

Yale University

EliScholar – A Digital Platform for Scholarly Publishing at Yale

Yale Graduate School of Arts and Sciences Dissertations

Spring 2022

Systematic Screen of Histone H4 in *Arabidopsis thaliana*

Emma Tung Corcoran

Yale University Graduate School of Arts and Sciences, emma.tung.corcoran@gmail.com

Follow this and additional works at: https://elischolar.library.yale.edu/gsas_dissertations

Recommended Citation

Corcoran, Emma Tung, "Systematic Screen of Histone H4 in *Arabidopsis thaliana*" (2022). *Yale Graduate School of Arts and Sciences Dissertations*. 577.

https://elischolar.library.yale.edu/gsas_dissertations/577

This Dissertation is brought to you for free and open access by EliScholar – A Digital Platform for Scholarly Publishing at Yale. It has been accepted for inclusion in Yale Graduate School of Arts and Sciences Dissertations by an authorized administrator of EliScholar – A Digital Platform for Scholarly Publishing at Yale. For more information, please contact elischolar@yale.edu.

Abstract

Systematic Screen of Histone H4 in *Arabidopsis thaliana*

Emma Tung Corcoran

2022

Histones regulate diverse processes in eukaryotes and consequently, can have widespread effects on organismal fitness and development. Histones are a dynamic target for a variety of post-translational modifications (PTMs) and the assessment of histone function has typically been accomplished by mutating enzymes that catalyze and/or recognize these PTMs (i.e., writers and readers, respectively). Although considerable information has been gained in the past several decades by using this strategy, multiple issues such as writer/reader redundancy, unidentified writers/readers of histone PTMs, and writers/readers with additional non-histone targets can preclude the identification of new roles for histones and complicate the assessment of mutant phenotypes. To bypass these issues and provide a complementary strategy to study histones, large-scale histone replacement systems have been developed and optimized in yeast and fly model systems. However, such systems have never been implemented in plants in part due to the difficulty in eliminating endogenous histone genes that are typically present in many copies and different locations in plant genomes.

Here, we present the development of a genetic strategy for the plant model organism *Arabidopsis thaliana* in which the expression of endogenous histone H4 can be completely replaced with modified H4 transgenes. We use histone H4, which is a single variant histone in plants that is encoded by the largest number of genes (8) among all functionally-distinct histone proteins, as a proof-of-concept for an experimental system allowing the direct assessment of histone function in plants. Our CRISPR/Cas9-based strategy allows for the simultaneous targeting of many histone genes for the generation of a background depleted of endogenous histone expression. We validated

our platform by showing that a single transformation with our modified H4 transgenes can restore a wild-type phenotype, demonstrating that our system can be used for the rapid establishment of histone replacement in plants. Using this strategy, we established a collection of plants expressing different H4 point mutants targeting residues that may be post-translationally modified *in vivo*. To demonstrate the utility of this new H4 mutant collection, we screened it to uncover substitutions in H4 that alter flowering time, rosette morphology, DNA replication, chromatin structure, and gene silencing. We identified different mutations in the tail (H4R17A) and the globular domain (H4R36A, H4R39K, H4R39A, and H4K44A) of H4 that strongly accelerate the floral transition. Additionally, we used machine learning to identify H4 mutations that alter different morphometric traits in vegetative tissue. Finally, we identified several novel roles for H4 tail and globular domain residues in the regulation of endoreduplication, chromatin condensation, and transposon silencing.

After these broad screens for histone function, we then performed targeted analyses of H4R17A mutants to determine a molecular mechanism responsible for the early flowering displayed by these mutants. We found that a conserved regulatory relationship between H4R17 and the ISWI chromatin remodeling complex in plants is responsible for the phenotypes observed in H4R17A mutants. Similar to other biological systems, H4R17 regulates nucleosome spacing via ISWI, and mutation of H4R17 results in large-scale changes to global nucleosome positioning and gene expression, leading to altered development. Overall, this work provides a large set of H4 mutants to the plant epigenetics community that can be used to systematically assess histone H4 function in *A. thaliana* and a blueprint to replicate this strategy for studying other histone proteins in plants. As this resource represents the largest collection of H4 point mutants in a multicellular organism, our work will enable new insights into the regulation of chromatin by histone H4 in multicellular eukaryotes.

Systematic Screen of Histone H4 in *Arabidopsis thaliana*

A Dissertation
Presented to the Faculty of the Graduate School
Of
Yale University
in Candidacy for the Degree of
Doctor of Philosophy

by
Emma Tung Corcoran

Dissertation Director: Yannick Jacob

May 2022

© 2022 by Emma Tung Corcoran
All rights reserved.

Table of Contents

Abstract	i
Table of Contents	v
Acknowledgements	x
List of Figures and Tables	xii
List of Abbreviations	xvi
Chapter 1: Introduction	1
Epigenetics	2
Chromatin	5
Histone post-translational modifications	10
Nucleosome positioning	17
Maintenance of epigenetic signals	26
Genome stability	30
Flowering time	37
Assessing histone function	44
Dissertation goals	57
Chapter 2: Establishment of a histone H4 replacement system in <i>Arabidopsis thaliana</i>	59
Overview	60
Results	60
Generation of an endogenous H4 depletion background.....	60
Histone H4 replacement plasmids	68
Characterization of histone H4 replacement lines.....	74
Discussion	78
A CRISPR/Cas9-based strategy for histone replacement in <i>Arabidopsis thaliana</i> .	78

Depletion of endogenous histone H4 causes genome instability and abnormal morphology	78
Histone H4 depletion triggers a dosage compensation mechanism	81
Overexpression of histone H4 in H4 replacement lines may trigger dosage compensation.....	88
Varying expression of histone H4 in H4 replacement lines as a consequence of T-DNA integration.....	90
Low percentage of histone H4 mutations result in lethality	93
Second generation transgenic H4 replacement plants display a robust replacement of endogenous histone H4	98
Multiple independent transgenic lines mitigate off-target effects	100
Perspectives.....	101
Chapter 3: Phenotypic screens of histone H4 mutants	102
Overview	103
Results	103
Systematic screens of histone H4 mutants	103
Flowering time.....	104
Rosette morphology	114
DNA content.....	122
Chromatin structure.....	129
Transcriptional silencing.....	132
Discussion	138
Many novel functions uncovered for histone H4 residues.....	138
Phenotype variability observed between transgenic lines.....	139
Flowering time analyses identify several early flowering mutants.....	140
H4 mutations cause diverse effects on rosette development.....	143

H4 septuple mutant plants exhibit early flowering time	144
rH4 plants exhibit minor phenotypes related to flowering time and rosette morphology	145
No H4 mutants exhibit consistent <i>BRCA1</i> upregulation.....	147
Assays of H4K16 and H4K20 mutants support previously described roles for these residues.....	148
H4R17 mutants demonstrate unique phenotypes relative to H4 N-terminal tail residues.....	153
Several mutations of the nucleosome entry site (H4R35, H4R36, H4R39, H4R40, and H4K44) cause early flowering	155
Mutations corresponding to the yeast H4 low sporulation patch (H4R77 and H4T80) impact chromatin structure.....	160
C-terminal H4 mutants (H4K91 and H4Y98) demonstrate altered DNA content ..	164
Perspectives.....	168
Chapter 4: Regulation of ISWI chromatin remodeling by histone H4 arginine 17	170
Overview	171
Results	171
Generation of complete H4R17A replacement backgrounds.....	171
Dominance of H4R17A mutation.....	172
No relationship between PRMT7 and H4R17 in plants.....	176
Functional relationship between H4R17 and ISWI in the regulation of flowering..	177
Impact of H4R17A mutation on global nucleosome positioning.....	201
Proposed mechanism for action of H4R17 on the regulation of flowering time and development.....	217
Discussion	223
Utility of histone H4 replacement system for assessing function of H4 residues ..	223

H4R17A mutation shows incomplete dominance.....	224
Histone dosage compensation observed in rH4 plants.....	225
Genome-wide transcriptional analyses uncover multiple biological processes affected by H4R17A mutation	228
Altered PRMT7 function not responsible for H4R17 mutant phenotypes.....	229
H4R17 regulates nucleosome remodeling in plants.....	231
Perspectives.....	239
Chapter 5: Conclusions.....	241
Summary.....	242
A novel system for studying histone function in plants	244
Materials and Methods	249
Plant materials.....	249
Generation of transgenic <i>Arabidopsis</i> plants	249
Flowering time.....	250
Rosette measurements.....	250
Plasmid construction.....	250
DNA extraction, PCR and sequencing analyses	251
RT-qPCR	252
Flow cytometry.....	252
Immunostaining	253
Dimensionality reduction and clustering.....	253
Endoreduplication index	254
Next-generation sequencing library preparation	254
RNA-seq processing and analysis	255
MNase-seq processing and analysis	256
Model building.....	257

Primers	257
Statistical analyses	257
Data availability statement	257
Accession numbers	258
Supplemental Tables	259
References	281

Acknowledgements

I am sincerely grateful for all of the support that has made this dissertation possible. First, I would like to thank my advisor, Yannick Jacob, for his guidance throughout this project. I am thankful for his advice and support during this ambitious project and am honored to be the first graduate student in his lab. I would also like to thank all of the past and present members of the Jacob Lab who have created a collaborative research environment, including discussions about data, troubleshooting advice, and help with experiments. Chantal LeBlanc, Mia Arias Tsang, Anthony Sarkiss, Anisa Iqbal, Benoit Mermaz, Axel Poulet, Valentin Joly, Gonzalo Villarino, Yi-Chun Huang, Lauren Dickinson, Geoffrey Thomson, Wenxin Yuan, Josefina Mendez, and Maximo Kesselhaut have all provided some form of assistance during my six years in the lab that has been instrumental in the success of my project. I would also like to thank all members of the Plant Molecular Biology community at Yale for their advice and feedback over my time here.

My thesis committee members Josien van Wolfswinkel, Joshua Gendron, and Matthew Simon were a reservoir of knowledge providing instrumental guidance during this project. I am also thankful to the core facilities at Yale, without whom many of the experiments described in this dissertation would not have been possible, including the Marsh Gardens Staff (especially Chris Bolick), the Yale Center for Genome Analysis, the Yale Center for Research and Computing, the Bioinformatics Support Program at the Cushing/Whitney Medical Library (especially Nur-Taz Rahman), and the flow cytometry facility (and guidance from Ken Nelson). I am also grateful to the Yale Science Building and Osborn Memorial Laboratories Facilities Services team for maintaining our research spaces.

I was supported by several graduate student-focused grants during my PhD, including the Gruber Science Fellowship, the National Institutes of Health (NIH)-funded

Training Program in Cellular, Molecular, and Quantitative Biology, and the National Science Foundation Graduate Research Fellowship. I am grateful for the financial support and opportunities that these programs have provided me.

I have had the honor of being a part of multiple communities outside of research during my time at Yale. First, the Office for Graduate Student Development and Diversity (OGSDD) was an incredible fountain of support throughout my PhD journey, and I would like to acknowledge the incredible amount of work that Dean Michelle Nearon, Assistant Dean Danica Tisdale Fisher, and former Assistant Dean Denzil Streete poured into this institution to create a more inclusive and supportive community at Yale. The graduate fellows at the OGSDD were a pleasure to work alongside for the past three years, and I have learned so much from each of them. My friends and colleagues in oSTEM and the Graduate Student Disability Alliance have also been such an amazing source of community and support through the difficult times that I have experienced in graduate school.

Most importantly, I would like to thank my friends and family for helping me in countless ways during my PhD. My parents and brother have always been supportive of my goals and have listened to many of my descriptions of my research project. They know a lot about histones and epigenetics now! My mom and dad have also made many trips to New Haven to cook me delicious food and bring me my favorite Chinese snacks that I cannot get in Connecticut. Additionally, my large extended family has been so encouraging, sending me messages of support throughout my PhD. Finally, my fiancée, Cynthia, has helped me through some of the toughest times of my life. She made sure that I had healthy, home-cooked food when I was too sick to cook for myself, and she made sure that I was always laughing even when things were going very poorly. I would not have been able to do this without her.

