to fungi but exists in a variety of eukaryotic lineages.

**CHAPTER V**  
**EVOLUTIONARILY ANCIENT BAH-PHD PROTEIN**  
**MEDIATES POLYCOMB SILENCING**

This work was a collaborative effort. I mapped and complemented one of the four identified *epr-1* mutant alleles, analyzed linear growth and sexual development phenotypes of  $\Delta epr-1$  strains, performed H3K27me3 westerns and H3K27me2/3 ChIP-seq on  $\Delta epr-1$  strains, created GFP-EPR-1-expressing strains and performed some microscopy, created EPR-1-Dam-expressing strains and performed DamID-seq and Southern, as well as wrote initial draft of the main body of the manuscript. Elizabeth Wiles created the forward-genetic selection strain and generated the *epr-1* UV alleles, verified the mRNA-seq results, verified the H3K27me2/3 ChIP-seq results, performed and verified GFP-EPR-1 ChIP-seq results, cloned BAH domain of EPR-1, and contributed significantly to manuscript. Saumya De Silva, under supervision of Catherine Musselman, performed NMR titration experiments. Gurmeet Kaur, under supervision of L. Aravind, performed the phylogenetic analysis. Jeanne Selker performed some microscopy experiments. Eric Selker was the principal investigator for this work.

**Introduction**

The establishment and maintenance of transcriptionally repressive chromatin is critical for the development of multicellular organisms<sup>121-124</sup>. Polycomb group (PcG) proteins, originally discovered in *Drosophila melanogaster*<sup>31</sup>, form multiple complexes

that maintain such chromatin repression<sup>125</sup>. Although the composition of PcG complexes varies, a few core constituents define two major classes of chromatin-modifying complexes, namely Polycomb repressive complex 1 (PRC1) and Polycomb repressive complex 2 (PRC2)<sup>7</sup>. According to the ‘classical model,’ PcG-mediated gene silencing is initiated by targeting of PRC2 to chromatin<sup>80</sup>, which catalyzes methylation of H3K27<sup>34</sup>. Canonical PRC1, which contains a chromodomain protein (*e.g.* Polycomb in *Drosophila melanogaster* and CBX2/4/6-8 in mammals), recognizes tri-methylated H3K27<sup>126</sup>, catalyzes monoubiquitination of neighboring histone H2A lysine 119 by RING1A/B<sup>35</sup>, and promotes chromatin compaction<sup>127,128</sup>. In reality, this hierarchical recruitment model is an oversimplification, as PRC1 can be recruited to PcG targets irrespective of PRC2 activity<sup>129</sup> and PRC1 presence is required for stable PRC2 association at many Polycomb Response Elements in *Drosophila melanogaster*<sup>130</sup>. Interdependence of these complexes has limited our understanding of their respective roles and the function of their associated chromatin ‘marks’ on gene repression.

While plants and animals utilize distinct sets of accessory proteins to recognize methylated H3K27<sup>126,131-133</sup>, they are generally thought to mediate repression in the context of a canonical PRC1 complex<sup>7,134-136</sup>. In fungal lineages that employ H3K27 methylation as a repressive chromatin mark, however, core PRC1 components are notably absent<sup>7</sup>. This raises the question of how H3K27 methylation mediates repression in the absence of PRC1. It suggests that either: 1. H3K27 methylation *per se* may be repressive, or 2. There is a ‘reader’ of H3K27 methylation that functions outside the context of canonical PRC1.

To elucidate the repressive mechanism of H3K27 methylation in fungi, we developed and employed a forward genetics approach to identify effectors of Polycomb repression using *Neurospora crassa*. H3K27 methylation covers approximately 7% of the *N. crassa* genome and is responsible for the repression of scores of genes<sup>9,58</sup>. We found four mutant alleles of an undescribed gene (*NCU07505*) that we show is critical for H3K27 methylation-mediated silencing and therefore named it effector of Polycomb repression 1 (*epr-1*). It encodes a protein with a bromo-adjacent homology (BAH) domain and plant homeodomain (PHD) finger. Although *epr-1* mutants display phenotypic and gene expression changes similar to strains lacking PRC2 components, H3K27 methylation is essentially unaffected. We demonstrate that EPR-1 forms nuclear foci, reminiscent of Polycomb bodies<sup>137</sup>, and its genomic distribution is limited to, and dependent upon, H3K27-methylated chromatin, which it recognizes directly through its BAH domain. Finally, we discover that EPR-1 orthologs are widely distributed across eukaryotes, contrary to previous reports<sup>136,138,139</sup>, suggesting an ancient role of EPR-1 homologs in Polycomb repression that was then lost on multiple occasions in certain lineages.

## Results

### Genetic selection for factors necessary for H3K27 methylation-mediated repression

In an effort to identify factors required for H3K27 methylation-mediated repression, we engineered a strain of *N. crassa* in which we replaced the open reading frames of two PRC2-repressed genes<sup>58</sup>, *NCU05173* and *NCU07152*, with the antibiotic-

resistance genes *hph* and *nat-1*, respectively (Fig. 22A). Strains that bear these gene replacements and lack the H3K27 methyltransferase (SET-7) are resistant to Hygromycin B and Nourseothricin, whereas a wild-type strain with these gene replacements is sensitive to these drugs (Fig. 22C). We subjected conidia collected from such an antibiotic-sensitive, otherwise wild-type strain to ultraviolet (UV) mutagenesis and selected for mutants that derepressed both the *hph* and *nat-1* genes (Fig. 22B). One mutant isolated in this manner and characterized here is effector of polycomb repression 1 (*epr-1*) (Fig. 22C).

### **Mapping and identification of *epr-1* as *NCU07505***

In order to map and identify the causative mutation in the *epr-1<sup>UV1</sup>* mutant, we crossed *epr-1<sup>UV1</sup>*, which is in an Oak Ridge genetic background, to a highly polymorphic wild-type strain named “Mauriceville”<sup>107</sup>. We then pooled the genomic DNA from Hygromycin B-resistant progeny and subjected it to whole-genome sequencing (~15x coverage; Fig. 22D). When we scored the percentage of Oak Ridge single nucleotide polymorphisms (SNPs) across the genome<sup>48</sup>, we found a region on linkage group (LG) I that was enriched for Oak Ridge SNPs and included an early stop mutation (Q206\*, CAG->TAG) in *NCU07505* (Fig. 22E).

To verify that the early stop in *NCU07505* is the causative mutation in *epr-1<sup>UV1</sup>*, we targeted a wild-type copy of *NCU07505* to the *his-3* locus in the *epr-1<sup>UV1</sup>* mutant background. This ectopic copy of *NCU07505* complemented the mutation, *i.e.*, it restored drug sensitivity (Fig. 22F). In addition, we found that deletion of *NCU07505* resulted in resistance to Hygromycin B, similar to the *epr-1<sup>UV1</sup>* strain (Fig. 22F). We subsequently

isolated and characterized three additional alleles of *epr-1* generated in the mutagenesis, further supporting the notion that mutations in *NCU07505* support drug resistance (Fig. 23).

### **EPR-1 and SET-7 repress an overlapping set of H3K27-methylated genes**

Although our selection was designed to isolate mutants with specific defects in Polycomb repression, we could conceivably recover mutants that globally altered transcription or led to antibiotic resistance independent of *hph* or *nat-1* upregulation. To determine if EPR-1 was specifically required for repression of H3K27-methylated genes, we performed mRNA-seq on  $\Delta epr-1$  siblings and compared the gene expression profile to previously published wild-type and  $\Delta set-7$  data sets<sup>58</sup>. We found that 632 genes were upregulated and 974 genes were downregulated greater than two-fold in  $\Delta epr-1$  strains compared to wild-type strains ( $P < 0.05$ ) (Fig. 25A). The upregulated gene set in  $\Delta epr-1$  was significantly enriched for H3K27-methylated genes ( $\chi^2_{(1, N = 632)} = 40.8$ ,  $P = 1.684 \times 10^{-10}$ ) (Fig. 25A), and H3K27-methylated genes upregulated in both  $\Delta epr-1$  and  $\Delta set-7$  significantly overlapped ( $P = 1.436 \times 10^{-16}$ ) (Fig. 24A). To verify our mRNA-seq results, we performed reverse transcription followed by quantitative polymerase chain reaction (RT-qPCR) on RNA isolated from biological triplicates of wild-type,  $\Delta set-7$ , and  $\Delta epr-1$  strains. Five of six examined H3K27-methylated genes found upregulated in both  $\Delta set-7$  and  $\Delta epr-1$  strains by mRNA-seq were confirmed by RT-qPCR (Fig. 24B and Fig. 25B). In contrast, only one out of six H3K27-methylated genes found exclusively upregulated in  $\Delta set-7$  or  $\Delta epr-1$  by mRNA-seq was confirmed by RT-qPCR (Fig. 25C-D). Thus, these

data show that loss of EPR-1 derepresses a significant number of H3K27-methylated genes that are also upregulated in strains lacking H3K27 methylation.

### ***Δepr-1* and *Δset-7* strains share a defect in sexual development**

Since *Δepr-1* and *Δset-7* strains exhibit similar transcriptional profiles, we wondered if they also shared vegetative growth and sexual development phenotypes as well. To assess if *Δepr-1* strains have an altered vegetative growth rate, we measured linear growth rates of wild-type, *Δset-7* and *Δepr-1* strains with ‘race tubes’<sup>140</sup>. We confirmed that *Δset-7* strains do not have a linear growth defect<sup>9</sup> and found that *Δepr-1* strains also grow at wild-type rates (Fig. 25E). Loss of SET-7 has been implicated in promoting sexual development in mutants that are homozygous sterile<sup>141</sup>. To determine if *Δset-7* and/or *Δepr-1* strains aberrantly promote sexual differentiation in the absence of a mating partner, we singly inoculated crossing plates<sup>30</sup> with wild-type, *Δset-7* or *Δepr-1* strains. After two weeks of unfertilized growth at 25 °C, we observed the development of few false perithecia with wild-type controls, whereas *Δepr-1* and *Δset-7* developed approximately 10- and 100-fold more false perithecia than wild type, respectively (Fig. 24C and Fig. 25F). These data suggest that EPR-1, and SET-7 to a greater extent, repress premature sexual development, which is reminiscent of fertilization-independent seed development observed in plant Polycomb mutants<sup>142,143</sup>.

### **H3K27 methylation is essentially normal in *Δepr-1***

As a first step to assess if the transcriptional silencing and sexual development defects shared between *Δepr-1* and *Δset-7* strains were due to a common global loss of

H3K27 methylation, we performed a western blot on whole cell lysates to detect H3K27me3 in wild-type,  $\Delta set-7$ , and  $\Delta epr-1$  strains. The total levels of H3K27me3 in  $\Delta epr-1$  strains were comparable to that in wild type (Fig. 26A and Fig. 27A-C). However, because only a subset of H3K27-methylated genes are derepressed in both  $\Delta epr-1$  and  $\Delta set-7$  strains, we wanted to know if H3K27 methylation might be specifically lost at upregulated genes in  $\Delta epr-1$  strains. To examine this possibility, we performed H3K27me2/3 chromatin immunoprecipitation followed by sequencing (ChIP-seq) on two  $\Delta epr-1$  siblings and compared the data to that for wild type<sup>110</sup>. We found that the global distribution of H3K27me2/3 in  $\Delta epr-1$  appeared to mirror that of wild type (Fig. 26B). Comparison of H3K27me2/3 levels associated with individual genes showed good agreement between the averaged wild-type and  $\Delta epr-1$  data sets ( $R^2 = 0.9105$ ), and within replicate data for wild-type ( $R^2 = 0.9272$ ) and  $\Delta epr-1$  ( $R^2 = 0.8941$ ) strains (Fig. 26C and Fig. 27D-E). We did, however, identify 29 genes with a greater than two-fold decrease and nine genes with a greater than two-fold increase in H3K27me2/3 levels in  $\Delta epr-1$  compared to wild type. Interestingly, none of these 38 genes with altered H3K27 methylation were classified among the upregulated or downregulated gene sets in the mRNA-seq analysis of  $\Delta epr-1$ . To validate the H3K27me2/3 ChIP-seq results, we performed H3K27me2/3 ChIP followed by quantitative Polymerase Chain Reaction (ChIP-qPCR) on wild-type,  $\Delta set-7$ , and  $\Delta epr-1$  strains in biological triplicate. These data confirmed wild-type levels of H3K27me2/3 at a subtelomere (Tel IL) and at the genes replaced by the antibiotic-resistance markers (*NCU07152* and *NCU05173*), and also corroborated the loss (*NCU08834*) and gain (*NCU02856*) of H3K27me2/3 observed in the ChIP-seq of  $\Delta epr-1$  strains (Fig. 26D). Altogether, these data show that the



derepression of H3K27-methylated genes in  $\Delta epr-1$  strains is not due to concomitant loss of H3K27 methylation.

### **EPR-1 forms telomere-associated foci dependent on EED**

To localize EPR-1 *in vivo*, we used the *ccg-1* promoter to drive expression of wild-type EPR-1 fused with GFP at its N-terminus (EPR-1<sup>WT</sup>)<sup>120</sup> in a strain that had fluorescent markers for the nuclear membrane (ISH1), telomeres (TRF1), and centromeres (CenH3)<sup>58</sup>. We found that EPR-1<sup>WT</sup> was restricted to the nucleus and formed distinct foci that were typically closely associated with TRF1 foci (Fig. 28A). EPR-1<sup>WT</sup> foci were significantly closer to telomeres as compared to centromeres (negative control;  $P = 0.0403$ ) (Fig. 29A), and the number of EPR-1<sup>WT</sup> and TRF1 foci per nucleus were not statistically different ( $P = 0.7422$ ), although the majority of nuclei examined had more EPR-1<sup>WT</sup> than TRF1 foci (Fig. 29B-C). Considering that H3K27 methylation is predominantly present near chromosome ends in *N. crassa*<sup>9</sup>, it is not unexpected that a putative PcG protein, such as EPR-1, would co-localize with the telomere marker, TRF1.

To determine if the formation of EPR-1<sup>WT</sup> foci was dependent on H3K27 methylation, we introduced a deletion of *eed*, a gene encoding a component of PRC2 necessary for catalytic activity<sup>9</sup>. Strains bearing this deletion lacked distinct nuclear foci of EPR-1<sup>WT</sup> and instead displayed a diffuse nuclear distribution of EPR-1<sup>WT</sup> (Fig. 28B). Thus, an intact PRC2 complex or H3K27 methylation is required for proper EPR-1<sup>WT</sup> subnuclear localization.

### **An intact BAH domain is required for normal nuclear distribution of EPR-1**

EPR-1 is predicted<sup>144</sup> to have a BAH domain and PHD finger, protein modules implicated in chromatin engagement<sup>145,146</sup>. To determine if these domains are necessary for the formation of the EPR-1<sup>WT</sup> foci, we created GFP-EPR-1 constructs in which a previously identified critical tryptophan in either the BAH domain (EPR-1<sup>BAH</sup>) or PHD finger (EPR-1<sup>PHD</sup>) was replaced with an alanine<sup>147,148</sup>. We found that EPR-1<sup>BAH</sup> displayed the same diffuse nuclear distribution as EPR-1<sup>WT</sup> in a *Δeed* background, consistent with the possibility that the BAH domain mediates interaction with H3K27-methylated chromatin (Fig. 28C). In contrast, EPR-1<sup>PHD</sup> still formed nuclear foci that were equivalent to EPR-1<sup>WT</sup> in number and proximity to TRF1 foci (Fig. 28D and Fig. 29D-E), demonstrating the nonessential nature of this conserved tryptophan residue for the normal nuclear distribution of EPR-1.

### **EPR-1 localizes to H3K27 methylation genome-wide**

Since proper subnuclear localization of EPR-1<sup>WT</sup> appeared to require H3K27 methylation, we performed ChIP-seq to determine if EPR-1<sup>WT</sup> genomic targets coincided with H3K27me<sub>2/3</sub> throughout the genome (Fig. 30A). Results of EPR-1<sup>WT</sup> ChIP-seq appeared to match the distribution of H3K27me<sub>2/3</sub> in wild type (Fig. 30A) and we found good correlation between EPR-1<sup>WT</sup> and H3K27me<sub>2/3</sub> relative sequencing coverage over each gene ( $R^2 = 0.8675$ ) (Fig. 31A). We also examined the genomic distribution of EPR-1<sup>PHD</sup> and EPR-1<sup>BAH</sup> mutant alleles (Fig. 30A), which were expressed at comparable levels (Fig. 31B-C). The ChIP-seq coverage of EPR-1<sup>PHD</sup> was still enriched at H3K27-methylated genes, albeit less so than EPR-1<sup>WT</sup>, while the EPR-1<sup>BAH</sup> ChIP-seq did not

show enrichment (Fig. 31D). To validate the ChIP-seq of GFP-EPR-1-expressing strains, we performed ChIP-qPCR for representative regions (Fig. 30B). Consistent with the ChIP-seq data, EPR-1<sup>WT</sup>, and EPR-1<sup>PHD</sup> to a lesser degree, were enriched at examined regions bearing H3K27 methylation. In addition, the ChIP-qPCR confirmed that EPR-1<sup>BAH</sup>, as well as EPR-1<sup>WT</sup> in a  $\Delta eed$  background, lack such enrichment.

As an orthogonal approach to ChIP, we determined the chromatin targets of EPR-1 by fusing an *E. coli* DNA adenine methyltransferase (Dam)<sup>149</sup> to the C-terminus of endogenous EPR-1 and assayed adenine-methylated DNA fragments by sequencing (DamID-seq)<sup>150</sup> (Fig. 30A). Using DamID-seq of EPR-1<sup>WT</sup>-Dam and methyl-sensitive restriction enzyme Southern blots, we found that EPR-1<sup>WT</sup>-Dam localizes to H3K27-methylated genes and this is dependent upon EED (Fig. 30A and Fig. 31E-F). Mutation of the PHD finger in the EPR-1-Dam fusion did not abolish its targeting to H3K27-methylated chromatin (Fig. 30A). These results were consistent with our ChIP-seq findings. We conclude that EPR-1 localizes to H3K27-methylated regions of the genome and that proper recruitment of EPR-1 to chromatin requires both an intact BAH domain and the integral PRC2 component, EED.

### **Both the BAH domain and PHD finger of EPR-1 are necessary for gene repression**

Our localization studies of EPR-1<sup>PHD</sup> and EPR-1<sup>BAH</sup>, while suggestive, did not directly test the role of the PHD finger and BAH domain of EPR-1 in H3K27 methylation-mediated silencing. We therefore utilized our antibiotic-resistance reporters, used in the initial selection, to test more directly their possible involvement in gene repression. We targeted ectopic copies of *epr-1*<sup>WT</sup>, *epr-1*<sup>BAH</sup> or *epr-1*<sup>PHD</sup> to the *his-3* locus

in an *epr-1<sup>UV1</sup>* strain bearing the antibiotic-resistance genes and scored drug resistance. Whereas *epr-1<sup>WT</sup>* restored sensitivity to Hygromycin B and Nourseothricin, *epr-1<sup>BAH</sup>* remained resistant to both drugs (Fig. 30C). In contrast, introduction of *epr-1<sup>PHD</sup>* apparently re-silenced the *hph*, but not the *nat-1*, antibiotic-resistance gene (Fig. 30C). This suggests that while the PHD finger of EPR-1 is not essential for recruitment to H3K27 methylated chromatin, it is not entirely dispensable for gene silencing.

### **BAH domain of EPR-1 binds to H3K27me3 *in vitro***

To test, directly, whether the BAH domain of EPR-1 recognizes H3K27 methylation, we cloned, expressed, and purified a HIS-SUMO-BAH<sub>EPR-1</sub> fusion protein for use in nuclear magnetic resonance (NMR) spectroscopy experiments. We found that the BAH<sub>EPR-1</sub> domain alone was unstable once cleaved from the HIS-SUMO tag and therefore we used the HIS-SUMO-tagged BAH<sub>EPR-1</sub> in all subsequent experiments. An initial <sup>1</sup>H-<sup>15</sup>N-heteronuclear single quantum coherence (<sup>1</sup>H-<sup>15</sup>N HSQC) spectrum revealed that the fusion protein was well folded (Fig. 31H). An overlay with an <sup>1</sup>H-<sup>15</sup>N HSQC spectrum of <sup>15</sup>N labelled HIS-SUMO tag alone was used to determine which peaks belonged to BAH<sub>EPR-1</sub> (Fig. 31H). Addition of increasing concentrations of an H3K27me3 peptide to the HIS-SUMO-BAH<sub>EPR-1</sub> fusion protein led to significant chemical shift perturbations (CSPs), indicating binding (Fig. 30D-E). Importantly, the perturbed resonances belonged exclusively to BAH<sub>EPR-1</sub>. We conclude that the BAH domain of EPR-1 associates with an H3K27me3 peptide. Due to protein stability problems, we were unable to calculate an accurate dissociation constant (K<sub>d</sub>); however, the CSPs appeared consistent with a high micromolar K<sub>d</sub>.

## **EPR-1 is a homolog of plant EBS/SHL and widely distributed across eukaryotes**

To determine if EPR-1 orthologs exist outside of *N. crassa*, we performed sequence similarity searches to identify homologs, followed by phylogenetic and domain architectural analysis of those to identify genuine orthologs. Consequently, we were able to identify orthologs in various fungal species as well as a wide range of other eukaryotes (Fig. 32A). Notably, we determined the *Arabidopsis thaliana* paralogs EBS and SHL as orthologs of EPR-1 in our analyses, which have been erroneously reported as plant-unique proteins<sup>136,138,139</sup>. Similar to EPR-1, the plant paralogs, EBS and SHL, bind H3K27 methylation and have been implicated in gene repression<sup>132,133,136</sup>. As one approach to investigate if other EPR-1 orthologs may have roles independent of H3K27 methylation, we checked if any species has an EPR-1 ortholog but lacks a SET-7 (H3K27 methyltransferase) ortholog. With the exception of Chytridiomycota and Fonticula lineages, in which the presence of SET-7 homologs was deemed ambiguous due to lack of a definitive pre-SET domain, all examined species with EPR-1 homologs had clear SET-7 homologs (Fig. 32A). This result is consistent with EPR-1 orthologs mediating H3K27 methylation-based repression in a wide variety of species.

In contrast to EPR-1, *A. thaliana* EBS is not entirely restricted to H3K27-methylated genes. Indeed, the majority of EBS-bound genes are devoid of H3K27me3 and instead are associated with an ‘active’ chromatin mark, H3K4me3<sup>151</sup>, via the PHD finger of EBS<sup>132</sup>. To understand this discrepancy, we examined the underlying protein sequence of the PHD finger of EPR-1. Whereas aromatic residues implicated in methylated histone recognition in the BAH domain<sup>152</sup> are present in both EPR-1 and the

plant paralogs, comparable residues in the PHD finger<sup>133,153,154</sup> are present in EBS and SHL, but lacking in EPR-1 (Fig. 32B, highlighted in red). Furthermore, a single amino acid substitution replacing an aromatic tyrosine residue with an alanine residue in the PHD finger of a plant SHL was sufficient to diminish *in vitro* H3K4me3-binding greater than two-fold<sup>133</sup>. Intriguingly, this residue corresponds to the naturally occurring alanine 275 in EPR-1 (Fig. 32B). This suggests that while many species have EPR-1 homologs, they may not necessarily be bivalent histone readers like plant EBS and SHL.

## Discussion

Deciphering the basic mechanism(s) of Polycomb repression has been difficult, in part, due to the diversity<sup>103</sup>, redundancy<sup>155</sup>, as well as interdependence<sup>130</sup> of protein players involved. For this reason, the model organism *N. crassa*, which employs H3K27 methylation catalyzed by PRC2 for gene repression<sup>9</sup>, yet conspicuously lacks PRC1 components<sup>7</sup>, represents an ideal organism to uncover fundamental aspects of Polycomb silencing. Here we have identified, to the best of our knowledge, the first known reader and effector of H3K27 methylation in fungi, EPR-1. This provides insight into how Polycomb silencing can function in the absence of PRC1.

Our ChIP and DamID results demonstrated that EPR-1 co-localizes with H3K27 methylation genome-wide and cytological examination revealed that GFP-EPR-1 forms approximately 3-5 foci per nucleus. This implies that domains of H3K27 methylation within and between the seven *N. crassa* chromosomes generally self-associate. This is consistent with the observed intra- and inter-chromosomal contacts among H3K27-

methylated regions of the genome in Hi-C experiments<sup>36</sup>. Similar, and perhaps equivalent, higher-order chromatin structures, referred to as Polycomb bodies<sup>137</sup>, are known to be mediated by PRC1 components in both plant and animal cells<sup>156-159</sup>. Our group has previously shown that SET-7, the catalytic component of PRC2, is required for normal 3D genome organization<sup>58</sup>. It would be interesting to learn if EPR-1, the only known effector of H3K27 methylation in *N. crassa*, is also essential for this wild-type chromatin organization.

Despite the loss of transcriptional silencing observed in *epr-1* mutants, they do not exhibit appreciably altered H3K27 methylation – this is striking for a few reasons. First, it suggests that H3K27 methylation-mediated silencing is a unidirectional pathway in *N. crassa*, in which EPR-1 acts downstream of PRC2. This is in contrast to findings in plants and animals, in which PRC1 components can affect the recruitment or activity of PRC2<sup>136,160,161</sup>, a fact that has hampered the elucidation of the direct role of PRC components in gene repression. Second, since  $\Delta epr-1$  strains de-repress H3K27-methylated genes without loss of H3K27 methylation, it suggests that active transcription does not necessarily preclude PRC2 activity. This was surprising since transcriptional shut-off is thought to precede PRC2 activity during normal animal development<sup>41</sup> and because artificial gene repression can be sufficient to recruit PRC2<sup>84,85</sup>. Finally, the presence of H3K27 methylation on de-repressed genes in  $\Delta epr-1$  strains demonstrates that H3K27 methylation *per se* is not sufficient for effective gene repression, consistent with previous reports<sup>162,163</sup>. It is noteworthy, however, that while  $\Delta set-7$  and  $\Delta epr-1$  strains share transcriptional and sexual defects, the phenotype of  $\Delta set-7$  strains is

generally more pronounced. This suggests that PRC2 or H3K27 methylation may have additional roles in gene repression that go beyond recruitment of EPR-1 to chromatin.

Our investigation into the phylogenetic distribution of EPR-1 homologs indicates that an ancestral EPR-1 emerged prior to the divergence of plants, animals, and fungi. This ancestral EPR-1 may have been an integral component of an early eukaryotic Polycomb silencing system. While animals are a notable exception, regarding the absence of EPR-1 homologs with a BAH-PHD structure, a human BAH domain-containing protein, BAHD1, has been reported to ‘read’ H3K27me<sup>3</sup><sup>147</sup> and promote gene silencing<sup>164</sup>, although apparently not interact with known PRC1 components<sup>165</sup>. It is therefore conceivable that BAHD1 homologs present in animal lineages are actually divergent orthologs of EPR-1 that lack a PHD finger. Regardless of their ancestry, both human BAHD1 and *N. crassa* EPR-1 represent novel forms of H3K27 methylation-mediated repression that do not rely on PRC1 components.

## **Materials and Methods**

### **Strains, media and growth conditions**

All *N. crassa* strains used in this study are listed in Appendix B. Liquid cultures were grown with shaking at 32 °C in Vogel’s minimal medium (VMM) with 1.5% sucrose<sup>29</sup>. Crosses were performed at 25 °C on modified Vogel’s with 0.1% sucrose<sup>30</sup>. Spot tests were performed at 32 °C on VMM with 0.8% sorbose, 0.2% fructose, and 0.2% glucose (FGS). When appropriate, plates included 200 µg/mL Hygromycin B Gold (InvivoGen) or 133 µg/mL Nourseothricin (Gold Biotechnology). Linear growth rates



were determined as previously described except 25 mL serological pipettes were used in place of glass tubes<sup>140</sup>. Genomic DNA was isolated as previously described<sup>166</sup>.

Nutritional supplements required for auxotrophic strains were included in all growth media when necessary.

### **Selection for mutants defective in H3K27 methylation-mediated silencing**

Ten thousand conidia of strain N6279 (created with primers in Appendix C) were plated on VMM supplemented with FGS and 500 µg/mL histidine and subjected to 0, 3, 6, or 9 seconds of ultraviolet light (Spectrolinker XL-1500 UV Crosslinker, Spectronics Corporation) in a dark room and plates were wrapped in aluminum foil to prevent photoreactivation<sup>167</sup>. Plates were incubated at 32 °C for 16 hours before being overlaid with 1% top agar containing VMM, FGS, 500 µg/mL histidine, Hygromycin B and Nourseothricin. Drug-resistant colonies were picked after an additional 48-72 hours at 32 °C. Initial mutant strains were crossed to a *Sad-1* mutant<sup>168</sup> strain (N3756) and resultant progeny were germinated on Hygromycin B- and Nourseothricin-containing medium to obtain homokaryotic mutants.

### **Whole genome sequencing, mapping and identification of *epr-1* alleles**

Homokaryotic mutants resistant to Hygromycin B and Nourseothricin were crossed to the genetically polymorphic Mauriceville strain (FGSC 2225)<sup>107</sup> and resultant progeny were germinated on medium containing Hygromycin B and/or Nourseothricin to select for strains bearing the causative mutation. Genomic DNA from approximately 15-20 progeny per mutant were pooled and sequencing libraries were prepared with a

Nextera kit (Illumina, FC-121-1030). All whole genome sequencing data is available on NCBI SRA (accession #PRJNA526508). To map the approximate location of a particular causative mutation, the fraction of Oak Ridge (versus Mauriceville) SNPs across the *N. crassa* genome was determined as previously described<sup>48</sup> and visualized as a moving average (window size = 10 SNPs, step size = 1 SNP) with Matplotlib<sup>169</sup>. We utilized FreeBayes and VCFtools to identify genetic variants present in our pooled mutant genomic DNA but absent in the original mutagenized strain (N6279) and the Mauriceville strain<sup>118,119</sup>. Only genetic variants of high probability that were consistent with the mapping data were considered further.

### **Western blotting**

*N. crassa* tissue from a 16 hour liquid culture of germinated conidia was collected by filtration, washed with 1x phosphate-buffered saline (PBS; 137 mM NaCl, 10 mM phosphate, 2.7 mM KCl, pH 7.5), and suspended in 500  $\mu$ L of ice-cold lysis buffer (50 mM HEPES [pH 7.5], 150 mM NaCl, 10% glycerol, 0.02-0.2% NP-40, 1 mM EDTA) supplemented with 1x Halt<sup>TM</sup> Protease Inhibitor Cocktail (Thermo Scientific). Tissue was sonicated (Branson Sonifier-450) for three sets of 10 pulses (Output = 2, Duty cycle = 80), keeping the sample on ice between sets. Insoluble material was pelleted by centrifugation at 14,000 RPM at 4 °C for 10 minutes and the supernatant used as the western sample. Anti-H3K27me3 (Cell Signaling Technology, 9733) and anti-hH3 (Abcam, ab1791) primary antibodies were used with IRDye 680RD goat anti-rabbit secondary (LI-COR, 926-68071). Anti-GFP (Thermo Fisher, A10262) primary antibody was used with goat anti-chicken HRP conjugated secondary antibody (Abcam, 6877).

Images were acquired with an Odyssey Fc Imaging System (LI-COR) and analyzed with Image Studio software (LI-COR).

### **Chromatin immunoprecipitation**

Liquid cultures were grown for approximately 18 hours with shaking at 32 °C. Tissue samples for H3K27me2/3 ChIP were cross-linked with 0.5% formaldehyde for 10 minutes and GFP ChIP samples were cross-linked with 1% formaldehyde for 10 minutes in 1x PBS. Cross-linking was quenched with 125mM glycine, tissue was washed with 1x PBS and collected. Cells were lysed in ChIP lysis buffer (50 mM HEPES [pH 7.5], 90 mM NaCl, 1 mM EDTA, 1% Triton X-100, 0.1% Deoxycholate) supplemented with 1x Halt protease inhibitor cocktail (Thermo Scientific) using a Branson Sonifier 450. Chromatin was sheared using a Bioruptor (Diagenode) and 2 µl of appropriate antibody (H3K27me2/3, Active Motif 39536, or GFP, MBL 598) was added and incubated with rotation at 4 °C overnight. Protein A/G agarose (40 µL; Santa Cruz Biotechnologies) was added to H3K27me2/3 ChIP samples and Protein A agarose (40 µL; Sigma) was added to GFP ChIP samples and incubated for 3 hours, rotating at 4 °C. Beads were washed twice with ChIP lysis buffer supplemented with 140 mM NaCl, once with ChIP lysis buffer with 0.5 M NaCl, once with LiCl wash buffer (10 mM Tris-HCl [pH 8.0], 250 mM LiCl, 0.5% NP-40, 0.5% Deoxycholate, 1 mM EDTA), and once with TE (10 mM Tris-HCl, 1 mM EDTA), all rotating at 4 °C for 10 minutes each. DNA was eluted from beads by incubation in TES (10 mM Tris-HCl, 1 mM EDTA, 1% sodium dodecyl sulfate) at 65 °C. Crosslinking was reversed by incubation at 65 °C for 16 hours and then samples were treated with proteinase K for 2 hours at 50 °C. DNA was purified using Minelute

columns (Qiagen) and subsequently used for qPCR with the PerfCTa SYBR Green FastMix (QuantBio, 95071-012) on a Step One Plus Real Time PCR System (Life Technologies) using primer pairs in Appendix C, or prepared for sequencing using the NEBNext DNA Library Prep Master Mix Set for Illumina (New England BioLabs). ChIP-seq data are available on the NCBI GEO database (GSE128317).

### **ChIP-seq mapping and analysis**

The suite of tools available on the open-source platform Galaxy<sup>170</sup> was used to map ChIP-sequencing reads<sup>93</sup> against the corrected *N. crassa* OR74A (NC12) genome<sup>36</sup> and to create bigWig coverage files normalized to reads per kilobase per million mapped reads (RPKM)<sup>171</sup>. ChIP-seq tracks were visualized with the Integrative Genomics Viewer<sup>96</sup>. MAnorm was used to compare H3K27me2/3 ChIP-seq coverage on all genes (designated as ‘peaks’) between samples and data were visualized with Matplotlib<sup>169</sup>. Genes were scored by their normalized H3K27me2/3 ChIP-seq coverage in wild type and the top-ranking genes (873) were designated as H3K27 methylated.

### **RNA isolation, mRNA-seq library prep, and RT-qPCR**

RNA was extracted from germinated conidia grown for 16-18 hours with a 1:1:1 glass beads, NETS (300mM NaCl, 1mM EDTA, 10mM Tris-HCl, 0.2% SDS), acid phenol:chloroform mixture (5:1; [pH 4.5]) using a bead beater and ethanol precipitated. RNA was treated with DNase I (Amplification grade; Thermo Fisher Scientific). DNase I-treated RNA was used for RNA-seq library preparation as previously described<sup>58</sup> or cDNA was synthesized using the SuperScript III First Strand-Synthesis System (Thermo

Fisher Scientific) with poly-dT primers. cDNA was used for qPCR using the PerfCTa SYBR Green FastMix (QuantBio) on a Step One Plus Real Time PCR System (Life Technologies) using primer pairs in Appendix C. mRNA-seq data are available on the NCBI GEO database (GSE128317).

### **mRNA-seq mapping and analysis**

Tools available on Galaxy<sup>170</sup> were used to map mRNA-sequencing reads (intron size < 1kb)<sup>172</sup> against the corrected *N. crassa* OR74A (NC12) genome<sup>36</sup>, to count the number of reads per gene<sup>172</sup> and to identify differentially expressed genes<sup>173</sup>. *P* values reported are adjusted for false discovery rates<sup>173</sup>. The Chi-squared test was used to determine if upregulated genes in  $\Delta epr-1$  strains were enriched for genes marked with H3K27 methylation. Significance of gene set intersections were calculated with a hypergeometric distribution ([http://nemates.org/MA/progs/overlap\\_stats\\_prog.html](http://nemates.org/MA/progs/overlap_stats_prog.html)).

### **DamID Southern hybridizations and sequencing**

Southern hybridizations were carried out as previously described<sup>98</sup>, except probes were made with PCR products amplified from wild-type *N. crassa* genomic DNA (*NCU05173*, Tel VIII) or plasmid pBM61 (*his-3*) using primer pairs from Appendix C. Preparation of N6-methyladenine-containing DNA for sequencing was performed using a previously reported procedure<sup>150</sup> using primers in Appendix C with the following modifications: 5  $\mu$ g of genomic DNA from *N. crassa* strains expressing a Dam fusion was digested with 1  $\mu$ L of DpnI (NEB, 20 units/ $\mu$ L); ligation to primer 5050 was carried out overnight at 16 °C; amplification reactions of ligated DNA with primer 5051 were

performed in triplicate with 5  $\mu$ L dNTPs, and an additional PCR cycle was added to the 2<sup>nd</sup> and 3<sup>rd</sup> phase of the PCR protocol (4 and 18 cycles respectively); 3  $\mu$ g of pooled, amplified DNA was sheared using a Bioruptor (Diagenode) twice on high for 10 minutes (30 seconds on/off) at 4 °C; biotinylated DNA was purified using 250  $\mu$ L slurry of streptavidin-agarose beads (Sigma). DNA was cleaved from the beads with DpnII and libraries were prepared for sequencing using the NEBNext DNA Library Prep Master Mix Set for Illumina (New England BioLabs).

### **False perithecia assay and image analysis**

To assay of the development of false perithecia, strains were grown on modified Vogel's media<sup>30</sup> as described above without a fertilizing strain. Images of plates were acquired after two weeks of growth at 25 °C and false perithecia were detected and quantified using the Laplacian of Gaussian blob detection algorithm from scikit-image<sup>174</sup>.

### **Microscopy image acquisition and analysis**

Live conidia were suspended in water and placed on a poly-L-lysine (Sigma) coated coverslip (No. 1.5; VWR) and mounted on a glass slide. Single plane images for distance measurements were captured with the ELYRA S.1 system (Zeiss) mounted on an AXIO Observer Z1 inverted microscope stand (Zeiss) equipped with a 63x, 1.4 NA Plan-Apochromat oil-immersion lens (Zeiss) and analyzed using Imaris (version 9.2.1). Images for volume renderings and max projections were collected with the DeltaVision Ultra microscope system (GE) equipped with a 100x, 1.4 NA UPlanSApo objective (Olympus). Three-dimensional Z-stack wide-field fluorescent (eGFP, TagRFP, iRFP, and

TagBFP) images were captured with an sCMOS camera controlled with Acquire Ultra software. Images were processed using 10 cycles of enhanced ratio deconvolution and max projections were made using softWoRx (GE, version 7.1.0). Imaris (version 9.2.1) was used to make volume renderings and TFR1 and EPR1 foci were counted by hand.

### **Cloning, expression and purification of BAH domain from EPR-1**

The BAH domain of EPR-1 (amino acids 134-264) was PCR-amplified from a plasmid containing *epr-1* cDNA with primers 6681 and 6682 (Appendix C), and cloned into pE-SUMO (LifeSensors), resulting in plasmid 3410 (Appendix D). A HIS-SUMO-only control construct was generated by adding a stop codon immediately before the BAH<sub>EPR-1</sub> domain in plasmid 3410. The HIS-SUMO-BAH<sub>EPR-1</sub> fusion and HIS-SUMO tag constructs were expressed in *Escherichia coli* BL21 DE3 cells (New England BioLabs). Both HIS-SUMO-BAH<sub>EPR-1</sub> and HIS-SUMO cultures were initially grown in 4 liters of LB media at 37 °C at 200 RPM until cultures reached an optical density of ~ 1.0 at 600 nm and then cultures were pelleted. Cells were then re-suspended separately in 1 liter of M9 medium supplemented with <sup>15</sup>N-NH<sub>4</sub>Cl. HIS-SUMO-BAH<sub>EPR-1</sub> cultures were induced with 0.4 mM IPTG and grown for 16-18 hours at 18 °C. HIS-SUMO cultures were induced with 1 mM IPTG and grown for 16-18 hours at 25 °C. Both cultures were collected separately by centrifugation at 6,000 RPM for 20 minutes, frozen in liquid N<sub>2</sub>, and stored at -80 °C. The same protocol outlined below was used for the purification of both HIS-SUMO-BAH<sub>EPR-1</sub> and HIS-SUMO alone.

Cells were re-suspended in lysis buffer (20 mM Tris-HCl [pH 7.5], 500 mM NaCl, 5 mM imidazole) with DNase I, EDTA-free protease inhibitors (Roche mini-

tablets) and lysozyme. The cell suspension was sonicated over a period of 1 minute, alternating 1 second on and 2 seconds off. The lysate was cleared by centrifugation at 15,000 RPM for one hour at 4 °C, and the soluble fraction was loaded onto a column packed with Ni(II)-nitrilotriacetic acid agarose (Qiagen) pre-equilibrated in lysis buffer. The column was washed with 100 mL of elution buffer containing 5 mM imidazole, and the protein was eluted with buffer containing 100 mM - 200 mM imidazole. Elution fractions were then analyzed using SDS-PAGE, were pooled and concentrated using 10,000 MWCO membrane (Millipore). The protein was further purified using fast protein liquid chromatography (FPLC). Concentrated protein was loaded onto a pre-equilibrated Superdex 75 (GE Healthcare) column containing 20 mM Tris-HCl pH 7.5, 150 mM NaCl, 2 mM DTT. Size exclusion fractions were analyzed using SDS-PAGE, concentrated, and stored at -80 °C.

### **NMR spectroscopy**

H3K27me3 (amino acids 23–34) peptide was obtained from Anaspec. For NMR studies, peptides were re-suspended in H<sub>2</sub>O to a final concentration of 20 mM, and pH was adjusted to 7.0. Titration experiments of peptide into HIS-SUMO-BAH<sub>EPR-1</sub> were carried out by collecting <sup>1</sup>H-<sup>15</sup>N HSQC spectra on <sup>15</sup>N labeled protein at 0.09 mM in the presence of increasing peptide concentrations. Titration points were taken at protein:peptide molar ratios of 1:0, 1:1, 1:5 and 1:10. In addition, an <sup>1</sup>H-<sup>15</sup>N HSQC was collected on <sup>15</sup>N labeled 0.1 mM HIS-SUMO tag alone. All NMR data were collected on a Bruker Avance II 800 MHz NMR spectrometer equipped with a cryoprobe. Data were processed using NMRPipe, and further analyzed using CcpNmr.



### **Bioinformatic identification of EPR-1 and SET-7 homologs**

PSI-BLAST<sup>175</sup> was used to initiate sequence similarity searches with representative sequences of EPR-1 (accession number XP\_965052.2) and SET-7 (accession number XP\_965043.2, residues 577-833) from *N. crassa* against the National Center for Biotechnology Information (NCBI) non-redundant (NR) database and locally maintained databases of proteins from representative species from different branches of the tree of life. HHpred<sup>176</sup> was used to perform profile-profile comparisons against the Protein Data Bank (PDB), Pfam and locally maintained protein sequence profiles. Sequences were clustered using BLASTCLUST (<ftp://ftp.ncbi.nih.gov/blast/documents/blastclust.html>). Multiple sequence alignments were generated using Kalign<sup>177</sup> and then adjusted manually. FastTree<sup>178</sup> was used to assess phylogenetic relationships among the proteins retrieved after sequence similarity searches. Customized PERL scripts were used for the analysis of domain architectures and other contextual information about protein sequences. Stringent criteria of domain-architectural concordance were used to retrieve only orthologs of all proteins under study along with grouping in phylogenetic tree analysis. In the case of SET-7, only those protein sequences that have a complete pre-SET domain followed by a SET domain were considered for analysis.

### **Replacement of *NCU05173* and *NCU07152* ORFs with *hph* and *nat-1***

To delete *NCU07152*, the 5' and 3' regions flanking the open reading frame (ORF) were amplified from wild-type genomic DNA with primers 6385-6388 (Appendix

C). The *nat-1* gene was amplified from plasmid 3237 with primers 6269 and 6270. The resulting three pieces of DNA were stitched by overlap extension using primers 6385 and 6388 and the final product was transformed into N4840. Primary transformants were selected on Nourseothricin-containing medium and crossed to a wild-type strain (N3753) to remove  $\Delta set-7$  and  $\Delta mus-52$  from the genetic background, resulting in strain N5808. To delete *NCU05173*, the 5' and 3' regions flanking the ORF were amplified from wild-type genomic DNA with primers 6605-6608. The *hph* gene was amplified from plasmid 2283. The 5' flank was stitched to the 5' portion of *hph* by overlap extension using primers 6605 and 2955, and the 3' flank was stitched to the 3' portion of *hph* using primers 6608 and 2954. The resulting two pieces of 'split marker' DNA were transformed into strain N4840. Primary transformants were selected on Hygromycin B-containing medium and crossed to N5808 to generate a homokaryotic strain with both marker genes (N6233). N6233 was crossed to N623 to introduce *his-3*, resulting in the strain used for the mutant hunt (N6279).

### **Creation and targeting of *his-3<sup>+</sup>::pCCG::N-GFP::EPR-1* plasmids**

The ORF and 3' UTR of *epr-1* were PCR-amplified from wild-type genomic DNA with primers 6416 and 6417 (Appendix C) and cloned into plasmid 2406<sup>120</sup> using *PacI* and *XbaI* restriction sites. For the BAH point mutant, *epr-1<sup>W184A</sup>*, two PCR products amplified from wild-type genomic DNA with primers pairs 6416 and 6368, and 6367 and 6417 were PCR-stitched together with primers 6416 and 6417, and similarly cloned into plasmid 2406. For the PHD point mutant, *epr-1<sup>W292A</sup>*, two PCR products amplified from wild-type genomic DNA with primer pairs 6416 and 6400, and 6399 and 6417 were

PCR-stitched together with primers 6416 and 6417, and cloned into plasmid 2406. Plasmids were linearized with NdeI and targeted to *his-3* in either N7451 (for complementation) or N7567 (for microscopy), as previously described<sup>91</sup>. Primary transformants were then crossed to N7549 or N7552, respectively.

### **Replacement of *epr-1* with *trpC::nat-1***

The 5' and 3' flanks of *epr-1* were PCR-amplified from wild-type genomic DNA with primer pairs 6401 and 6402, and 6350 and 6351, respectively (Appendix C). The 5' and 3' flanks were separately PCR-stitched with plasmid 3237 (source of *nat-1*) using primer pairs 6401 and 4883, and 4882 and 6351, respectively. These two 'split-marker' PCR products were transformed into strain N7537 and *epr-1* replacements were selected on Nourseothricin-containing medium. Primary transformants were crossed to N3752 (to generate N7567) and N6234 (to generate N7576).

### **Endogenous C-terminal-tagging of EPR-1 with 10xGly::Dam**

The regions immediately upstream (5') and downstream (3') of *epr-1*'s stop codon were PCR-amplified from wild-type genomic DNA with primer pairs 6348 and 6349, and 6350 and 6351, respectively (Appendix C). The 5' and 3' regions were separately PCR-stitched with plasmid 3131 (source of 10xGly::Dam::*trpC::nat-1*) using primer pairs 6348 and 4883, and 4882 and 6351, respectively. These two 'split-marker' PCR products were transformed into N2718 and 'knock-ins' were selected on Nourseothricin-containing medium. Primary transformants were crossed to N3752 to obtain homokaryons. To make *epr-1*<sup>W292A</sup> fusions with Dam, the same approach was

taken except the 5' region was a PCR product stitched from two PCR products amplified from wild-type genomic DNA with primer pairs 6346 and 6400, and 6399 and 6349.

## CHAPTER VI

### CONCLUDING SUMMARY

In Chapter II, I investigated the role of *Neurospora* SIR2 homologues in the establishment of H3K27 methylation. Deletion of *nst-1,4,5* (*neurospora sir two-1,4,5*) had minimal to no effect on the distribution of H3K27 methylation, but deletion of *nst-3* unexpectedly resulted in accumulation of H3K27 methylation at a subset of centromeres. Since NST-3 is a putative H3K56 deacetylase, I wondered if this effect was mediated by hyper H3K56 acetylation (H3K56ac). Double mutants with  $\Delta nst-3$  that either lacked lysine 56 on histone H3 (*H3K56R*) or lacked the putative H3K56 acetyltransferase (*rtt109*) fully suppressed the ectopic, centromeric H3K27 methylation. Centromeric H3K27 methylation observed in strains lacking HP1<sup>12</sup> was also partially suppressed when combined with *H3K56R*. Finally, Vincent Bicocca looked at the distribution of H3K56ac in wild-type and  $\Delta nst-3$  strains. While he did detect hyper H3K56ac at centromeres in  $\Delta nst-3$  strains that was fully suppressed by *H3K56R*, the residual signal in strains bearing *H3K56R* suggests that the antibody used is not specific to H3K56ac.

These results implicate hyper H3K56ac in ectopic accumulation of H3K27 methylation. The fact that the redistribution of H3K27 methylation in strains lacking HP1 appears to be partially suppressed by *H3K56R* suggests that the ectopic H3K27 methylation in  $\Delta hpo$  and  $\Delta nst-3$  strains occurs through a shared pathway. However, the milder phenotype observed in  $\Delta nst-3$  strains, compared to  $\Delta hpo$  strains, suggests that additional, parallel pathways promoting ectopic H3K27 methylation may also be triggered. My speculation is that hyper H3K56ac *per se* is not sufficient for accumulation

of H3K27 methylation at centromeres. I think that downstream effectors, components of an activated DNA damage response machinery, mediate the ectopic H3K27 methylation. In wild type strains, ectopic H3K27 methylation at sites of perceived DNA damage may be beneficial to block local transcription<sup>28</sup>. But, in the case of  $\Delta hpo$  strains, in which all constitutive heterochromatin appears damaged, the aberrant activity of PRC2 is toxic.

In Chapter III, I collaborated with Kirsty Jamieson and others to determine the effect of chromosome position on H3K27 methylation. We uncovered two classes of H3K27 methylation by examining the distribution of H3K27 methylation in strains bearing chromosomal rearrangement: position-dependent and position-independent. We found that proximity to a chromosome end was necessary to maintain and sufficient to induce subtelomeric H3K27 methylation, and that the observed changes in H3K27 methylation were reflected in altered gene expression. I further demonstrated that subtelomeric H3K27 methylation required telomerase and that ectopic telomere repeats (TTAGGG)<sub>n</sub> can induce large domains of H3K27 methylation.

To the best of our knowledge, telomere repeats represent the first known *cis*-acting signal that direct H3K27 methylation in fungi. However, the mechanism by which these sequences induce H3K27 methylation, and not just locally, is unresolved. One can imagine that PRC2 is directly recruited to these sequences, either through an interaction with DNA, or a transcribed RNA. Alternatively, PRC2 could be recruited indirectly through another protein complex. Evidence in mammals<sup>73</sup> and plants<sup>111</sup> suggest the former and latter, respectively. In addition to the identification of telomere repeats as a potent signal for H3K27 methylation, our chromosomal rearrangement strains elucidated

regions of the *N. crassa* genome that are ripe for dissection to further uncover *cis*-acting signals. Indeed, much of the H3K27 methylation in *N. crassa* is telomere-independent.

In Chapter IV, I collaborated with Elizabeth Wiles to identify factors required for H3K27 methylation-mediated silencing using a forward genetic selection that she implemented. We identified an undescribed gene (*NCU04278*) that is necessary for subtelomeric H3K27 methylation and silencing of underlying genes.

Immunoprecipitation followed by mass spectrometry analysis of NCU04278-associated proteins demonstrated that NCU04278 is PRC2 complex member, and we therefore deemed it PRC2 accessory subunit (PAS). Phylogenetic analysis of PAS homologs demonstrated that they are likely limited to the fungal class Sordariomycetes and Leotiomyces.

These results help elucidate the molecular mechanism that differentiates the two classes of H3K27 methylation uncovered in Chapter III. PAS, an accessory subunit of PRC2, is responsible for the establishment of H3K27 methylation that is dependent on telomere repeats. How exactly PAS recruits or activates PRC2 specifically at subtelomeric regions is unclear. However, work with mammalian PRC2 has demonstrated that recruitment to chromatin is dependent on the N-terminus of the PRC2 member SUZ12, and that SUZ12 directly interacts with PRC2 accessory subunits that are essential for targeting<sup>116</sup>. Yeast two-hybrid experiments that examine the possible interaction between PAS and SUZ12, and subsequent *in vivo* targeted deletions of SUZ12 will elucidate whether the general mechanism of PRC2 targeting is conserved between animals and fungi.

In Chapter V, I further collaborated with Elizabeth Wiles and others, using a forward genetic selection that she implemented, to identify factors required for H3K27 methylation-mediated silencing. In this case, we found a mutation in an undescribed gene (*NCU07505*), whose deletion resulted in transcriptional and developmental defects shared with strains lacking all H3K27 methylation, *i.e.*  $\Delta set-7$ , yet H3K27 methylation in  $\Delta NCU07505$  strains was essentially normal. We determined that EPR-1 (*NCU07505*) formed nuclear foci and localized to H3K27-methylated chromatin dependent on its BAH domain. Further *in vitro* work demonstrated that the BAH domain of EPR-1 directly interacts with H3K27me3 peptides. Finally, we demonstrated that homologs of EPR-1 exist in a variety of eukaryotic lineages, suggesting an ancient origin.

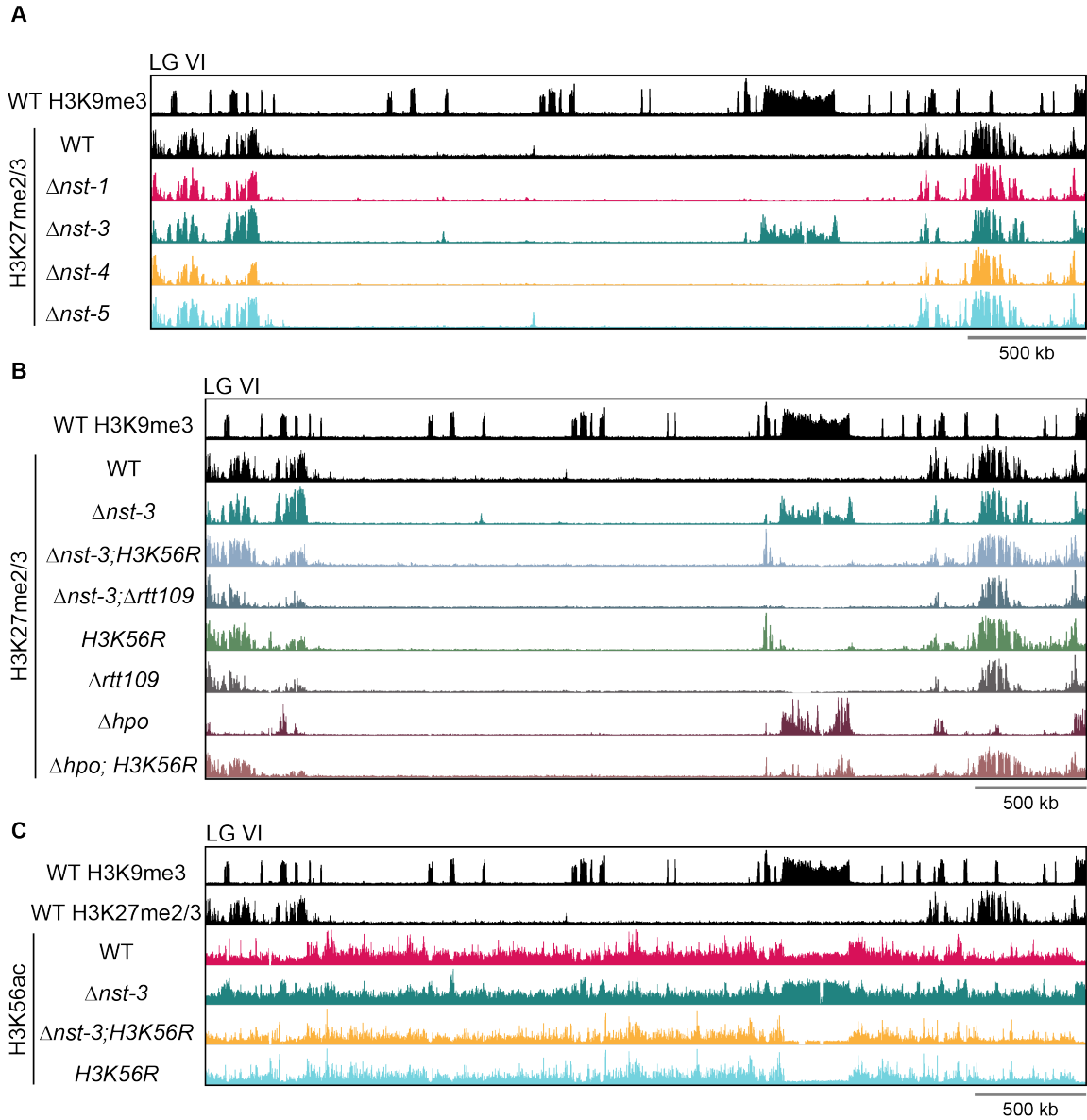
These results fill in a major gap of our understanding regarding the mechanism of Polycomb silencing in fungi and beyond. Prior to the identification of EPR-1 in *N. crassa*, it was mysterious as to how H3K27 methylation could be repressive when canonical PRC1 components were absent<sup>7</sup>. Choanoflagellates, the closest living relatives to animals, also lack canonical PRC1 members<sup>7</sup>, but our phylogenetic analyses demonstrate that they have an EPR-1 homolog. In plants, EPR-1 paralogs EBS and SHL interact with PRC1 components<sup>136</sup>, but whether EPR-1 forms a complex in other lineages that lack PRC1 components is unresolved. Future biochemical work may help resolve the mechanism of silencing enacted by EPR-1 in *N. crassa* and beyond.



# APPENDIX A

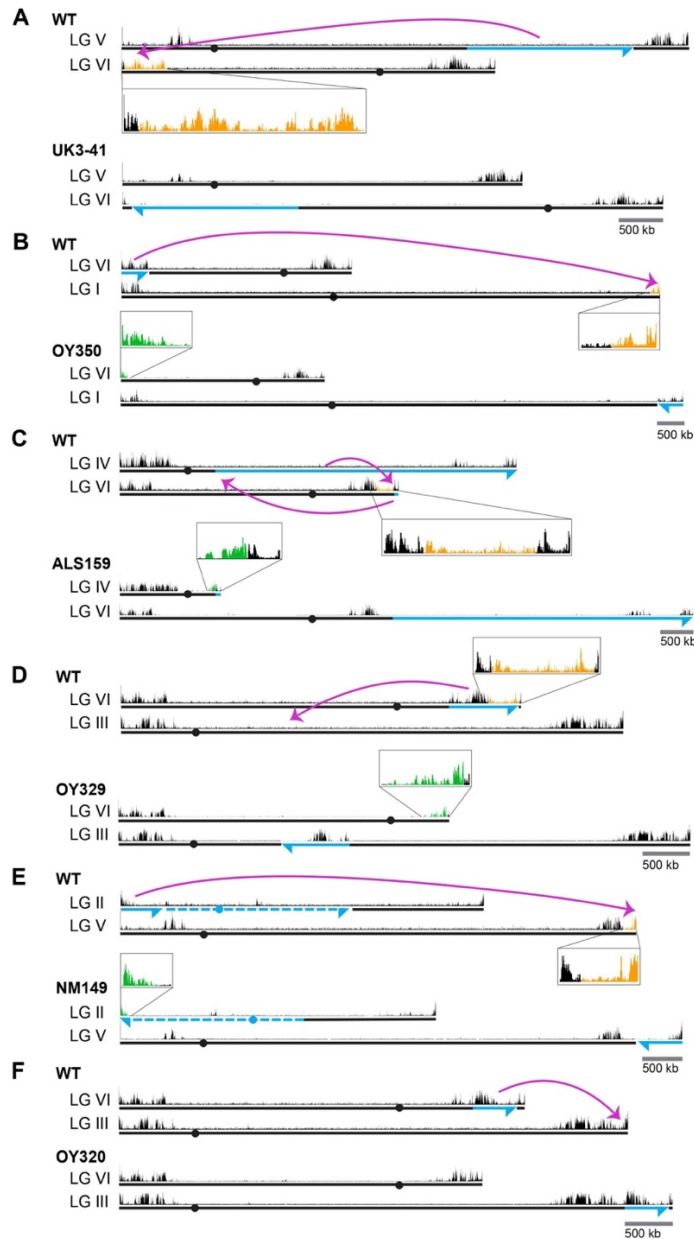
## FIGURES

### Figures for Chapter II

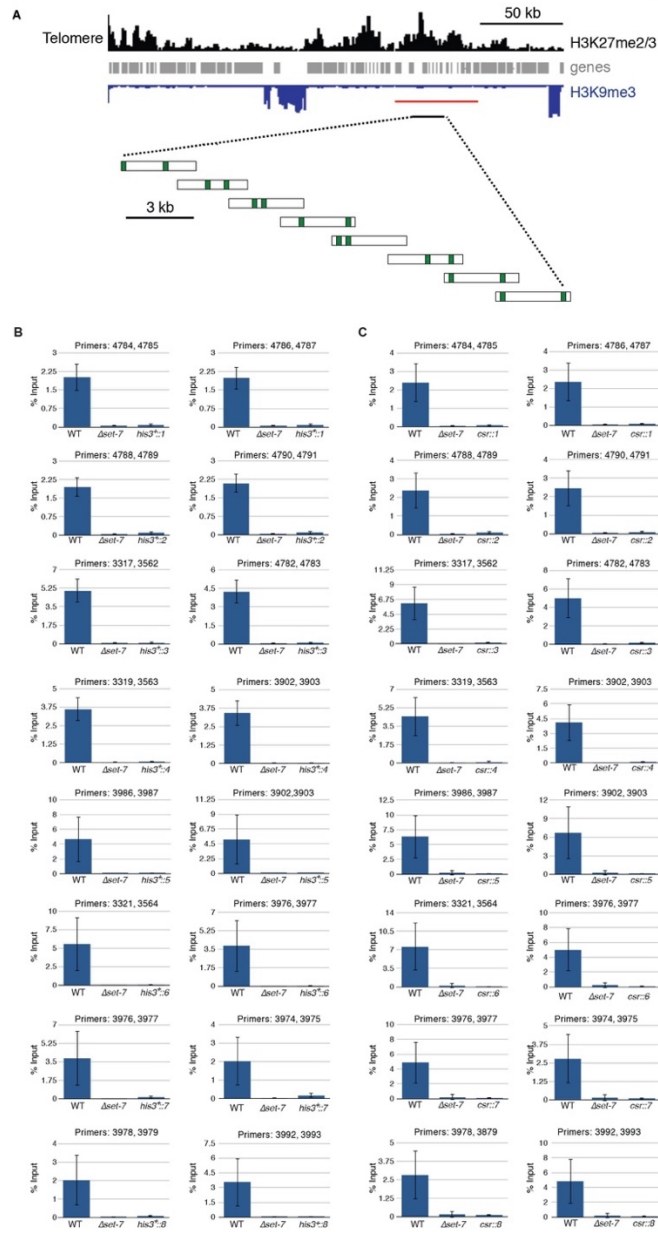


**Fig. 1. Aberrant H3K56ac implicated in ectopic H3K27me2/3.** (A-C) ChIP-seq tracks of indicated histone modification distribution in indicated genotypes on linkage group VI. Grey bar represents 500 kilobases.

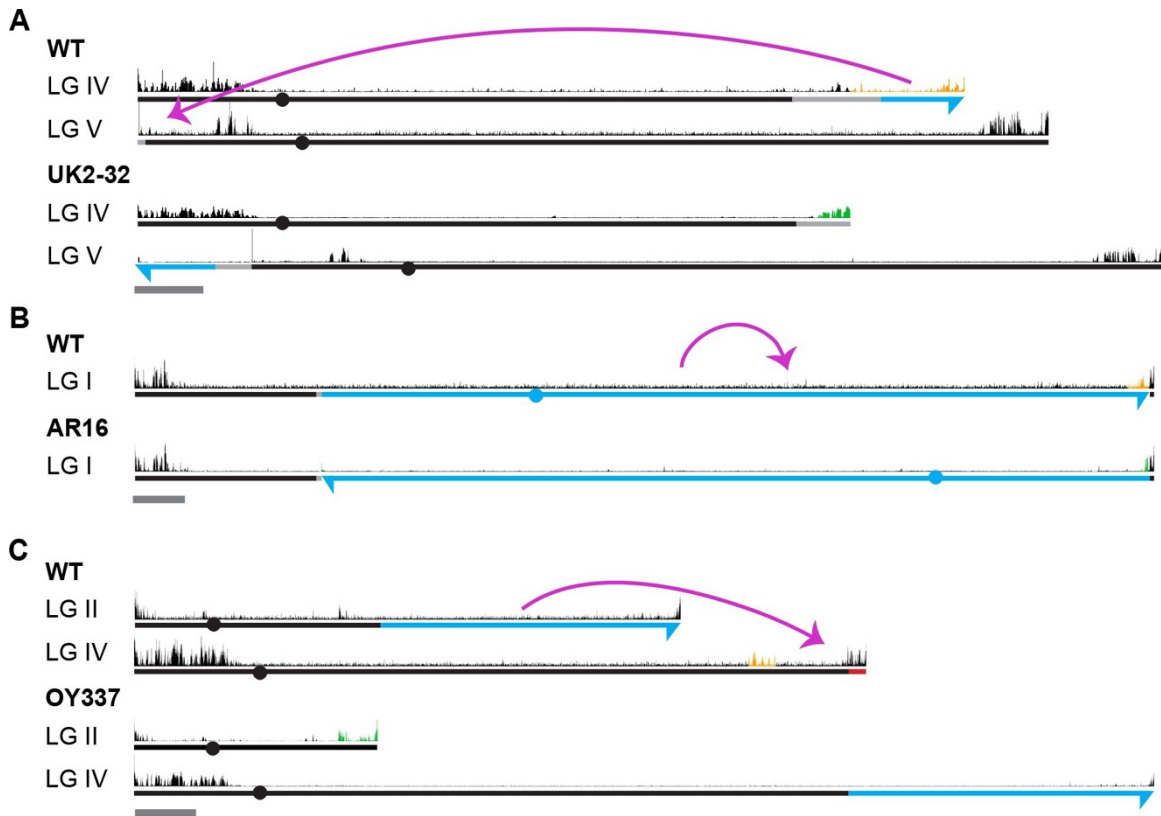
## Figures for Chapter III



**Fig. 2. Chromosomal rearrangements are associated with altered H3K27me2/3.** (A-F) Schematics show the movement (magenta curved arrows) of translocated segments and the resulting chromosomal rearrangements for six translocation strains. ChIPs were done on biological triplicates and pooled for sequencing. Blue and dashed lines distinguish source of chromosome while arrowheads indicate directionality of rearranged segments. H3K27me2/3 ChIP-seq tracks of WT and translocation strains are displayed above chromosome diagrams with zoom-in sections in boxes. H3K27me2/3 signals that were lost in translocation strains are shown in orange, H3K27me2/3 signals gained are indicated in green and invariant H3K27me2/3 signals are shown in black. Circles indicate centromeres. Gains and losses of H3K27me2/3 were confirmed by qPCR (Fig. 6).