List of Figures and Tables

Figure 1.1 Simplified scheme of chromatin organization within the cell nucleus.	7
Figure 1.2 Epigenetic modifications are reversible and dynamic.	13
Figure 1.3 Nine chromatin states identified in <i>Arabidopsis thaliana</i>	15
Figure 1.4 Stereotypical nucleosome positioning around gene bodies.	19
Figure 1.5 Three major classes of nucleosome remodelers.	22
Figure 1.6 Major flowering time regulatory pathways.	41
Figure 2.1 Protein sequence similarity of histone H4 across species.	62
Figure 2.2 Design of guide RNAs to generate H4 septuple mutant background.	64
Table 2.1 Effect of H4 septuple mutations on H4 messenger RNA (mRNA).	65
Figure 2.3 Morphological phenotypes of the H4 septuple mutant and H4 replacement plants.	66
Figure 2.4 Expression levels of H4 (<i>At3g53730</i>), <i>BRCA1</i> , and <i>TSI</i> in H4 septuple mutant and rH4 lines.	67
Table 2.2 Mutations in endogenous H4 genes present in H4 depletion backgrounds.	69
Figure 2.5 Morphological phenotypes of H4 sextuple mutant and H4 septuple mutant plants.	70
Figure 2.6 Design of H4 replacement plasmid.	72
Figure 2.7 63 mutations generated for systematic screens of histone H4.	73
Figure 2.8 Workflow for histone H4 replacement system.	75
Figure 2.9 H4 expression levels in rH4S1A mutant lines.	76
Figure 2.10 Efficiency of CRISPR/Cas9 in H4 replacement plants.	77
Table 3.1 Phenotypic analysis of plants with mutations in histone H4 residues.	106
Figure 3.1 Flowering time screen of rH4 mutants.	108
Figure 3.2 Multiple rH4 mutants display early flowering phenotypes.	109

Figure 3.3 Morphological phenotypes of early flowering rH4 mutants in long-day growth conditions.....	110
Figure 3.4 Morphological phenotypes of early flowering rH4 mutants in short-day growth conditions.....	111
Figure 3.5 Mutations in specific residues of histone H4 generate early flowering phenotypes in <i>A. thaliana</i>	113
Figure 3.6 Mutations in specific residues of histone H4 generate morphological phenotypes in <i>A. thaliana</i>	116
Figure 3.7 Rosette phenotypes of histone H4 mutants.....	117
Figure 3.8 Principal component analysis of morphological features.....	118
Figure 3.9 Phenotypic traits of rH4 mutant rosettes.....	120
Figure 3.10 rH4R35K, rH4K91R, and rH4Y98F mutants exhibit altered DNA content.	123
Figure 3.11 Ploidy levels of rH4R35K, rH4K91R, rH4K91A, and rH4Y98F mutants. ...	125
Figure 3.12 Endogenous H4 mutations in rH4R35K and rH4R35A mutant plants.	127
Figure 3.13 Endoreduplication indices for rH4R35K, rH4K91R, rH4K91A, and rH4Y98F mutants.	128
Figure 3.14 rH4 mutants exhibit altered chromatin structure.....	131
Figure 3.15 Expression levels of <i>TSI</i> in rH4 mutant lines.....	133
Figure 3.16 Expression levels of <i>BRCA1</i> in rH4 mutant lines.....	134
Figure 3.17 rH4R39K and rH4R39A mutants exhibit transcriptional derepression.....	136
Figure 4.1 Generation of complete H4 replacement lines.....	173
Table 4.1 Effect of H4 (<i>At3g53730</i>) mutations on mRNA.....	174
Figure 4.2 Phenotypic analyses of H4R17A mutants.....	175
Figure 4.3 Validation of <i>PRMT7</i> mutations.....	178
Figure 4.4 The effect of <i>PRMT7</i> mutations on the floral transition.....	179

Figure 4.5 The effect of ISWI and rH4R17A mutations on the floral transition and development.	180
Figure 4.6 ISWI and rH4R17A mutations cause early flowering.	181
Figure 4.7 Bioanalyzer electropherograms of RNA-seq replicates.	183
Figure 4.8 Spearman correlation of RNA-seq replicates.	184
Figure 4.9 H4 expression levels in rH4 lines.	186
Figure 4.10 Expression levels of most highly expressed H2A, H2B, and H3 genes in rH4 lines.	187
Figure 4.11 Overlap of differentially expressed genes in rH4R17A, ISWI and SWR1 mutants.	189
Figure 4.12 Patterns of differential gene expression in rH4R17A, ISWI and SWR1 mutants.	190
Figure 4.13 Principal component analysis of differentially expressed genes.	191
Figure 4.14 Correlation of flowering regulatory gene expression in rH4R17A, ISWI and SWR1 mutants.	193
Figure 4.15 Expression of flowering regulatory genes in rH4R17A, ISWI and SWR1 mutants.	194
Figure 4.16 Genome browser views of key flowering regulatory genes in rH4R17A, ISWI and SWR1 mutants.	195
Figure 4.17 GO term enrichment analysis for rH4R17A and ISWI mutants.	197
Figure 4.18 Expression of DNA damage response, cell cycle, and anthocyanin pathway regulatory genes in rH4R17A, ISWI and SWR1 mutants.	198
Figure 4.19 TE expression in rH4R17A, ISWI, and SWR1 mutants.	200
Figure 4.20 Spearman correlation of MNase-seq replicates.	202
Figure 4.21 Determination of the function of H4R17 on regulating nucleosome positioning.	203

Figure 4.22 Nucleosome occupancy profiles at individual genes.	206
Figure 4.23 Dimensionality reduction and clustering analysis of nucleosome occupancy profiles at all protein-coding genes.	208
Figure 4.24 Clustering analysis of nucleosome occupancy profiles at active genes and inactive genes in Col and rH4R17A plants.	212
Figure 4.25 Clustering analysis of nucleosome occupancy profiles at active genes. ...	213
Figure 4.26 Clustering analysis of nucleosome occupancy profiles at inactive genes.	214
Figure 4.27 Dimensionality reduction of nucleosome positioning profiles at differentially expressed genes.....	216
Figure 4.28 Conservation of ISWI proteins.	218
Figure 4.29 Homology model of <i>Arabidopsis thaliana</i> CHR11.....	220
Figure 4.30 Model for the regulation of flowering and plant development by H4R17. ...	222
Supplemental Table 1 Selected biological pathways identified with GO term enrichment analysis in rH4R17A-1 mutants.	260
Supplemental Table 2 Selected biological pathways identified with GO term enrichment analysis in rH4R17A-2 mutants.	262
Supplemental Table 3 Selected biological pathways identified with GO term enrichment analysis in <i>arid5</i> mutants.	266
Supplemental Table 4 Selected biological pathways identified with GO term enrichment analysis in <i>rtt1/2</i> mutants.	270
Supplemental Table 5 Selected biological pathways identified with GO term enrichment analysis in <i>chr11/17</i> mutants.	272
Supplemental Table 6 Cloning and PCR Primers.....	279
Supplemental Table 7 Statistics for mapping and coverage of the NGS data.	280

List of Abbreviations

- 53BP1: p53-Binding Protein
- AP1/2: Apetala1/2
- APL: Altered Phloem Development
- AGO9: Argonaute 9
- ARP6: Actin Related Protein 6
- ATAC-seq: Assay for Transposase-Accessible Chromatin with high-throughput sequencing
- ATM: Ataxia-Telangiectasia Mutated
- ATP: Adenosine Triphosphate
- ATX1/2: Arabidopsis Trithorax 1/2
- ATXR5/6: Arabidopsis Trithorax-Related Protein 5/6
- bp: base pairs
- BARD1: BRCA1-Associated RING Domain Protein 1
- BRCA1: Breast Cancer Susceptibility 1
- CAF1: Chromatin Assembly Factor 1
- Cas9: CRISPR-Associated Protein 9
- CDF: Cycling DOF Factor
- CG: cytosine followed by guanine
- CHD: Chromodomain Helicase DNA-binding
- CHR11/17: Chromatin Remodeling 11/17
- ChIP-seq: Chromatin Immunoprecipitation followed by sequencing
- CLF: Curly Leaf
- CO: Constans
- Col: Columbia ecotype
- COP1: Constitutively Photomorphogenic

CDD: CHR11/17-DDR1/3/4/5-DDW1
CDM: CHR11/17-DDP1/2/3-MSI3
CRA: CHR11/17-RLT1/2-ARID5
CRISPR: Clustered Regularly Interspaced Short Palindromic Repeats
CYCD3;2: Cyclin D3;2
DAPI: 4',6-diamidino-2-phenylindole
DDM1: Decrease in DNA Methylation 1
DDP1/2/3: DDT-PHD Protein 1/2/3
DDR1/3/4/5: DDT-Related Protein 1/3/4/5
DDW1: DDT-WAC Protein 1
DEL3: DP-E2F-Like Protein 3
DNA: Deoxyribonucleic Acid
DNMT1: DNA Methyltransferase 1
DSB: Double-Strand Break
EAF3: ESA1 Associated Factor 3
EDTA: ethylenediaminetetraacetic acid
EFM: Early Flowering MYB Protein
ELF6: Early Flowering 6
FAS1: Fasciata 1
FBL17: F Box-Like 17
FKF1: Flavin-Binding, Kelch Repeat, F-Box 1
FISH: Fluorescent *In Situ* Hybridization
FLC: Flowering Locus C
FLK: Flowering Locus KH Domain
FRET: Fluorescence Resonance Energy Transfer
FRI: Frigida

FUL: Fruitfull

FT: Flowering Locus T

GA: Gibberellic Acid

GI: Gigantea

GO: Gene Ontology

gRNA: guide RNA

HAM1/2: Histone Acetyltransferase of The MYST Family 1/2

HIRA: Histone Regulator A

HJURP: Holliday Junction Recognition Protein

HP1: Heterochromatin Protein 1

HR: Homologous Recombination

HU: Hydroxyurea

indel: insertion-deletion mutation

ISWI: Imitation Switch

JMJ13: Jumonji 13

KRP: KIP-Related Protein

LD: Long-day

LFY: Leafy

LHP1: Like Heterochromatin Protein 1

LOS: Low Sporulation

LRS: Loss of rDNA Silencing

MAF2/3: Mads Affecting Flowering 2/3

MCM: Minichromosome Maintenance Complex

MET1: Methyltransferase 1

MMS: Methyl Methanesulfonate

MNase-seq: Micrococcal Nuclease digestion followed by high-throughput sequencing

miR156/172: microRNA 156/172

mRNA: messenger Ribonucleic Acid

MS: Murashige-Skoog

MSI1: Multicopy Suppressor of IRA 1

MYBD: MYB-Like Domain Transcription Factor

MYBH: MYB Hypocotyl Elongation-Related

NegC: Negative Regulator of Coupling

NFR: Nucleosome-Free Region

NHEJ: Non-Homologous End Joining

NuA4: Nucleosome Acetyltransferase of Histone H4

NURF: Nucleosome Remodeling Factor

NuRD: Nucleosome Remodeling and Deacetylase

ORC: Origin Recognition Complex

PCNA: Proliferating Cell Nuclear Antigen

PCR: Polymerase Chain Reaction

PRC2: Polycomb Repressive Complex 2

PRMT5/7/10: Protein Arginine Methyltransferase 5/7/10

PTM: Post-Translational Modification

RAC: Overreplication-Associated Center

RAD17: Radiation Sensitive 17

RBR1: Retinoblastoma-Related 1

rDNA: ribosomal DNA

REF6: Relative of Early Flowering 6

RNA-seq: whole-transcriptome analysis via RNA sequencing

RSC: Remodels the Structure of Chromatin

RT-qPCR: quantitative Reverse Transcription PCR

SD: Short-day

SIR: Silent Information Regulator

SKB1: SHK1 Binding Protein 1

SLBP: Stem–Loop Binding Protein

SMZ: Schlafmütze

SOC1: Suppressor of Overexpression Of Constans 1

SPA: Suppressor Of PHYA-105

STUbL4: SUMO-Targeted Ubiquitin E3 Ligase 4

SPL: Squamosa-Promoter Binding Protein-Like

SUV39H1: Suppressor of Variegation 3-9 Homolog 1

SVP: Short Vegetative Phase

SWI/SNF: Switch/Sucrose Non-Fermentable

SWR1: SWI/SNF2-Related 1

t-SNE: t-Distributed Stochastic Neighbor Embedding

T1/ T2/ T3: First-Generation Transformant/ Second-Generation Transformant/ Third-Generation Transformant

TE: Transposable Element

TIRR: Tudor-Interacting Repair Regulator

TPM: Transcripts Per Million

TSI: Transcriptionally Silent Information

TSS: Transcription Start Site

TTS: Transcription Termination Site

UHRF1: Ubiquitin Like with PHD And Ring Finger Domains 1

VIM1: Variant In Methylation 1

VIN3: Vernalization Insensitive 3

VRN1/2: Vernalization 1/2

Chapter 1: Introduction

Author's Note: Portions of this chapter have been submitted for publication as:

Emma Tung Corcoran¹, Chantal LeBlanc¹, Mia Arias Tsang¹, Anthony Sarkiss¹, Yuzhao Hu², Ullas V. Pedmale², and Yannick Jacob¹. 2021. Systematic histone H4 replacement in *Arabidopsis thaliana* reveals a role for H4R17 in regulating flowering time.

Affiliations:

1. Yale University, Department of Molecular, Cellular and Developmental Biology, Faculty of Arts and Sciences; 260 Whitney Avenue, New Haven, Connecticut 06511, United States.

2. Cold Spring Harbor Laboratory; Cold Spring Harbor, New York 11724, United States.

Epigenetics

The modern study of inheritance began in the 19th century with the works of scientists including Charles Darwin, Jean-Baptiste Lamarck, and Gregor Mendel. In 1859, Charles Darwin introduced the theory of evolution by natural selection, arguing that small variation between organisms arises by chance and more favorable random variations are inherited to accumulate in a population. When Darwin introduced his theory of evolution by natural selection, it was quite controversial and took more than half a century to become widely supported. Many contemporaneous critics of Darwin's work turned to an earlier theory—the theory of inheritance of acquired characteristics—described in 1809 by Jean-Baptiste Lamarck as an alternative. While Lamarck also argued that variation is inherited, he contended that this variation is acquired through use or disuse rather than arising stochastically.

Support for Darwin's work began to grow in the 20th century as scientists rediscovered another piece of work published less than one decade later than Darwin's *On the Origin of Species*, in which he first introduced his theory of evolution by natural selection. In 1866, Gregor Mendel laid out a genetic model now known as Mendelian inheritance based on his work studying the inheritance of pea flower color, but similar to Darwin, his work received little support from the scientific community when it was first released. In the late 19th and early 20th century, several biologists and statisticians merged Mendel's and Darwin's theories to create the framework for the modern study of genetics. As molecular biology initiated as a field in the mid-20th century, deoxyribonucleic acid (DNA) was established as the carrier of hereditary information and mutations in an organism's DNA sequence were identified as the basis for how random variations arise. Elucidating the mechanisms by which mutations cause variations in phenotype and how they are passed onto subsequent generations became two of the foundational themes in modern genetics.

In recent decades, a new field has developed studying heritable changes in gene expression or cellular phenotype caused by mechanisms other than changes in the underlying DNA sequence. This field was termed epigenetics due to its focus on factors “on top of” traditional genetics. When C.H. Waddington first coined the term epigenetics in the 1940s, he used it to describe how the interactions between the environment and an organism’s genes lead to phenotypes during development (Tronick and Hunter, 2016; Waddington, 1942). Since this initial usage, however, epigenetics as a term has undergone changes in its widely used definition. Most broadly, epigenetics has been defined as “the structural adaptation of chromosomal regions so as to register, signal or perpetuate altered activity states” (Bird, 2007), while a recent consensus definition from 2009 more narrowly defined an epigenetic trait as “a stably heritable phenotype resulting from changes in a chromosome without alterations in the DNA sequence” (Berger et al., 2009).