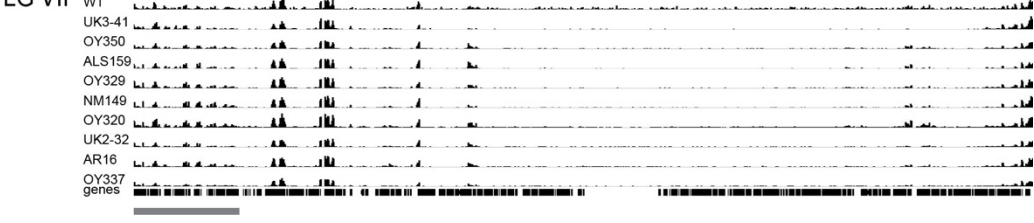
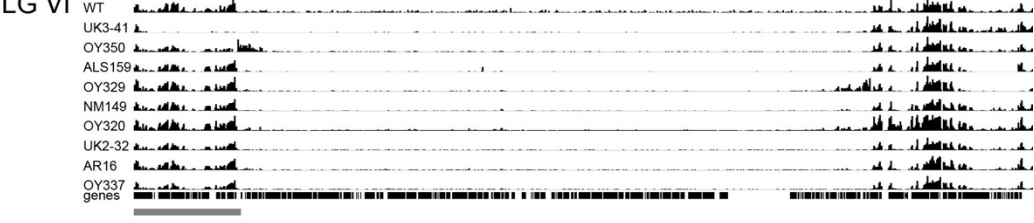
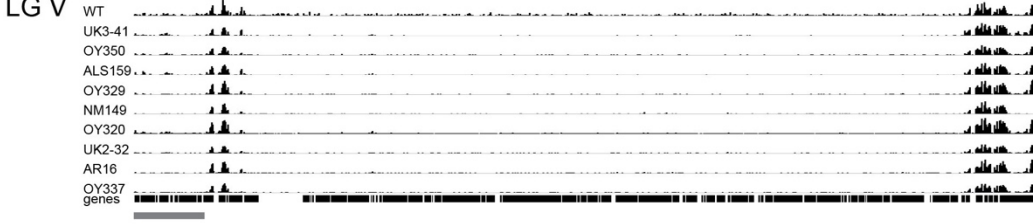
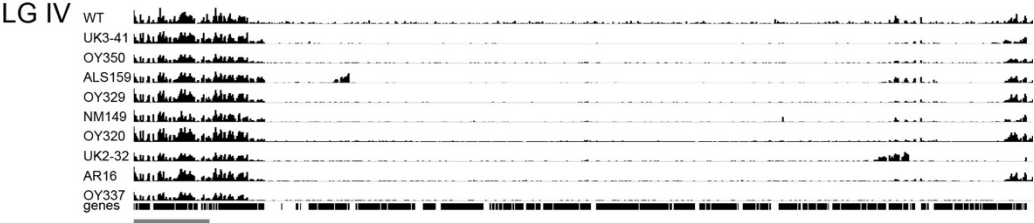
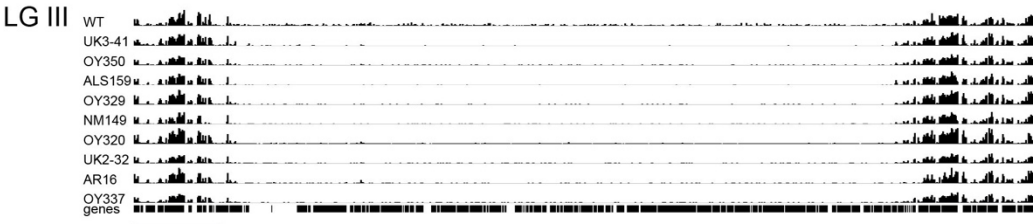
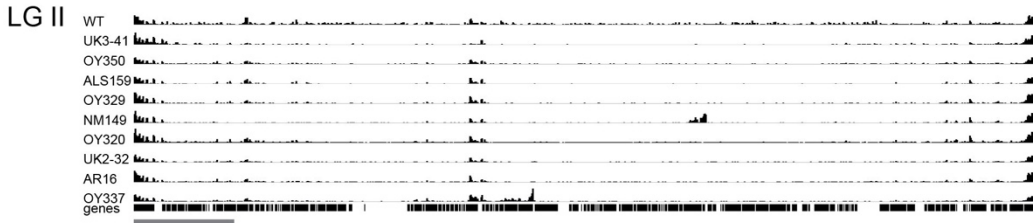
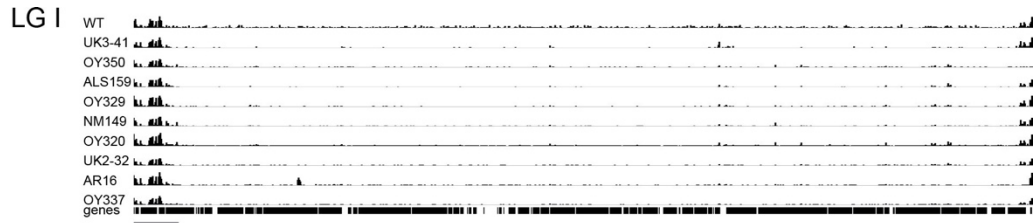


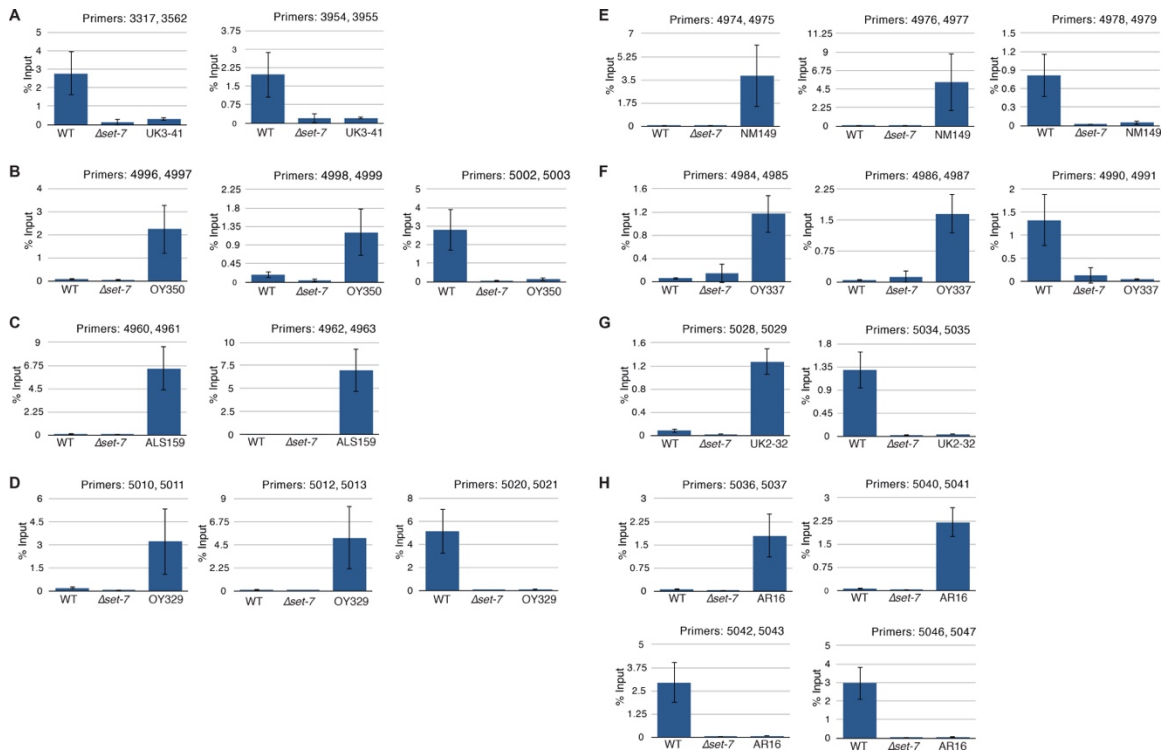
**Fig. 3. Three kilobase segments from a natural H3K27me2/3 domain are insufficient to trigger *de novo* H3K27me2/3 at ectopic loci.** (A) Distribution of H3K27me2/3 (black), genes (gray) and H3K9me3 (blue, inverted) are displayed for the left subteleromere of LG VI. The red bar indicates the region deleted for the transplantation experiments. The eight three kb segments (expanded below the black bar) were targeted to the *his-3* and *csr-1* loci. Green vertical bars represent the locations of qPCR primer pairs used in H3K27me2/3 ChIP experiments. H3K27me2/3 ChIP-qPCR indicates the absence of ectopic H3K27me2/3 in strains containing regions 1-8, each separately targeted to (B) *his-3* and (C) *csr-1*. Each ChIP experiment was performed in biological triplicate; qPCR analyses were performed on each sample in technical triplicate. Bars represent means of biological triplicates and error bars show standard deviation.



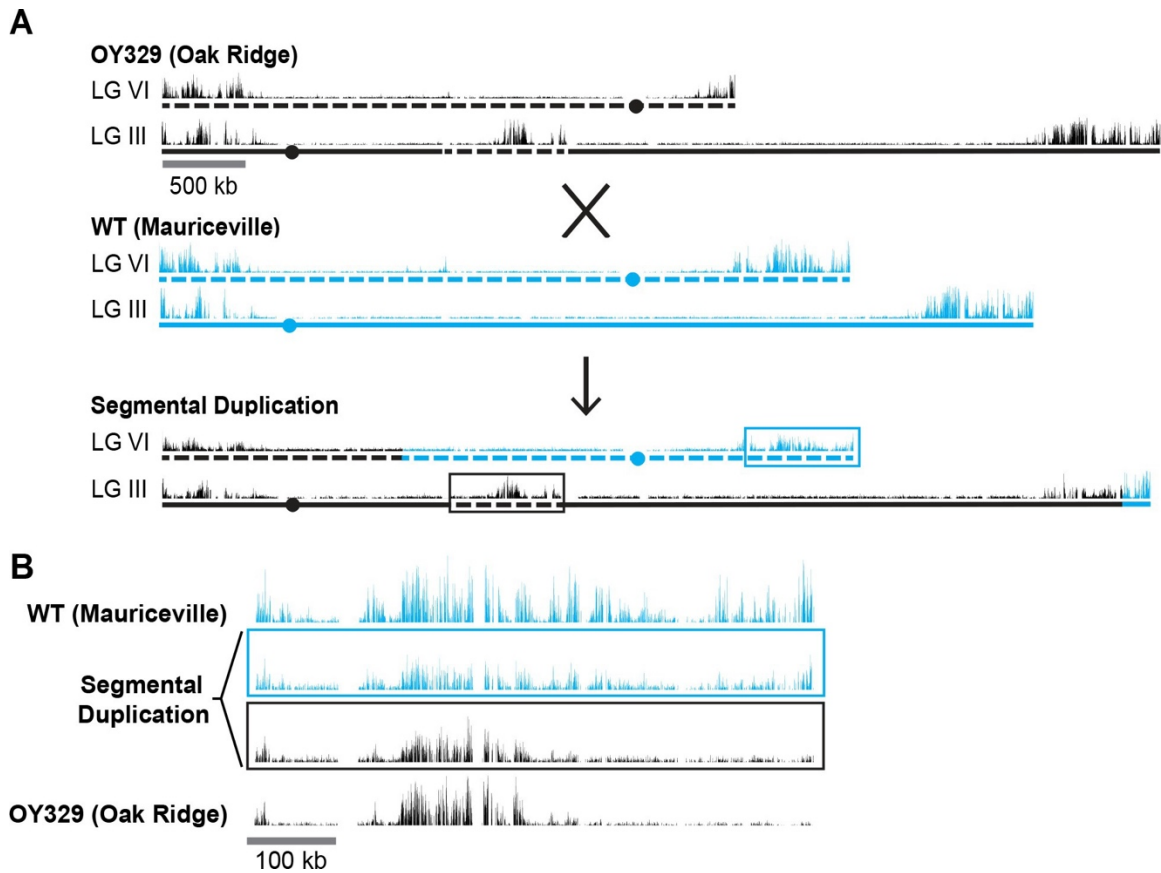
**Fig. 4. H3K27me2/3 profiles of additional strains containing chromosomal rearrangements.** Chromosome rearrangements and H3K27me2/3 ChIP-seq (pooled biological triplicates) profiles are shown for translocations UK2-32 (A), AR16 (B) and OY337 (C). Magenta curved arrows illustrate the movement of chromosomal segments in the translocations. Blue and dashed lines distinguish chromosome sources while arrows indicate the directionality of translocated segments. Light gray lines highlight regions of uncertainty with respect to the locations of breakpoints based on available genetic evidence. Solid red line indicates chromosome deletions. H3K27me2/3 ChIP-seq tracks of WT and translocation strains are displayed above chromosome diagrams. H3K27me2/3 that was lost in a translocation strain is highlighted in orange in the corresponding WT track while new H3K27me2/3 is indicated in green. Conserved H3K27me2/3 is shown in black. Circles indicate centromeres. Dark gray bars indicate 500 kb.

**Fig. 5 (see next page). Whole genome view of H3K27me2/3 ChIP-seq in wild-type and chromosomal translocation strains.** H3K27me2/3 ChIP-seq for all seven *N. crassa* chromosomes is displayed for WT and nine chromosomal translocation strains mapped onto the WT genome. Genes are indicated below the ChIP-seq track as vertical black bars. Grey bar represents 500 kb.



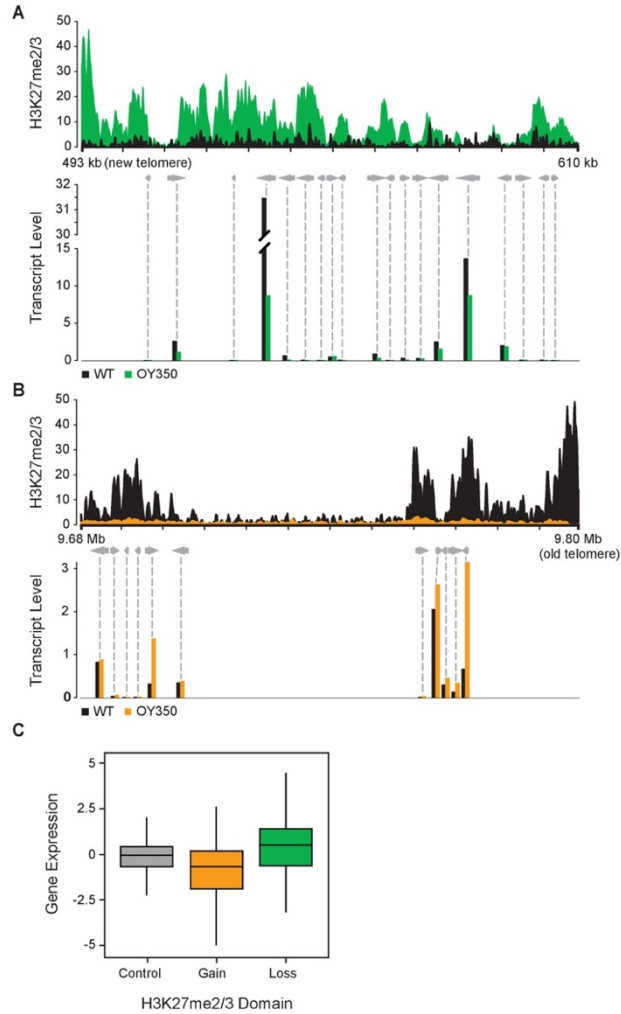


**Fig. 6. qPCR validation of H3K27me2/3 ChIP-seq data from translocation strains.** The gains and losses of H3K27me2/3 in ChIP-seq experiments were confirmed by ChIP-qPCR at representative regions for UK3-41 (A), OY350 (B), ALS159 (C), OY329 (D), NM149 (E), OY337 (F), UK2-32 (G), and AR16 (H). Each ChIP experiment was performed in biological triplicate; qPCR analyses were performed on each sample in technical triplicate. Bars represent means of biological triplicates and error bars show standard deviation.



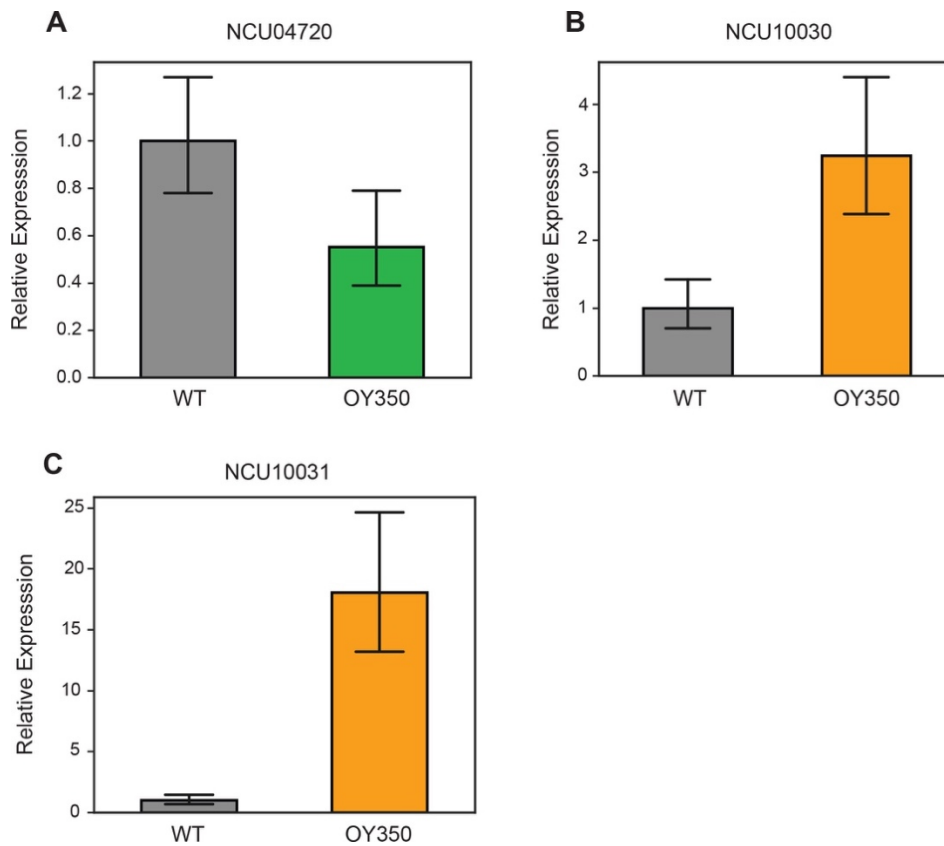
**Fig. 7. A segmental duplication confirms H3K27me2/3 position effect.** (A) Diagram of a cross between the OY329 (Oak Ridge) insertional translocation strain and a polymorphic wild-type strain (Mauriceville), resulting in a strain bearing a duplicated chromosomal segment (outlined by rectangles). H3K27me2/3 ChIP-seq tracks (single experiment) and chromosomes are shown in black for the Oak Ridge strain and blue for Mauriceville. Solid and dashed lines indicate chromosome source. (B) Expanded view of the duplicated chromosome segment in parental strains and duplication-containing offspring along with associated SNP-parsed H3K27me2/3 profiles.





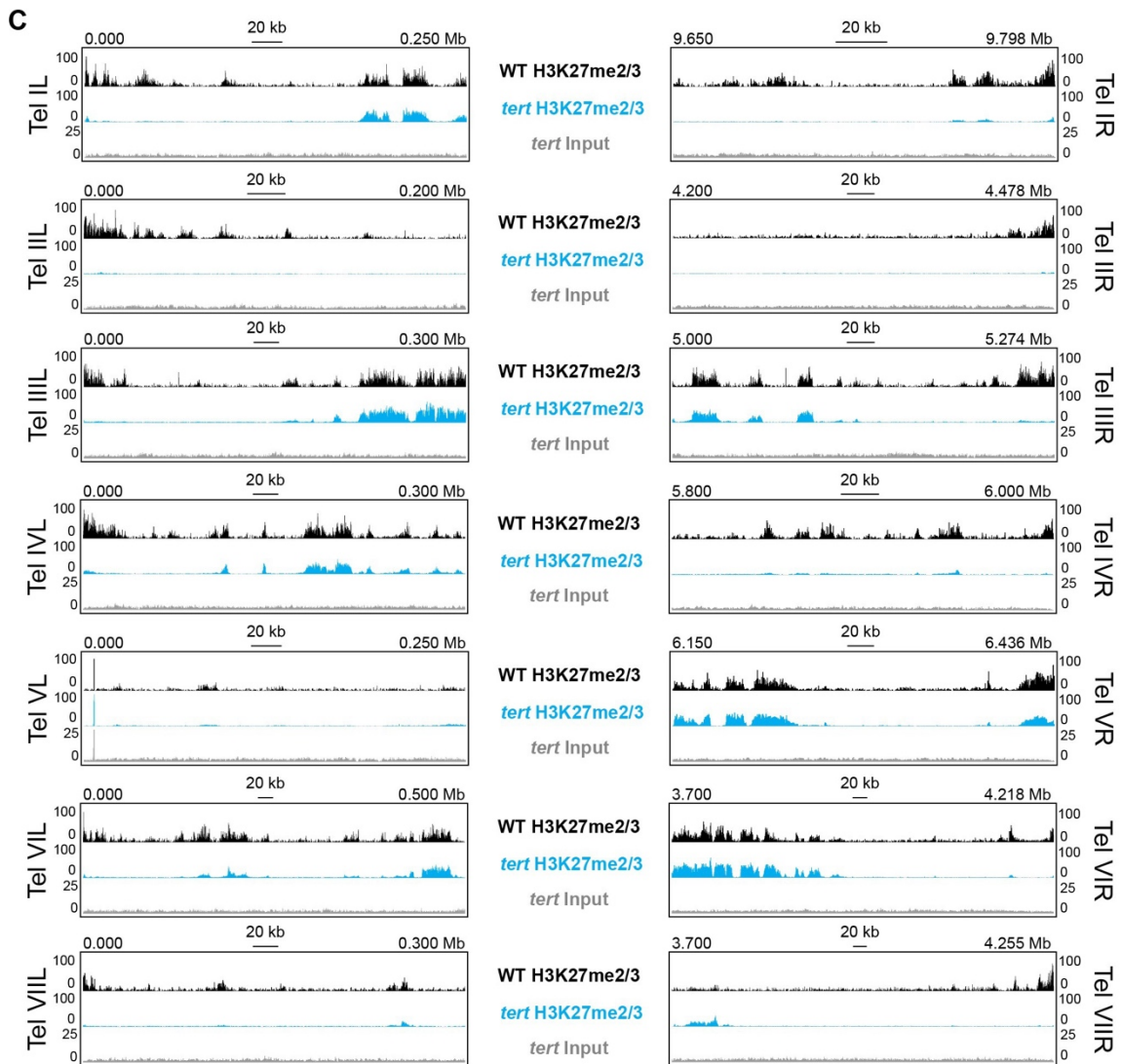
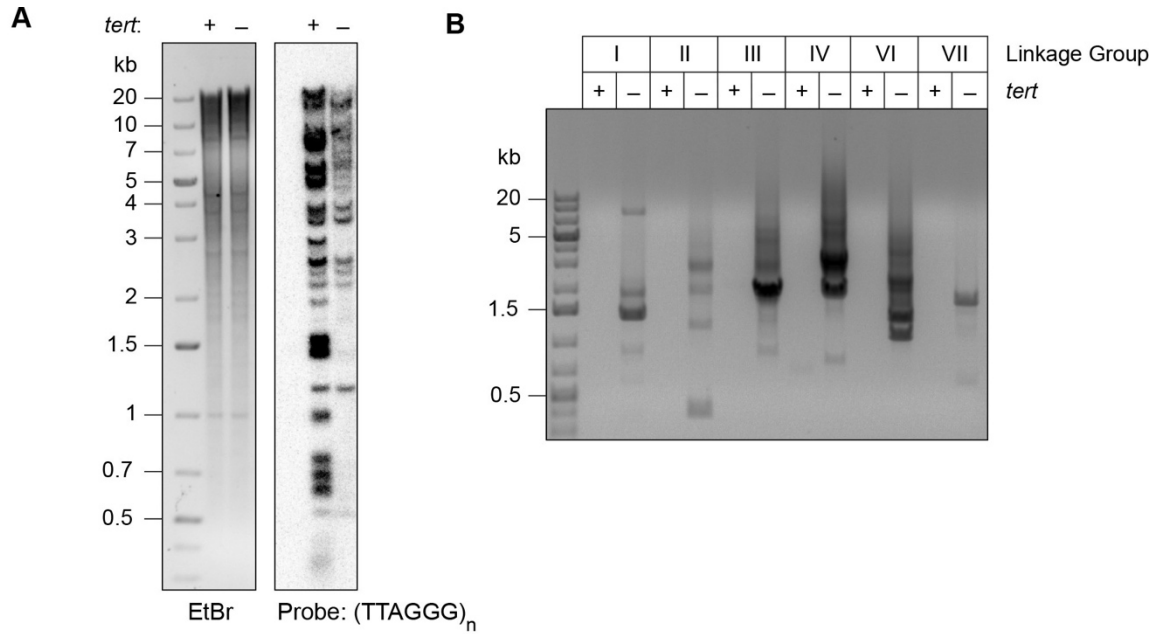
**Fig. 8. Altered gene expression reflects changes in H3K27me2/3.** (A) Reduced gene expression in novel subtelomeric region (LG VIL) that gains H3K27me2/3 in translocation OY350. The top panel displays H3K27me2/3 ChIP-seq reads of OY350 (green) normalized to the corresponding the non-translocated segment in WT (black) using a 350 bp sliding window. The bottom panel shows normalized mRNA-seq read counts (counts/1000) for genes (gray arrows) in the region for WT (black) and OY350 (green). ChIP-seq data are from pooled biological triplicates; RNA-seq data are from biological duplicates (B) Shifting a normally subtelomeric region to an internal position in translocation OY350 results in loss of H3K27me2/3 on LG IR and increased gene expression. The top panel displays H3K27me2/3 ChIP-seq reads of OY350 (orange) normalized to WT (black) using a 500 bp sliding window. The bottom panel shows normalized mRNA-seq read counts (counts/1000) for genes (gray arrows) in the region for WT (black) and OY350 (orange). (C) A box plot summarizes the relationship between gain (green) or loss (orange) of H3K27me2/3 in seven chromosomal translocation strains (ALS159, AR16, OY329, OY337, OY350, UK2-32, and UK3-41) and associated gene expression changes ( $\log_2[\text{translocation}/\text{WT}]$ ) within the corresponding domains. The control (gray) represents gene expression in regions that do not exhibit changes in H3K27me2/3.

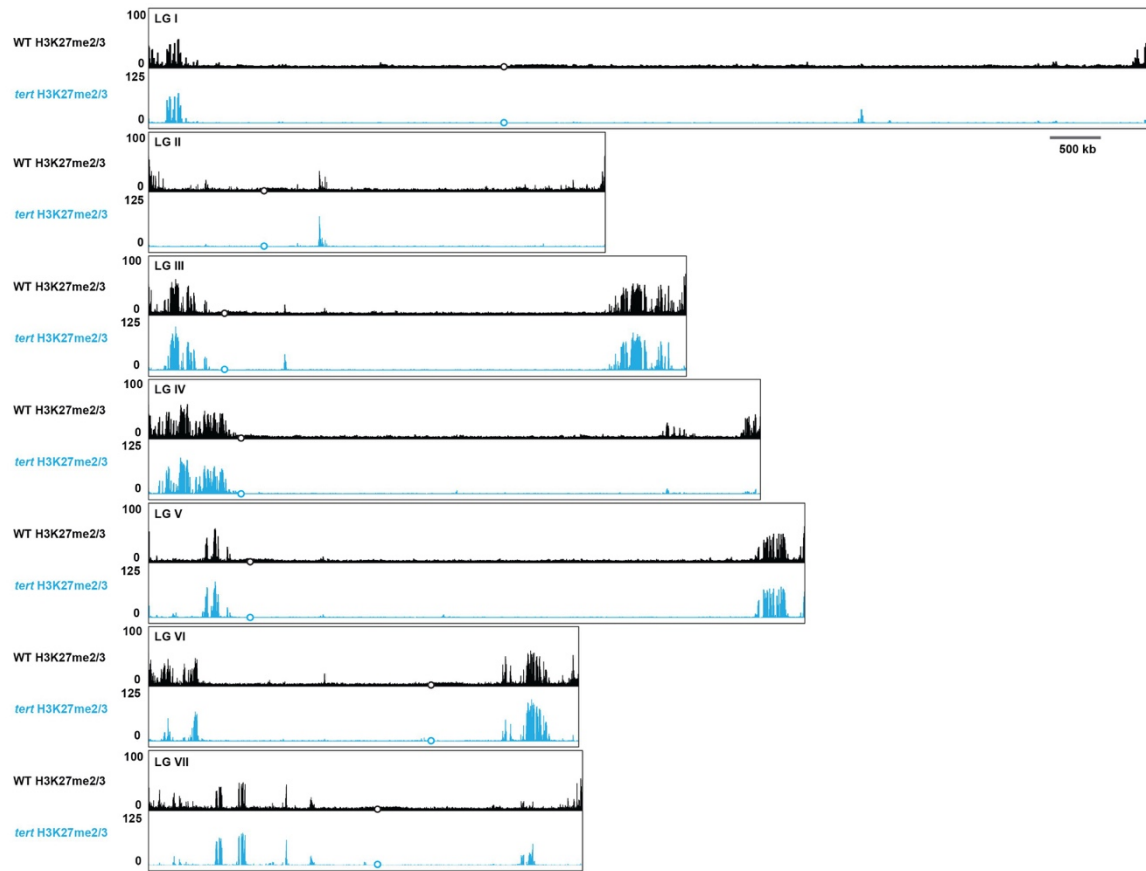




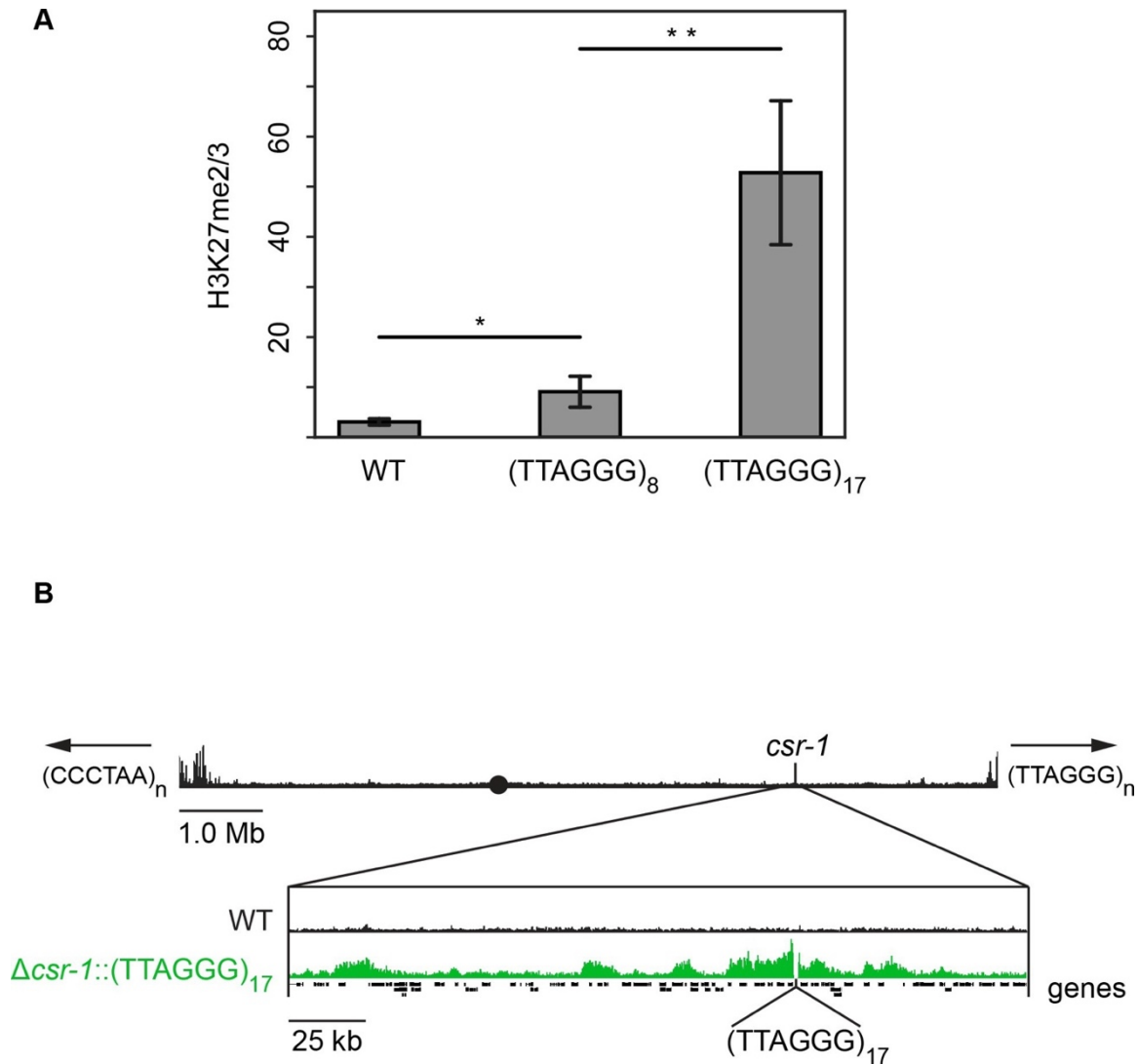
**Fig. 9. qPCR validation of RNA-seq expression changes in translocation OY350.** The effects of gain of H3K27me2/3 (A) and loss of H3K27me2/3 (B,C) on associated gene expression levels were verified by qPCR. Gene expression levels were normalized to housekeeping gene, NCU02840. qPCR analyses were performed on each sample in technical triplicate. Bars represent means of biological triplicates and error bars show standard deviation.

**Fig. 10. (see next page) Loss of *tert* disrupts chromosome ends and abolishes subtelomeric H3K27me2/3.** (A) Southern hybridization of genomic DNA from WT and *tert* strains digested with *Hind*III/*Not*I reveals loss of chromosome ends marked with the telomere repeats (TTAGGG)<sub>n</sub>, which was used as the probe. The ethidium bromide (EtBr) image demonstrates equal loading of WT and *tert* genomic DNA. (B) Circularization of chromosomes was demonstrated by generation of PCR products with outwardly directed primers near chromosome ends in a *tert* strain, but not a WT strain. (C) H3K27me2/3 ChIP-seq of *tert* (single sample; blue track) compared to WT (pooled biological triplicates; black track). Sequence coverage of the *tert* ChIP input (gray track) is also shown.

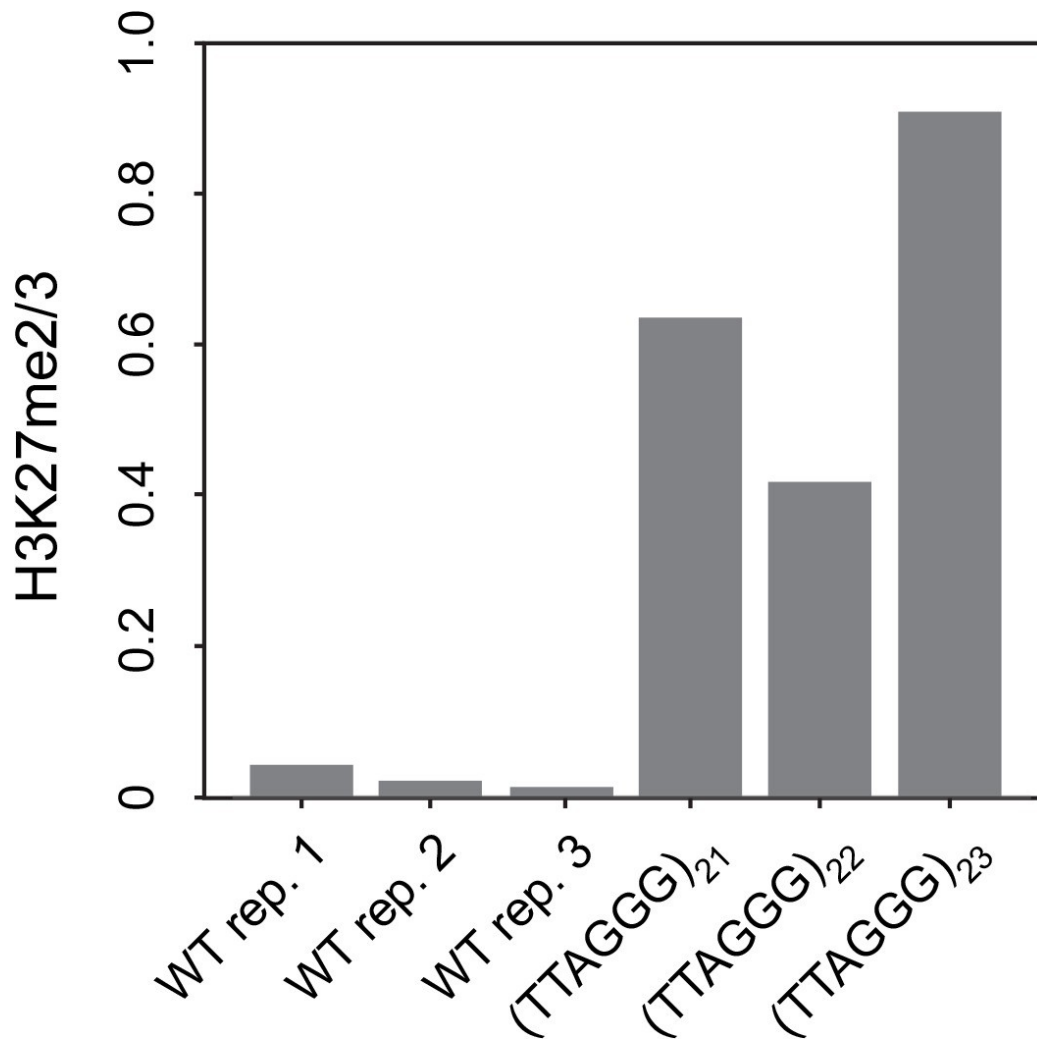




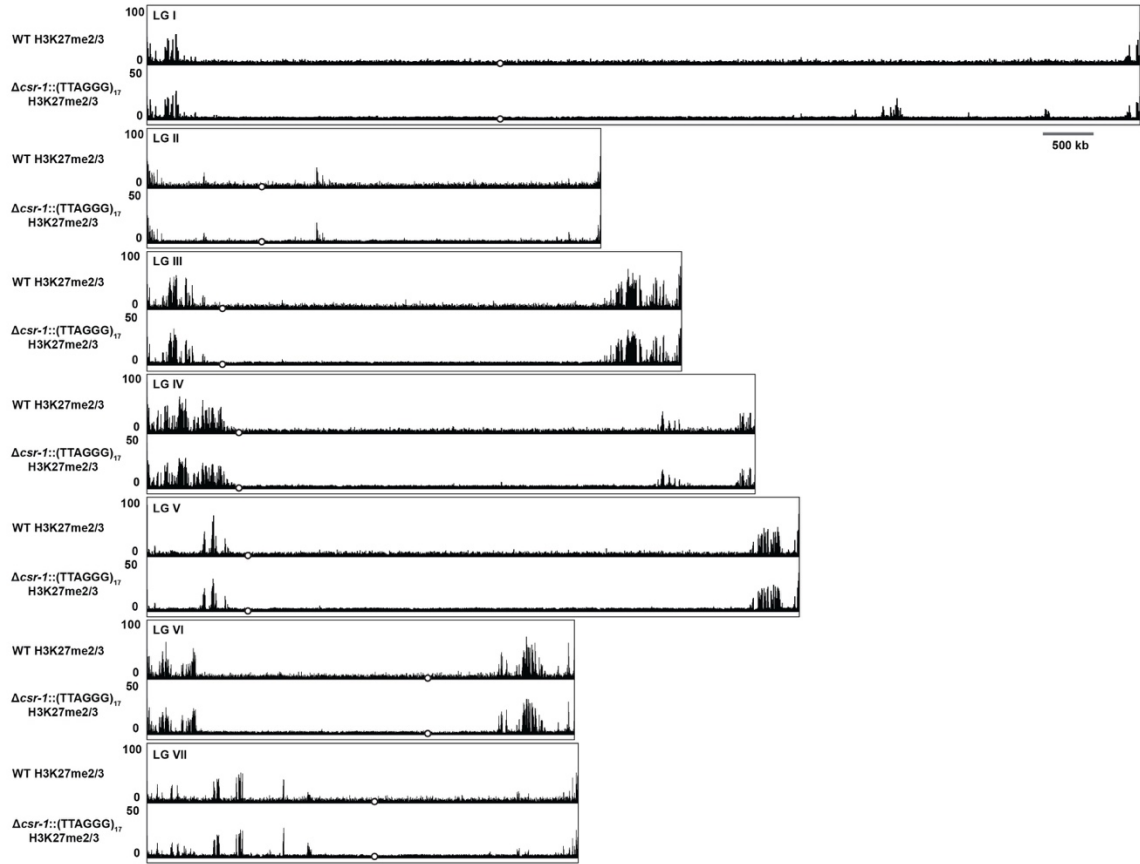
**Fig. 11. Whole genome view of H3K27me2/3 ChIP-seq in wild type and *tert*.** H3K27me2/3 ChIP-seq for all seven *N. crassa* chromosomes is displayed for WT (pooled biological triplicates; black track) and *tert* (single sample; blue). The y-axis is number of reads averaged over 25 bp windows. Open black or blue circles indicate location of centromeres.



**Fig. 12. Telomere repeats targeted to the *csr-1* locus induce an approximately 225 kb H3K27me2/3 domain.** (A) qPCR analyses (biological triplicates) of H3K27me2/3 ChIP with strains containing 8 or 17 telomere repeats at the *csr-1* locus, with WT strain as a control (means displayed; error bars show standard deviation; single and double asterisks represent  $p = 0.01488$  and  $p = 0.00336$  respectively). Level of H3K27me2/3 normalized to Telomere 1L. (B) H3K27me2/3 ChIP-seq of full WT LG I (pooled biological triplicates; black track). Black circle indicates the centromere, *csr-1* position is indicated by vertical line, and arrows at chromosome ends indicate 5' to 3' polarity of telomere repeats. Expansion below shows the extent and shape of the H3K27me2/3 domain induced in the  $\Delta csr-1::(TTAGGG)_{17}$  strain (single sample; green track) compared to WT (black track). Site of *csr-1* replacement with telomere repeats indicated below ChIP-seq tracks. Genes are displayed as black bars.

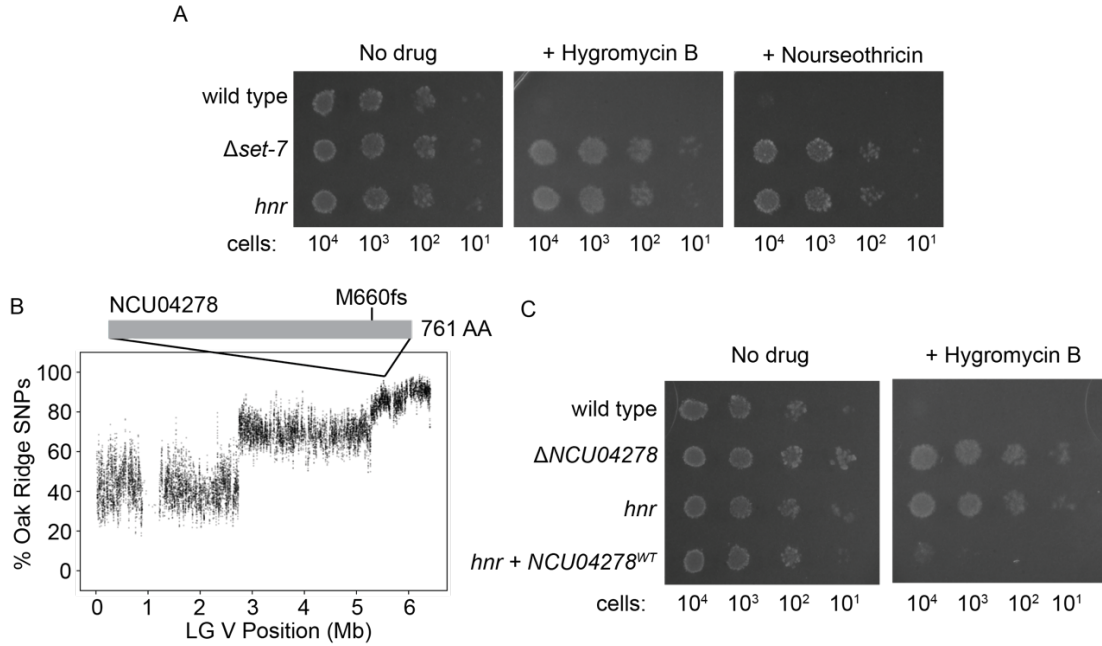


**Fig. 13. Telomere repeats targeted to the *his-3* locus induce local H3K27me2/3 domain.** qPCR analyses of H3K27me2/3 ChIP with strains containing 21, 22, or 23 telomere repeats at the *his-3* locus, with WT biological triplicates as a control. Level of H3K27me2/3 normalized to Telomere 1L.



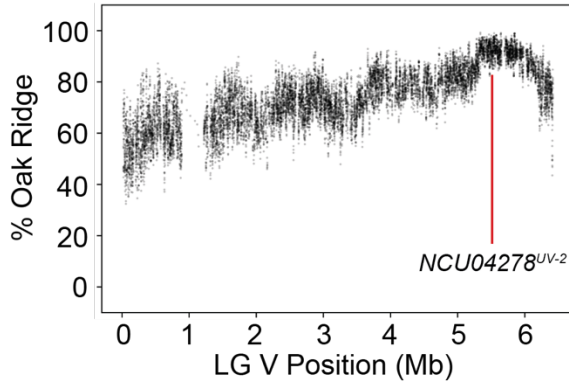
**Fig. 14. Whole genome view of H3K27me2/3 ChIP-seq in wild type and  $\Delta csr-I::(TTAGGG)_{17}$ .** H3K27me2/3 ChIP-seq for all seven *N. crassa* chromosomes is displayed for WT (pooled biological triplicates) and  $\Delta csr-I::(TTAGGG)_{17}$  (single sample). The y-axis is number of reads averaged over 25 bp windows. Open black circles indicate location of centromeres.

## Figures for Chapter IV

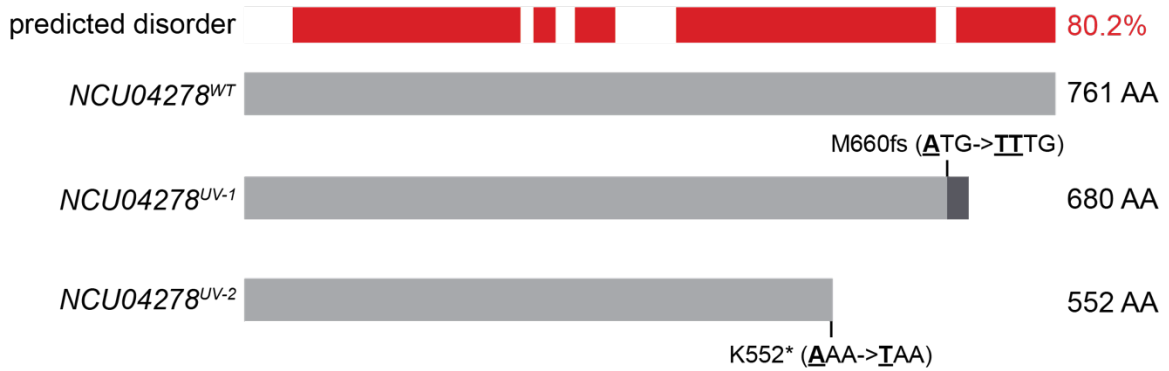


**Fig. 15. Identification of NCU04278 as a potential component of the Polycomb silencing pathway.** (A) Serial dilution spot test silencing assay for the indicated strains plated on the indicated media. All strains harbor  $P_{NCU05173}::hph$  and  $P_{NCU07152}::nat-1$ . (B) Whole genome sequencing of pooled *hnr* mutant genomic DNA identified a region on the right arm of LG V which is enriched for Oak Ridge SNPs and contained an insertional frameshift in *NCU04278* (structure of which is represented as a solid gray bar, as no known domains predicted). Each translucent point represents a running average (window size = 10 SNPs, step size = 1 SNP). (C) Serial dilution spot test silencing assay for the indicated strains. *hnr* +  $NCU04278^{WT}$  has a wild-type copy of *NCU04278* at the *his-3* locus. All strains harbor  $P_{NCU05173}::hph$ .

A

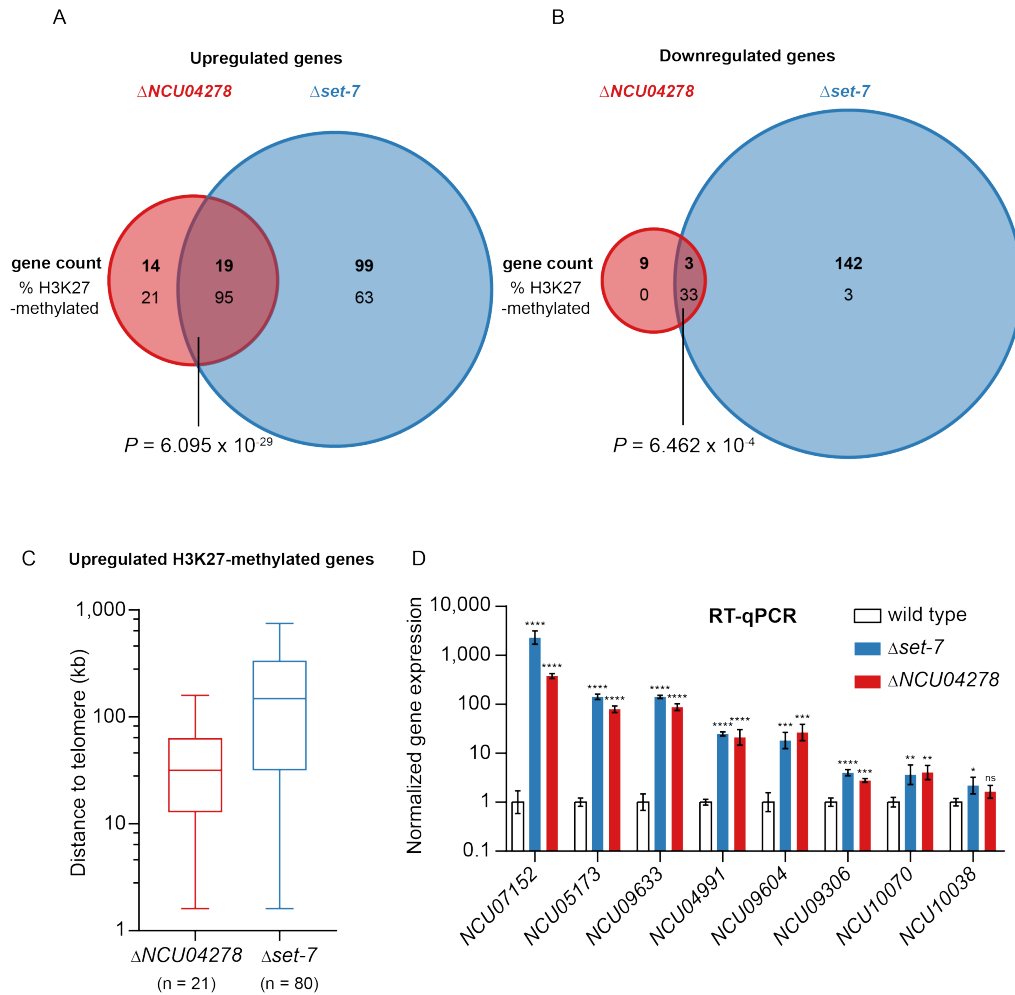


B

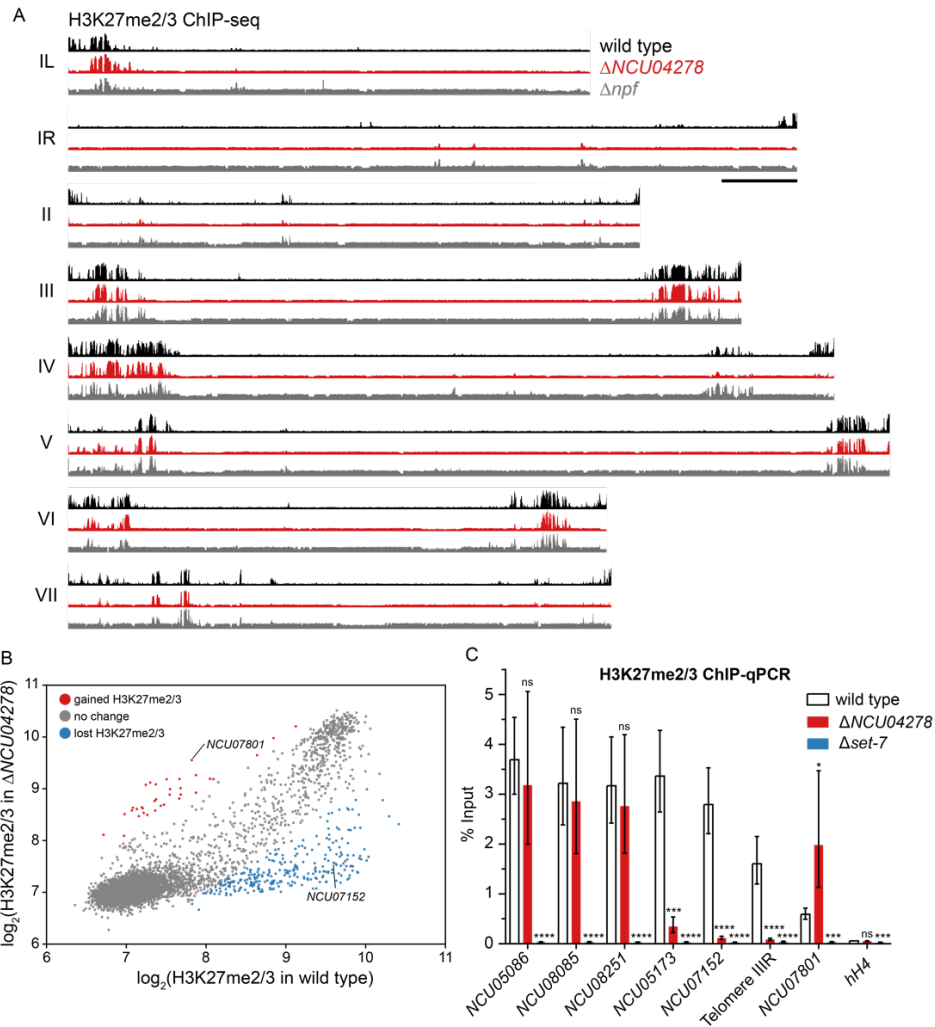


**Fig. 16. Mapping and characterization of an additional *NCU04278* allele.** (A) Whole genome sequencing of pooled mutant genomic DNA identified a region on the right arm of LG V which is enriched for Oak Ridge SNPs and contained a nonsense mutation in *NCU04278*. Each translucent point represents a running average (window size = 10 SNPs, step size = 1 SNP). (B) Diagram of the primary structure of each indicated *NCU04278* allele (light gray bars are wild-type sequence; dark gray bar is frameshifted sequence). Predicted intrinsically disordered regions within *NCU04278* are labelled above (red bars).

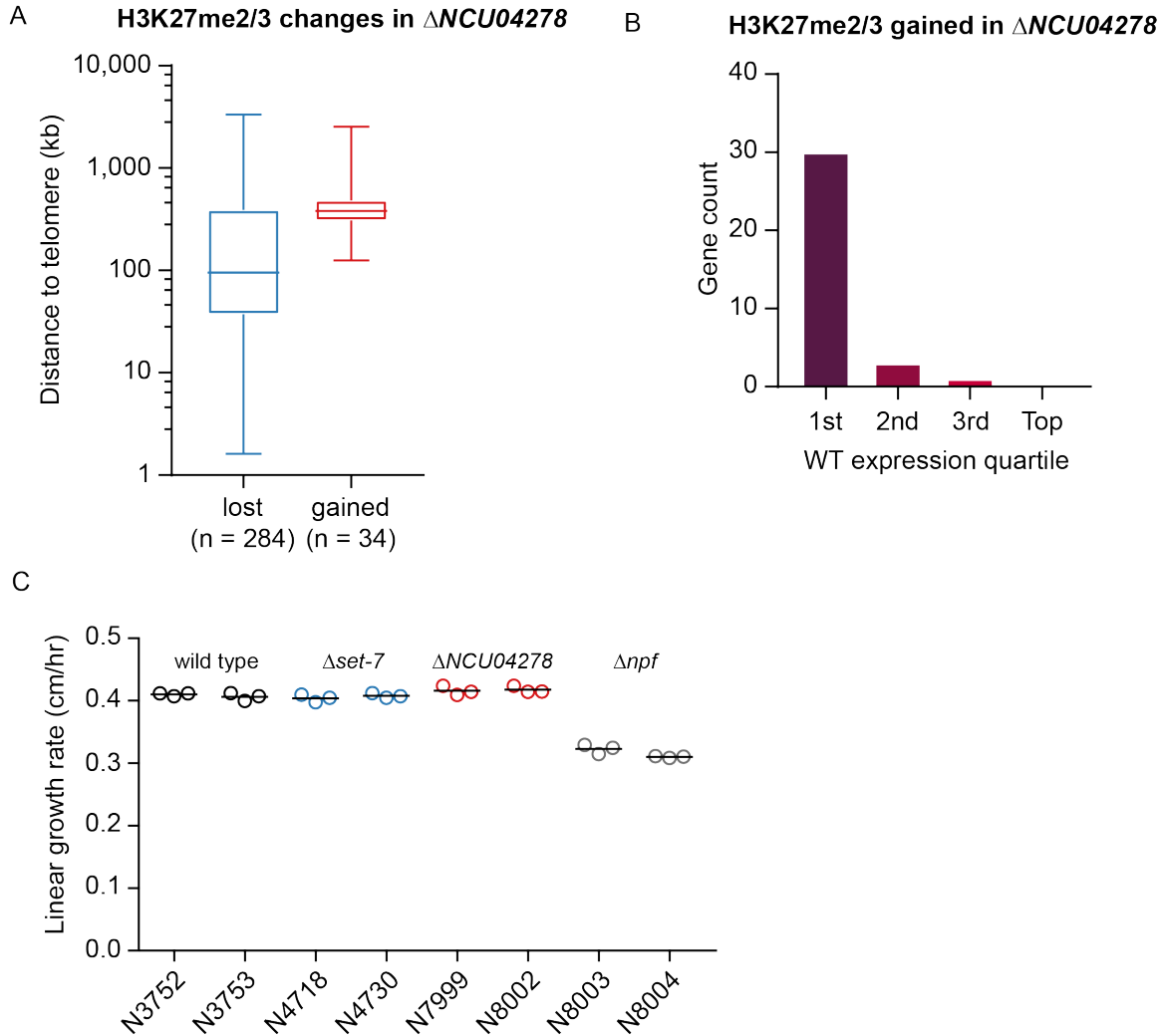




**Fig. 17. Loss of NCU04278 upregulates a subset of SET-7 repressed genes near chromosome ends.** (A) Venn diagram depicting genes (numbers in bold) that appear upregulated by mRNA-seq in both  $\Delta NCU04278$  and  $\Delta set-7$  strains, only in  $\Delta NCU04278$  strains, or only in  $\Delta set-7$  strains, using a significance cutoff of  $\log_2(\text{mutant/wild type}) > 1$  and  $P$  value  $< 0.05$ . The percentage of total upregulated genes that are H3K27 methylated for each gene set is indicated below the total gene count. Significance of overlapping gene sets was determined using a hypergeometric test. (B) Same as in A, but for downregulated genes ( $\log_2(\text{mutant/wild type}) < -1$ ). (C) Box and whisker plot of the distances from the telomere of H3K27-methylated genes that appear upregulated by mRNA-seq in the indicated genotypes. Boxes represent interquartile range, horizontal lines represent the median, and whiskers represent minimum and maximum values. (D) RT-qPCR of H3K27-methylated genes that were replaced with antibiotic resistance genes (*NCU07152* and *NCU05173*) and used for initial selection of mutants, and H3K27-methylated genes that appeared upregulated in both  $\Delta NCU04278$  and  $\Delta set-7$  strains by mRNA-seq (*NCU09633*, *NCU04991*, *NCU09604*, *NCU09306*, *NCU10070*, *NCU10038*). Each value was normalized to expression of *NCU02840* and presented relative to wild type. Filled bars represent the mean from biological triplicates and error bars show standard deviation (\*\*\*\* for  $P < 0.0001$ , \*\*\* for  $P < 0.001$ , \*\* for  $P < 0.01$ , \* for  $P < 0.05$ , and ns for not significant; all relative to wild type by two-tailed, unpaired t-test).



**Fig. 18. Genome-wide distribution and level of H3K27 methylation in wild-type,  $\Delta NCU04278$ , and  $\Delta npf$  strains.** (A) H3K27me2/3 ChIP-seq of merged biological replicates of wild type (black),  $\Delta NCU04278$  (red), and  $\Delta npf$  (gray) strains displayed for all seven LGs of *N. crassa*. LG I is split at the right end of its centromere into IL and IR. Y-axes are 900, 850, and 450 RPKM for wild type,  $\Delta NCU04278$ , and  $\Delta npf$  respectively. Black bar represents 500 kb. (B) Scatter plot showing the correlation of H3K27me2/3 levels at all genes (dots) in wild type and  $\Delta NCU04278$  based on merged biological replicates of ChIP-seq data. Dots representing genes are colored based on their relative H3K27me2/3 levels in  $\Delta NCU04278$  compared to wild type (red for  $\log_2(\Delta NCU04278/wild\ type) > 1$ , gray for  $-1 < \log_2(\Delta NCU04278/wild\ type) < 1$ , and blue for  $\log_2(\Delta NCU04278/wild\ type) < -1$ ). Representative genes that gained (*NCU07801*) or lost (*NCU07152*) H3K27me2/3 in  $\Delta NCU04278$  are indicated. (C) H3K27me2/3 ChIP-qPCR to validate ChIP-seq data at eight regions – *NCU05086*, *NCU08085*, *NCU08251* (positive controls); *NCU05173*, *NCU07152*, Telomere IIIR (lost H3K27me2/3 in  $\Delta NCU04278$ ); *NCU07801* (gained H3K27me2/3 in  $\Delta NCU04278$ ); *hH4* (negative control). Filled bars represent the mean of biological triplicates and error bars show standard deviation (\*\*\*\* for  $P < 0.0001$ , \*\*\* for  $P < 0.001$ , \* for  $P < 0.05$ , ns for not significant; all relative to wild type by two-tailed, unpaired t-test).

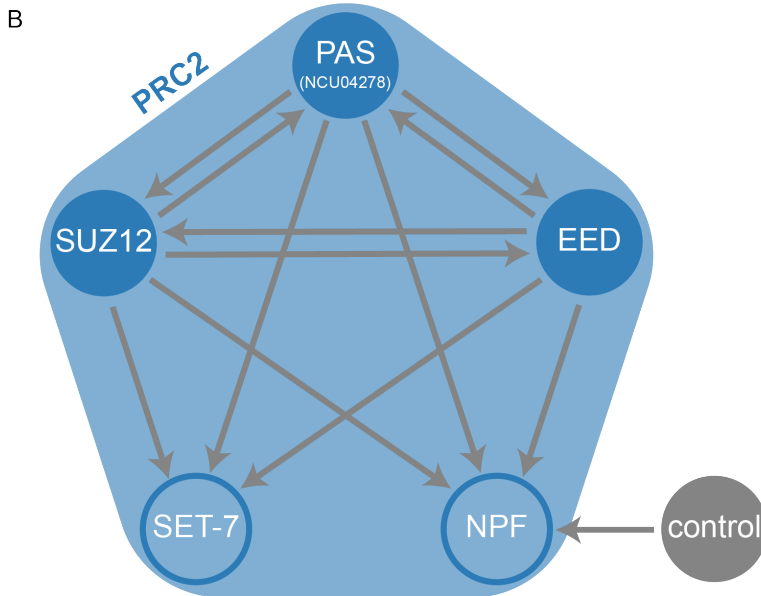


**Fig. 19. The effect of loss of NCU04278 on H3K27 methylation and growth rate.**

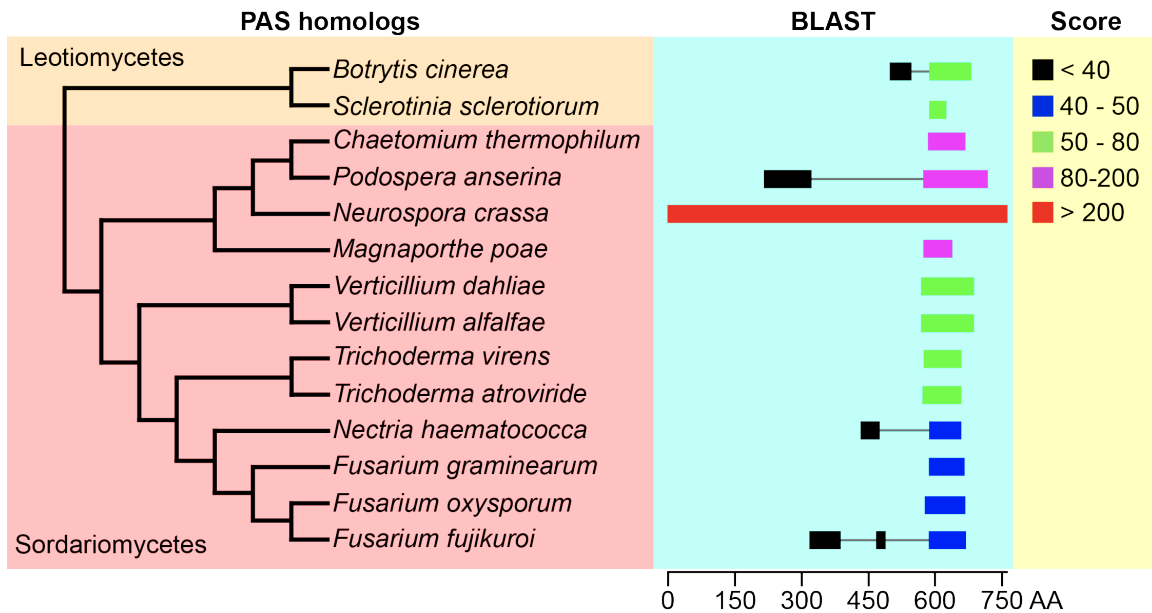
(A) Box and whisker plot of distance from the telomere of genes that lost ( $\log_2(\Delta NCU04278/\text{wild type}) < -1$ ,  $n = 284$ ) and gained ( $\log_2(\Delta NCU04278/\text{wild type}) > 1$ ,  $n = 34$ ) H3K27me2/3 in  $\Delta NCU04278$ . Box represents interquartile range, horizontal line is median, and whiskers represent minimum and maximum values. (B) Analysis of gene expression level of genes that gain H3K27me2/3 in  $\Delta NCU04278$  strains. (C) Linear growth rates measured by ‘race tubes’ are shown for two biological replicates of wild-type (N3752, N3753),  $\Delta set-7$  (N4718, N4730),  $\Delta NCU04278$  (N7999, N8002), and  $\Delta npf$  (N8003, N8004) strains. Horizontal lines represent the mean of three technical replicates (open circles).

A

		IP-MS 'bait'			
		NCU04278-3xFLAG	SUZ12-3xFLAG	3xFLAG-EED	3xFLAG-control
IP-MS 'prey'	NCU04278 (81 kDa)	<b>138</b>	5	6	0
	SET-7 (176 kDa)	64	28	10	0
	SUZ12 (97 kDa)	56	<b>31</b>	8	0
	EED (74 kDa)	24	10	<b>32</b>	0
	NPF (50 kDa)	19	8	5	3
	control (60 kDa)	0	0	0	<b>142</b>



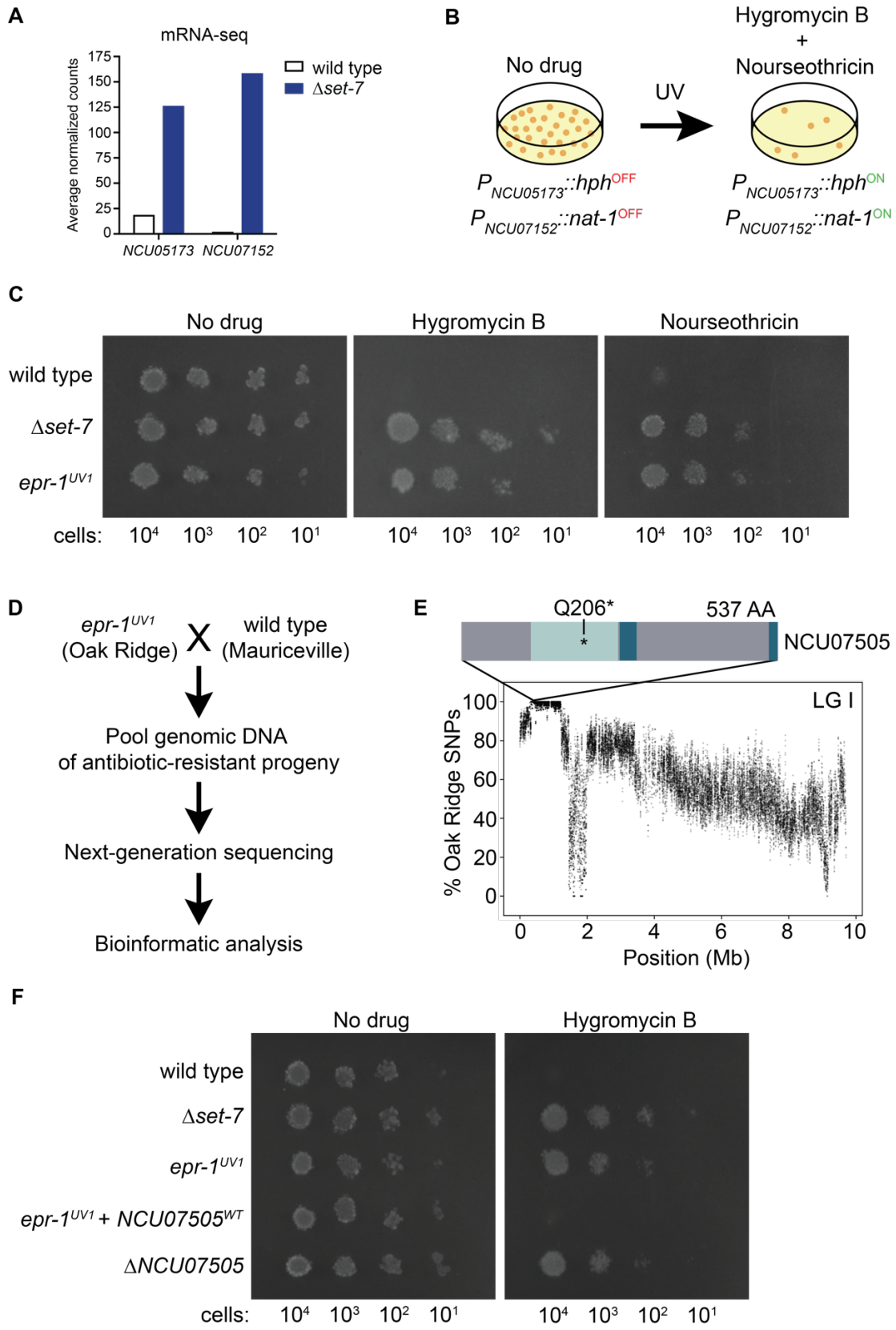
**Fig. 20. NCU04278 associates with known *N. crassa* PRC2 members.** (A) Exclusive unique peptide counts from immunoprecipitation followed by mass spectrometry (IP-MS) experiments. Each column represents the results of a single IP-MS with the indicated 'baits.' The 3xFLAG-EED dataset is previously published. The 3xFLAG-control is a nuclear protein that localizes to H3K27 methylated chromatin (NCU07505). (B) Protein interaction network of NCU04278-associated proteins. A separate IP-MS experiment was performed for each PAS (NCU04278), SUZ12, EED (solid blue circles) and the negative nuclear control (solid grey circle). IP-MS was not performed for SET-7 or NPF (empty blue circles). Arrows point to each 'baits' (solid blue circles) respective associated proteins.