On an organismal level, the field of epigenetics has elucidated how multicellular organisms such as humans can generate many different types of cells all from the same genome. In fact, Waddington initially invented the word epigenetics by fusing the words “genetics” and “epigenesis”—the process by which the adult organism develops from the embryo through a sequence of steps of differentiation and organ formation—as he was referring to the genetic control of developmental processes (Bonasio et al., 2010; Waddington, 1942). Every cell in a human’s body is derived from the same initial cell, a one-cell embryo, which eventually differentiates into stable and inherited cell types. Epigenetic mechanisms explain how the exact same genome can give rise to vastly different and specialized cell types, each with their own specific and stable phenotypes, during embryonic development (Cheedipudi et al., 2014).

On a population level, interesting questions have arisen about the nature of transgenerational epigenetic inheritance and the origination of epigenetic variation

between individuals. Many studies in the past two decades have provided substantial support to the idea that an individual's environment can influence the epigenetic regulation occurring within their cells. In certain cases using model systems, epigenetic changes resulting from an environmental stimulus have been demonstrated to be heritable in subsequent generations. For example, traumatic exposure to the odor acetophenone in mice was shown to be inherited via a transgenerational epigenetic mechanism (Dias and Ressler, 2014). Based on similar studies showing epigenetic mechanisms of heritability, some have noted similarities between Lamarck's theory of evolution—that heritable traits can be acquired through use or disuse—and transgenerational epigenetic inheritance, and some have even argued that epigenetics has provided evidence to support a Lamarckian theory of evolution (Jablonka et al., 1998). While it has been disputed how closely epigenetic mechanisms of inheritance resemble Lamarck's theory, the field of epigenetics has raised very interesting questions about the relationship between an individual's environment and the traits they pass on to subsequent generations.

Although these questions are outside the scope of this dissertation, they are nonetheless important to consider as background to why scientists and non-scientists alike are so fascinated with the field of epigenetics in the current moment. Discussions of epigenetics have reached the mainstream, in large part due to the perception that epigenetics provides a way for individuals to influence not only their own genetics, but the genetics of their descendants as well. While this portrayal of epigenetics may be sensationalist, epigenetic regulation has undeniable impacts on an organism's health and fitness. One notable field in which there is a strong drive to develop epigenetic therapies is in the treatment of human cancers, and several drugs targeting epigenetic pathways have been developed with promising clinical outcomes (Allis and Jenuwein, 2016). Additionally, due to the high level of conservation of many epigenetic

mechanisms, studying epigenetics can illuminate fundamental biological processes that occur in eukaryotic organisms as divergent as plants and animals. Therefore, elucidating how epigenetic mechanisms work to maintain the health and fitness of an organism has implications not just for improving the quality of life for humans, but also for generating better crops to improve agriculture and food yield for our society.

This chapter will introduce background and motivation for the work described in this dissertation. First, this chapter describes how DNA is organized into chromatin in the nuclei of eukaryotic cells. Second, this chapter details how histone post-translational modifications regulate processes occurring on chromatin and how nucleosomes are positioned within chromatin. Third, this chapter describes how epigenetic information is maintained, and how epigenetic signals regulate processes such as genome stability and flowering time. Finally, this chapter highlights and assesses previous work studying the functional significance of histones, and describes the motivation for the research performed in chapters 2-4, as well as the goals of this dissertation.

Chromatin

In eukaryotic cells, genomic DNA is wrapped around histone octamers—composed of the positively charged core histones H2A, H2B, H3 and H4—to form nucleosomes, the repeating unit of chromatin (Luger et al., 1997). Each histone contains a histone fold domain, which allows the histones to associate into the heterodimers H2A-H2B and H3-H4, and comprises approximately 75% of the histone protein mass (Arents et al., 1991; Zheng and Hayes, 2003). In addition to the structured histone fold domains, each core histone also contains an unstructured N-terminal and/or C-terminal tail—comprising approximately 25-30% of the histone protein mass—that protrudes from the nucleosome (Zheng and Hayes, 2003). The histone N- and C-terminal tails are thus accessible to chromatin-binding proteins and form targets for a variety of post-

translational modifications (PTMs), encompassing a wide array of functions in the nucleus.

In the absence of DNA, the H3-H4 heterodimers associate to form a tetramer while the H2A-H2B heterodimers remain independent. In the presence of DNA, the H3-H4 tetramer and H2A-H2B heterodimers form the histone octamer (Khorasanizadeh, 2004). Approximately 146 base pairs (bp) of negatively charged DNA is wrapped around the histone octamer, with between 10-60 bp of linker DNA connecting each nucleosome, to form a repeating structure called the 10-nm fiber (Maeshima et al., 2020) (Figure 1.1). Additional non-histone proteins, such as heterochromatin protein 1 (HP1), and the linker histone H1 associate with the 10-nm fiber to organize the DNA into higher-order chromatin structures (Machida et al., 2018; Woodcock et al., 2006).

Active research in the past decade has evolved the understanding of how higher-order chromatin structures are organized within a live nucleus. The long-standing supposition that the 10-nm fiber organizes into a uniform 30-nm fiber was recently challenged in favor of a new, less ordered model of chromatin organization, supported by comprehensive data from diverse techniques including electron microscopy and super-resolution imaging (Nozaki et al., 2017; Ou et al., 2017; Ricci et al., 2015). These data support the assessment that the 10-nm fiber forms irregular loop structures, organized by cohesin complexes, and these loops subsequently form chromatin domains in animals (Nozaki et al., 2017). In contrast, plants seem to lack large interactive chromatin domains, but do appear to have small interactive regions between heterochromatin contacts (Feng et al., 2014a). Despite the enormous degree of compaction that the DNA undergoes in order to form chromatin, chromatin as a structure remains very dynamic and mobile, while factors such as the cohesin complex and the transcriptional machinery are able to constrain local chromatin movement (Nagashima et al., 2019).

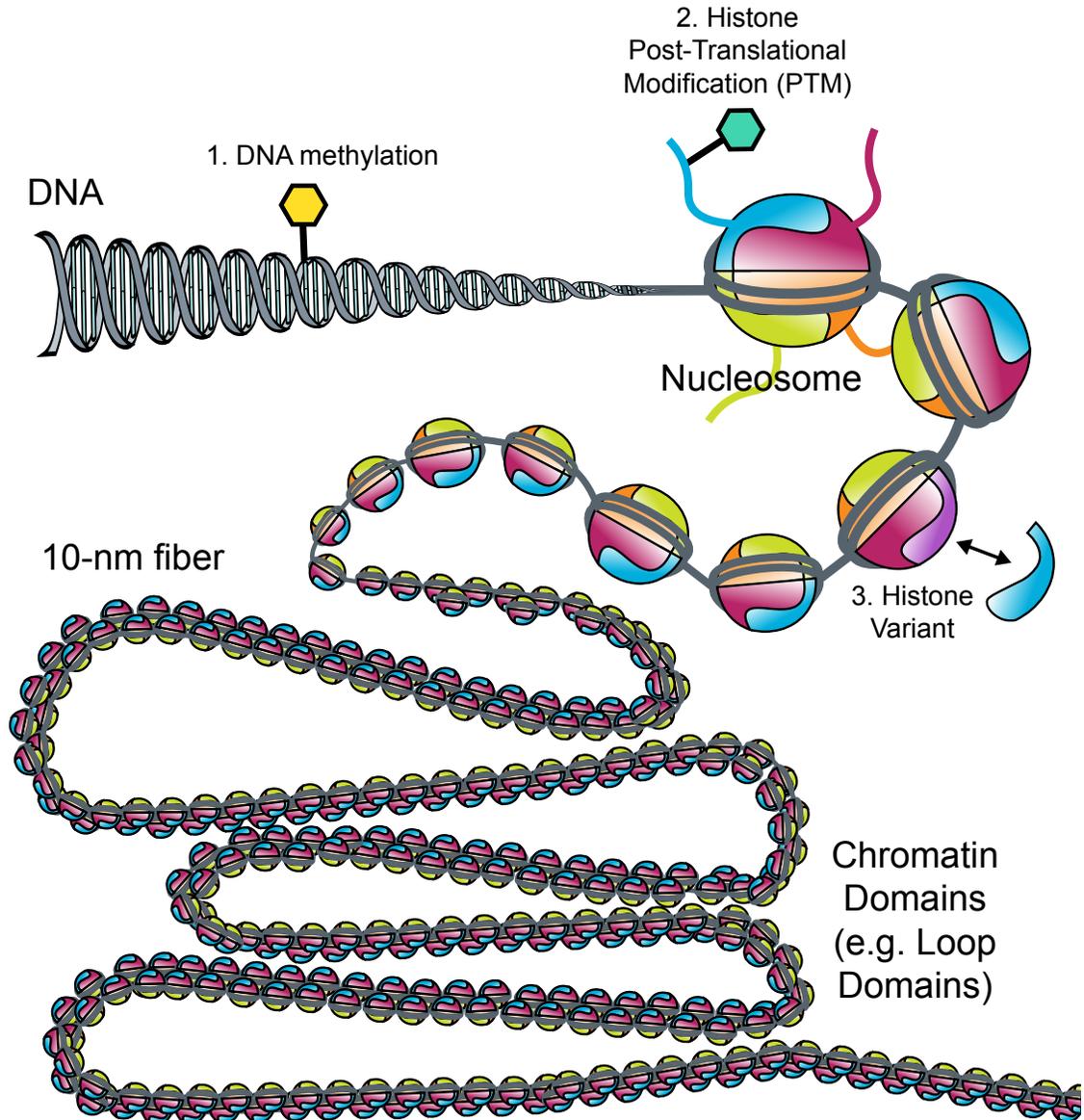


Figure 1.1 Simplified scheme of chromatin organization within the cell nucleus. DNA is wrapped around the histone octamer to form the nucleosome. Each nucleosome is connected with 10-60 bp of linker DNA to comprise the 10-nm fiber. The 10-nm fiber organizes into higher order chromatin structures, such as loops and chromatin domains (Maeshima et al., 2020; Nozaki et al., 2017). Three sources of variation within chromatin are shown: 1.) DNA methylation 2.) histone post-translational modifications (PTMs) and 3.) histone variants.

Chromatin is broadly divided into two categories: less condensed, more accessible euchromatin and more condensed, less accessible heterochromatin. Heterochromatin and euchromatin were first differentiated into these two categories in 1928 by Emil Heitz, who observed that chromosomes were composed of regions not stained after telophase, indicating decondensation (euchromatin), and regions that remained stained throughout the cell cycle, indicating that they remained condensed (heterochromatin) (Allshire and Madhani, 2018). Heitz also noted that some heterochromatic regions are only stained in certain cells, while he observed that other heterochromatic regions are always stained. These two types of heterochromatin were later named facultative heterochromatin and constitutive heterochromatin, respectively. Constitutive heterochromatin remains condensed throughout the cell cycle and preferentially assembles at repetitive elements such as transposons, while facultative heterochromatin often assembles at developmentally regulated genes and can adapt an open conformation in response to cellular signals and gene activity (Grewal and Jia, 2007; Wang et al., 2016).

Since these early experiments dividing chromatin into three different states, the identification of several sources of variation in the nucleosome coupled with new genome-wide sequencing techniques have allowed researchers to refine their view of chromatin state and distinguish many more specialized categories. In many eukaryotic species, DNA methylation contributes to the formation of heterochromatic regions and is associated with transcriptional repression (Baubec and Defossez, 2020; Finnegan et al., 1998). High-throughput bisulfite sequencing experiments have revealed genome-wide DNA methylation patterns, with characteristic high levels of DNA methylation at transposons and repeats that likely suppress the expression of these elements, while the role of DNA methylation in the context of genes appears more nuanced (Jones, 2012; Zemach et al., 2010). For example, DNA methylation around the transcription start site

(TSS) blocks transcription initiation, while DNA methylation in the gene body does not seem to block transcription elongation and may even have a positive effect on elongation.

Regions of chromatin also vary in their incorporation of histone variants. Whereas the assembly of the canonical histones into the nucleosome is coupled to sites of DNA replication, histone variants such as H2A.Z and H3.3 can replace individual histones within the nucleosome throughout the cell cycle in a replication-independent fashion (Talbert and Henikoff, 2017). Some histone variants can differ substantially in structure from the canonical histones they replace, while others may vary by only a few amino acids, but in either case the incorporation of a histone variant to a region of chromatin can have significant consequences on the chromatin state. For example, the centromeric H3 variant (called CENPA in vertebrates and CENH3 in plants) allows tighter DNA wrapping and forms the foundation of centromeric chromatin, allowing it to be recognized by components of the kinetochore (McKinley and Cheeseman, 2016). Moreover, the H2A variant H2A.Z is present at promoters and is thought to recruit Pol II to poise genes for transcriptional activation (Adam et al., 2001; Talbert and Henikoff, 2017).

In addition to DNA methylation status and the presence of histone variants, another major factor that distinguishes chromatin states is the presence of chromatin-binding proteins and distinct histone PTMs, the latter of which is discussed in greater detail in the next section (Sequeira-Mendes et al., 2014). Due to the accessibility of the histone N-terminal tails in chromatin, these domains form a major target for chromatin-binding proteins and the addition of PTMs, although PTMs can also be added to the core histone fold domain. Some PTMs are specific to certain histone variants, while others can be added to both a histone variant and its canonical counterpart. Additionally, histone PTMs and DNA methylation can influence each other's addition during

development (Cedar and Bergman, 2009). The interplay between all of these sources of variation creates a rich tapestry of epigenetic regulation throughout the genome.