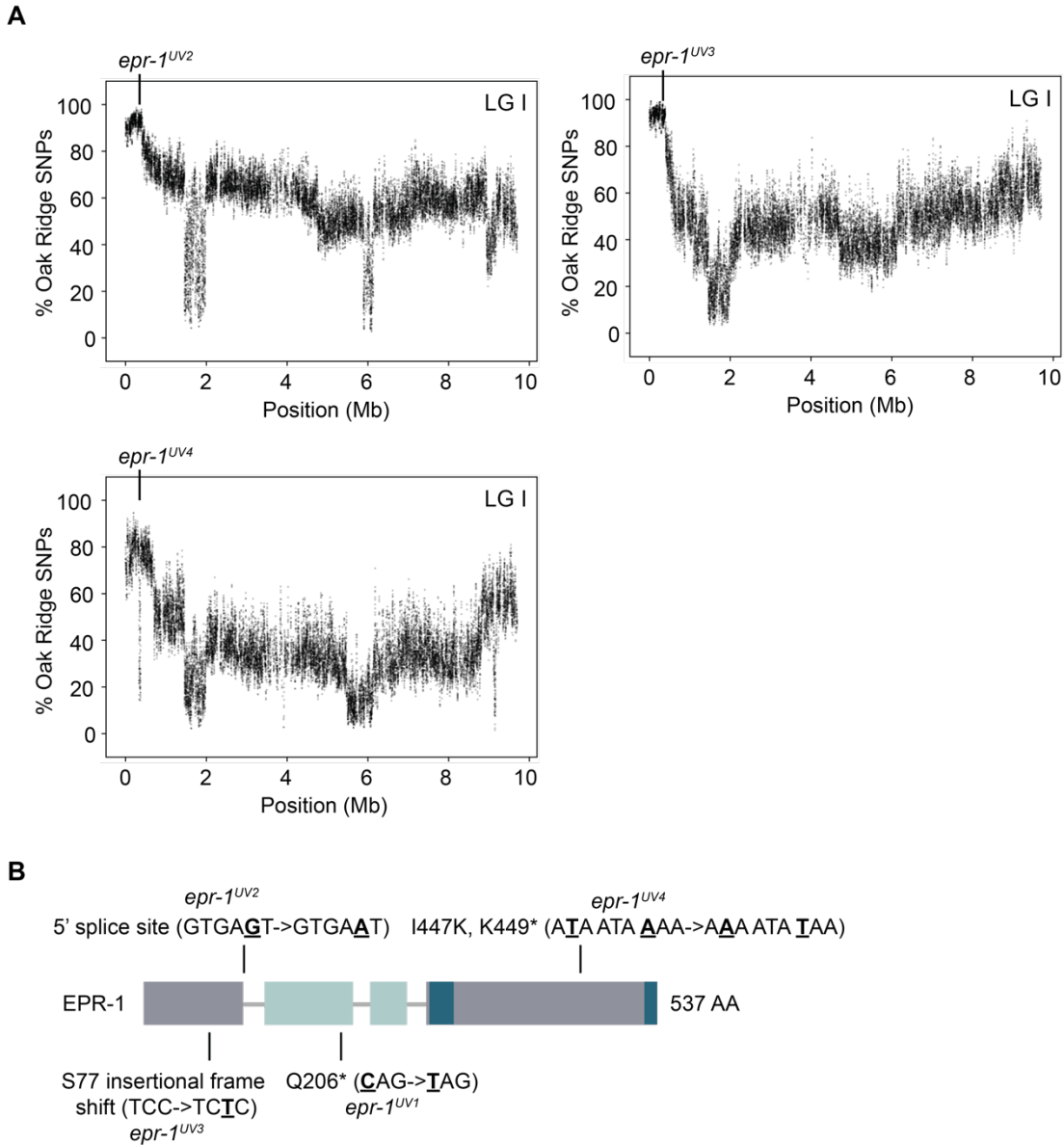


**Fig. 21. PAS homologs appear restricted to fungal lineages.** Phylogenetic tree highlighting select fungal species in class Sordariomyces (red background) or Leotiomyces (orange background) that contain predicted PAS homologs. Results of the Basic Local Alignment Search Tool (BLAST, light blue background), using *Neurospora crassa* PAS as the query, for each of the indicated species, is displayed to the right of the tree. Solid colored bars indicated an aligned region. Grey lines are for visual reference. Amino acid (AA) position in PAS is indicated below the alignments. Alignments are color coded according to their alignment score (yellow background).

## Figures for Chapter V

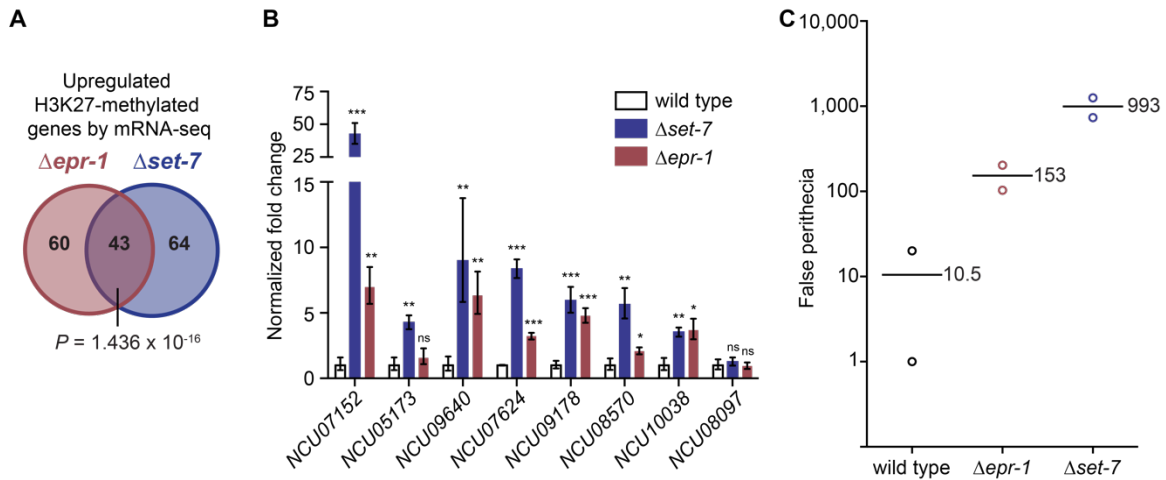
**Fig. 22 (see next page). Forward genetics identifies a novel gene, *epr-1*, required for H3K27 methylation-mediated repression.** (A) mRNA-seq results for two genes repressed by the *N. crassa* H3K27 methyltransferase, encoded by *set-7*. (B) Selection scheme, utilizing reporter genes illustrated in A, to identify factors required for H3K27 methylation-mediated silencing. (C) Serial dilution spot test silencing assay for the indicated strains plated on the indicated media. All strains harbor *P<sub>NCU05173</sub>::hph* and *P<sub>NCU07152</sub>::nat-1*. (D) Scheme for genetic mapping of critical mutation in *epr-1<sup>UVI</sup>*. (E) Whole genome sequencing of pooled *epr-1<sup>UVI</sup>* mutant genomic DNA identified a region on the left arm of linkage group I that is enriched for Oak Ridge single nucleotide polymorphisms (SNPs) and contains a premature stop codon in the BAH domain of NCU07505 (BAH domain, light blue; PHD finger (split), dark blue; no annotated domains, gray). Each translucent point represents a running average of SNPs (window size = 10 SNPs, step size = 1 SNP). (F) Serial dilution spot test silencing assay for the indicated strains. *epr-1<sup>UVI</sup>* + *NCU07505<sup>WT</sup>* has a wild-type copy of *NCU07505* at the *his-3* locus. All strains harbor *P<sub>NCU05173</sub>::hph*.





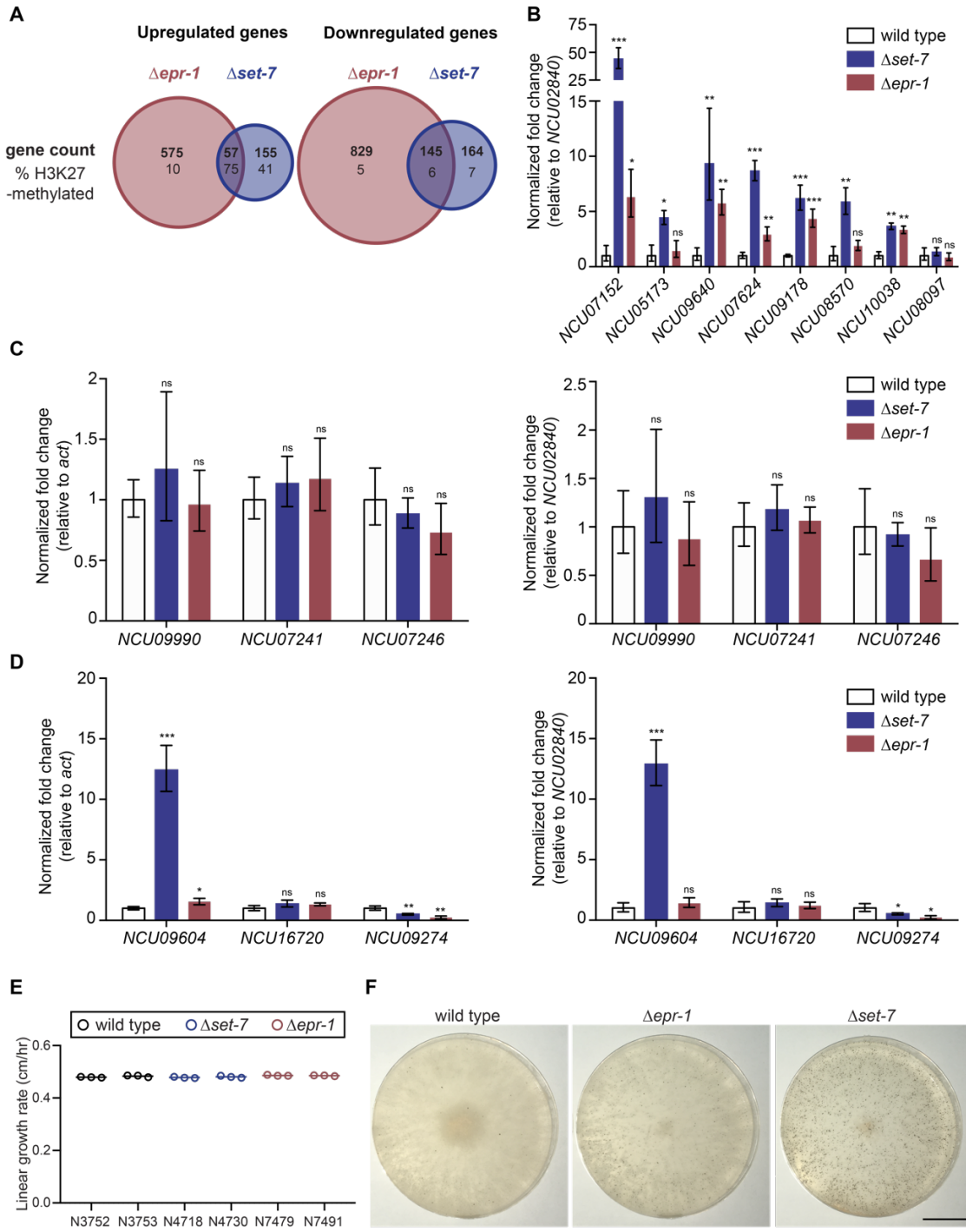
**Fig. 23. Three additional alleles of *epr-1* were identified during selection for mutants defective in H3K27 methylation-mediated repression.** (A) Whole genome sequencing of pooled genomic DNAs of mutant progeny, isolated from three additional mutants crossed with wild-type Mauriceville, independently identified a region on the left arm of LG I that was enriched for Oak Ridge SNPs and contained mutations in *epr-1* (*NCU07505*). (B) Domain structure of *NCU07505* showing the location and nature of the mutations in the *epr-1* alleles. Exons are indicated as boxes, with the BAH domain in light blue, the PHD finger (split) in dark blue, and exons with no known domains in gray. Gray lines connecting exons indicate the introns.

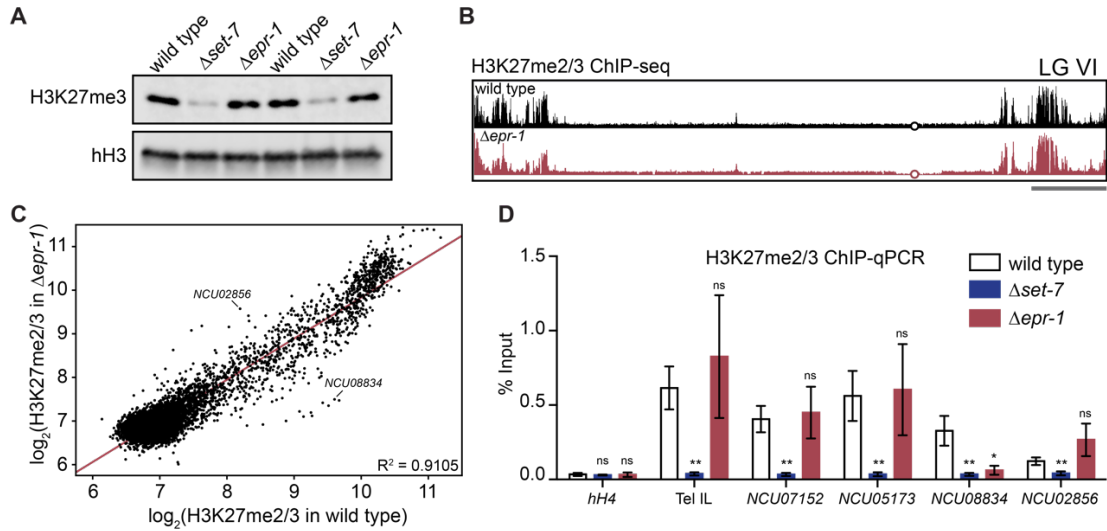




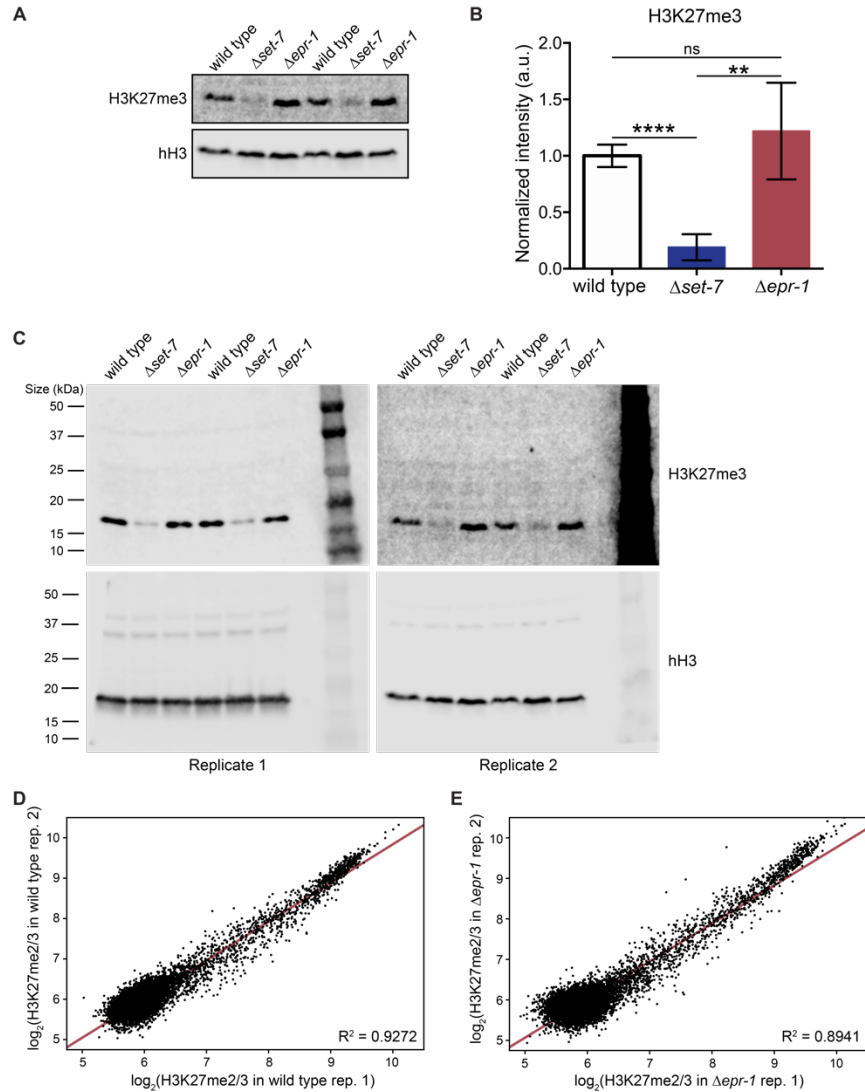
**Fig. 24.  $\Delta epr-1$  and  $\Delta set-7$  strains share defects in transcriptional silencing and sexual development.** (A) Venn diagram depicting H3K27-methylated genes that appear upregulated by mRNA-seq in both  $\Delta epr-1$  and  $\Delta set-7$  strains, only in  $\Delta epr-1$  strains, or only in  $\Delta set-7$  strains, using a significance cutoff of  $\log_2(\text{mutant/wild type}) > 1$  and a  $P$  value  $< 0.05$ . Significance of genes upregulated in both  $\Delta epr-1$  and  $\Delta set-7$  strains was determined using a hypergeometric test. (B) RT-qPCR of H3K27-methylated genes that were replaced with antibiotic resistance genes (*NCU07152*, *NCU05173*) and used for initial selection of mutants, and H3K27-methylated genes that appeared upregulated in both  $\Delta epr-1$  and  $\Delta set-7$  strains by mRNA-seq (*NCU09640*, *NCU07624*, *NCU09178*, *NCU08570*, *NCU10038*, *NCU08097*). Each value was normalized to expression of actin gene (*act*) and presented relative to wild type. Filled bars represent the mean from biological triplicates and error bars show standard deviation. (\*\*\*) for  $P < 0.001$ , \*\* for  $P < 0.01$ , \* for  $P < 0.05$ , and ns for not significant; all relative to wild type by two-tailed, unpaired t-test). (C) Quantification of false perithecia developed in a Petri dish (85 mm diameter) after two weeks of unfertilized growth are shown for the indicated strains. Horizontal lines and numbers indicate the mean of two biological replicates (open circles).

**Fig. 25 (see next page). Gene expression and phenotypic analysis of  $\Delta epr-1$  and  $\Delta set-7$  strains.** (A) Bold numbers indicate genes upregulated (left) or downregulated (right) in only  $\Delta epr-1$  strains, only  $\Delta set-7$  strains, or in both  $\Delta epr-1$  and  $\Delta set-7$  strains by mRNA-seq. Significance was determined using a cutoff of  $\log_2(\text{mutant/wild type}) > 1$  for upregulated genes and  $< -1$  for downregulated genes with a  $P$  value  $< 0.05$ . The percentage of upregulated genes that are marked by H3K27 methylation for each gene set is indicated below the gene count. (B) RT-qPCR of H3K27-methylated genes that were replaced with antibiotic resistance genes (*NCU07152*, *NCU05173*), and used for initial selection of mutants, and H3K27-methylated genes that appeared upregulated in both  $\Delta epr-1$  and  $\Delta set-7$  strains by mRNA-seq (*NCU09640*, *NCU07624*, *NCU09178*, *NCU08570*, *NCU010038*, *NCU08097*). Each value was normalized to gene expression of *NCU02840* (an alternative housekeeping gene) and presented relative to wild type. (C) RT-qPCR for three genes (*NCU09990*, *NCU07241*, *NCU07246*) that appeared upregulated only in  $\Delta epr-1$  strains by mRNA-seq. Each value was normalized to gene expression of *act* (left) or *NCU02840* (right) and presented relative to wild type. (D) RT-qPCR for three genes (*NCU09604*, *NCU16720*, *NCU09274*) that appeared upregulated only in  $\Delta set-7$  strains by mRNA-seq. Each value was normalized to gene expression of *act* (left) or *NCU02840* (right) and presented relative to wild type. For all RT-qPCR data, filled bars represent the mean from biological triplicates and error bars show standard deviation (\*\*\*) for  $P < 0.001$ , \*\* for  $P < 0.01$ , \* for  $P < 0.05$ , and ns for not significant; all relative to wild type by two-tailed, unpaired t-test) (E) Linear growth rates measured by ‘race tubes’ are shown for two biological replicates of wild-type (N3752, N3753),  $\Delta set-7$  (N4718, N4730), and  $\Delta epr-1$  (N7497, N7491) strains. Horizontal lines represent the mean of three technical replicates (open circles). (F) Representative images of biological replicates graphed in Fig. 24C of wild-type,  $\Delta epr-1$  and  $\Delta set-7$  strains grown unfertilized at 25 °C for two weeks on modified Vogel’s with 0.1% sucrose. Overlaid black bar represents 2 cm.

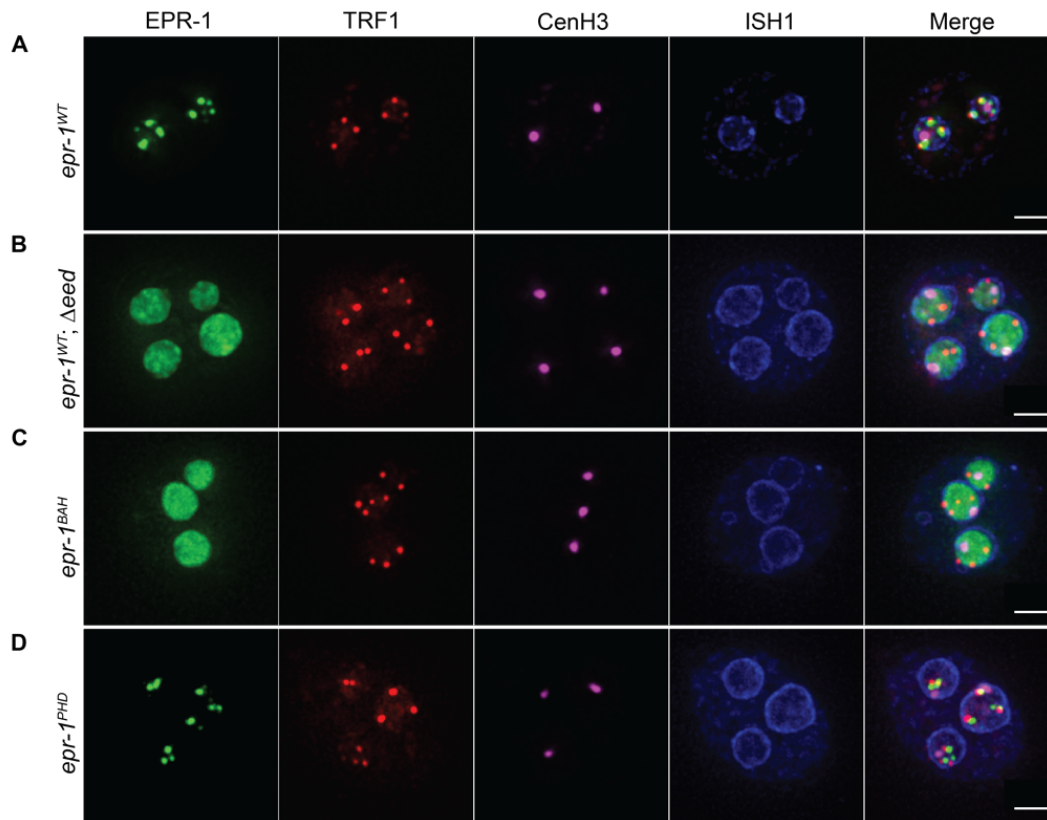




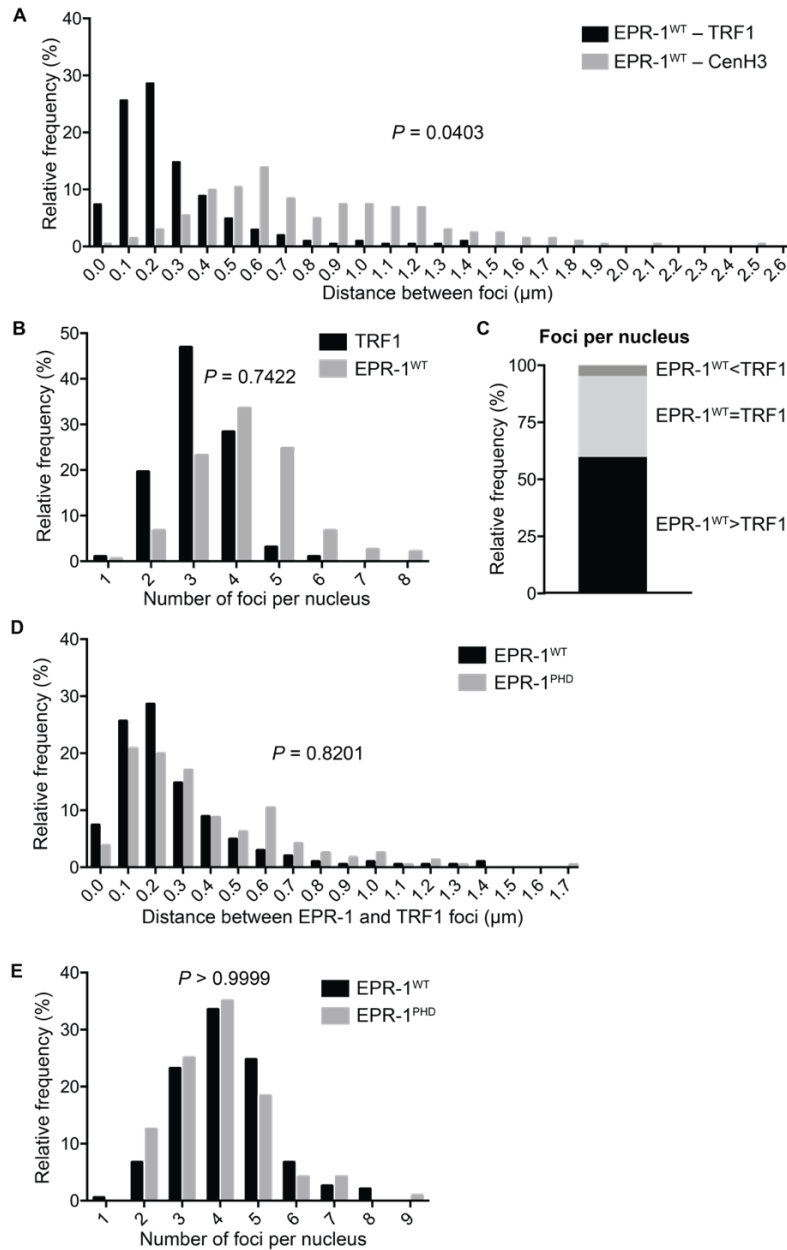
**Fig. 26. *epr-1* is not required for H3K27 methylation.** (A) Western blot showing H3K27me3 and total histone H3 (hH3) in the indicated strains. Biological replicates are shown. The same lysate was run on separate gels, and hH3 was used as a sample processing control. (B) ChIP-seq track showing average levels of H3K27me2/3 from two biological replicates of wild-type and  $\Delta epr-1$  strains on LG VI. Open circle indicates the middle of the centromere region. Gray bar represents 500 kb. Y-axis is 0-800 RPKM for wild type and 0-1200 RPKM for  $\Delta epr-1$ . (C) Scatter plot showing the correlation of H3K27me2/3 levels at all genes (black dots) in wild type and  $\Delta epr-1$  strains based on biological replicates of ChIP-seq data. Line of best fit displayed in red ( $R^2 = 0.9105$ ). Representative genes that gained (NCU02856) or lost (NCU08834) H3K27me2/3 in  $\Delta epr-1$  are indicated. (D) H3K27me2/3 ChIP-qPCR to confirm ChIP-seq data at six regions in wild-type,  $\Delta set-7$  and  $\Delta epr-1$  strains: *hH4* (negative control), Tel IL (unchanged H3K27me2/3 in  $\Delta epr-1$ ), NCU07152 (unchanged H3K27me2/3 in  $\Delta epr-1$ ), NCU05173 (unchanged H3K27me2/3 in  $\Delta epr-1$ ), NCU08834 (loss of H3K27me2/3 in  $\Delta epr-1$ ) and NCU02856 (gain of H3K27me2/3 in  $\Delta epr-1$ ). Filled bars represent the mean of biological triplicates and error bars show standard deviation (\*\* for  $P < 0.01$ , \* for  $P < 0.05$ , and ns for not significant; all relative to wild type by two-tailed, unpaired t-test).



**Fig. 27. H3K27 methylation levels in  $\Delta$ epr-1 are comparable to wild type.** (A) Western blot showing H3K27me3 and total histone H3 (hH3) in the indicated strains (additional replicate of experiment illustrated in Fig. 26A). The same lysate was run on separate gels, and hH3 was used as a sample processing control. (B) Quantification of the H3K27me3 band intensity averaged from 4 biological replicates. Each band intensity is normalized to the corresponding hH3 band and to the wild-type average. Filled bars represent the mean and error bars show standard deviation (a.u. signifies arbitrary units; \*\*\*\* for  $P < 0.0001$ , \*\* for  $P < 0.01$ , and ns for not significant; all relative to wild type by two-tailed, unpaired t-test). (C) Full blots with molecular weight markers in kilodaltons (kDa) for the cropped images in Fig. 26A and panel A (above) are shown. (D) Scatter plot showing the correlation of H3K27me2/3 levels at all genes (black dots) for the two wild-type ChIP-seq biological replicates. Line of best fit displayed in red ( $R^2 = 0.9272$ ). (E) Scatter plot showing the correlation of H3K27me2/3 levels at all genes (black dots) for the two  $\Delta$ epr-1 ChIP-seq biological replicates. Line of best fit displayed in red ( $R^2 = 0.8941$ ).

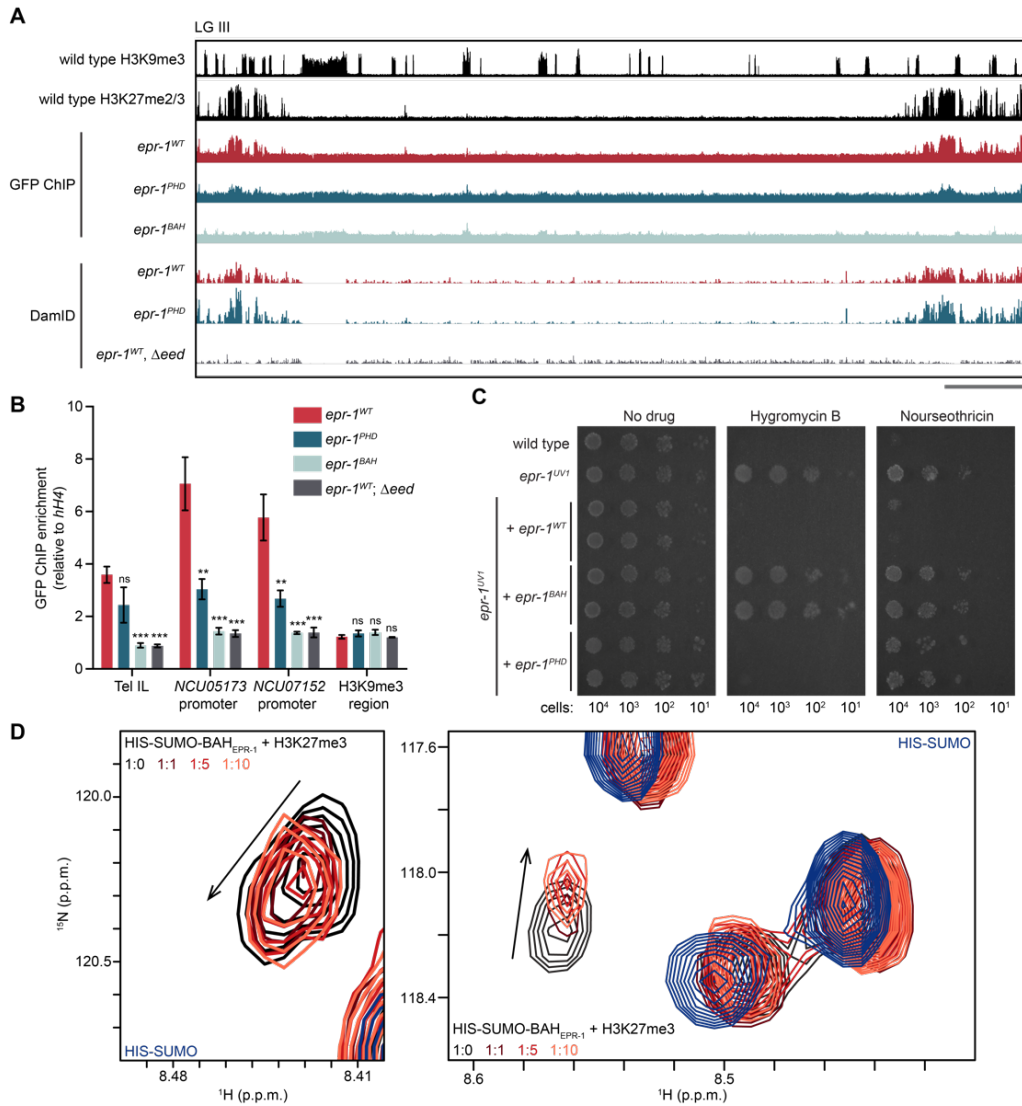


**Fig. 28. EPR-1 forms telomere-associated foci that are dependent on EED and the BAH domain of EPR-1.** Maximum intensity projection images of fluorescence microscopy Z-stacks showing EPR-1 (GFP-EPR-1, green) for *epr-1<sup>WT</sup>* (A), *epr-1<sup>WT</sup>; Δeed* (B), *epr-1<sup>BAH</sup>* (C), and *epr-1<sup>PHD</sup>* (D). Telomeres (TRF1-TagRFP-T, red), centromeres (CenH3-iRFP670, magenta), the nuclear membrane (ISH1-TagBFP2, blue), and merged images are shown for reference. Each image shows a single conidium with multiple nuclei. Overlaid white bar represents 2  $\mu$ m.