Histone post-translational modifications

The establishment of the connection between chromatin state and transcriptional activity marked some of the first experiments in the modern study of epigenetics, beginning in the late 20th century. In 1982, Vavra et al. reported that the ciliate *Tetrahymena* has a transcriptionally active macronucleus, with high levels of histone acetylation, and a transcriptionally inactive micronucleus, with nearly undetectable levels of histone acetylation (Vavra et al., 1982). In 1996, Brownell et al. identified the enzyme responsible for histone acetylation in *Tetrahymena* and demonstrated that this histone acetyltransferase was highly homologous to the yeast transcriptional activator Gcn5 (Brownell et al., 1996). Moreover, in the same year, Taunton et al. demonstrated that a mammalian histone deacetylase was related to the yeast transcriptional regulator Rpd3p (Taunton et al., 1996). With these experiments, direct evidence was provided linking histone acetylation to gene expression.

A mechanistic understanding of how acetylation of one specific histone residue can impact gene expression arose when more structural and biochemical studies of the nucleosome were performed. When Luger et al. solved the crystal structure of the nucleosome core particle at 2.8 Å resolution in 1997, they noted that the basic patch on the H4 N-terminal tail contacts an acidic patch on the H2A-H2B surface of the adjacent nucleosome (Luger et al., 1997). This observation led to the hypothesis that acetylation of lysine 16 on H4 (H4K16ac) neutralizes the positive charge of the lysine residue to directly block nucleosome-nucleosome interaction and modulate chromatin structure. In 2006, Shogren-Knaak et al. generated histone H4 homogeneously acetylated at K16 and found that the formation of higher order chromatin structure was indeed impeded

after the acetylated histones' incorporation into nucleosomal arrays (Shogren-Knaak et al., 2006). This experiment confirmed that a histone PTM—specifically H4K16ac—can influence higher order chromatin structure. Moreover, it provided a mechanistic explanation for the transcriptional activation linked to H4K16ac: that the chromatin decompaction caused by H4K16 acetylation increases the accessibility of factors that promote transcription.

Since these fundamental experiments demonstrating the link between histone acetylation and transcriptional activity, there has been an explosion in the identification of histone PTMs and the analysis of their functions. Many different classes of histone PTMs have been identified, including acetylation, methylation, phosphorylation, ubiquitinylation, sumoylation, ADP ribosylation, deamination, propionylation, and butyrylation (Kebede et al., 2015; Kouzarides, 2007). These PTMs can be added in different combinations to residues on both the histone tails and bodies, and even more complexity derives from the fact that some modifications can be added in different forms. For example, methylation can be added as mono-, di-, or trimethylation at lysine residues and mono- or dimethylation at arginine residues. As in the case of H4K16ac, some histone PTMs have been shown to have a direct effect on chromatin compaction. Notably, H4K20me₂ and H4K20me₃ were shown to have a contrasting role to H4K16ac and enhance *in vitro* chromatin condensation (Lu et al., 2008). However, unlike these PTMs, which were demonstrated to have a direct impact on chromatin structure, most histone PTMs have been shown to act in an indirect fashion to modulate chromatin regulation.

Histone PTMs can be recognized either individually or in a combinatorial manner to recruit “reader” proteins, such as chromatin remodelers and transcription factors, and indirectly achieve various outcomes within a cell (Lawrence et al., 2016). For example, the PTM H3K4me₃, which is associated with active promoters, has been shown to lead

to transcriptional activation by binding subunits of the transcription factor TFIID complex (Heintzman et al., 2007; Vermeulen et al., 2007). Additionally, histone PTMs can block the access of chromatin-binding proteins and chromatin remodeling complexes to regions of chromatin. In addition to its role in binding TFIID, methylation of H3K4 has also been shown to disrupt the binding of the Nucleosome Remodeling and Deacetylase (NuRD) repressor complex, consistent with this modification's association with regions of active chromatin (Zegerman et al., 2002).

Histone PTMs are also capable of recruiting epigenetic “writer” or “eraser” proteins, which can add or remove epigenetic modifications from a region of chromatin, respectively (Figure 1.2). In this way, epigenetic modifications such as histone PTMs and DNA methylation are reversible and intrinsically dynamic. In the fission yeast *Schizosaccharomyces pombe*, H3K9 methylation, a mark of constitutive heterochromatin, binds the chromodomain of the Clr4 methyltransferase complex. The writer protein Clr4 in turn methylates H3K9 of neighboring nucleosomes to spread the heterochromatic state (Zhang et al., 2008). In contrast, methylation at H3K36 has been found to suppress inappropriate transcription initiation within the coding region of genes by recruiting the Rpd3 histone deacetylase complex, an eraser of histone acetylation, through the subunit Esa1 associated factor 3 (Eaf3). The subsequent deacetylation of chromatin by Rpd3 prevents spurious intragenic transcription by inhibiting access to internal initiation sites (Carrozza et al., 2005; Joshi and Struhl, 2005; Keogh et al., 2005). Clearly, the pattern of histone PTMs within the nucleus is highly complex, and crosstalk between PTMs further complicates analysis of their function.

A recent study by Sequeira-Mendes et al. analyzing diverse genome-wide data of histone PTMs, DNA methylation, histone variants, nucleosome occupancy, and DNA sequence identified nine distinct chromatin states in the plant model system *Arabidopsis thaliana* (Sequeira-Mendes et al., 2014). Interestingly, chromatin states were shown to

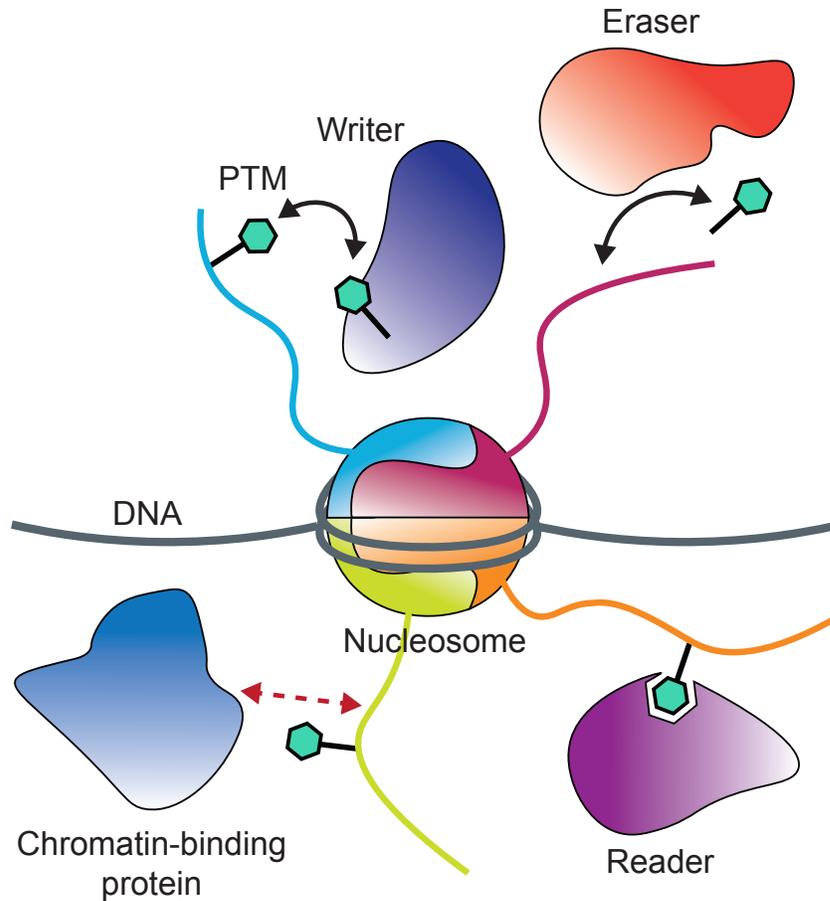


Figure 1.2 Epigenetic modifications are reversible and dynamic. Three classes of proteins—reader, writer, and eraser proteins—recognize, add, and remove epigenetic modifications from regions of chromatin. Reader proteins recognize epigenetic modifications such as histone PTMs to achieve various outcomes within a cell (e.g., transcriptional activation). Writer and eraser proteins add or remove epigenetic modifications, respectively. In addition to recruiting reader proteins, epigenetic modifications are also capable of blocking the access of chromatin-binding proteins and chromatin remodeling complexes to regions of chromatin.

have differential preferences in how they associate with each other in the linear organization of the genome. For example, two states corresponding to constitutive heterochromatin, defined by enrichment in histone H3.1, CG methylation, H3K9me2, and H3K27me1, were identified that differ in their C+G content: one GC-rich state that predominantly corresponded to regions of pericentromeric heterochromatin (state 9), and one AT-rich state (state 8) that was often found to be interspersed within regions of state 9 heterochromatin (Figure 1.3).

Additionally, four chromatin states were identified to be enriched in genes, and these four states typically followed specific euchromatic motifs depending on the local context. Each of these four states was proposed to have its own individual functional role, as well as a combinatorial role when comprising a motif. For example, long genes were characterized by the four chromatin states in the following order: (1) chromatin around the TSS, with high levels of H3K4me2/3, H3 acetylation, H3K36me3, H2Bub, and enriched in H3.3 and H2A.Z (state 1), (2) highly accessible chromatin that colocalized with the start of coding sequences and represented a transcription elongation signature, characterized by high levels of H3K4me1/2/3, H2Bub, and H3K36me3 (state 3), (3) intragenic chromatin with prominent H3K4me1, H2Bub, and H3K36me3 (state 7), and finally (4) chromatin that colocalized with the transcription termination site (TTS), characterized by a slight enrichment of H2A.Z and H3K4me1 (state 6). It is notable that while many of the PTMs were present in multiple neighboring chromatin states, each chromatin state still presented subtle differences distinguishing it from other euchromatic states and often representing a unique component of the gene.

Finally, three chromatin states were found to be enriched in the Polycomb mark H3K27me3. H3K27me3 is classically considered a repressive PTM involved in the regulation of development, and thus enrichment of H3K27me3 alongside depletion of the

Genic euchromatin

Transcription Start Site	Transcriptional Elongation	Intragenic Chromatin	Transcription Termination Site
1	3	7	6
H3K4me2/3 H3 acetylation H3K36me3 H2Bub H3.3 H2A.Z	H3K4me1/2/3 H2Bub H3K36me3	H3K4me1 H2Bub H3K36me3	H2A.Z H3K4me1

Intergenic chromatin

Proximal promoter	Distal promoter
2	4
H3K4me2/3 H3K27me3 H3.3 H2A.Z	H3.3 H2A.Z H3K27me3

Polycomb heterochromatin

5
H3K27me3 H2A.Z H3.1

Constitutive heterochromatin

More accessible	Less accessible
8	9
H3K9me2 H3K27me1 H3.1 DNA methylation AT-rich	H3K9me2 H3K27me1 H3.1 DNA methylation GC-rich

Figure 1.3 Nine chromatin states identified in *Arabidopsis thaliana*. Each chromatin state shown with characteristic enriched signatures below (Sequeira-Mendes et al., 2014). Four genic euchromatic states correspond to the TSS (state 1), transcriptional elongation (state 3), intragenic chromatin (state 7), and the TTS (state 6). Two intergenic states correspond to proximal (state 2) and distal (state 4) promoter elements. One state corresponds to classical Polycomb-regulated heterochromatin (state 5). Finally, two states correspond to constitutive heterochromatin that differ in their GC-content and accessibility. State 8 is AT-rich and more accessible relative to the GC-rich state 9.

activating PTM H3K36me3 is a signature of facultative heterochromatin (Goodrich et al., 1997; Xiong et al., 2016). One state was denoted to be classical Polycomb chromatin, as it was depleted in all other PTMs assayed and enriched in the canonical histone H3.1 (state 5). The two other states (state 2 and state 4) displayed enrichment in the histone variants H3.3 and H2A.Z, associated with active chromatin, alongside the presence of H3K27me3. Additionally, state 2 also showed an abundance of the activating PTMs H3K4me2/3, and sequential chromatin immunoprecipitation experiments demonstrated that H3K27me3 and H3K4me3 co-occurred within a two to three nucleosome size range. The two states that showed a coexistence of active and repressive signatures were prevalently found in intergenic regions and also appeared to form a border that physically separated the facultative and constitutive heterochromatin from the euchromatic domains.

Coexistence of the activating PTM H3K4me3 and the repressive PTM H3K27me3 in the same region of chromatin suggests that similar to mammalian cells, at least a subpopulation of cells in plants contain bivalent domains (Bernstein et al., 2006). In mammalian pluripotent cells, bivalent domains are present on many developmental gene promoters and are thought to poise these genes to be rapidly activated for transcription upon the appropriate cues (Voigt et al., 2013). In addition to this identification of bivalent domains in *Arabidopsis*, in which Sequeira-Mendes et al. utilized sequencing data generated from whole seedlings, a previous study found overlap between the profiles of H3K4me3 and H3K27me3 for two root epidermis cell types (Deal and Henikoff, 2010). Although this latter study did not assess whether these two PTMs were present on the same chromatin fiber, it did provide initial evidence for bivalent domains existing in two specific cell types in plants. Further experiments studying the global distribution of bivalent domains in different cell types would serve to clarify the function of bivalent chromatin in plant systems such as *Arabidopsis*.

While the aforementioned study by Sequiera-Mendes et al. provided an expansive assessment of chromatin states across the *Arabidopsis* genome, likely much more complexity in chromatin state exists than was identified, as the authors only examined PTMs covering six distinct histone residues. This study focused on previously well-studied PTMs with established functions, but a substantially larger variety of histone PTMs exists than the nine PTMs evaluated. Therefore, opportunities remain to discover greater complexity imparted to chromatin state by other histone PTMs. Furthermore, tissue for these analyses was taken from whole seedlings, and thus the sequencing data represents an average over many different cell types. Further analysis of global PTM distribution in specific cell types using novel single-cell epigenomic techniques, as opposed to analyzing whole seedlings, could reveal more nuanced information about chromatin states in specialized tissues. Finally, it is important to consider that histone PTMs have many functions outside of transcriptional regulation that add additional elements to the analysis of chromatin state.

Nucleosome positioning

In addition to modifications and variants of the histones that comprise the nucleosome, nucleosome position and density can also affect whether proteins such as transcription factors can bind a region of DNA (Jiang and Pugh, 2009). Moreover, the presence of nucleosomes can differentially inhibit the ability of transcription factors to bind chromatin. While many transcription factors cannot occupy their target sites on nucleosomal DNA, “pioneer transcription factors” comprise a special class that can bind to nucleosomal DNA and enable the binding of other transcription factors (Zaret and Carroll, 2011). In contrast, the presence of nucleosomes blocks RNA polymerase II binding, and thus, RNA polymerase II requires a nucleosome-free region (NFR) to bind promoters and initiate transcription (Struhl and Segal, 2013; Workman and Kingston,

1998). Genome-wide nucleosome positioning studies in *Saccharomyces cerevisiae* first detected a stereotyped 5' NFR approximately 200 bp upstream of the start codon at RNA polymerase II promoters, and genomic studies of other species including *Drosophila melanogaster* and humans also identified nucleosome-free core promoter regions (Mavrich et al., 2008b; Schones et al., 2008; Yuan et al., 2005).

The 5' NFR is flanked by two well-positioned nucleosomes, called the -1 and +1 nucleosomes (Jiang and Pugh, 2009). The -1 nucleosome regulates the accessibility of promoter regulatory elements, and is evicted after pre-initiation complex formation. The +1 nucleosome is found downstream of the TSS and displays a high level of phasing, as measured by the low level of delocalization of nucleosome position (Mavrich et al., 2008a). The +1 nucleosome often contains histone variant H2A.Z, which improves the accessibility of DNA by destabilizing the H2A-H2B and H3-H4 interface to promote looser packaging of the DNA (Jiang and Pugh, 2009; Suto et al., 2000). Downstream of the +1 nucleosome, subsequent nucleosomes (+2 nucleosome, +3 nucleosome, etc.) display a successive reduction in phasing, with an increasing tendency for random nucleosome positions found further than 1 kb from the TSS (Jiang and Pugh, 2009). Finally, the 3' end of a gene typically contains a 3' NFR, corresponding to the region where transcription is terminated (Figure 1.4).

Genome-wide nucleosome positioning is determined by the combined effects of many factors, including DNA sequence, transcription factors, and nucleosome remodelers (Struhl and Segal, 2013). Nucleosomes are not highly specific for certain DNA sequences to the same degree that many DNA-binding proteins are, but they do show a preference towards some 147 bp-stretches of DNA over others (Thastrom et al., 1999). The ability of the DNA sequence to bend around the nucleosome affects nucleosome positioning, and optimally, more bendable sequences contact the histones while less bendable sequences are exposed (Struhl and Segal, 2013). When well-

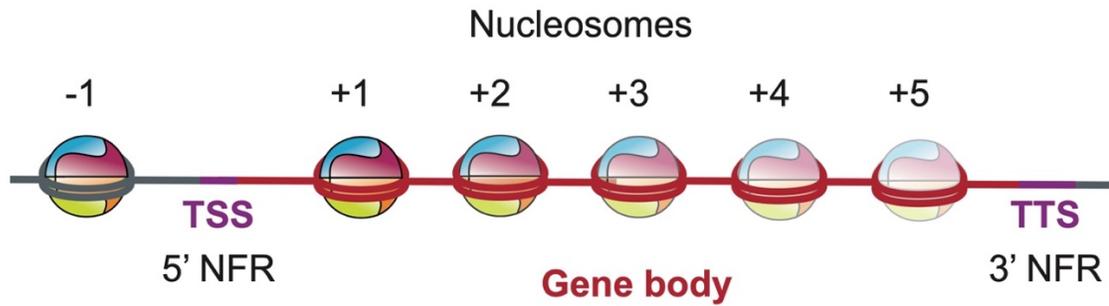


Figure 1.4 Stereotypical nucleosome positioning around gene bodies. The 5' nucleosome-free region (NFR) is flanked by the -1 and +1 nucleosomes (Jiang and Pugh, 2009). The +1 nucleosome is found downstream of the TSS and often contains histone variant H2A.Z. The +2 nucleosome, +3 nucleosome, and subsequent nucleosomes are found downstream of the +1 nucleosome and display a successive reduction in phasing. Finally, a 3' NFR is typically found at the 3' end of a gene, corresponding to the region around the TTS.

defined nucleosomal DNA sequences were analyzed beginning in the 1980s in an attempt to determine a nucleosome positioning sequence, it was found that they contained an enrichment of AA, TT, and TA dinucleotides occurring at 10 bp intervals, while GC dinucleotides also appeared at 10 bp intervals at an offset of 5 bp from the first pattern (Ioshikhes et al., 1996; Ioshikhes et al., 2006). The periodic dinucleotide patterns are thought to contribute to the rotational phasing of the nucleosome by altering the major groove of the DNA to facilitate DNA wrapping around the histone core (Struhl and Segal, 2013). However, this preference for DNA sequence remains just one of many factors that determine nucleosome positioning, as the enrichment of these periodic dinucleotide patterns in nucleosomal DNA sequences is modest and predominantly applies to the -1 and +1 nucleosomal sequences (Mavrigh et al., 2008a).

The 5' NFR also shows DNA sequence-specificity and is often enriched in poly(dA:dT) tracts (Anderson and Widom, 2001; Radwan et al., 2008). Poly(dA:dT) tracts are intrinsically rigid and disfavor nucleosome formation by decreasing the stability of DNA wrapping. While *in vitro* experiments mapping the assembly of nucleosomes on purified genomic DNA reconstituted nucleosome depletion at many promoter and terminator regions, nucleosomes assembled *in vitro* often did not show the same pattern of translational positioning as observed *in vivo* (Kaplan et al., 2009; Zhang et al., 2009). Therefore, while DNA sequence appears to play a significant role in governing nucleosome density (the average amount of nucleosomes on a given region of DNA), the translational positioning (the precise position of an individual nucleosome on a DNA sequence) of nucleosomes does not seem to primarily be determined by DNA sequence (Zhang et al., 2009).

The translational positioning of the +1 nucleosome and nucleosome spacing constraints are two critical factors in the positioning of downstream nucleosomes (Struhl and Segal, 2013). Based on evidence including the strong relationship between the

position of the +1 nucleosome and the TSS, transcription initiation has been proposed to be crucial in establishing the translational positioning of the +1 nucleosome, with the pre-initiation complex in particular serving as a major candidate to perform that role (Zhang et al., 2009). Additionally, the elongating RNA polymerase II machinery has been suggested to play a role in establishing the pattern of downstream nucleosome positioning via the recruitment of nucleosome remodeling complexes (Hughes et al., 2012). This mechanism would explain why nucleosome arrays predominantly occur in the transcribed direction, as well as the inability of nucleosome remodelers to position more downstream nucleosomes in the absence of transcriptional activity (Zhang et al., 2011).

Adenosine triphosphate (ATP)-dependent nucleosome remodelers (also known as chromatin remodelers) comprise a major regulator of nucleosome positioning by regulating nucleosome density and spacing, and ejecting nucleosomes to enable transcription factor binding to DNA (Figure 1.5). In addition to their role in regulating nucleosome positioning, nucleosome remodelers also create specialized regions where canonical histones are replaced with histone variants (Clapier et al., 2017). On the basis of the structure of their catalytic ATPases and associated subunits, nucleosome remodelers can be divided into four subfamilies: switch/sucrose non-fermentable (SWI/SNF), imitation switch (ISWI), chromodomain helicase DNA-binding (CHD), and INO80 (Tyagi et al., 2016). Most nucleosome remodelers are specialized to preferentially perform one of the three functions mentioned above (nucleosome assembly and spacing, chromatin access, and incorporation/ removal of histone variants). While the different subfamilies are not strictly divided by functionality, many remodelers of the same subfamily share the same general function (Clapier et al., 2017).

Most ISWI subfamily remodelers assemble and regularly space nucleosomes, although some ISWI remodelers such as the nucleosome remodeling factor (NURF)

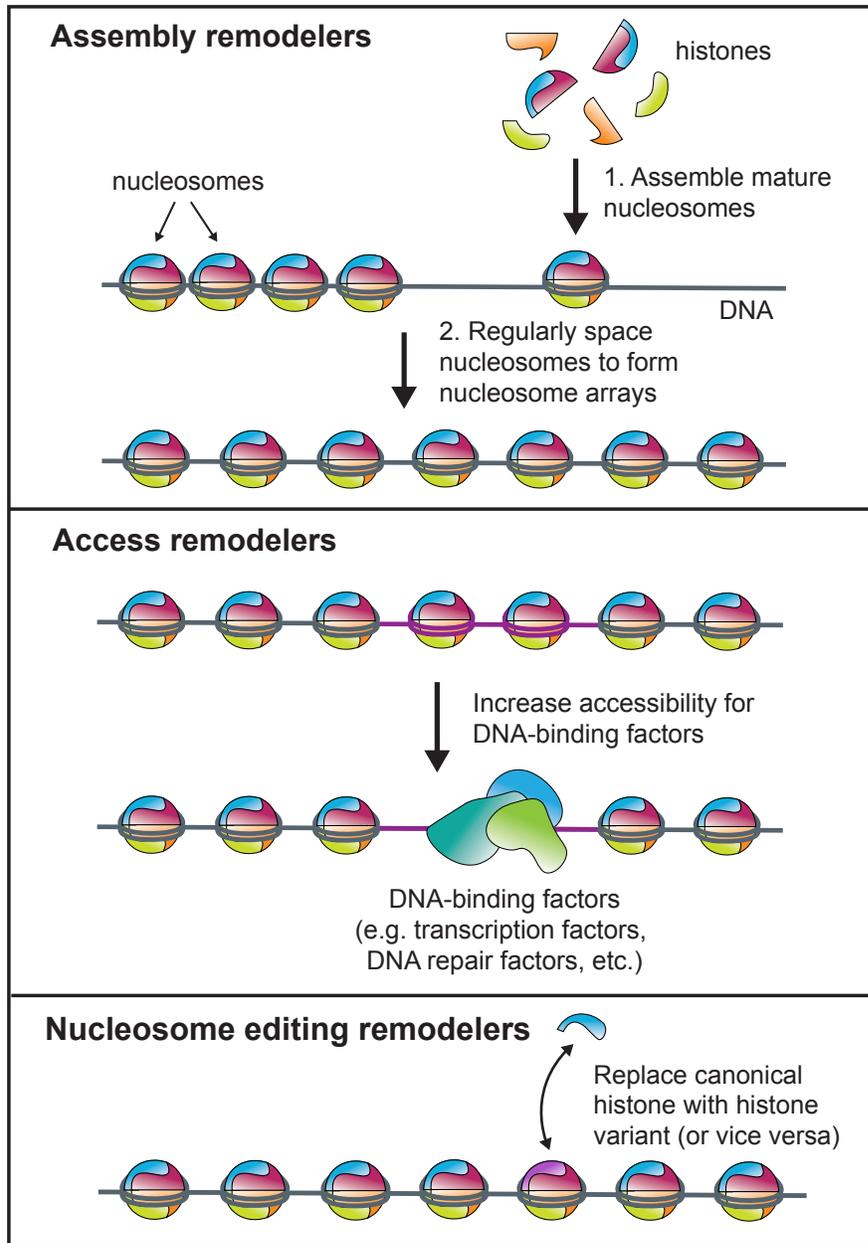


Figure 1.5 Three major classes of nucleosome remodelers. Assembly remodelers first help assemble the mature nucleosome, and then regularly space nucleosomes to form nucleosome arrays (Corona et al., 1999; Fei et al., 2015). Access remodelers increase accessibility for factors involved in DNA repair, recombination, and transcriptional activation or repression (Clapier et al., 2017). Nucleosome editing remodelers replace particular histones with canonical histones or histone variants, such as H2A.Z and H3.3 (Goldberg et al., 2010; Mizuguchi et al., 2004).

complex function to promote chromatin access (Erdel and Rippe, 2011; Langst and Becker, 2001; Xiao et al., 2001). During replication and at certain genomic locations after transcription, assembly remodelers first help the pre-nucleosome—the initial histone-DNA complex that is randomly deposited—mature into the canonical nucleosome, and then regularly space nucleosomes to form nucleosome arrays (Corona et al., 1999; Fei et al., 2015). The CHD subfamily shares many similarities to the ISWI subfamily, including structural similarities of the ATPase domains, negative regulator of coupling (NegC) domains, and DNA-binding domains that are used to measure the distance between nucleosomes (Clapier and Cairns, 2012; Hauk et al., 2010; McKnight et al., 2011). Moreover, some CHD subfamily remodelers also function as assembly remodelers similar to the ISWI subfamily, but other members of the CHD subfamily perform the other two general remodeling processes (e.g., opening up chromatin structure at promoters or depositing the histone variant H3.3) (Konev et al., 2007; Lusser et al., 2005; Murawska and Brehm, 2011).

In contrast to ISWI and CHD remodelers, SWI/SNF remodelers slide and eject nucleosomes and typically act as access remodelers, which can increase accessibility for factors involved in DNA repair, recombination, and transcriptional activation or repression (Clapier et al., 2017). Finally, many INO80 subfamily remodelers perform nucleosome editing functions by replacing particular histones with canonical histones or histone variants, such as H2A.Z and H3.3 (Goldberg et al., 2010; Mizuguchi et al., 2004). Nucleosome remodelers are often targeted to specific sites by histone PTMs or histone variants, and members of all four subfamilies contain domains that bind to histone PTMs (Clapier et al., 2017; Sukanuma and Workman, 2011). Additionally, histones can regulate the activity of nucleosome remodelers, as in the example of the histone H4 tail being shown to stimulate ISWI remodeling activity in yeast and animals (Clapier et al., 2001; Clapier et al., 2002; Dann et al., 2017; Fazzio et al., 2005; Hamiche

et al., 2001; Ludwigsen et al., 2017; Mueller-Planitz et al., 2013; Racki et al., 2014; Yan et al., 2016). Nucleosome remodelers also cooperate with chromatin modifiers and transcription factors to alter chromatin structure and gene expression (Mitra et al., 2006; Yudkovsky et al., 1999). Moreover, the interaction of certain nucleosome remodelers with the transcriptional elongation machinery appears to be vital for the efficiency of establishing nucleosome positioning in gene bodies (Gkikopoulos et al., 2011; Hughes et al., 2012).