**Fig. 29. EPR-1<sup>WT</sup> and EPR-1<sup>PHD</sup> exhibit no difference in distance to telomeres or EPR-1 foci number.** (A) Histogram shows the relative frequency of distances between an EPR-1<sup>WT</sup> focus and the closest TRF1 focus (black bars) or CenH3 focus (gray bars) (n = 203,  $P = 0.0403$ , two-tailed Mann-Whitney test). (B) Histogram shows the relative frequency of nuclei with the indicated number of TRF1 or EPR-1<sup>WT</sup> foci (n = 194,  $P = 0.7422$ , Wilcoxon matched-pairs test). (C) Stacked bar graph comparing the number of TRF1 and EPR-1<sup>WT</sup> foci within a single nucleus (n = 194). (D) Histogram shows the relative frequency of distances between an EPR-1<sup>WT</sup> focus and the closest TRF1 focus (black bars; n = 203) or an EPR-1<sup>PHD</sup> focus and the closest TRF1 focus (gray bars; n = 241) ( $P = 0.8201$ , two-tailed Mann-Whitney test). (E) Histogram shows the relative frequency of EPR-1 foci per nucleus for EPR-1<sup>WT</sup> (n = 194) or EPR-1<sup>PHD</sup> (n = 120) ( $P > 0.9999$ , Wilcoxon matched-pairs test).

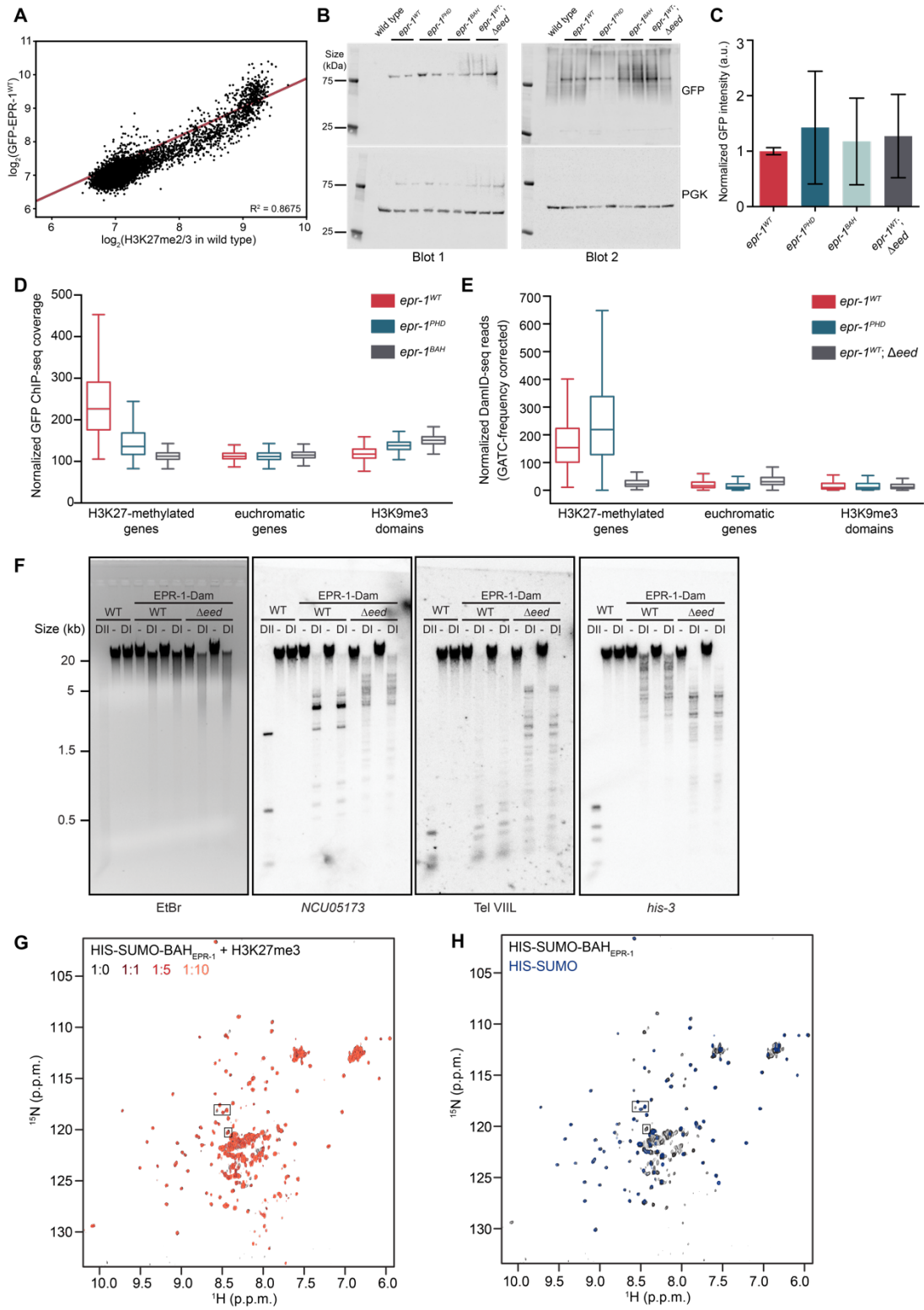


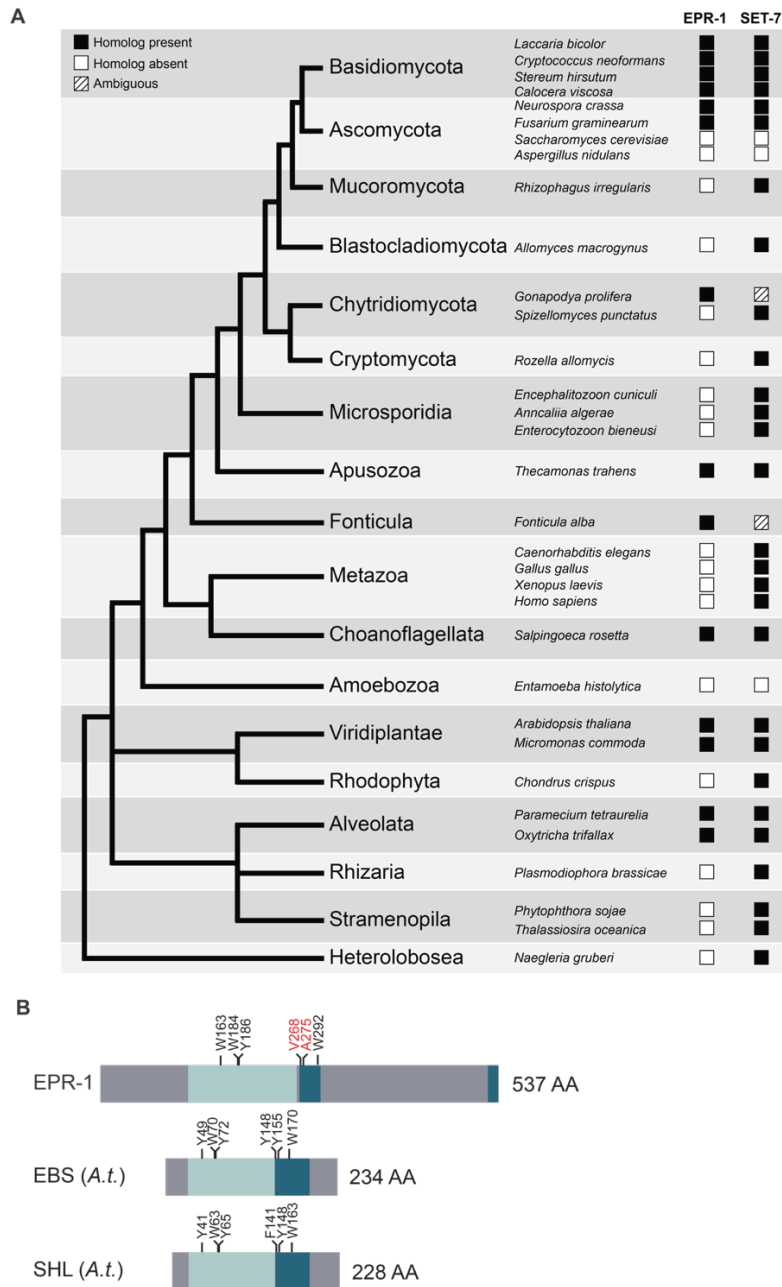


**Fig. 30. EPR-1 directly interacts with H3K27 methylated chromatin through its BAH domain.** (A) ChIP-seq and DamID-seq tracks showing average levels from two biological replicates for the indicated genotypes on LG III. Y-axis is 0-800 RPKM for H3K9me3 and H3K27me2/3 ChIP-seq, 0-500 RPKM for all GFP ChIP-seq, and 0-3500 RPKM for all DamID-seq. Gray bar represents 500 kb. (B) GFP ChIP-qPCR to validate GFP-EPR-1 ChIP-seq data at four genomic regions: Tel IL (H3K27-methylated), *NCU05173* promoter (H3K27-methylated), *NCU07152* promoter (H3K27-methylated), and H3K9me3 region (negative control, LG VI centromere). All data are normalized to a negative, euchromatic control, *hH4*. Filled bars represent the mean and error bars show standard deviation from three biological replicates (\*\*\*) for  $P < 0.001$ , \*\* for  $P < 0.01$ , and ns for not significant; all relative to wild type by two-tailed, unpaired t-test). (C) Serial dilution spot test silencing assay for the indicated strains. All strains harbor *P<sub>NCU05173</sub>::hph* and *P<sub>NCU07152</sub>::nat-1*. (D) Two regions of an overlay of <sup>1</sup>H-<sup>15</sup>N HSQC spectra of <sup>15</sup>N-labelled HIS-SUMO-BAH<sub>EPR-1</sub> fusion in the presence of increasing concentrations of H3K27me3 peptide. Spectra are color coded as indicated. <sup>15</sup>N-labelled HIS-SUMO is included for reference.



**Fig. 31 (see next page). EPR-1 associates with H3K27 methylation *in vivo* and *in vitro*.** (A) Scatter plot showing the correlation of levels of H3K27me<sub>2/3</sub> and GFP-EPR-1<sup>WT</sup> for all genes (black dots), as determined by ChIP-seq. Line of best fit displayed in red ( $R^2=0.8675$ ). (B) Western blot shows GFP-EPR-1 expression in the indicated strains. Phosphoglycerate kinase (PGK) is used as a loading control. Each genotype (except wild type, negative control) was run in biological duplicate and repeated. (C) Quantification of the GFP band intensity averaged over 4 biological replicates. The intensities are relative to the corresponding PGK band and normalized to the wild-type average from the same blot. Filled bars represent the mean and error bars show standard deviation (a.u. signifies arbitrary units). (D) Box and whisker plot of normalized GFP-EPR-1 ChIP-seq coverage from *epr-1*<sup>WT</sup>, *epr-1*<sup>PHD</sup> and *epr-1*<sup>BAH</sup> strains is shown for the indicated regions of the genome. Box represents interquartile range, horizontal line is median, and whiskers represent minimum and maximum values. (E) Box and whisker plot of normalized EPR-1 DamID-seq coverage from *epr-1*<sup>WT</sup>, *epr-1*<sup>PHD</sup> and *epr-1*<sup>WT</sup>;  $\Delta$ *eed* strains is shown for the indicated regions of the genome. Reads have been corrected for the frequency of GATC sites. Box represents interquartile range, horizontal line is median, and whiskers represent minimum and maximum values. (F) DamID Southern blot of genomic DNA from the indicated strains digested with DpnI (DI), DpnII (DII) or left undigested (-). DpnII, which digests GATC sites without methylated adenines, reveals pattern of complete digestion in wild type. DpnI, which fails to digest GATC sites bearing adenine methylation, reveals the extent of methylation by Dam at probed regions (*NCU05173* and Tel VIII, H3K27-methylated; *his-3*, euchromatin). Ethidium bromide (EtBr) shows total DNA. Biological replicates are shown. (G) An overlay of <sup>1</sup>H-<sup>15</sup>N HSQC spectra of <sup>15</sup>N-labelled HIS-SUMO-BAH<sub>EPR-1</sub> fusion in the presence of increasing concentrations of H3K27me<sub>3</sub> peptide. (H) An overlay of <sup>15</sup>N-labelled HIS-SUMO-BAH<sub>EPR-1</sub> fusion with the <sup>15</sup>N-labelled HIS-SUMO alone. Boxed areas are shown enlarged in Fig. 30D. Spectra are color coded as indicated





**Fig. 32. EPR-1 homologs are present in species beyond fungi.** (A) The presence and absence of EPR-1 and SET-7 protein homologs across major species divisions is depicted in a representative tree of eukaryotes. The leaves of the tree are labelled with the names of the divisions. Representative species are featured and their associated squares in the EPR-1 and SET-7 columns indicate the presence or absence of homologs, as indicated. (B) Protein domain structure of EPR-1, as well as EBS and SHL from *Arabidopsis thaliana* (*A.t.*). BAH domains are indicated by light blue, the PHD fingers by dark blue, and regions with no known domains are gray. Aromatic amino acid residues involved in methylated histone recognition in the BAH domain and PHD finger are indicated above the protein structure diagram (black text). Red text above the PHD finger in EPR-1 highlights the absence of aromatic residues at these amino acid positions.

## APPENDIX B

### STRAIN TABLES

All strains are *N. crassa*.

#### Chapter II

Strain	Genotype
N2931	<i>mat a; Δmus-52::bar</i>
N3752 (FGSC 2489)	<i>mat A; Oak Ridge</i>
N4922	<i>mat a; Δhpo::hph</i>
N6227	<i>mat a; Δnst-3::nat-1</i>
N6355	<i>mat a; Δnst-4::hph</i>
N6420	<i>mat A; Δrtt109::hph</i>
N6421	<i>mat a; Δnst-3::nat-1; Δrtt109::hph</i>
N6445	<i>mat a; H3K56R::nat-1</i>
N6539	<i>mat a; Δnst-1::hph</i>
N6543	<i>mat a; Δnst-5::hph</i>
N6585	<i>Δhpo::hph; H3K56R::nat-1</i>
N6734	<i>Δnst-3::nat-1; H3K56R::nat-1</i>

#### Chapter III

Strain	Genotype
N51 (FGSC 2225)	<i>mat A; Mauriceville</i>
N625	<i>mat a; his-3</i>
N2834	<i>mat A; his-3; Δmus-52::hph<sup>+</sup></i>
N2931	<i>mat a; Δmus-52::bar<sup>+</sup></i>
N3752 (FGSC 2489)	<i>mat A; Oak Ridge</i>
N4730	<i>mat A; Δset-7::bar<sup>+</sup></i>
N4933	<i>mat a; Δ47.4 kb::hph<sup>+</sup></i>
N5100 (FGSC 1614)	<i>mat a; In(IL-&gt;IR)AR16</i>
N5101 (FGSC 2100)	<i>mat A; T(IV-&gt;VI)ALS159</i>
N5102 (FGSC 3670)	<i>mat A; T(VI-&gt;III)OY329</i>
N5547	<i>mat a; Δmus-52::bar<sup>+</sup>; Δ47.4 kb::hph<sup>+</sup></i>
N5683	<i>mat A; his-3<sup>+</sup>::1; Δ47.4 kb::hph<sup>+</sup></i>
N5684	<i>mat A; his-3<sup>+</sup>::2; Δ47.4 kb::hph<sup>+</sup></i>
N5685	<i>mat A; his-3<sup>+</sup>::3; Δ47.4 kb::hph<sup>+</sup></i>
N5686	<i>mat A; his-3<sup>+</sup>::4; Δ47.4 kb::hph<sup>+</sup></i>
N5687	<i>mat A; his-3<sup>+</sup>::5; Δ47.4 kb::hph<sup>+</sup></i>
N5688	<i>mat A; his-3<sup>+</sup>::6; Δ47.4 kb::hph<sup>+</sup></i>
N5689	<i>mat A; his-3<sup>+</sup>::7 Δ47.4 kb::hph<sup>+</sup></i>
N5690	<i>mat A; his-3<sup>+</sup>::8; Δ47.4 kb::hph<sup>++</sup></i>

N5695	<i>mat a</i> ; $\Delta csr-1::1$ ; $\Delta 47.4$ kb:: <i>hph</i> <sup>+</sup>
N5696	<i>mat a</i> ; $\Delta csr-1::2$ ; $\Delta 47.4$ kb:: <i>hph</i> <sup>+</sup>
N5697	<i>mat a</i> ; $\Delta csr-1::3$ ; $\Delta 47.4$ kb:: <i>hph</i> <sup>+</sup>
N5698	<i>mat A</i> ; $\Delta csr-1::4$ ; $\Delta 47.4$ kb:: <i>hph</i> <sup>+</sup>
N5699	<i>mat A</i> ; $\Delta csr-1::5$ ; $\Delta 47.4$ kb:: <i>hph</i> <sup>+</sup>
N5700	<i>mat a</i> ; $\Delta csr-1::6$ ; $\Delta 47.4$ kb:: <i>hph</i> <sup>+</sup>
N5701	<i>mat a</i> ; $\Delta csr-1::7$ ; $\Delta 47.4$ kb:: <i>hph</i> <sup>+</sup>
N5702	<i>mat A</i> ; $\Delta csr-1::8$ ; $\Delta 47.4$ kb:: <i>hph</i> <sup>+</sup>
N5739	<i>mat A</i> ; <i>his-3</i> ; $\Delta 47.4$ kb:: <i>hph</i> <sup>+</sup>
N5857 (FGSC 1483)	<i>mat A</i> ; T(II->V)NM149
N5858 (FGSC 3668)	<i>mat A</i> ; T(II->IV)OY337; <i>al-2</i>
N5859 (FGSC 7294)	<i>mat A</i> ; T(IVR->VL)UK2-32
N5862 (FGSC 4641)	<i>mat A</i> ; T(VIL->IR)OY350
N5863 (FGSC 3635)	<i>mat A</i> ; T(VI->III)OY320
N5866 (FGSC 6869)	<i>mat A</i> ; <i>inl</i> ; T(VR )UK3-41
N6089	<i>mat a</i> ; Dp(VI->III)OY329
N6093 (FGSC 3671)	<i>mat a</i> ; T(VI->III)OY329
N6228	<i>mat A</i> ; $\Delta tert::nat-1$ <sup>+</sup>
N6381	<i>mat a</i> ; $\Delta mus-52::bar$ <sup>+</sup> ; $\Delta csr-1::(TTAGGG)$ <sub>8</sub>
N6383	<i>mat a</i> ; $\Delta mus-52::bar$ <sup>+</sup> ; $\Delta csr-1::(TTAGGG)$ <sub>17</sub>
N6984	<i>mat a</i> ; $\Delta mus-52::bar$ <sup>+</sup> ; $\Delta csr-1::(TTAGGG)$ <sub>8</sub>
N6985	<i>mat a</i> ; $\Delta mus-52::bar$ <sup>+</sup> ; $\Delta csr-1::(TTAGGG)$ <sub>8</sub>
N6986	<i>mat a</i> ; $\Delta mus-52::bar$ <sup>+</sup> ; $\Delta csr-1::(TTAGGG)$ <sub>17</sub>
N6987	<i>mat a</i> ; $\Delta mus-52::bar$ <sup>+</sup> ; $\Delta csr-1::(TTAGGG)$ <sub>17</sub>
N7515	<i>mat A</i> ; $\Delta mus-52::hph$ <sup>+</sup> ; <i>his-3</i> <sup>+</sup> ::(TTAGGG) <sub>23</sub> (heterokaryon)
N7516	<i>mat A</i> ; $\Delta mus-52::hph$ <sup>+</sup> ; <i>his-3</i> <sup>+</sup> ::(TTAGGG) <sub>22</sub> (heterokaryon)
N7517	<i>mat A</i> ; $\Delta mus-52::hph$ <sup>+</sup> ; <i>his-3</i> <sup>+</sup> ::(TTAGGG) <sub>21</sub> (heterokaryon)

#### Chapter IV

Strain	Genotype
N51 (FGSC 2225)	<i>mat A</i> ; Mauriceville
N2930	<i>mat A</i> ; <i>his-3</i> ; $\Delta mus-52::bar$
N3752 (FGSC 2489)	<i>mat A</i> ; Oak Ridge
N3753 (FGSC 4200)	<i>mat a</i> ; Oak Ridge
N3756	<i>mat A</i> ; <i>Sad-1</i> ; <i>his-3</i>
N4666	<i>mat a</i> ; <i>suz12::C-Gly::3xFLAG::hph</i>
N4718	<i>mat a</i> ; $\Delta set-7::hph$
N4730	<i>mat A</i> ; $\Delta set-7::bar$
N4879	<i>mat a</i> ; <i>hph::P<sub>qa-2</sub>::3xFLAG::N-Gly::eed</i>
N6232	<i>mat a</i> ; <i>pNCU05173::hph</i>
N6279	<i>mat a</i> ; <i>pNCU05173::hph</i> ; <i>pNCU07152::nat-1</i> ; <i>his-3</i>
N6762	<i>mat a</i> ; <i>pNCU05173::hph</i> ; <i>pNCU07152::nat-1</i> ; <i>his-3</i> ; <i>NCU04278<sup>UVI</sup></i>

N7600	<i>mat A; pNCU05173::hph; pNCU07152::nat-1; Δset-7::bar</i>
N7617	<i>mat a; ΔNCU04278::nat-1</i>
N7667	<i>mat a; pNCU05173::hph; ΔNCU04278::nat-1</i>
N7680	<i>mat a; ΔNCU04278::nat-1</i>
N7742	<i>mat A; pNCU05173::hph; pNCU07152::nat-1; his-3; NCU04278<sup>UV1</sup></i>
N7807	<i>mat a; pNCU05173::hph; pNCU07152::nat-1; his-3<sup>+</sup>::P<sub>ccg-1</sub>::NCU04278<sup>WT</sup>::C-Gly::3xFLAG; NCU04278<sup>UV1</sup></i>
N7950	<i>mat a; pNCU05173::hph; pNCU07152::nat-1; his-3; NCU04278<sup>UV2</sup></i>
N7999	<i>mat a; ΔNCU04278::nat-1</i>
N8002	<i>mat A; ΔNCU04278::nat-1</i>
N8003	<i>mat a; Δnpf::hph</i>
N8004	<i>mat a; Δnpf::hph</i>

## Chapter V

Strain	Genotype
N51 (FGSC 2225)	<i>mat A; Mauriceville</i>
N623	<i>mat A; his-3</i>
N625	<i>mat a; his-3</i>
N2718	<i>mat a; Δmus-52::hph</i>
N3752 (FGSC 2489)	<i>mat A; Oak Ridge</i>
N3753 (FGSC 4200)	<i>mat a; Oak Ridge</i>
N3756	<i>mat A; Sad-1; his-3</i>
N4718	<i>mat a; Δset-7::hph</i>
N4730	<i>mat A; Δset-7::bar</i>
N4840	<i>mat A; Δset-7::bar; Δmus-52::bar</i>
N5808	<i>mat A; pNCU07152::nat-1</i>
N6223	<i>mat A; ish-1::TagBFP2::hph; hpo::TagRFP-T::hph; cenH3::iRFP670::hph</i>
N6233	<i>mat a; pNCU05173::hph; pNCU07152::nat-1</i>
N6234	<i>mat A; pNCU05173::hph; Δset-7::bar</i>
N6279	<i>mat a; pNCU05173::hph; pNCU07152::nat-1; his-3</i>
N6761	<i>mat a; pNCU05173::hph; pNCU07152::nat-1; his-3; epr-1<sup>UV3</sup></i>
N7451	<i>mat a; pNCU05173::hph; pNCU07152::nat-1; his-3; epr-1<sup>UV1</sup></i>
N7552	<i>mat A; ish-1::TagBFP2::hph; cenH3::iRFP670::hph; his-3</i>
N7567	<i>mat a; Δepr-1::nat-1; trf-1::TagRFP-T::hph; his-3</i>
N7596	<i>mat a; pNCU05173::hph; pNCU07152::nat-1; his-3; epr-1<sup>UV4</sup></i>
N7549	<i>mat A; pNCU05173::hph; pNCU07152::nat-1; his-3; epr-1<sup>UV1</sup></i>
N7479	<i>mat A; Δepr-1::hph</i>
N7491	<i>mat a; Δepr-1::hph</i>
N7525	<i>mat A; epr-1<sup>WT</sup>::10xGly::Dam::nat-1</i>
N7526	<i>mat a; epr-1<sup>WT</sup>::10xGly::Dam::nat-1</i>

N7537	<i>mat a; trf-1::TagRFP-T::hph; his-3</i>
N7538	<i>mat a; epr-1<sup>WT</sup>::10xGly::Dam::nat-1; Δeed::hph</i>
N7539	<i>mat a; epr-1<sup>WT</sup>::10xGly::Dam::nat-1; Δeed::hph</i>
N7600	<i>mat A; pNCU05173::hph; pNCU07152::nat-1; Δset-7::bar</i>
N7900	<i>mat a; epr-1<sup>W292A</sup>::10xGly::Dam::nat-1</i>
N7901	<i>mat a; epr-1<sup>W292A</sup>::10xGly::Dam::nat-1</i>
N7576	<i>mat A; Δepr-1::nat-1; pNCU05173::hph</i>
N7689	<i>pNCU05173::hph; pNCU07152::nat-1; his-3<sup>+</sup>::pCCG::N-GFP::epr-1<sup>WT</sup>; epr-1<sup>UV1</sup></i>
N7690	<i>pNCU05173::hph; pNCU07152::nat-1; his-3<sup>+</sup>::pCCG::N-GFP::epr-1<sup>WT</sup>; epr-1<sup>UV1</sup></i>
N7722	<i>mat a; Δepr-1::nat-1; trf-1::TagRFP-T::hph; ish-1::TagBFP2::hph; cenH3::iRFP670::hph; his-3<sup>+</sup>::pCCG::N-GFP::epr-1<sup>WT</sup></i>
N7743	<i>mat a; Δepr-1::nat-1; trf-1::TagRFP-T::hph; ish-1::TagBFP2::hph; cenH3::iRFP670::hph; his-3<sup>+</sup>::pCCG::N-GFP::epr-1<sup>WT</sup>; Δeed::hph</i>
N7752	<i>mat a; pNCU05173::hph; pNCU07152::nat-1; his-3<sup>+</sup>::pCCG::N-GFP::epr-1<sup>W184A</sup>; epr-1<sup>UV1</sup></i>
N7753	<i>mat A; pNCU05173::hph; pNCU07152::nat-1; his-3<sup>+</sup>::pCCG::N-GFP::epr-1<sup>W184A</sup>; epr-1<sup>UV1</sup></i>
N7754	<i>mat A; pNCU05173::hph; pNCU07152::nat-1; his-3<sup>+</sup>::pCCG::N-GFP::epr-1<sup>W292A</sup>; epr-1<sup>UV1</sup></i>
N7755	<i>mat a; pNCU05173::hph; pNCU07152::nat-1; his-3<sup>+</sup>::pCCG::N-GFP::epr-1<sup>W292A</sup>; epr-1<sup>UV1</sup></i>
N7756	<i>mat a; Δepr-1::nat-1; trf-1::TagRFP-T::hph; ish-1::TagBFP2::hph; cenH3::iRFP670::hph; his-3<sup>+</sup>::pCCG::N-GFP::epr-1<sup>W184A</sup></i>
N7922	<i>Δepr-1::nat-1; trf-1::TagRFP-T::hph; ish-1::TagBFP2::hph; cenH3::iRFP670::hph; his-3<sup>+</sup>::pCCG::N-GFP::epr-1<sup>W292A</sup></i>
N8039	<i>mat a; pNCU05173::hph; pNCU07152::nat-1; his-3; epr-1<sup>UV2</sup></i>

## APPENDIX C

### PRIMER TABLES

#### Chapter II

Primer	Description	Sequence (5'→3')
3498	hH3_5FP	CCCAATAAGTTCAAACCG
4883	<i>nat-1</i> RP	AACCCCATCCGCCGGTACGCG
4884	<i>nat-1</i> FP	TCCTTCACCACCGACACCGTCTTCC
5303	hH3K56R_FP	GAGCAGCTCAGTGGACCTCTGGTATCTCCTGAT
5304	hH3K56R_RP	ATCAGGAGATACCAGAGGTCCACTGAGCTGCTC
5305	hH3K56Q_FP	AGCAGCTCAGTGGACTGCTGGTATCTCCTGATC
5306	hH3K56Q_RP	GATCAGGAGATACCAGCAGTCCACTGAGCTGCT
5307	hH3_5RP	GTCTGCTTGGTCCTGGCCATTGTGATGGTTTATGT GTGTTGATAAC
5333	hH3_3UTR_FP	GGTGAGAGAAATTAAGCGACTCTTCGATATGGAG TAG
5334	hH3_3UTR_RP	CCTCCGCCTCCGCCTCCGCCGCCTCCGCCGCGGGT CACAGTGATGTTTATG
5335	hH3_3FP	GAGCTCGGTACCAAGCTTGATGCATAGCTGAGAC GCGTCCGTTCGC
5336	hH3_3RP	GGGTTACCTTGGTTCACATGC
5337	hH3_codon_RP	ATCGAAGAGTCGCTTAATTTCTCTCACCACGCAGT CTTCG

#### Chapter III

Primer	Description	Sequence (5'→3')
3325	Construction of <i>his-3</i> <sup>+</sup> ::1 (N5683)	CGGGATCCCCGGATCGAACGGCG GATGG (BamHI)
3326	Construction of <i>his-3</i> <sup>+</sup> ::1 (N5683)	GCTCTAGATGAGCTCCTTCCCC ACGGT (XbaI)
3327	Construction of <i>his-3</i> <sup>+</sup> ::2 (N5684)	CGGGATCCCGTGGGTTCCAGTGC GTCCC (BamHI)
3328	Construction of <i>his-3</i> <sup>+</sup> ::2 (N5684)	GCTCTAGATGGCAAGCGCCGACA TGTGA (XbaI)
3329	Construction of <i>his-3</i> <sup>+</sup> ::3 (N5685)	GGGCCAGGGAACAGGCGCAGT GCAG (ApaI)
3330	Construction of <i>his-3</i> <sup>+</sup> ::3 (N5685)	GACTAGTTGGCGCGCTTGACCAG CAA (SpeI)
3331	Construction of <i>his-3</i> <sup>+</sup> ::4 (N5686)	GACTAGTCCCGGAACGGCCACT CCAT (SpeI)
3332	Construction of <i>his-3</i> <sup>+</sup> ::4 (N5686)	GCTCTAGAGTGCTGCCAAGGCC GACAT (XbaI)



3297	Construction of <i>his-3</i> <sup>+</sup> ::5 (N5687)	GGAATTCCAAAGGATCGCGCCCG GAGG (EcoRI)
3333	Construction of <i>his-3</i> <sup>+</sup> ::5 (N5687)	CGGGATCCACGACGGGAAGACG AGGGGT (BamHI)
3334	Construction of <i>his-3</i> <sup>+</sup> ::6 (N5688)	GGGCCCGGGAGTGTGGCGCTGTG GAC (ApaI)
3335	Construction of <i>his-3</i> <sup>+</sup> ::6 (N5688)	GACTAGTAGCAGCATCTCGGGTC CGGT (SpeI)
3301	Construction of <i>his-3</i> <sup>+</sup> ::7 (N5689)	GGAATTCGGCATGCCAGGGATT GGGG (EcoRI)
3336	Construction of <i>his-3</i> <sup>+</sup> ::7 (N5689)	GCTCTAGAGACGGAACGGTGAC AGGCGG (XbaI)
3303	Construction of <i>his-3</i> <sup>+</sup> ::8 (N5690)	GGAATTCTGGTTCGCGGACGAGG GCTA (EcoRI)
3337	Construction of <i>his-3</i> <sup>+</sup> ::8 (N5690)	CGGGATCCGGCTATTGCTGCCAG CCGGT (BamHI)
4654	3' <i>csr-1</i> RP	AACACCTCCGTCGCCATAAACTC C
4655	5' <i>csr-1</i> FP	GGCCCCTGGTTTACTGAGGGC
4721	5' <i>csr-1</i> RP	TGCAGCCATTGACGACATTGC
4722	3' <i>csr-1</i> FP	TGGATTTCCTGCGCTGCACAC
4723	Construction of $\Delta$ <i>csr-1</i> ::1 (N5695)	GCAATGTCGTCAATGGCTGCACC GGATCGAACGGCGGATGG
4724	Construction of $\Delta$ <i>csr-1</i> ::1 (N5695)	GTGTGCAGCGCAGGAAATCCATG AGCTCCTTCCCCACGGT
4725	Construction of $\Delta$ <i>csr-1</i> ::2 (N5696)	GCAATGTCGTCAATGGCTGCACG TGGTTCCAGTGCGTCC
4726	Construction of $\Delta$ <i>csr-1</i> ::2 (N5696)	GTGTGCAGCGCAGGAAATCCATG GCAAGCGCCGACATGTGA
4727	Construction of $\Delta$ <i>csr-1</i> ::3 (N5697)	GCAATGTCGTCAATGGCTGCAAG GGAACAGGCGCAGTGCAG
4728	Construction of $\Delta$ <i>csr-1</i> ::3 (N5697)	GTGTGCAGCGCAGGAAATCCATG GCGCGCTTGACCAGCAA
4729	Construction of $\Delta$ <i>csr-1</i> ::4 (N5698)	GCAATGTCGTCAATGGCTGCATC CCGGAACGGCCACTCCAT
4730	Construction of $\Delta$ <i>csr-1</i> ::4 (N5698)	GTGTGCAGCGCAGGAAATCCAGT GCTGCCAAGGCCCGACAT
4731	Construction of $\Delta$ <i>csr-1</i> ::5 (N5699)	GCAATGTCGTCAATGGCTGCACA AAGGATCGCGCCCGGAGG
4732	Construction of $\Delta$ <i>csr-1</i> ::5 (N5699)	GTGTGCAGCGCAGGAAATCCAAC GACGGGAAGACGAGGGGT
4733	Construction of $\Delta$ <i>csr-1</i> ::6 (N5700)	GCAATGTCGTCAATGGCTGCAGG GAGTGTGGCGCTGTGGAC
4734	Construction of $\Delta$ <i>csr-1</i> ::6 (N5700)	GTGTGCAGCGCAGGAAATCCAA GCAGCATCTCGGGTCCGGT