As such, nucleosome remodelers play a major role in determining nucleosome spacing in positioned nucleosome arrays located in genes. *Saccharomyces cerevisiae* mutants lacking both Isw1 and Chd1 remodelers display a prominent loss in nucleosome positioning through coding regions, with the positions of the +3 nucleosomes onward being largely lost, while the +1 and -1 nucleosomes remain predominantly unaffected (Gkikopoulos et al., 2011). Genetic ablation of nucleosome remodelers in other species similarly affects nucleosome spacing in gene coding regions. For example, while the loss of the CHD1 remodelers Hrp1 and Hrp3 in *Schizosaccharomyces pombe* does not alter the average distance between nucleosomes outside of their genomic context (i.e., the nucleosome spacing), it does compromise the formation of genic arrays due to the unlinking of these arrays from the TSS (Hennig et al., 2012; Pointner et al., 2012; Shim et al., 2012). Similarly, *Arabidopsis thaliana* mutants lacking the two ISWI proteins CHROMATIN REMODELING 11 (CHR11) and CHR17 show a loss in the evenly spaced nucleosome pattern found in gene bodies, while the nucleosome density is unaffected (Li et al., 2014). Moreover, loss of *ISWI* in *Drosophila melanogaster* causes global changes in nucleosome spacing as well as a reduction in chromatin-bound histone H1 levels (Corona et al., 2007; Sala et al., 2011).

In contrast to their role in positioning more downstream nucleosomes in gene bodies, nucleosome remodelers are able to reconstitute some facets of *in vivo*

nucleosome positioning in the absence of transcription. The addition of yeast cell-free extract and ATP to purified histones and genomic DNA recapitulates nucleosome depletion at promoters, as well as +1 and -1 nucleosomes flanking this region, although the precise positioning of these two nucleosomes poorly matches the *in vivo* data (Struhl and Segal, 2013; Zhang et al., 2011). This reconstitution, which does not require other nucleotide triphosphates, argues against a transcription-dependent mechanism for generating NFRs and points instead to ATP-dependent nucleosome remodeling complexes. Genetic experiments have identified mutants lacking specific nucleosome remodelers that show altered nucleosome positioning near promoters. For example, experiments analyzing yeast mutants lacking the ISWI subfamily remodeler *Isw2* indicate that this enzyme functions near promoter regions to direct nucleosomes to DNA sequences that are intrinsically inhibitory to nucleosome formation (Whitehouse et al., 2007; Whitehouse and Tsukiyama, 2006). Additionally, yeast mutants lacking the remodels the structure of chromatin (RSC) complex display a diminishment in the size of NFRs due to nucleosomes repositioning to more favorable DNA sequences (Hartley and Madhani, 2009).

While many aspects of nucleosome positioning seem to be conserved between single-celled eukaryotic species and multicellular species, there do also appear to be differences. For example, the linker histone H1 does not have a major effect on nucleosome spacing in *Saccharomyces cerevisiae* due to its low endogenous levels, but it does seem to play a role in determining nucleosome spacing in metazoans and plants (Freidkin and Katcoff, 2001; Struhl and Segal, 2013). Depletion of histone H1 in mouse embryonic stem cells and chicken lymphocyte cells decreases global nucleosome spacing, and genome-wide studies of nucleosome spacing and H1 occupancy in *Drosophila melanogaster* embryos found correlations between H1 placement and nucleosome spacing, as well as potential relationships between H1 occupancy and NFR

formation (Fan et al., 2005; Hashimoto et al., 2010; Hu et al., 2018). Moreover, deletion of histone H1 in *Arabidopsis thaliana* causes greater variability in nucleosome distribution in heterochromatin, suggesting that histone H1 enforces the regularity of nucleosome spacing in heterochromatic regions in this organism (Rutowicz et al., 2019). Regardless of species-specific differences on various mechanisms of determining nucleosome positioning, nucleosome positioning in all cases is vital for the regulation of diverse processes in eukaryotes, with its impact on gene expression being the most well-studied.

Maintenance of epigenetic signals

While the precise position of nucleosomes must be reset after every round of DNA replication, with the DNA sequence, transcription factors, and nucleosome remodeling complexes together ensuring the correct positioning pattern, certain aspects of chromatin are directly inherited to daughter cells after replication to maintain a memory of the parent chromatin. A well-established way that this process occurs is via the semiconservative replication of DNA, in which each daughter cell receives one DNA strand from the original double stranded DNA helix (Meselson and Stahl, 1958). In addition to the genetic information in the cell, however, epigenetic information can also be transmitted through replication and cell division to maintain epigenetic states. As mentioned previously, there have been several definitions of epigenetics proposed that vary in the degree to which they include heritability as a requirement (Bonasio et al., 2010). In its loosest definition, epigenetics encompasses aspects of chromatin biology, such as histone PTMs, DNA methylation, RNA interference, and nucleosome positioning, even if these elements are not heritable (Bird, 2007). However, many chromatin modifications are heritable—mitotically and/or meiotically—and thus provide a mechanism for the maintenance of epigenetic signals. These heritable features of

chromatin state comprise true epigenetic information in the strictest definition of the term (Probst et al., 2009).

Symmetrical CG (cytosine followed by guanine) methylation represents one example of a true epigenetic modification that has a clear mechanism for its faithful propagation. After DNA replication, the two daughter helices are in a hemimethylated state, as they each received one of the methylated parental DNA strands. In mammals, DNA methyltransferase 1 (DNMT1) preferentially targets hemimethylated DNA sites following DNA replication to methylate the unmethylated cytosine in the newly synthesized strand and maintain the methylation pattern (Greenberg and Bourc'his, 2019; Hermann et al., 2004). To further target its specificity to hemimethylated DNA, DNMT1 acts in concert with Ubiquitin Like With PHD And Ring Finger Domains 1 (UHRF1), an E3 ubiquitin-protein ligase (Bostick et al., 2007; Sharif et al., 2007). UHRF1 specifically binds hemimethylated CG dinucleotides and methylated H3K9, and recruits and relieves the autoinhibition of DNMT1 (Ishiyama et al., 2017; Rothbart et al., 2012). Similarly to mammals, the DNMT1 ortholog METHYLTRANSFERASE 1 (MET1) serves as the maintenance CG methylator in plants, and as in mammals, is thought to be recruited by UHRF1 orthologs, the VARIANT IN METHYLATION (VIM) proteins (Woo et al., 2008; Zhang et al., 2018). In this way, CG methylation is copied to daughter cells with high fidelity.

The inheritance of DNA methylation is one way in which parent cells propagate chromatin state to their daughter cells, but chromatin state is also propagated through the transmission of histone PTMs and chromatin-binding proteins. Unlike DNA, nucleosomes are not thought to be inherited semiconservatively, and thus the mechanism by which histone PTMs are propagated is less clear. Following DNA replication, during which nucleosomes become disrupted, current models propose that parental histones are randomly distributed among new histones along both sides of the

replication fork to form three types of nucleosomes: parental, mixed, and new (Almouzni and Cedar, 2016). One way in which the maintenance of histone PTMs on nascent chromatin can be achieved is by using neighboring histones, which contain the parental PTMs, as a template (Probst et al., 2009). Evidence for this mechanism supports its use in maintaining chromatin states where repetitive regions contain long arrays of nucleosomes with the same histone PTMs.

For example, H3K9 methylation provides a binding site to recruit HP1, which in turn recruits the H3K9 methyltransferase suppressor of variegation 3-9 homolog 1 (SUV39H1) (Bannister et al., 2001; Lachner et al., 2001). Moreover, HP1 promotes its own inheritance by interacting with chromatin assembly factor 1 (CAF1) during replication to ensure its delivery at heterochromatic sites (Quivy et al., 2004). In this way, H3K9 methylation and HP1 form a self-sustaining loop to propagate a silenced chromatin state and maintain the pericentromeric heterochromatin structure (Probst et al., 2009). Additionally, a model has been proposed whereby the Polycomb repressive mark H3K27me3 forms a self-propagating loop in which polycomb repressive complex 2 (PRC2), the histone lysine methyltransferase responsible for catalyzing H3K27me3, binds H3K27me3 to maintain this PTM after its establishment (Hansen et al., 2008).

The rapid exchange of H2A and H2B outside of S phase compared to that of H3.1 and H4 suggests that PTMs on histones H3 and H4 are more likely to provide a long-term source of memory (Kimura and Cook, 2001; Probst et al., 2009). In most organisms, histone H4 lacks variants that differ in their amino acid sequence, while histone H3 has functionally distinct variants including the canonical replicative H3.1, the centromeric variant CenH3 (or CENPA) and the replication-independent variant H3.3, which is associated with active genomic regions (Holmes et al., 2005; Yuan and Zhu, 2013). Histone variants can serve as a mechanism to transmit epigenetic information in either a replication-dependent or replication-independent manner. Unlike the canonical

H3.1, the deposition of new CENPA occurs outside of S phase and is facilitated by Holliday junction recognition protein (HJURP), which specifically recognizes prenucleosomal CENPA (Foltz et al., 2009; Jansen et al., 2007). HJURP is recruited to centromeres by Mis18, a complex that is stably recruited to centromeres prior to new CENPA deposition in early G1 (Nardi et al., 2016; Wang et al., 2014). Moreover, existing CENPA nucleosomes are inherited during S phase through a mechanism where HJURP associates with pre-existing CENPA and interacts with the Minichromosome Maintenance Complex (MCM)2-7 helicase complex to ensure CENPA deposition at centromeres (Zasadzinska et al., 2018).

Histone chaperones such as HJURP are key players that assist in the reassembly of nucleosomes (Almouzni and Cedar, 2016). Certain chaperones can distinguish between nucleosome variants, providing another layer of regulation in the maintenance of epigenetic information (De Koning et al., 2007). CAF1 is recruited to the replication fork by interacting with proliferating cell nuclear antigen (PCNA) to deposit newly synthesized H3.1-H4 dimers during DNA replication and DNA repair (Gaillard et al., 1996; Probst et al., 2009; Smith and Stillman, 1989). In contrast, histone regulator A (HIRA) specifically incorporates H3.3-H4 dimers in a replication-independent fashion, and has been proposed to combat the dilution of H3.3 that occurs after replication to maintain genomic regions containing H3.3 (Probst et al., 2009; Ray-Gallet et al., 2002; Tagami et al., 2004). In addition to chaperones, components of the replication machinery including MCM2 and Pol ϵ ensure that parental histones remain associated with their original genomic location after DNA replication (Petryk et al., 2018; Stewart-Morgan et al., 2020; Yu et al., 2018).

While epigenetic information is capable of being maintained through replication, a major distinction that differentiates epigenetic from genetic information is its ability to be reprogrammed. For example, epigenetic information is extensively reprogrammed during

both mammalian early embryonic development and primordial germ cell development (Probst et al., 2009). In plants, there is evidence for a less extensive partial reprogramming occurring during gametogenesis and embryogenesis (Borg et al., 2020; Kawashima and Berger, 2014). Epigenetic information can be removed either actively (e.g., by eraser enzymes or histone exchange) or passively by dilution. In spite of this reprogramming, however, transgenerational epigenetic inheritance is observed in both animals and plants (Hauser et al., 2011; Morgan et al., 1999; Rakyan et al., 2003). Therefore, there appear to be mechanisms to ensure that certain epigenetic information can be retained through germline and embryonic reprogramming.

Genome stability

Although the mechanism for transmitting genetic information to the next generation is elegant and clear, this process is not error-free. DNA replication is coordinated with DNA damage sensing and repair to ensure genome integrity through cell divisions (Aguilera and Gomez-Gonzalez, 2008). Moreover, during the lifespan of an organism, its genome is exposed to a multitude of DNA damaging agents including UV and ionizing radiation, alkylating agents, and crosslinking agents (Aguilera and Garcia-Muse, 2013; Chatterjee and Walker, 2017). While some level of mutation is adaptive for long-term species survival, major genetic alterations including a high number of mutations, chromosome rearrangements, and chromosomal instability are associated with pathological disorders such as cancer and neurodegeneration in humans, as well as ageing (Aguilera and Gomez-Gonzalez, 2008; Chatterjee and Walker, 2017; Feng and Riddle, 2020). Therefore, maintaining a sufficient level of genome stability is essential for the health and fitness of an organism.

Cells utilize a large network of pathways to regulate genome stability, and different facets of genome instability are associated with the malfunction or misregulation

of specific pathways. Errors during replication, failures of the base excision repair or mismatch repair pathways, and error-prone translesion synthesis can all cause the accumulation of mutations such as base substitutions, micro-insertions or micro-deletions (Aguilera and Gomez-Gonzalez, 2008; Chatterjee and Walker, 2017). In contrast, chromosomal instability, characterized by changes in chromosome number, is caused by defects in chromosome transmission or the spindle mitotic checkpoint (Aguilera and Gomez-Gonzalez, 2008; Draviam et al., 2004). Finally, large chromosomal rearrangements such as translocations, duplications, inversions, or deletions are often caused by DNA breaks associated with replication stress, leading to processes including ectopic homologous recombination (HR) or end-joining between non-homologous DNA fragments (Aguilera and Gomez-Gonzalez, 2008).

Certain genomic regions are more susceptible to undergoing these large chromosomal rearrangements. For example, fragile sites are DNA sequences that preferentially exhibit gaps or breaks when DNA synthesis is partially impaired, and they are associated with hotspots for chromosomal rearrangements (Aguilera and Gomez-Gonzalez, 2008; Durkin and Glover, 2007; Sutherland, 1977). Fragile sites are often repetitive regions that can form secondary structures such as hairpins, leading to impaired replication fork progression and DNA breaks (Zhang and Freudenreich, 2007). Additionally, highly transcribed DNA regions show high recombination frequencies and are linked to replication stress (Aguilera, 2002; Aguilera and Gomez-Gonzalez, 2008). In addition to hindering replication fork progression, potentially leading to stalling and ultimately breakage, the movement of transcription complexes along DNA also produces topological constraints that can cause DNA breakage if not resolved properly (Gaillard and Aguilera, 2016). To avoid genome instability-inducing conflicts between the transcription and replication machinery, transcription and replication are tightly coordinated both temporally and spatially (Magdalou et al., 2014). In summary, diverse

pathways related to transcription, replication, DNA damage response, DNA repair, chromosome segregation, and cell cycle progression all contribute to the regulation of genome stability (Aguilera and Gomez-Gonzalez, 2008; Feng and Riddle, 2020; Magdalou et al., 2014). All of these processes depend on the chromatin environment, and thus epigenetic mechanisms play a major role in regulating genome stability. The relationship between epigenetics and genome stability is a vibrant area of research, and many histone PTMs have been implicated in regulating this process (Deem et al., 2012).

One histone PTM that ensures genome integrity through multiple mechanisms is the methylation of H4K20 (Jorgensen et al., 2013). In animals, H4K20me2 directly recruits the DNA repair factor p53-binding protein (53BP1) to double-strand breaks (DSBs) by binding the tandem Tudor domain of 53BP1 (Botuyan et al., 2006). While H4K20me2 is a highly abundant histone PTM that is not specifically present at sites of DNA damage, the binding of 53BP1 to H4K20me2 is repressed in the absence of DNA damage due to the Tudor-interacting repair regulator (TIRR) masking the binding motif of 53BP1 (Drane et al., 2017; Pesavento et al., 2008). Consequently, 53BP1 is only recruited to H4K20me2 under DNA damage conditions when the 53BP1-TIRR complex is dissociated, partially as a result of ataxia-telangiectasia mutated (ATM) phosphorylation of 53BP1 (Drane et al., 2017). When 53BP1 is bound to chromatin, it recruits additional DNA signaling and repair proteins and plays an important role in the decision of which DSB repair pathway to utilize, serving as a positive regulator of non-homologous end joining (NHEJ)-mediated DSB repair (Panier and Boulton, 2014).

In addition to its role in recruiting 53BP1 to regulate the DNA damage response, H4K20 methylation has also recently been shown to regulate genome stability by maintaining a chromatin compaction threshold in order to limit origin of replication licensing in G1 phase (Lu et al., 2008; Shoaib et al., 2018). Mutation of H4K20 or loss of the H4K20 monomethyltransferase SET8 leads to chromatin decompaction and

excessive loading of the origin recognition complex (ORC), subsequently causing single-stranded DNA formation and DNA damage (Shoaib et al., 2018). H4K20 methylation is thus able to regulate genome stability in animals both in its roles in the DNA damage response pathway and in its impact on replication and chromatin compaction.

One major histone PTM regulating genome stability in plants is monomethylation of histone H3.1 lysine 27 (H3.1K27me1) (Jacob et al., 2009). This modification is maintained during DNA replication by the lysine monomethyltransferases ARABIDOPSIS TRITHORAX-RELATED PROTEIN 5 (ATXR5) and ATXR6, which are recruited to the replication fork by interacting with PCNA to specifically monomethylate histone variant H3.1 (Jacob et al., 2014; Jacob et al., 2009; Raynaud et al., 2006). The *atxr5 atxr6* double mutant (*atxr5/6*), which has significantly lower levels of H3.1K27me1, exhibits a loss of silencing of repressed elements such as DNA repeats and transposons, heterochromatin decondensation, and amplification of heterochromatic DNA (Jacob et al., 2009; Jacob et al., 2010). It has been proposed that amplification of heterochromatic DNA in *atxr5/6* occurs due to the loss of transcriptional silencing of heterochromatin during S-phase, which causes transcription-replication collisions and/or R loop formation. These mutagenic events can cause DSBs, leading to heterochromatin amplification during repair (Aguilera and Garcia-Muse, 2012; Hale et al., 2016).

In a genetic screen for suppressors of the heterochromatin amplification defect in the *atxr5/6* mutant, the authors found that every suppressor that they isolated reduced both the transposon derepression phenotype and the DNA amplification defects in the mutant (Hale et al., 2016). However, previous work had found that mutations in DNA methyltransferases suppressed the extra DNA defects of the *atxr5/6* mutant while enhancing the transcriptional defects (Stroud et al., 2012). Therefore, the transcriptional derepression phenotype of *atxr5/6* is not dependent on the heterochromatic amplification phenotype, and the replication defects do not cause the transcription defects. Since

every mutation that suppresses the transcription defects also suppresses the extra DNA defects, it is also unlikely that the *atxr5/6* mutations affect transcription and replication independently (Hale et al., 2016). The most likely remaining model is that the transcription defects in the *atxr5/6* mutant cause the heterochromatic amplification (Hale et al., 2016). In this model, transcriptional derepression of heterochromatic elements during S-phase generates transcription machinery bound to genomic regions that the replication machinery is not equipped to evade. Collisions between the transcription and replication machineries then cause the replication forks to stall or collapse, which leads to amplification of heterochromatic regions (Hale et al., 2016).

Further support for the model of transcription-replication collisions in *atxr5/6* comes from examining DNA damage in this mutant. DNA damage response genes, such as *BREAST CANCER SUSCEPTIBILITY 1 (BRCA1)* and *RAD51*, are upregulated in the *atxr5/6* mutant and comet assays detect evidence of breaks as well (Feng et al., 2017). Additionally, the nuclei of *atxr5/6* have “overreplication-associated centers” (RACs), in which the heterochromatic chromocenters exhibit partial decondensation and a distinct hollow appearance (Feng et al., 2017). DNA methylation mutants, which suppress the heterochromatic amplification defects but not the transcriptional derepression of *atxr5/6*, also suppress the formation of these structures. Moreover, the centers of the RACs contain foci of phosphorylated H2AX (γ -H2AX), a marker of DSBs, and the DNA repair protein RAD51. Amplified pericentromeric heterochromatin localizes to the RACs as well (Feng et al., 2017). Finally, *atxr5/6* mutants are hypersensitive to mutations in DNA damage response pathways, especially in ATR (Feng et al., 2017). Together, these data suggest that the *atxr5/6* mutant has increased DNA damage and DSBs, which could arise from transcription-replication collisions due to derepression of silent elements.

Recent work has elucidated a potential mechanism for how H3.1K27me1 regulates transcription in heterochromatin. In the proposed model, H3.1K27me1

interferes with transcription by preventing another histone PTM, specifically the activating PTM H3K27ac, from being added to the H3K27 residue. Consistent with this model, H3K27ac was found to be enriched at the TSS of protein-coding genes in wild-type plants, and additional H3K27ac peaks exist at heterochromatic regions, especially the 5' end of transposons, in the *atxr5/6* mutant (Dong et al., 2021). These additional heterochromatic H3K27ac peaks appear to be almost entirely dependent on the histone acetyltransferase GCN5, as the vast majority of these peaks are missing in the *atxr5/6 gcn5* triple mutant, and moreover, the *atxr5/6 gcn5* triple mutant displays suppression of the heterochromatin amplification and heterochromatin decondensation phenotypes. Together, these data support the hypothesis that the histone acetyltransferase GCN5 acetylates H3.1K27 in the absence of H3.1K27me1 to activate transcription, leading to inappropriate transcription in heterochromatic regions in *atxr5/6* mutant plants. H3.1K27me1 is thus an important histone PTM that likely regulates genome stability in plants by maintaining silencing of repressed elements in order to prevent replication stress.

A related role has been proposed for H3K4 methylation in yeast, in which this histone PTM regulates genome stability by mitigating transcription-replication conflicts during replication stress; however, unlike H3.1K27me1, H3K4 methylation is thought to prevent transcription-replication conflicts in highly active, rather than repressed, regions (Chong et al., 2020). In *Saccharomyces cerevisiae*, active genes demonstrate an H3K4 methylation gradient, where H3K4me3 peaks at the 5' ends of genes, H3K4me2 is enriched in the middle portion of genes, and H3K4me1 is deposited at the 3' end of genes (Liu et al., 2005; Pokholok et al., 2005). This H3K4 methylation gradient pattern appears to be conserved in mammals as well (Barski et al., 2007). While H3K4me is broadly considered an activating mark and has been shown to promote transcription of specific genes, there is no conserved mechanism for how methylation of H3K4 activates

transcription, and loss of H3K4me3 has in fact been shown to minimally impact global transcription (Clouaire et al., 2012; Howe et al., 2017; Lauberth et al., 2013) Moreover, some evidence has arisen suggesting that the H3K4me gradient can be the product, not the cause, of transcription (Soares et al., 2017). Therefore, the precise mechanistic role of the H3K4me gradient in transcriptional regulation requires further clarification, but a recent study proposed an exciting role for this gradient in protecting from replication stress.

Prior to this study, H3K4 methylation had been implicated in regulating the DNA damage response and DNA replication, but the molecular mechanism of its activity was unclear (Faucher and Wellinger, 2010; Higgs et al., 2018; Rizzardi et al., 2012). In the “speed bump” model proposed by Chong et al., the H3K4me gradient hinders replication fork progression in areas of high transcriptional activity (Chong et al., 2020). This model asserts that in regions with high levels of H3K4 methylation as a consequence of high levels of transcription, the H3K4me “speed bumps” protect regions at high risk of transcription-replication conflicts by slowing down replication, while allowing faster replication in areas with lower transcriptional activity. The authors’ model is supported by genetic evidence indicating that H3K4me and the S-phase checkpoint kinase Rad53 function in the same pathway, biochemical evidence showing that H3K4 methylation causes fork stalling in replication-stressed *rad53* mutants, and data demonstrating that H3K4 methylation decelerates fork progression and prevents mutagenic events on highly transcribed regions. Under hydroxyurea-induced replication stress, the evidence suggests that Rad53 stabilizes replication forks while H3K4 methylation slows down the replication forks to relieve transcription-replication conflicts at highly transcribed regions. In the absence of H3K4 methylation, the replication forks may progress too quickly and collide with the transcription machinery, leading to mutagenesis and genome instability.

To conclude, these histone PTMs (H4K20me, H3.1K27me1, and H3K4me)

represent a few examples demonstrating the diverse ways in which epigenetic mechanisms regulate genome stability, ranging from promoting DNA repair to modulating chromatin compaction to protecting from replication stress (Chong et al., 2020; Dong et al., 2021; Drane et al., 2017; Feng et al., 2017; Hale et al., 2016; Jacob et al., 2010; Shoaib et al., 2018). Determining precisely how the chromatin environment contributes to processes that counteract genome instability has profound implications for understanding disease and ageing. While many promising breakthroughs have been made, more research is needed to further elucidate the ways in which histones support the maintenance of genome stability.

Flowering time

Epigenetic mechanisms also play a major role in regulating developmental transitions in multicellular eukaryotes. One of the most important developmental processes in angiosperms is the transition from vegetative growth to reproductive development (Andres and Coupland, 2012; Song et al., 2015), and thus the varied and complex pathways that regulate flowering time have been extensively studied. Determining how epigenetic mechanisms in turn play a role in these different pathways is a newer field, but one that has nonetheless seen several important advances in the two decades of its existence. Understanding how epigenetic pathways regulate the floral transition is exceedingly important for the improvement of agricultural crops in a rapidly changing environment. In crops that require harvesting of seeds or fruits, modulating flowering time is essential for maximizing crop yield, while flowering avoidance is often optimal for crops that require the harvesting of vegetative tissue (Jung et al., 2016). Further developing our understanding of flowering time regulation, and what role epigenetic mechanisms play in this process, will thus advance our ability to enhance agricultural crops and food yield.

Flowering time is tightly coordinated with environmental cues, including the day length or photoperiod, the quality of light (including the red/far-red ratio and the intensity of blue light), and the temperature (Amasino, 2005; Robson et al., 1993; Srikanth and Schmid, 2011). The model system *Arabidopsis thaliana* is classified as a facultative long-day (LD) plant, as long days (16 hours light/ 8 hours dark) promote flowering, but plants will eventually flower under short-day (SD) conditions (8 hours light/ 16 hours dark) (Srikanth and Schmid, 2011). One of the first *Arabidopsis thaliana* mutants that was identified to be insensitive to inductive day length was *constans* (*co*) (Redei, 1962). CONSTANS (CO) is a transcription factor whose expression is modulated by the circadian clock, oscillating with a phase of 24 hours, and is essential for photoperiod-dependent flowering induction (An et al., 2004; Putterill et al., 1995).

Three circadian clock-regulated proteins, called GIGANTEA (GI), FLAVIN-BINDING, KELCH REPEAT, F-BOX 1 (FKF1), and CYCLING DOF FACTOR (CDF), have been demonstrated to regulate the transcription of CO in order to establish the proper CO phasing (Fornara et al., 2009; Imaizumi et al., 2005; Imaizumi et al., 2003; Sawa et al., 2007). The expression of CO is further regulated at the posttranslational level by CONSTITUTIVELY PHOTOMORPHOGENIC (COP1), an E3 ubiquitin ligase, and the SUPPRESSOR OF PHYA-105 (SPA) protein family (Laubinger et al., 2006; Liu et al., 2008b). The modulation of CO expression is complex, and other pathways, including alternative splicing to produce two CO protein variants, also regulate its accumulation and consequently, the photoperiod-induced floral transition (Gil et al., 2017). CO in turn regulates flowering time by activating the production of FLOWERING LOCUS T (FT), a protein that is thought to stimulate flowering by acting as a long-distance signal between the leaves and the shoot meristem to induce floral meristem identity genes (An et al., 2004; Jaeger and Wigge, 2007; Srikanth and Schmid, 2011).