4735	Construction of $\Delta csr-1::7$ (N5701)	GCAATGTCGTCAATGGCTGCAGG CATGCCAGGGATTGGGG
4736	Construction of $\Delta csr-1::7$ (N5701)	GTGTGCAGCGCAGGAAATCCAG ACGGAACGGTGACAGGCGG
4737	Construction of $\Delta csr-1::8$ (N5702)	GCAATGTCGTCAATGGCTGCATG GTTCGCGGACGAGGGCTA
4738	Construction of $\Delta csr-1::8$ (N5702)	GTGTGCAGCGCAGGAAATCCAG GCTATTGCTGCCAGCCGGT
4784	qChIP <i>his-3</i> & <i>csr-1</i> targeting P1Fwd	AATGCAAGGTCCCGAACACT
4785	qChIP <i>his-3</i> & <i>csr-1</i> targeting P1Rev	TGGCTGTCGCAATTACCAGT
4786	qChIP <i>his-3</i> & <i>csr-1</i> targeting P2Fwd	GACCAAGCATGCGTTAGCTG
4787	qChIP <i>his-3</i> & <i>csr-1</i> targeting P2Rev	ACCCAAGGTGGGTGTGTTTT
4788	qChIP <i>his-3</i> & <i>csr-1</i> targeting P3Fwd	CCGTTTGAGCTGGTCTTCCT
4789	qChIP <i>his-3</i> & <i>csr-1</i> targeting P3Rev	TGACGGATGCTCTTTGTCCC
4790	qChIP <i>his-3</i> & <i>csr-1</i> targeting P4Fwd	CAACCAGCTTGACGGCTTTC
4791	qChIP <i>his-3</i> & <i>csr-1</i> targeting P4Rev	TACCGTAGGTGCCCTGTGTA
3317	qChIP <i>his-3</i> & <i>csr-1</i> targeting P5Fwd and deletion analysis	CCCCTTCCTGCCGTGGGAGA
3562	qChIP <i>his-3</i> & <i>csr-1</i> targeting P5Rev and deletion analysis	TCAGCAGGCATAGTCAAGACTGG T
4782	qChIP <i>his-3</i> & <i>csr-1</i> targeting P6Fwd	CCCGCTTCAGCAACCAAGTT
4783	qChIP <i>his-3</i> & <i>csr-1</i> targeting P6Rev	AACTTTAGCCCGCGTTACGG
3319	qChIP <i>his-3</i> & <i>csr-1</i> targeting P7Fwd and deletion analysis	TGGGTGCATGGAGTACCTTCCCC
3563	qChIP <i>his-3</i> & <i>csr-1</i> targeting P7Rev and deletion analysis	TGCACTATCCTTTTCAGGGGCTT GT
3902	qChIP <i>his-3</i> & <i>csr-1</i> targeting P8Fwd	GACCTACACGGCCCGGGGAA
3903	qChIP <i>his-3</i> & <i>csr-1</i> targeting P8Rev	ACCGACGAGACTTGACTGCCCA
3986	qChIP <i>his-3</i> & <i>csr-1</i> targeting P9Fwd	GTGGCGGCGTGAACGGTCAT
3987	qChIP <i>his-3</i> & <i>csr-1</i> targeting P9Rev	AGTCAAGCCTCGCGATCGTGA
3321	qChIP <i>his-3</i> & <i>csr-1</i> targeting P10Fwd	TGAACAGGTGACGGCGGGAGT

3564	qChIP <i>his-3</i> & <i>csr-1</i> targeting P10Rev	CGGGTCCGGAGTCCATCACCA
3976	qChIP <i>his-3</i> & <i>csr-1</i> targeting P11Fwd	AGTGGTCCAGAGTGGGATCGGT
3977	qChIP <i>his-3</i> & <i>csr-1</i> targeting P11Rev	ACCGCCAATATGGCATCGCCC
3974	qChIP <i>his-3</i> & <i>csr-1</i> targeting P12Fwd	TGACGGCGCGCAGATTGGAG
3975	qChIP <i>his-3</i> & <i>csr-1</i> targeting P12Rev	TCGCTTCCCTTCTCCCACCATCC
3978	qChIP <i>his-3</i> & <i>csr-1</i> targeting P13Fwd	TTGACGGCGCGCAGATTGGAG
3979	qChIP <i>his-3</i> & <i>csr-1</i> targeting P13Rev	CCACCATCCTTCCCTCTGCCACA
3992	qChIP <i>his-3</i> & <i>csr-1</i> targeting P14Fwd	CATCGCAGCTCAACCGCAGA
3993	qChIP <i>his-3</i> & <i>csr-1</i> targeting P14Rev	GCCAGCCGGTGTCAAGACAGA
3904	qChIP <i>his-3</i> & <i>csr-1</i> targeting P15Fwd	AACAAAGACGCTCTTCTGGTGGC C
3905	qChIP <i>his-3</i> & <i>csr-1</i> targeting P15Rev	ACTACCAAACCTGCCGACGGCT
3943	Construction of $\Delta$ 47.4 kb:: <i>hph</i> <sup>+</sup> (N4933)	GTAACGCCAGGGTTTTCCAGTC ACGACGTGTGGCTTGCAGGCACG CAA
3944	Construction of $\Delta$ 47.4 kb:: <i>hph</i> <sup>+</sup> (N4933)	ACCGGGATCCACTTAACGTTACT GAAATCAGTCCGAGTGGGCCTGC CTC
3945	Construction of $\Delta$ 47.4 kb:: <i>hph</i> <sup>+</sup> (N4933)	GCTCCTTCAATATCATCTTCTGTG GACGGACCACCACCAGCGTGG AAAG
3946	Construction of $\Delta$ 47.4 kb:: <i>hph</i> <sup>+</sup> (N4933)	GCGGATAACAATTTACACAGGA AACAGCTTGCCGCCGGCTGAGAA ACC
4960	ALS159 (N5101) qPCR	TTGGGATGATTTGGGACGGG
4961	ALS159 (N5101) qPCR	TCCAAGCTGACAGTTCCAC
4961	ALS159 (N5101) qPCR	TGCATGCTCTCCCCCTTTG
4963	ALS159 (N5101) qPCR	TCTGAGGGATGTGCCAAACC
4974	NM149 (N5857) qPCR	CGCCATTTCTACCCCGATGA
4975	NM149 (N5857) qPCR	TGCCAAGCCATCTTTTTGCC
4976	NM149 (N5857) qPCR	CTACGGGTTGCTGCCAAGTA
4977	NM149 (N5857) qPCR	CCTCAGAGAATCGGGGCATC
4978	NM149 (N5857) qPCR	GGGCTCAGTCACTTGCTACA
4979	NM149 (N5857) qPCR	GATATACCCGCACCAGCACA
4984	OY337 (N5858) qPCR	CTTCGCCTCTCACTCCGATG

4985	OY337 (N5858) qPCR	GGCAGCTAGCAATCGGTTTT
4986	OY337 (N5858) qPCR	GAGCCTGTCCAAGACGACAA
4987	OY337 (N5858) qPCR	CGGTGACGGTAGTGTGTAGG
4990	OY337 (N5858) qPCR	GTTACTGGCGGGAAATGGGA
4991	OY337 (N5858) qPCR	TTGGGACCAGGTTTGTCCAC
4996	OY350 (N5862) qPCR	GCCTTGGACCCTCGAATGAA
4997	OY350 (N5862) qPCR	TGGGAAAACGTGGGGGAAAA
4998	OY350 (N5862) qPCR	TGGGTGAGGTCTTTGGAGGA
4999	OY350 (N5862) qPCR	AAGAGTTCCTGAACGTCGCC
5002	OY350 (N5862) qPCR	AGGGTTGCTGGTAATCCGTG
5003	OY350 (N5862) qPCR	CAAGGCTTGGGGAAAGGGAA
3317	UK3-41 (N5866) qPCR	CCCCTTCCTGCCGTGGGAGA
3562	UK3-41 (N5866) qPCR	TCAGCAGGCATAGTCAAGACTGG T
3954	UK3-41 (N5866) qPCR	GTAGCTAGCGGGTGCTGCCG
3955	UK3-41 (N5866) qPCR	AGGCGCCAGGAAGAGTATAGCC C
5012	OY329 (N5102) qPCR	AAATCCACTCATCCTCGGCG
5013	OY329 (N5102) qPCR	CTCGGATCACCGTCAACAGG
5020	OY329 (N5102) qPCR	ATGGTAACGTGGACAGGTGC
5021	OY329 (N5102) qPCR	TTGAACGCCGTAGAGGGATG
5028	UK2-32 (N5859) qPCR	AGGAAGTACGCCTTGCAGTC
5029	UK2-32 (N5859) qPCR	CCTGTATAATGGCGGTCCCC
5034	UK2-32 (N5859) qPCR	TCGAACCATGTGAGCTGCTT
5035	UK2-32 (N5859) qPCR	GAGAACGCCGAATCGCTCTA
5036	AR16 (N5100) qPCR	GAGAACGCCGAATCGCTCTA
5037	AR16 (N5100) qPCR	GCCCCCTTTTTGTGCTTAGC
5040	AR16 (N5100) qPCR	CGTCAACGGTAGCTGGAAGA
5041	AR16 (N5100) qPCR	CCTCTTTGTGTCGAAGCCCA
5042	AR16 (N5100) qPCR	TGCCAAAGCACAACAAGCTG
5043	AR16 (N5100) qPCR	TTGATACCACGGGCTTCGAC
5046	AR16 (N5100) qPCR	GGATCATCGGTAGGTTGGGT
5047	AR16 (N5100) qPCR	TCAGATCCAGCTAGTTTCGCC
5316	Telomere repeat FP	TTAGGGTTAGGGTTAGGG
5317	Telomere repeat RP	CCCTAACCCCTAACCC
5318	Targeting tel repeats to <i>csr-1</i> FP1	GCAATGTCGTCAATGGCTGCAAT TAACCCTCACTAAAGGGA
5319	Targeting tel repeats to <i>csr-1</i> RP1	GTGTGCAGCGCAGGAAATCCATA ATACGACTCACTATAGGG
5320	Targeting tel repeats to <i>csr-1</i> FP2	GTGTGCAGCGCAGGAAATCCAAT TAACCCTCACTAAAGGGA
5321	Targeting tel repeats to <i>csr-1</i> RP2	GCAATGTCGTCAATGGCTGCATA ATACGACTCACTATAGGG
5322	Verify LG I circularization FP	AGAGGAGTCCGTAGGCGAAT
5323	Verify LG I circularization RP	TCGTTCCGTTGACAGCTTGA

5324	Verify LG II circularization FP	TGTTTCGGCGATGGGAAGAA
5325	Verify LG II circularization RP	ACTTCGAGTATGTAGCGGCG
5326	Verify LG III circularization FP	CGAGGCTCCATAATGCTCGT
5327	Verify LG III circularization RP	TATTATAGGGCGCGCGGAAG
5328	Verify LG IV circularization FP	GGCGCAAAAACCTTCCTACC
5329	Verify LG IV circularization RP	ACGACAGGGCCTAGGGTAAT
5330	Verify LG VI circularization FP	TAGGTTGAAGGCTATCGGCG
5331	Verify LG VI circularization RP	CCTTGTTGCATTTGGTGGG
5332	Verify LG VII circularization FP	GCCTTCGGCTACCTTTCCTT
5346	Verify LG VII circularization RP	CTCCCTTTCAGCTCGTGTGT
5353	3' <i>csr-1</i> qPCR FP	CGCCGTTAATGCAGTTGTGAT
5354	3' <i>csr-1</i> qPCR RP	CCCCAGCAACTGCGTCTATT
5133	NM149 (N5857) breakpoint 1 FP	TTGCGGCAAGTTTGAAGTCG
5134	NM149 (N5857) breakpoint 1 RP	TGAAGCGTAAGCTCGTGTGT
5130	NM149 (N5857) breakpoint 2 FP	GCTCAAAGTGGGGACTGACA
5154	NM149 (N5857) breakpoint 2 RP	ATCCTTCTCCGCTGTTTCGG
5123	OY329 (N5102) breakpoint 1 FP	GTTGTTGTGGTTTTCCTCGCC
5124	OY329 (N5102) breakpoint 1 RP	ATATAGGCGTAGCGTTGCC
5125	OY329 (N5102) breakpoint 2 FP	TGTTGCCTGGACTGCTAGTG
5126	OY329 (N5102) breakpoint 2 RP	AGCCTAAACCTCGGCTAGGA
5127	OY329 (N5102) breakpoint 3 FP	CCGGTATCACGAGCTTCTCC
5128	OY329 (N5102) breakpoint 3 RP	GGGCGGAAGTTGAGCTGTAT
5155	UK3-41 (N5866) breakpoint 1 FP	GAACGGGACGTTCAAGGCTA
5157	UK3-41 (N5866) breakpoint 1 RP	TGCTTGTCTCGTTTTGCAGC
5158	UK3-41 (N5866) breakpoint 2 FP	CGGGAGAGGGGGATAGTTGA
5473	UK3-41 (N5866) breakpoint 2 RP	CGCACTCACATGCTGCATAC
5156	UK3-41 (N5866) breakpoint 3 FP	AAGGCGTAATGGACACGAGG
5472	UK3-41 (N5866) breakpoint 3 RP	GCATCGTATTTGCACCGTCC
5476	ALS159 (N5101) breakpoint 1 FP	TCCGCAGCCGAAGTTACAAT
5477	ALS159 (N5101) breakpoint 1 RP	ACGTCACTCTCTGCCCTAT
3902	PtpC-nat1 FP	CAACTGATATTGAAGGAGCA
1369	nat-1 RP	GGGCATGCTCATGTAGAGCG
3406	TERT KO FP1	AGCGAACTACTACCAACTACG
3407	TERT KO RP2	CCCAAAAATGCTCCTTCAATAT CAGTTGGCATGTGAGACACTATC ACG
3408	TERT KO FP3	CGCTCTACATGAGCATGCCCTGC CCCTGAAAGTCTGGTACTGGGAT TGG
3409	TERT KO RP4	CTTTCACGACTACCTTCCAAG
6271	6271_NCU02840_qPCR_F	CCCTCTCAGACGAGGATATTCA
6272	6272_NCU02840_qPCR_R	GCTCTGCTGCTTCTCCTTTAT
6277	6277_NCU04720_qPCR_F	TCAGGTTGGAGGAGAGGTATAA
6278	6278_NCU04720_qPCR_R	CTATGAGGCCGAAATCCTTGT
6281	6281_NCU10030_qPCR_F	CTGCGATCGTAACACTGGATTA

6282	6282_NCU10030_qPCR_R	CCGTCCGACATGTAATTA
6283	6283_NCU10031_qPCR_F	CCGATGTCCAGAAGCAGTATATT A
6284	6284_NCU10031_qPCR_R	CAGAGCAACTGAGTGGATAGTC
3565	Telomere 1L qPCR FP	AGCGTTCAAATGCCGTGACCTGT
3566	Telomere 1L qPCR RP	AGTCCAATGGTGCTAACGGCGA
5408	5_his-3_FP	AAGCTTGCCATCTCCACCATC
5409	5_his-3_RP	GGGCGTGCACGGCTATGG
5410	3_his-3_FP	GGGCCAAGCTACCCCGTC
5411	3_his-3_RP	GATCCAATGCGGATGGATTTCGC
5412	T3_his-3_FP1	CCATAGCCGTGCACGCCCATTA CCCTCACTAAAGGGA
5413	T7_his-3_RP1	GACGGGGTAGCTTGCCCTAATA CGACTCACTATAGGG
4776	his-3 qPCR FP	AACAGCTGAGGGAGCCAATG
4777	his-3 qPCR RP	GTTTCAGGGGTTTCGTTTCGC

#### Chapter IV

Primers for replacement of <i>NCU04278</i> with <i>trpC::nat-1</i>		
Primer	Description	Sequence (5'→3')
6405	NCU04278_5del_FP	TTTTCGTGCTTCACACGCTG
6406	NCU04278_5del_RP	GCCTCCGCCTCCGCCTCCGCCGCCTCCGCCC CTGAGGAGGGCCTTCT
6409	NCU04278_3del_FP	GAGCTCGGTACCAAGCTTGATGCATAGCCAG AGGGCACAGCTTCGTGG
6410	NCU04278_3del_RP	TTGCCTGCGCCAAGATAGAG
4882	<i>nat-1</i> split FP	GTACAAGTAACAACCTGATATTGAAGGAGC
4883	<i>nat-1</i> split RP	AACCCCATCCGCCGGTACGCG

Primers for generation and verification of <i>his-3<sup>+</sup>::P<sub>CCG-1</sub>::NCU04278::3xFLAG</i> construct		
Primer	Description	Sequence (5'→3')
6397	NCU04278_his3_FP	ACGTACGTTCTAGATGGGCTACTCTGTCTCCT CTTCC (XbaI)
6398	NCU04278_his3_RP	ACGTACGTTCTTAATTAACGTGGCGGTCCTG CCG (PacI)
6715	<i>Pccg</i> seq FP	CAGGCAGCCAGGAACACAAGC
6407	NCU04278_seq1_FP	GAACCGGCACCAGTTCCAGA
6426	NCU04278_seq2_FP	CTTCACTCAAGCCTAAGACCCTCTC
6427	NCU04278_seq3_FP	CGAATGGGGCCGAAGTAACTC

RT-qPCR primers		
Primer	Description	Sequence (5'->3')
6271	NCU02840 RTqPCR FP	CCCTCTCAGACGAGGATATTCA
6272	NCU02840 RTqPCR RP	GCTCTGCTGCTTCTCCTTTAT
6687	NCU04991 RTqPCR FP	GCCATTTTGGATCGGTCACAAGTACT
6688	NCU04991 RTqPCR RP	CGACGAAGATAGCCCTGGGA
6581	NCU05173 RTqPCR FP	CGAGTGTGTTGGACCTGACG
6568	NCU05173 RTqPCR RP	CCTGTTCGAGTTATCGGTGTTG
6583	NCU07152 RTqPCR FP	GGTGACCCCAAACCTTATGTGCG
6584	NCU07152 RTqPCR RP	GGCTCGAATCTGCCTCCAGC
6685	NCU09306 RTqPCR FP	GCGGTGATGGAGTATGGCTT
6686	NCU09306 RTqPCR RP	GCAAGTGCTCAGCGCCAT
6595	NCU09604 RTqPCR FP	GGCTAGCCTCCGTCGAGCATG
6596	NCU09604 RTqPCR RP	AATGGCGGGTGCAGATGTAG
6683	NCU09633 RTqPCR FP	ACAGCCTCACCACAACATACTAATC
6684	NCU09633 RTqPCR RP	CATTGTTTGTGCGGGTCTGCG
6589	NCU10038 RTqPCR FP	CCTGGCTGGCACTCGTATGGG
6590	NCU10038 RTqPCR RP	GTCACTGGTGCAGCACTTGG
6689	NCU10070 RTqPCR FP	AGATTACCGGGTCAAGAGTCATTTATAC
6690	NCU10070 RTqPCR RP	TTGGCTGTTTCTCGCAGGAT

ChIP-qPCR primers		
Primer	Description	Sequence (5'->3')
4082	hH4 qPCR FP	CATCAAGGGGTCATTAC
4083	hH4 qPCR RP	TTTGGAATCACCCCTCCAG
5455	Tel IIIR qPCR FP	CGGGCTGTTGGTCAGTAGCG
5456	Tel IIIR qPCR RP	CAAGCCGCGCCGTCATTACC
6609	NCU07152 qPCR FP	GGCTCTTGGAACCTTACTAGCG
6610	NCU07152 qPCR RP	GGGTCACCTCTTCTGGAAGGC
6611	NCU05173 qPCR FP	GGCGCACACATTGCTCATTG
6612	NCU05173 qPCR RP	TGGTGGAAAACAGCATCACCC
6707	NCU08251 qPCR FP	TCTGTCCAGCACCAAGTTATC
6708	NCU08251 qPCR RP	TGGCCAGGTACTGCATTAAG
6709	NCU08085 qPCR FP	CAGGAACAGCAGGTGGATATT
6710	NCU08085 qPCR RP	GGACTATTCGAAGGATGGTATGG
6711	NCU05086 qPCR FP	CAAGAGAATGCCAGACCTAGAG
6712	NCU05086 qPCR RP	GGTACTCCAGATGGGAAAGATAG
6713	NCU07801 qPCR FP	ACTGACGTGCCGAGATGAAG
6714	NCU07801 qPCR RP	TTGGGTGACATGATGAGGGC

## Chapter V

Replacement of NCU05173 and NCU07152 ORFs with <i>hph</i> and <i>nat-1</i> respectively		
Primer	Description	Sequence (5'→3')
6385	NCU07152_5del_FP	GAGCAGGTCGTGTCCTCAGG
6386	NCU07152_5del_RP	CGTGTCGTCGAGGGTGGCCATCTTGATGGAT TGTCTTGATCGGATGC
6387	NCU07152_3del_FP	TGAGCATGCCCTGCCCTGACGGGTAAATGA CGAGTGTCTGTTGC
6388	NCU07152_3del_RP	GGTTCACGAGACAGTCAGCATAGG
6605	NCU05173_5del_FP	GCGCTTCTTCAGGCTGAGCG
6606	NCU05173_5del_RP	TCGCGGTGAGTTCAGGCTTTTTTCATTGTCGCA GTGGTTGATGACTGAGG
6607	NCU05173_3del_FP	ACCGGGATCCACTTAACGTTACTGAAATCTC GTGCAGCTGATGACATTGTAACC
6608	NCU05173_3del_RP	GGATTGAGAAGAAGTAATAGAAGGCCG
6269	<i>nat-1</i> ORF RP	TCAGGGGCAGGGCATGCTCATG
6270	<i>nat-1</i> ORF FP	ATGGCCACCCTCGACGACACG
6625	<i>hph</i> ORF FP	ATGAAAAGCCTGAACTACCGCGA
2895	<i>hph</i> ORF RP	AGCTGACATCGACACCAACG
2954	<i>hph</i> _internal_FP	TCGCCTCGCTCCAGTCAATGACC
2955	<i>hph</i> _internal_RP	AAAAAGCCTGAACTACCGCGACG

ChIP-qPCR primer pairs		
Primer	Description	Sequence (5'→3')
3565	Telomere_IL_qPCR_FP	AGCGTTCAAATGCCGTGACCTGT
3566	Telomere_IL_qPCR_RP	AGTCCAATGGTGCTAACGGCGA
4082	hH4_qPCR_FP	CATCAAGGGGTCATTAC
4083	hH4_qPCR_RP	TTTGGAATCACCCCTCCAG
6487	NCU08834_qPCR_FP	TCCCGATCCTGCTCTAGCTT
6488	NCU08834_qPCR_RP	ACTAGTTCGGCTGAGTCCCT
6531	NCU02856_qPCR_FP	AGTTTCACACGGGACTGGAC
6532	NCU02856_qPCR_RP	TACGTCCACCACTGCTTGTC
6609	NCU07152_5_qPCR_FP	GGCTCTTGGAACCTTACTAGCG
6610	NCU07152_5_qPCR_RP	GGGTCACCTCTTCTGGAAGGC
6611	NCU05173_5_qPCR_FP	GGCGCACACATTGCTCATTG
6612	NCU05173_5_qPCR_RP	TGGTGGAAAACAGCATCACCC
6565	NCU07152_promoter_FP	CGGTTCCAAAAGTCCCCTGTG
6645	NCU07152_promoter_RP	CTCAGCGGGGTATATCAACGGC
6567	NCU05173_promoter_FP	GCATTACCCTCGACAGGGTCTG
6646	NCU05173_promoter_RP	GCTACCACCATGTGAAGCTCTGG
6647	BAH_K9_peak_1_FP	GAATAAAAAAAGGCTTTTTTATTACTTCC TCGTC
6648	BAH_K9_peak_1_RP	GTTCTATTTTATTTATTTAATTTAAGAG ATTGCGGC



RT-qPCR primer pairs		
Primer	Description	Sequence (5'->3')
6271	NCU02840_RTqPCR_FP	CCCTCTCAGACGAGGATATTCA
6272	NCU02840_RTqPCR_RP	GCTCTGCTGCTTCTCCTTTAT
3209	NCU04173_Actin_RTqPCR_FP	AATGGGTCGGGTATGTGCAA
3210	NCU04173_Actin_RTqPCR_RP	CTTCTGGCCCCATACCGATCAT
6581	NCU05173_RTqPCR_FP	CGAGTGTGTTGGACCTGACG
6568	NCU05173_RTqPCR_RP	CCTGTTCGAGTTATCGGTGTTG
6583	NCU07152_RTqPCR_FP	GGTGACCCCAAACCTTATGTGCG
6584	NCU07152_RTqPCR_RP	GGCTCGAATCTGCCTCCAGC
6599	NCU07241_RTqPCR_FP	CTGATACCGACATCTCCGTAACAG
6600	NCU07241_RTqPCR_RP	CCAATCATGGGACCAGCCCAAG
6593	NCU07246_RTqPCR_FP	GATACCAGGGCACCTGGATC
6534	NCU07246_RTqPCR_RP	GGCGTTCTTTGCCGACTTAC
6615	NCU07624_RTqPCR_FP	CCCAGGGGCGACAAGCAACC
6616	NCU07624_RTqPCR_RP	CAGAAATCATGTCAGCGCGTATGC
6591	NCU08097_RTqPCR_FP	GTTACGCAAGGGGCAACACCTAC
6592	NCU08097_RTqPCR_RP	GAGCTGCGGTATCATCGGTG
6585	NCU08570_RTqPCR_FP	CCACGAAGCGCCAACATGACG
6586	NCU08570_RTqPCR_RP	GATGTCGTGATGTGGCTCCACG
6587	NCU09178_RTqPCR_FP	GCTCGTATCATCCTTACGACATGG
6588	NCU09178_RTqPCR_RP	CTAGAACTCTCATCATCGCTCCC
6603	NCU09274_RTqPCR_FP	CCTATTATGCTTGGACCACAACG
6604	NCU09274_RTqPCR_RP	ACATTTGGATGCGTCTGCCC
6595	NCU09604_RTqPCR_FP	GGCTAGCCTCCGTCGAGCATG
6596	NCU09604_RTqPCR_RP	AATGGCGGGTGCAGATGTAG
6613	NCU09640_RTqPCR_FP	CTCGTCTTTTATCTTGCACCTTACTT CC
6614	NCU09640_RTqPCR_RP	GCCAAAATGTGGTGATGAGCC
6597	NCU09990_RTqPCR_FP	GCAAGGATAGCAAAGCATGGGAGG
6598	NCU09990_RTqPCR_RP	CCCGTTGTTGCAATTCTTCCAATCG
6589	NCU10038_RTqPCR_FP	CCTGGCTGGCACTCGTATGGG
6590	NCU10038_RTqPCR_RP	GTCACTGGTGCAGCACTTGG
6601	NCU16720_RTqPCR_FP	GGAGGTTTGGCAGTTCACCAAGG
6602	NCU16720_RTqPCR_RP	GCGGATGACTGGACGCTCTCC

Southern Probes		
Primer	Description	Sequence (5'->3')
1665	<i>his-3</i> FP	GACGGGGTAGCTTGGCCCTAATTAACC
3128	<i>his-3</i> RP	CGATTTAGGTGACACTATAG
5271	Tel VIII_L FP	GGCATCCGTGGGTGTCCCAG
5272	Tel VIII_L RP	TTCCCGTCCCTACCAGGCAT
6567	NCU05173 FP	GCATTACCCTCGACAGGGTCTG
6568	NCU05173 RP	CCTGTTCGAGTTATCGGTGTTG

DamID-seq protocol		
Primer	Description	Sequence (5'->3')
5048	AdR-top	CTAATAACGACTCACTATAGGGCAGCGTGG TCGCGGCCGAGGA
5049	AdR-bottom	TCCTCGGCCG
5050	AdR-top + AdR-bottom (annealed)	(see above)
5051	Bio-AdR	Biotin-GGTCGCGGCCGAGGATC

Cloning the BAH domain of EPR-1		
Primer	Description	Sequence (5'->3')
6681	BAH domain F	AGATCCGGATCCTTAGCATAAAGAAAAGGACCC AAAACAAGG
6682	BAH domain R	AGATCCGCGGCCGCTTAGGTTCGAGAGTTCTTGAGA CCG

Generation and verification of <i>his-3<sup>+</sup>::pCCG::N-GFP::EPR-1</i> constructs		
Primer	Description	Sequence (5'->3')
6416	<i>NCU07505_his-3_Ntag_FP</i>	ACGTACGTTTAATTAACGCCTCCTCGCG CAAGAGA (PacI)
6417	<i>NCU07505_his-3_Ntag_RP</i>	ACGTACGTCTCTAGATGCTCACTCGCTT GCATCAT (XbaI)
6367	<i>NCU07505_W184A_FP</i>	TCTTCGCCAGAGTTTACGCGATGTACTG GCCTGACG
6368	<i>NCU07505_W184A_RP</i>	CGTCAGGCCAGTACATCGCGTAAACTCT GGCGAAGA
6399	<i>NCU07505_W292A_FP</i>	CTGCTCTTCAGAATCATGTAAGAAAGCG TTACACGAAGAATGCATCAAGGACC
6400	<i>NCU07505_W292A_RP</i>	GGTCCTTGATGCATTCTTCGTGTAACGCT TTCTTACATGATTCTGAAGAGCAG
6425	N-GFP_seq_FP	TCCTGCTGGAGTTCGTGACC
6346	<i>NCU07505_seq_FP1</i>	TGGACCGACATGACGAGGTA
6348	<i>NCU07505_seq_FP2</i>	GGAAAGCAAAGCCAAGACGG
6366	<i>NCU07505_seq_FP3</i>	GACGAGGAGACCCAGGATAGTC

Replacement of <i>epr-1</i> with <i>trpC::nat-1</i>		
Primer	Description	Sequence (5'->3')
6401	<i>NCU07505_5del_FP</i>	TGGCTGAGTGCAACGATTCT
6402	<i>NCU07505_5del_RP</i>	GCCTCCGCCTCCGCCTCCGCCGCCTCCGCCTG TTGAGTAGTGTGAGTGTAGTAGAAGAG
6350	<i>NCU07505_3_FP</i>	GAGCTCGGTACCAAGCTTGATGCATAGCTAG TGATGGCGGACGATTACGG

6351	<i>NCU07505_3_RP</i>	TGCTCACTCGCTTGCATCAT
4882	<i>nat-1_split_FP</i>	GTACAAGTAACAACACTGATATTGAAGGAGC
4883	<i>nat-1_split_RP</i>	AACCCCATCCGCCGGTACGCG

Endogenous C-terminal tagging of EPR-1 with 10xGly::Dam (WT and PHD)		
Primer	Description	Sequence (5'→3')
6348	<i>NCU07505_5_FP</i>	GGAAAGCAAAGCCAAGACGG
6349	<i>NCU07505_5_RP</i>	<u>GCCTCCGCCTCCGCCTCCGCCGCCTCCGCC</u> ATTAACCAACACACCACACACCAA
6350	<i>NCU07505_3_FP</i>	<u>GAGCTCGGTACCAAGCTTGATGCATAGCT</u> AGTGATGGCGGACGATTACGG
6351	<i>NCU07505_3_RP</i>	TGCTCACTCGCTTGCATCAT
6346	<i>NCU07505_5alt_FP</i>	TGGACCGACATGACGAGGTA
6399	<i>NCU07505_W292A_FP</i>	CTGCTCTTCAGAATCATGTAAGAAAGCGTT ACACGAAGAATGCATCAAGGACC
6400	<i>NCU07505_W292A_RP</i>	GGTCCTTGATGCATTCTTCGTGTAAC <u>CGCTT</u> TCTTACATGATTCTGAAGAGCAG
4882	<i>nat-1_split_FP</i>	GTACAAGTAACAACACTGATATTGAAGGAGC
4883	<i>nat-1_split_RP</i>	AACCCCATCCGCCGGTACGCG

## APPENDIX D

### PLASMID TABLES

#### Chapter II

Plasmid	Description
3130	pZero::3xFLAG:: <i>nat-1</i> :: <i>loxP</i>
3167	histone H3 codon optimized
3168	histone H3 codon optimized (H3K56R)
3169	histone H3 codon optimized (H3K56Q)

#### Chapter III

Plasmid	Description
1991	pBM61 – <i>his-3</i> targeting vector
3110	LG VIL H3K27me2/3 segment #1 cloned into 1991
3111	LG VIL H3K27me2/3 segment #2 cloned into 1991
3112	LG VIL H3K27me2/3 segment #3 cloned into 1991
3113	LG VIL H3K27me2/3 segment #4 cloned into 1991
3114	LG VIL H3K27me2/3 segment #5 cloned into 1991
3115	LG VIL H3K27me2/3 segment #6 cloned into 1991
3116	LG VIL H3K27me2/3 segment #7 cloned into 1991
3117	LG VIL H3K27me2/3 segment #8 cloned into 1991
3172	(TTAGGG) <sub>76</sub> cloned into pCR4-TOPO-TA
FGSC #10598	pAL12-Lifeact – source of <i>trpC</i> :: <i>nat-1</i>

#### Chapter IV

Plasmid	Description
2401	<i>P<sub>CCG-1</sub></i> ::C-Gly::3xFLAG
3237 (FGSC 10598)	pAL12 – source of <i>nat-1</i>
3340	<i>P<sub>CCG-1</sub></i> :: <i>NCU04278<sup>WT</sup></i> ::C-Gly::3xFLAG

#### Chapter V

Plasmid	Description
1991	pBM61 – <i>his-3<sup>+</sup></i> targeting vector
2406	<i>his-3<sup>+</sup></i> :: <i>pCCG</i> ::N-GFP
3329	<i>his-3<sup>+</sup></i> :: <i>pCCG</i> ::N-GFP:: <i>NCU07505</i>
3345	<i>his-3<sup>+</sup></i> :: <i>pCCG</i> ::N-GFP:: <i>NCU07505<sup>W184A</sup></i>
3346	<i>his-3<sup>+</sup></i> :: <i>pCCG</i> ::N-GFP:: <i>NCU07505<sup>W292A</sup></i>

3130	pZero::10xGly::3xFLAG::trpC::nat-1
3131	pZero::10xGly::Dam::trpC::nat-1
3237 (FGSC 10598)	pAL12 – source of <i>nat-1</i>
2283	pCSN44 – source of <i>hph</i>
3230	pE-SUMO
3410	pE-SUMO + BAH <sub>EPR-1</sub>

## REFERENCES CITED

1. Herskowitz, I. Life cycle of the budding yeast *Saccharomyces cerevisiae*. *Microbiol. Rev.* **52**, 536–553 (1988).
2. Cabej, N. R. *Building the Most Complex Structure on Earth*. (Elsevier Science, 2013).
3. Schubert, M. *et al.* Perturbation-response genes reveal signaling footprints in cancer gene expression. *Nat Commun* **9**, 20 (2018).
4. Henikoff, S. & Shilatifard, A. Histone modification: cause or cog? *Trends Genet.* **27**, 389–396 (2011).
5. Kirmizis, A. *et al.* Silencing of human polycomb target genes is associated with methylation of histone H3 Lys 27. *Genes & development* **18**, 1592–1605 (2004).
6. Jürgens, G. A group of genes controlling the spatial expression of the bithorax complex in *Drosophila*. , *Published online: 11 July 1985*; | *doi:10.1038/316153a0* **316**, 153–155 (1985).
7. Schuettengruber, B., Bourbon, H.-M., Di Croce, L. & Cavalli, G. Genome Regulation by Polycomb and Trithorax: 70 Years and Counting. *Cell* **171**, 34–57 (2017).
8. Aramayo, R. & Selker, E. U. *Neurospora crassa*, a model system for epigenetics research. *Cold Spring Harbor perspectives in ...* (2013).
9. Jamieson, K., Rountree, M. R., Lewis, Z. A., Stajich, J. E. & Selker, E. U. Regional control of histone H3 lysine 27 methylation in *Neurospora*. *Proc. Natl. Acad. Sci. U.S.A.* **110**, 6027–6032 (2013).
10. Laugesen, A., Højfeldt, J. W. & Helin, K. Molecular Mechanisms Directing PRC2 Recruitment and H3K27 Methylation. *Molecular cell* **74**, 8–18 (2019).
11. Wiles, E. T. & Selker, E. U. H3K27 methylation: a promiscuous repressive chromatin mark. *Curr. Opin. Genet. Dev.* **43**, 31–37 (2016).
12. Jamieson, K. *et al.* Loss of HP1 causes depletion of H3K27me3 from facultative heterochromatin and gain of H3K27me2 at constitutive heterochromatin. *Genome Res.* **26**, 97–107 (2016).

13. Mathieu, O., Probst, A. V. & Paszkowski, J. Distinct regulation of histone H3 methylation at lysines 27 and 9 by CpG methylation in *Arabidopsis*. *The EMBO Journal* **24**, 2783–2791 (2005).
14. Deleris, A. *et al.* Loss of the DNA methyltransferase MET1 Induces H3K9 hypermethylation at PcG target genes and redistribution of H3K27 trimethylation to transposons in *Arabidopsis thaliana*. *PLOS Genet* **8**, e1003062 (2012).
15. Lindroth, A. M. *et al.* Antagonism between DNA and H3K27 methylation at the imprinted *Rasgrf1* locus. *PLOS Genet* **4**, e1000145 (2008).
16. Reddington, J. P. *et al.* Redistribution of H3K27me3 upon DNA hypomethylation results in de-repression of Polycomb target genes. *Genome Biol* **14**, R25 (2013).
17. Hagarman, J. A., Motley, M. P., Kristjansdottir, K. & Soloway, P. D. Coordinate regulation of DNA methylation and H3K27me3 in mouse embryonic stem cells. *PLoS ONE* **8**, e53880 (2013).
18. Zhang, Z. *et al.* Histone H3K56 acetylation is required for quelling-induced small RNA production through its role in homologous recombination. *J. Biol. Chem.* **289**, 9365–9371 (2014).
19. Smith, K. M. *et al.* The fungus *Neurospora crassa* displays telomeric silencing mediated by multiple sirtuins and by methylation of histone H3 lysine 9. *Epigenetics & Chromatin* **1**, 5 (2008).
20. Maas, N. L., Miller, K. M., DeFazio, L. G. & Toczyski, D. P. Cell Cycle and Checkpoint Regulation of Histone H3 K56 Acetylation by Hst3 and Hst4. *Molecular cell* **23**, 109–119 (2006).
21. Chen, C.-C. *et al.* Acetylated lysine 56 on histone H3 drives chromatin assembly after repair and signals for the completion of repair. *Cell* **134**, 231–243 (2008).
22. Kaplan, T. *et al.* Cell cycle- and chaperone-mediated regulation of H3K56ac incorporation in yeast. *PLOS Genet* **4**, e1000270 (2008).
23. Ferrari, P. & Strubin, M. Uncoupling histone turnover from transcription-associated histone H3 modifications. *Nucleic Acids Res.* **43**, 3972–3985 (2015).
24. Chory, E. J. *et al.* Nucleosome Turnover Regulates Histone Methylation Patterns over the Genome. *Molecular cell* **73**, 61–72.e3 (2019).
25. Simoneau, A. *et al.* Chromosome-wide histone deacetylation by sirtuins prevents hyperactivation of DNA damage-induced signaling upon replicative stress. *Nucleic Acids Res.* **44**, 2706–2726 (2016).

26. Celic, I. *et al.* The sirtuins Hst3 and Hst4p preserve genome integrity by controlling histone H3 lysine 56 deacetylation. *Curr. Biol.* **16**, 1280–1289 (2006).
27. Simoneau, A. *et al.* Chromosome-wide histone deacetylation by sirtuins prevents hyperactivation of DNA damage-induced signaling upon replicative stress. *Nucleic Acids Res.* **44**, gkv1537–2726 (2016).
28. Vissers, J. H. A., van Lohuizen, M. & Citterio, E. The emerging role of Polycomb repressors in the response to DNA damage. *J Cell Sci* **125**, 3939–3948 (2012).
29. Davis, R. H. *Neurospora: contributions of a model organism.* (Oxford University Press, 2000).
30. Russo, V., Sommer, T. & Chambers, J. A modified Vogel's medium for crossings, mating-type tests and the isolation of female-sterile mutants of *Neurospora crassa*. *Fungal Genetics Reports* **32**, (1985).
31. Lewis, E. B. A gene complex controlling segmentation in *Drosophila*. *Nature* **276**, 565–570 (1978).
32. Piunti, A. & Shilatifard, A. Epigenetic balance of gene expression by Polycomb and COMPASS families. *Science* **352**, aad9780 1–15 (2016).
33. Conway, E., Healy, E. & Bracken, A. P. PRC2 mediated H3K27 methylations in cellular identity and cancer. *Curr. Opin. Cell Biol.* **37**, 42–48 (2015).
34. Müller, J. *et al.* Histone methyltransferase activity of a *Drosophila* Polycomb group repressor complex. *Cell* **111**, 197–208 (2002).
35. Wang, H. *et al.* Role of histone H2A ubiquitination in Polycomb silencing. *Nature* **431**, 873–878 (2004).
36. Galazka, J. M. *et al.* *Neurospora* chromosomes are organized by blocks of importin alpha-dependent heterochromatin that are largely independent of H3K9me3. *Genome Res.* **26**, 1069–1080 (2016).
37. Schotanus, K. *et al.* Histone modifications rather than the novel regional centromeres of *Zymoseptoria tritici* distinguish core and accessory chromosomes. *Epigenetics & Chromatin* **8**, 1 (2015).
38. Dumesic, P. A. *et al.* Product Binding Enforces the Genomic Specificity of a Yeast Polycomb Repressive Complex. *Cell* **160**, 204–218 (2015).



39. Studt, L. *et al.* Knock-down of the methyltransferase Kmt6 relieves H3K27me3 and results in induction of cryptic and otherwise silent secondary metabolite gene clusters in *Fusarium fujikuroi*. *Environ. Microbiol.* **18**, 4037–4054 (2016).
40. Connolly, L. R., Smith, K. M. & Freitag, M. The *Fusarium graminearum* histone H3 K27 methyltransferase KMT6 regulates development and expression of secondary metabolite gene clusters. *PLOS Genet* **9**, e1003916 (2013).
41. Steffen, P. A. & Ringrose, L. What are memories made of? How Polycomb and Trithorax proteins mediate epigenetic memory. *Nat. Rev. Mol. Cell Biol.* **15**, 340–356 (2014).
42. Xiao, J. *et al.* Cis and trans determinants of epigenetic silencing by Polycomb repressive complex 2 in Arabidopsis. *Nature Genetics* **81**, 83 (2017).
43. Bauer, M., Trupke, J. & Ringrose, L. The quest for mammalian Polycomb response elements: are we there yet? *Chromosoma* **125**, 1–26 (2015).
44. Wu, H. *et al.* Dnmt3a-dependent nonpromoter DNA methylation facilitates transcription of neurogenic genes. *Science* **329**, 444–448 (2010).
45. Miao, V. P., Singer, M. J., Rountree, M. R. & Selker, E. U. A targeted-replacement system for identification of signals for *de novo* methylation in *Neurospora crassa*. *Molecular and cellular biology* **14**, 7059–7067 (1994).
46. Selker, E. U., Jensen, B. C. & Richardson, G. A. A portable signal causing faithful DNA methylation *de novo* in *Neurospora crassa*. *Science* **238**, 48–53 (1987).
47. Perkins, D. D. Chromosome rearrangements in *Neurospora* and other filamentous fungi. *Adv. Genet.* **36**, 239–398 (1997).
48. Pomraning, K. R., Smith, K. M. & Freitag, M. Bulk segregant analysis followed by high-throughput sequencing reveals the *Neurospora* cell cycle gene, *ndc-1*, to be allelic with the gene for ornithine decarboxylase, *spe-1*. *Eukaryotic Cell* **10**, 724–733 (2011).
49. Wu, C. *et al.* Characterization of chromosome ends in the filamentous fungus *Neurospora crassa*. *Genetics* **181**, 1129–1145 (2009).
50. Nakamura, T. M., Cooper, J. P. & Cech, T. R. Two modes of survival of fission yeast without telomerase. *Science* (1998). doi:10.1126/science.282.5388.493
51. Butler, D. K. & Metzberg, R. L. Expansion and contraction of the nucleolus organizer region of *Neurospora*: changes originate in both proximal and distal segments. *Genetics* **126**, 325–333 (1990).

52. Cunningham, M. D., Brown, J. L. & Kassis, J. A. Characterization of the Polycomb Group Response Elements of the *Drosophila melanogaster* inverted Locus. *Molecular and cellular biology* **30**, 820–828 (2010).
53. Horard, B., Tatout, C. & Poux, S. Structure of a polycomb response element and in vitro binding of polycomb group complexes containing GAGA factor. *Molecular and cellular ...* (2000). doi:10.1128/MCB.20.9.3187-3197.2000
54. De, S., Mitra, A., Cheng, Y., Pfeifer, K. & Kassis, J. A. Formation of a Polycomb-domain in the absence of strong Polycomb response elements. *PLoS Genet* **12**, e1006200 (2016).
55. Basu, A., Dasari, V., Mishra, R. K. & Khosla, S. The CpG island encompassing the promoter and first exon of human DNMT3L gene is a PcG/TrX response element (PRE). *PLoS ONE* **9**, e93561 (2014).
56. Sing, A. *et al.* A vertebrate Polycomb response element governs segmentation of the posterior hindbrain. *Cell* **138**, 885–897 (2009).
57. Woo, C. J., Kharchenko, P. V., Daheron, L., Park, P. J. & Kingston, R. E. A region of the human HOXD cluster that confers polycomb-group responsiveness. *Cell* **140**, 99–110 (2010).
58. Klocko, A. D. *et al.* Normal chromosome conformation depends on subtelomeric facultative heterochromatin in *Neurospora crassa*. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 15048–15053 (2016).
59. Kyrion, G., Liu, K., Liu, C. & Lustig, A. J. RAP1 and telomere structure regulate telomere position effects in *Saccharomyces cerevisiae*. *Genes & development* **7**, 1146–1159 (1993).
60. Nimmo, E. R., Cranston, G. & Allshire, R. C. Telomere-associated chromosome breakage in fission yeast results in variegated expression of adjacent genes. *The EMBO Journal* **13**, 3801–3811 (1994).
61. Castaño, I. *et al.* Telomere length control and transcriptional regulation of subtelomeric adhesins in *Candida glabrata*. *Mol. Microbiol.* **55**, 1246–1258 (2004).
62. Shaaban, M. *et al.* Involvement of transposon-like elements in penicillin gene cluster regulation. *Fungal Genetics and Biology* **47**, 423–432 (2010).
63. Doheny, J. G., Mottus, R. & Grigliatti, T. A. Telomeric position effect--a third silencing mechanism in eukaryotes. *PLoS ONE* **3**, e3864 (2008).

64. Pedram, M. *et al.* Telomere position effect and silencing of transgenes near telomeres in the mouse. *Molecular and cellular biology* **26**, 1865–1878 (2006).
65. Baur, J. A., Zou, Y., Shay, J. W. & Wright, W. E. Telomere position effect in human cells. *Science* **292**, 2075–2077 (2001).
66. Mefford, H. C. & Trask, B. J. The complex structure and dynamic evolution of human subtelomeres. *Nature Reviews Genetics* **3**, 91–102 (2002).
67. Arnoult, N., Van Beneden, A. & Decottignies, A. Telomere length regulates TERRA levels through increased trimethylation of telomeric H3K9 and HP1 $\alpha$ . *Nature Structural & Molecular Biology* **19**, 948–956 (2012).
68. Pryde, F. E. & Louis, E. J. Limitations of silencing at native yeast telomeres. *The EMBO Journal* **18**, 2538–2550 (1999).
69. Tennen, R. I., Bua, D. J., Wright, W. E. & Chua, K. F. SIRT6 is required for maintenance of telomere position effect in human cells. *Nat Commun* **2**, 433 (2011).
70. Baker, K. *et al.* Chromatin state analysis of the barley epigenome reveals a higher-order structure defined by H3K27me1 and H3K27me3 abundance. *Plant J.* **84**, 111–124 (2015).
71. Vaquero-Sedas, M. I., Luo, C. & Vega-Palas, M. A. Analysis of the epigenetic status of telomeres by using ChIP-seq data. *Nucleic Acids Res.* **40**, e163–e163 (2012).
72. Wirth, M. *et al.* HIS-24 linker histone and SIR-2.1 deacetylase induce H3K27me3 in the *Caenorhabditis elegans* germ line. *Molecular and cellular biology* **29**, 3700–3709 (2009).
73. Wang, X. *et al.* Targeting of Polycomb Repressive Complex 2 to RNA by short repeats of consecutive guanines. *Molecular cell* **65**, 1056–1067 (2017).
74. Chu, H.-P. *et al.* TERRA RNA Antagonizes ATRX and Protects Telomeres. *Cell* **170**, 86–101.e16 (2017).
75. Ruiz-Herrera, A., Nergadze, S. G., Santagostino, M. & Giulotto, E. Telomeric repeats far from the ends: mechanisms of origin and role in evolution. *Cytogenetic and Genome Research* **122**, 219–228 (2008).
76. Zofall, M., Smith, D. R., Mizuguchi, T., Dhakshnamoorthy, J. & Grewal, S. I. S. Taz1-Shelterin promotes facultative heterochromatin assembly at chromosome-internal sites containing late replication origins. *Molecular cell* **62**, 862–874 (2016).

77. Duan, Y.-M. *et al.* Molecular dynamics of de novo telomere heterochromatin formation in budding yeast. *J Genet Genomics* **43**, 451–465 (2016).
78. Majerová, E. *et al.* Chromatin features of plant telomeric sequences at terminal vs. internal positions. *Front Plant Sci* **5**, 593 (2014).
79. Kilburn, A. E., Shea, M. J., Sargent, R. G. & Wilson, J. H. Insertion of a telomere repeat sequence into a mammalian gene causes chromosome instability. *Molecular and cellular biology* **21**, 126–135 (2001).
80. Judith A Kassis, J. L. B. Polycomb Group Response Elements in *Drosophila* and Vertebrates. *Advances in genetics* **81**, 83–118 (2013).
81. Blastyák, A., Mishra, R. K., Karch, F. & Gyurkovics, H. Efficient and specific targeting of Polycomb group proteins requires cooperative interaction between Grainyhead and Pleiohomeotic. *Molecular and cellular biology* **26**, 1434–1444 (2006).
82. Nevil, M., Bondra, E. R., Schulz, K. N., Kaplan, T. & Harrison, M. M. Stable binding of the conserved transcription factor grainy head to its target genes throughout *Drosophila melanogaster* development. *Genetics* **205**, 605–620 (2017).
83. Paré, A., Kim, M., Juarez, M. T., Brody, S. & McGinnis, W. The functions of grainy head-like proteins in animals and fungi and the evolution of apical extracellular barriers. *PLoS ONE* **7**, e36254 (2012).
84. Riising, E. M. *et al.* Gene silencing triggers polycomb repressive complex 2 recruitment to CpG islands genome wide. *Molecular cell* **55**, 347–360 (2014).
85. Hosogane, M., Funayama, R., Shirota, M. & Nakayama, K. Lack of Transcription Triggers H3K27me3 Accumulation in the Gene Body. *Cell Rep* **16**, 696–706 (2016).
86. Hnisz, D. *et al.* Activation of proto-oncogenes by disruption of chromosome neighborhoods. *Science* **351**, 1454–1458 (2016).
87. Surace, C. *et al.* Telomere shortening and telomere position effect in mild ring 17 syndrome. *Epigenetics & Chromatin* **7**, 1 (2014).
88. Guilherme, R. S. *et al.* Position effect modifying gene expression in a patient with ring chromosome 14. *J. Appl. Genet.* **57**, 183–187 (2016).
89. Hogenbirk, M. A. *et al.* Defining chromosomal translocation risks in cancer. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E3649–56 (2016).

90. Colot, H. V. *et al.* A high-throughput gene knockout procedure for *Neurospora* reveals functions for multiple transcription factors. *PNAS* **103**, 10352–10357 (2006).
91. Margolin, B. S., Freitag, M. & Selker, E. U. Improved plasmids for gene targeting at the *his-3* locus of *Neurospora crassa* by electroporation. *Fungal Genetics Reports* 34–36 (1997).
92. Bardiya, N. & Shiu, P. K. T. Cyclosporin A-resistance based gene placement system for *Neurospora crassa*. *Fungal Genet. Biol.* **44**, 307–314 (2007).
93. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
94. Quinlan, A. R. & Hall, I. M. BEDTools: a flexible suite of utilities for comparing genomic features. *Bioinformatics* **26**, 841–842 (2010).
95. Hahne, F. & Ivanek, R. Visualizing genomic data using Gviz and Bioconductor. *Methods Mol. Biol.* **1418**, 335–351 (2016).
96. Thorvaldsdottir, H., Robinson, J. T. & Mesirov, J. P. Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinformatics* **14**, 178–192 (2013).
97. Layer, R. M., Chiang, C., Quinlan, A. R. & Hall, I. M. LUMPY: a probabilistic framework for structural variant discovery. *Genome Biol* **15**, 1–19 (2014).
98. Miao, V. P. W., Freitag, M. & Selker, E. U. Short TpA-rich segments of the  $\zeta$ - $\eta$  region induce DNA methylation in *Neurospora crassa*. *Journal of molecular biology* **300**, 249–273 (2000).
99. Filichkin, S. A. *et al.* Genome-wide mapping of alternative splicing in *Arabidopsis thaliana*. *Genome Res.* **20**, 45–58 (2010).
100. Laugesen, A., Højfeldt, J. W. & Helin, K. Role of the Polycomb Repressive Complex 2 (PRC2) in Transcriptional Regulation and Cancer. *Cold Spring Harb Perspect Med* **6**, a026575 (2016).
101. Margueron, R. & Reinberg, D. The Polycomb complex PRC2 and its mark in life. *Nature* **469**, 343–349 (2011).
102. Lewis, Z. A. Polycomb Group Systems in Fungi: New Models for Understanding Polycomb Repressive Complex 2. *Trends Genet.* **33**, 220–231 (2017).
103. Hauri, S. *et al.* A High-Density Map for Navigating the Human Polycomb Complexome. *Cell Rep* **17**, 583–595 (2016).

104. Herz, H.-M. *et al.* Polycomb repressive complex 2-dependent and -independent functions of Jarid2 in transcriptional regulation in *Drosophila*. *Molecular and cellular biology* **32**, 1683–1693 (2012).
105. Kang, H. *et al.* Sex comb on midleg (Scm) is a functional link between PcG-repressive complexes in *Drosophila*. *Genes & development* **29**, 1136–1150 (2015).
106. Nekrasov, M. *et al.* Pcl-PRC2 is needed to generate high levels of H3-K27 trimethylation at Polycomb target genes. *The EMBO Journal* **26**, 4078–4088 (2007).
107. Metzzenberg, R. L., Stevens, J. N., Selker, E. U. & Morzycka-Wroblewska, E. Identification and chromosomal distribution of 5S rRNA genes in *Neurospora crassa*. *PNAS* **82**, 2067–2071 (1985).
108. Gulshan, K. & Moye-Rowley, W. S. Multidrug resistance in fungi. *Eukaryotic Cell* **6**, 1933–1942 (2007).
109. Jamieson, K. *et al.* Telomere repeats induce domains of H3K27 methylation in *Neurospora*. *Elife* **7**, 948 (2018).
110. Bicocca, V. T., Ormsby, T., Adhvaryu, K. K., Honda, S. & Selker, E. U. ASH1-catalyzed H3K36 methylation drives gene repression and marks H3K27me2/3-competent chromatin. *Elife* **7**, 1455 (2018).
111. Zhou, Y. *et al.* Telobox motifs recruit CLF/SWN-PRC2 for H3K27me3 deposition via TRB factors in *Arabidopsis*. *Nature Genetics* **50**, 638–644 (2018).
112. Qüesta, J. I., Song, J., Geraldo, N., An, H. & Dean, C. *Arabidopsis* transcriptional repressor VAL1 triggers Polycomb silencing at FLC during vernalization. *Science* **353**, 485–488 (2016).
113. Tanay, A., O'Donnell, A. H., Damelin, M. & Bestor, T. H. Hyperconserved CpG domains underlie Polycomb-binding sites. *PNAS* **104**, 5521–5526 (2007).
114. Wachter, E. *et al.* Synthetic CpG islands reveal DNA sequence determinants of chromatin structure. *Elife* **3**, e03397 (2014).
115. Jermann, P., Hoerner, L., Burger, L. & Schübeler, D. Short sequences can efficiently recruit histone H3 lysine 27 trimethylation in the absence of enhancer activity and DNA methylation. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E3415–21 (2014).

116. Højfeldt, J. W. *et al.* Non-core Subunits of the PRC2 Complex Are Collectively Required for Its Target-Site Specificity. *Molecular cell* (2019). doi:10.1016/j.molcel.2019.07.031
117. Healy, E. *et al.* PRC2.1 and PRC2.2 Synergize to Coordinate H3K27 Trimethylation. *Molecular cell* (2019). doi:10.1016/j.molcel.2019.08.012
118. Garrison, E. & Marth, G. Haplotype-based variant detection from short-read sequencing. (2012).
119. Danecek, P. *et al.* The variant call format and VCFtools. *Bioinformatics* **27**, 2156–2158 (2011).
120. Honda, S. & Selker, E. U. Tools for fungal proteomics: multifunctional neurospora vectors for gene replacement, protein expression and protein purification. *Genetics* (2009).
121. He, Y. Control of the transition to flowering by chromatin modifications. *Molecular Plant* **2**, 554–564 (2009).
122. Ahringer, J. & Gasser, S. M. Repressive Chromatin in *Caenorhabditis elegans*: Establishment, Composition, and Function. *Genetics* **208**, 491–511 (2018).
123. Jégu, T., Aeby, E. & Lee, J. T. The X chromosome in space. *Nature Reviews Genetics* **18**, 377–389 (2017).
124. Freitag, M. Histone Methylation by SET Domain Proteins in Fungi. *Annu. Rev. Microbiol.* **71**, 413–439 (2017).
125. Kassis, J. A., Kennison, J. A. & Tamkun, J. W. Polycomb and Trithorax Group Genes in *Drosophila*. *Genetics* **206**, 1699–1725 (2017).
126. Min, J., Zhang, Y. & Xu, R.-M. Structural basis for specific binding of Polycomb chromodomain to histone H3 methylated at Lys 27. *Genes & development* **17**, 1823–1828 (2003).
127. Grau, D. J. *et al.* Compaction of chromatin by diverse Polycomb group proteins requires localized regions of high charge. *Genes & development* **25**, 2210–2221 (2011).
128. Cheutin, T. & Cavalli, G. Loss of PRC1 induces higher-order opening of Hox loci independently of transcription during *Drosophila* embryogenesis. *Nat Commun* **9**, 3898 (2018).

129. Tavares, L. *et al.* RYBP-PRC1 complexes mediate H2A ubiquitylation at polycomb target sites independently of PRC2 and H3K27me3. *Cell* **148**, 664–678 (2012).
130. Kahn, T. G. *et al.* Interdependence of PRC1 and PRC2 for recruitment to Polycomb Response Elements. *Nucleic Acids Res.* **44**, gkw701–10149 (2016).
131. Turck, F. *et al.* Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *PLoS Genet* **3**, e86 (2007).
132. Yang, Z. *et al.* EBS is a bivalent histone reader that regulates floral phase transition in Arabidopsis. *Nature Genetics* **352**, aad9780 (2018).
133. Qian, S. *et al.* Dual recognition of H3K4me3 and H3K27me3 by a plant histone reader SHL. *Nat Commun* **9**, 106 (2018).
134. Xu, L. & Shen, W.-H. Polycomb silencing of KNOX genes confines shoot stem cell niches in Arabidopsis. *Curr. Biol.* **18**, 1966–1971 (2008).
135. Bratzel, F., López-Torrejón, G., Koch, M., Del Pozo, J. C. & Calonje, M. Keeping cell identity in Arabidopsis requires PRC1 RING-finger homologs that catalyze H2A monoubiquitination. *Curr. Biol.* **20**, 1853–1859 (2010).
136. Li, Z., Fu, X., Wang, Y., Liu, R. & He, Y. Polycomb-mediated gene silencing by the BAH-EMF1 complex in plants. *Nature Genetics* **10**, 697 (2018).
137. Pirrotta, V. & Li, H.-B. A view of nuclear Polycomb bodies. *Curr. Opin. Genet. Dev.* **22**, 101–109 (2012).
138. López-González, L. *et al.* Chromatin-dependent repression of the Arabidopsis floral integrator genes involves plant specific PHD-containing proteins. *Plant Cell* **26**, 3922–3938 (2014).
139. Huang, Y. *et al.* Evolution and conservation of polycomb repressive complex 1 core components and putative associated factors in the green lineage. *BMC Genomics* 2015 16:1 **20**, 533 (2019).
140. Davis, R. H. & de Serres, F. J. [4] Genetic and microbiological research techniques for *Neurospora crassa*. *Methods in enzymology* **17**, 79–143 (1970).
141. Basenko, E. Y. *et al.* Genome-wide redistribution of H3K27me3 is linked to genotoxic stress and defective growth. *Proc. Natl. Acad. Sci. U.S.A.* **112**, E6339–48 (2015).
142. Luo, M. *et al.* Genes controlling fertilization-independent seed development in Arabidopsis thaliana. *PNAS* **96**, 296–301 (1999).



143. Ohad, N. *et al.* Mutations in FIE, a WD polycomb group gene, allow endosperm development without fertilization. *Plant Cell* **11**, 407–416 (1999).
144. Marchler-Bauer, A. *et al.* CDD: NCBI's conserved domain database. *Nucleic Acids Res.* **43**, D222–6 (2015).
145. Callebaut, I., Courvalin, J. C. & Mornon, J. P. The BAH (bromo-adjacent homology) domain: a link between DNA methylation, replication and transcriptional regulation. *FEBS Lett.* **446**, 189–193 (1999).
146. Schindler, U., Beckmann, H. & Cashmore, A. R. HAT3.1, a novel Arabidopsis homeodomain protein containing a conserved cysteine-rich region. *Plant J.* **4**, 137–150 (1993).
147. Zhao, D. *et al.* The BAH domain of BAHD1 is a histone H3K27me3 reader. *Protein Cell* **7**, 222–226 (2016).
148. Sanchez, R. & Zhou, M.-M. The PHD finger: a versatile epigenome reader. *Trends Biochem. Sci.* **36**, 364–372 (2011).
149. van Steensel, B. & Henikoff, S. Identification of in vivo DNA targets of chromatin proteins using tethered dam methyltransferase. *Nat. Biotechnol.* **18**, 424–428 (2000).
150. Bernstein, B. E. Methods for Global Characterization of Chromatin Regulators in Human Cells. (2012).
151. Hyun, K., Jeon, J., Park, K. & Kim, J. Writing, erasing and reading histone lysine methylations. *Exp. Mol. Med.* **49**, e324–e324 (2017).
152. Du, J. *et al.* Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. *Cell* **151**, 167–180 (2012).
153. Li, H. *et al.* Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* **442**, 91–95 (2006).
154. Boamah, D. *et al.* Characteristics of a PHD Finger Subtype. *Biochemistry* **57**, 525–539 (2018).
155. Scelfo, A. *et al.* Functional Landscape of PCGF Proteins Reveals Both RING1A/B-Dependent-and RING1A/B-Independent-Specific Activities. *Molecular cell* (2019). doi:10.1016/j.molcel.2019.04.002
156. Berry, S., Rosa, S., Howard, M., Bühler, M. & Dean, C. Disruption of an RNA-binding hinge region abolishes LHP1-mediated epigenetic repression. *Genes & development* **31**, 2115–2120 (2017).

157. Veluchamy, A. *et al.* LHP1 Regulates H3K27me3 Spreading and Shapes the Three-Dimensional Conformation of the Arabidopsis Genome. *PLoS ONE* **11**, e0158936 (2016).
158. Kundu, S. *et al.* Polycomb Repressive Complex 1 Generates Discrete Compacted Domains that Change during Differentiation. *Molecular cell* **65**, 432–446.e5 (2017).
159. Wani, A. H. *et al.* Chromatin topology is coupled to Polycomb group protein subnuclear organization. *Nat Commun* **7**, 10291 (2016).
160. Derkacheva, M. *et al.* Arabidopsis MSI1 connects LHP1 to PRC2 complexes. *The EMBO Journal* **32**, 2073–2085 (2013).
161. Blackledge, N. P. *et al.* Variant PRC1 complex-dependent H2A ubiquitylation drives PRC2 recruitment and polycomb domain formation. *Cell* **157**, 1445–1459 (2014).
162. Eskeland, R. *et al.* Ring1B compacts chromatin structure and represses gene expression independent of histone ubiquitination. *Molecular cell* **38**, 452–464 (2010).
163. Gao, Z. *et al.* PCGF homologs, CBX proteins, and RYBP define functionally distinct PRC1 family complexes. *Molecular cell* **45**, 344–356 (2012).
164. Bierne, H. *et al.* Human BAHD1 promotes heterochromatic gene silencing. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 13826–13831 (2009).
165. Lakisic, G. *et al.* Role of the BAHD1 Chromatin-Repressive Complex in Placental Development and Regulation of Steroid Metabolism. *PLOS Genet* **12**, e1005898 (2016).
166. Pomraning, K. R., Smith, K. M. & Freitag, M. Genome-wide high throughput analysis of DNA methylation in eukaryotes. *Methods* **47**, 142–150 (2009).
167. Kilbey, B. J. & De Serres, F. J. Quantitative and qualitative aspects of photoreactivation of premutational ultraviolet damage at the ad-3 loci of *Neurospora crassa*. *Mutat. Res.* **4**, 21–29 (1967).
168. Shiu, P. K. T., Raju, N. B., Zickler, D. & Metzenberg, R. L. Meiotic Silencing by Unpaired DNA. *Cell* **107**, 905–916 (2001).
169. Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering* **9**, 90–95 (2007).

170. Afgan, E. *et al.* The Galaxy platform for accessible, reproducible and collaborative biomedical analyses: 2018 update. *Nucleic Acids Res.* **46**, W537–W544 (2018).
171. Ramírez, F. *et al.* deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res.* **44**, W160–5 (2016).
172. Dobin, A. *et al.* STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* **29**, 15–21 (2012).
173. Love, M. I., Huber, W. & Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* **15**, 31 (2014).
174. van der Walt, S. *et al.* scikit-image: image processing in Python. *PeerJ* **2**, e453 (2014).
175. Altschul, S. F. *et al.* Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.* **25**, 3389–3402 (1997).
176. Söding, J., Biegert, A. & Lupas, A. N. The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.* **33**, W244–8 (2005).
177. Lassmann, T., Frings, O. & Sonnhammer, E. L. L. Kalign2: high-performance multiple alignment of protein and nucleotide sequences allowing external features. *Nucleic Acids Res.* **37**, 858–865 (2009).
178. Price, M. N., Dehal, P. S. & Arkin, A. P. FastTree 2--approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010).