Temperature regulates flowering time via multiple mechanisms, including

vernalization, in which a prolonged period of cold (during winter in natural conditions) induces flowering in the following spring, and ambient temperature sensing (Srikanth and Schmid, 2011). In the model plant *A. thaliana*, *FRIGIDA (FRI)* and *FLOWERING LOCUS C (FLC)* are both necessary for vernalization, with the *FRI* protein promoting the expression of *FLC* (Geraldo et al., 2009; Johanson et al., 2000; Michaels and Amasino, 1999). *FLC* encodes a MADS box protein that directly represses the transcription of flowering time genes, including *FT* and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS1 (SOC1)* (Helliwell et al., 2006; Hepworth et al., 2002). In addition to vernalization, plants can also sense ambient temperature during their vegetative development and modulate their flowering time in response (Samach and Wigge, 2005). For example, higher temperatures have been shown to induce flowering in *Arabidopsis thaliana* (Balasubramanian et al., 2006). One major regulator of the ambient temperature flowering pathway is *SHORT VEGETATIVE PHASE (SVP)*, a MADS box protein that acts as a floral repressor by binding to the *FT* and *SOC1* promoters (Hartmann et al., 2000; Lee et al., 2007).

In addition to exogenous signals, endogenous signals including gibberellic acid (GA) and sugars also function to regulate the floral transition (Srikanth and Schmid, 2011). GAs act in *A. thaliana* to regulate flowering time via the positive regulation of *FT* as well as the positive regulation of the floral meristem identity gene *LEAFY (LFY)* (Blazquez et al., 1998; Hisamatsu and King, 2008). Moreover, GAs promote flowering by positively regulating the expression of *SOC1* and downregulating floral repressors including *SVP* (Li et al., 2008; Liu et al., 2008a; Srikanth and Schmid, 2011). The concentrations of sugars, such as sucrose and trehalose-6-phosphate, have also been shown to have an effect on flowering time via multiple proposed interactions with the flowering network (Bolouri Moghaddam and Van den Ende, 2013; Matsoukas et al., 2012).

Two additional endogenous pathways regulating flowering time have been characterized: the autonomous pathway and the ageing pathway. Autonomous pathway mutants display delayed flowering regardless of day length and the expression of genes in the autonomous pathway have been shown to promote flowering by repressing *FLC* expression (Koornneef et al., 1998; Srikanth and Schmid, 2011). The ageing pathway is defined by microRNA 156 (miR156) and its SQUAMOSA-PROMOTER BINDING PROTEIN-LIKE (SPL) transcription factor targets (Wang, 2014; Wu et al., 2009). Expression of miR156 and SPL is temporally regulated; in young seedlings, miR156 levels are high and SPL levels are low, while miR156 levels decrease and SPL levels consequently increase with age (Wang et al., 2009a). Members of the SPL family activate floral meristem identity and organ identity genes, including *LFY*, *APETALA1* (*AP1*), and *FRUITFULL* (*FUL*), as well as the expression of microRNA 172 (miR172) (Wu et al., 2009; Yamaguchi et al., 2009). The gradual accumulation of miR172 in turn downregulates *APETALA2* (*AP2*)-like floral repressor target genes and upregulates *FT* expression via the repression of the *SCHLAFMÜTZE* (*SMZ*) repressor (Aukerman and Sakai, 2003; Jung et al., 2007; Mathieu et al., 2009).

In sum, six major pathways (photoperiod, vernalization, ambient temperature, gibberellic acid, autonomous, and ageing) regulate flowering time (Srikanth and Schmid, 2011) (Figure 1.6). Extensive crosstalk between pathways occurs, mediated by a small number of central floral pathway integrators including *FT* and *SOC1* (Simpson and Dean, 2002). *FT* expression is modulated by inputs from the photoperiod pathway through *CO*, the vernalization pathway through *FLC*, and the ageing pathway through *SMZ*, as well as others (Hepworth et al., 2002; Mathieu et al., 2009; Shim et al., 2017). *SOC1* similarly integrates inputs from multiple pathways, as it is regulated by *CO* in an *FT*-dependent manner, *SVP*, and *FLC*; moreover, *SOC1* is positively regulated by SPLs,

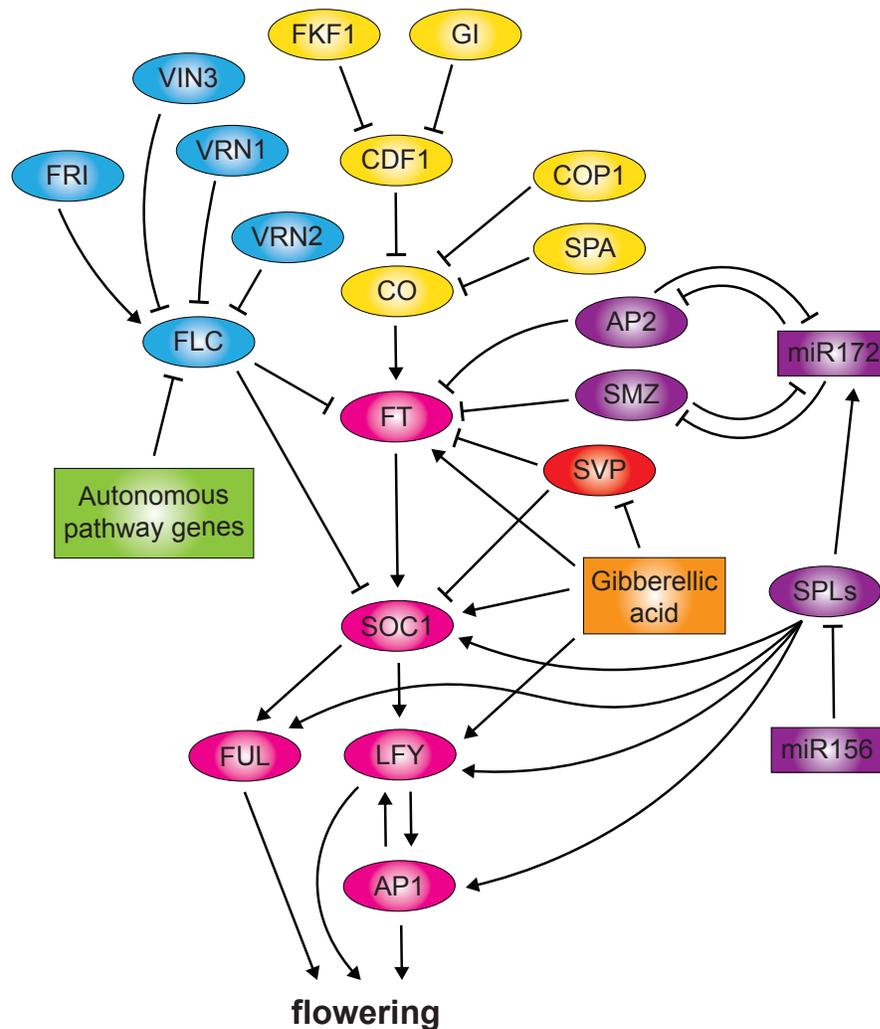


Figure 1.6 Major flowering time regulatory pathways. Simplified diagram of the six major pathways regulating flowering time (photoperiod, vernalization, ambient temperature, gibberellic acid, autonomous, and ageing) are shown. Floral integrators, floral meristem identity and organ identity genes are represented in pink. **Photoperiod (yellow):** Light regulates FKF1, GI, and CDF1 expression through the circadian clock. FKF1, GI, and CDF1 regulate the transcription of CO, and COP1 and SPA negatively regulate CO post-transcriptionally. CO is a transcriptional activator of FT, a floral integrator gene which in turn activates another floral integrator gene SOC1. SOC1 activates floral meristem identity genes such as LFY and FUL. **Vernalization (blue):** FRI promotes FLC expression and VRN1, VRN2, and VIN3 epigenetically silence FLC. FLC represses FT and SOC1. **Ambient temperature (red):** SVP represses FT and SOC1. **Gibberellic acid (orange):** GAs positively regulate FT, SOC1, and LFY, and negatively regulate SVP. **Autonomous (green):** Autonomous pathway genes repress FLC expression. **Ageing (purple):** miR156 represses expression of SPLs. SPLs activate the floral meristem and organ identity genes LFY, AP1, and FUL, as well as miR172 and SOC1. miR172 downregulates the expression of AP2 and SMZ, which are both negative regulators of FT and miR172.

a component of the ageing pathway, and GAs (Moon et al., 2003; Wang, 2014; Wang et al., 2009a; Yoo et al., 2005). SOC1 in turn regulates *LFY*, a floral pathway integrator that is also regulated by GAs and the ageing pathway through SPLs (Eriksson et al., 2006; Lee et al., 2008; Yamaguchi et al., 2009). Some floral pathway integrators such as *LFY* function not only to regulate the floral transition, but also to directly regulate floral meristem identity (Srikanth and Schmid, 2011; Weigel et al., 1992). Elucidating the transcriptional regulation of the flowering time pathways, and especially the floral pathway integrator genes, is essential for fully understanding the floral transition, and thus one active field of investigation is determining how the chromatin landscape contributes to this transcriptional regulation.

Epigenetic mechanisms regulate the complex pathways controlling flowering time in diverse ways. Many histone PTMs, including histone H3 lysine 4 (H3K4) methylation, H3K36 di- and trimethylation, H3K9 methylation, H3K27 methylation, H2B monoubiquitination, and acetylation of histones H3 and H4 have been shown to regulate the expression of key flowering time regulatory genes such as *FLC* and *FT* (Bastow et al., 2004; Bu et al., 2014; Cao et al., 2008; Choi et al., 2007; Crevillen et al., 2019; Crevillen et al., 2014; Cui et al., 2016; Deal et al., 2007; Deng et al., 2007; Gu et al., 2009; He, 2009; He et al., 2004; Jiang et al., 2008; Kim et al., 2005; Ning et al., 2019; Pajoro et al., 2017; Schmitz et al., 2008; Turck et al., 2007; Xu et al., 2009; Xu et al., 2008; Yu et al., 2011a; Zhang et al., 2007b; Zheng et al., 2019; Zhou et al., 2018). For example, three proteins VERNALIZATION 1 (VRN1), VERNALIZATION 2 (VRN2), and VERNALIZATION INSENSITIVE 3 (VIN3) contribute to the epigenetic silencing of *FLC* in response to vernalization via reduced histone acetylation and increased methylation of H3K27 and H3K9 (Bastow et al., 2004; Gendall et al., 2001; Levy et al., 2002; Sung and Amasino, 2004). VIN3 is induced in response to vernalization to establish the initial silencing of *FLC*, while VRN1, VRN2 and LIKE HETEROCHROMATIN PROTEIN 1

(LHP1) are later recruited to maintain this epigenetic state (De Lucia et al., 2008; Srikanth and Schmid, 2011; Wood et al., 2006). In contrast, the nuclear protein ACTIN RELATED PROTEIN 6 (ARP6) represses flowering by positively regulating *FLC* expression (Choi et al., 2005; Deal et al., 2005). ARP6 functions as a component of the SWI/SNF2-Related 1 (SWR1) chromatin remodeling complex, which introduces the variant H2A.Z into nucleosomes (Kumar and Wigge, 2010).

In addition to *FLC*, many other flowering time genes including *FT* are H3K27me3 targets (Turck et al., 2007; Zhang et al., 2007b). The H3K27me3 demethylases JUMONJI 13 (JM13), RELATIVE OF EARLY FLOWERING 6 (REF6), and EARLY FLOWERING 6 (ELF6) regulate flowering time by relieving the repression of flowering time genes such as *FT* and *SOC1*, as well as *FLC* (Crevillen et al., 2014; Cui et al., 2016; Zheng et al., 2019). Other histone-modifying enzymes have also been linked to the regulation of flowering time. Mutations in PROTEIN ARGININE METHYLTRANSFERASE 5 (PRMT5), also known as SHK1 BINDING PROTEIN 1 (SKB1), an enzyme which catalyzes the symmetric dimethylation of H4R3, have been shown to cause pleiotropic developmental phenotypes including late flowering (Pei et al., 2007). One proposed mechanism for the late flowering phenotype in *prmt5* mutants is that *FLC* repression is reduced due to decreased H4R3me2s (Schmitz et al., 2008; Wang et al., 2007). However, PRMT5 also methylates non-histone substrates including RNA processing factors, and mutations in *prmt5* have been demonstrated to cause splicing defects in several factors regulating flowering time, including transcripts of *FLOWERING LOCUS KH DOMAIN (FLK)*, encoding an autonomous pathway protein which represses *FLC* expression (Deng et al., 2010; Lim et al., 2004; Mockler et al., 2004). Therefore, histone-modifying enzymes such as PRMT5 may regulate flowering time both in their regulation of chromatin structure via histone PTMs as well as through the modification of non-histone substrates.

Assessing histone function

As demonstrated by the numerous examples above, histones and the chromatin landscape play a critical role in regulating various genomic and developmental processes, and consequently, can have widespread effects on organismal development and fitness. One of the major ways in which histones regulate these processes is through PTMs, or conversely, the absence of PTMs, although residues that cannot be post-translationally modified have also been demonstrated to function in significant roles. Traditionally, the functional significance of histone PTMs has mainly been deduced by the analysis of phenotypes resulting from the mutation of histone-modifying enzymes. For example, the PTM H3.1K27me1 was recognized to have a crucial role in maintaining genome stability in *Arabidopsis thaliana* through analysis of the *atxr5/6* double mutant, which has significantly reduced levels of H3.1K27me1 due to mutation of the monomethyltransferases ATXR5 and ATXR6 (Jacob et al., 2009; Jacob et al., 2010).

However, while this method has been successful in identifying functions for many histone residues, there are some limitations to this approach. As demonstrated by the case of PRMT5, analyzing the contributions of histone-modifying enzymes to the regulation of different phenotypes can be complicated by the fact that these enzymes often target non-histone substrates as well, as PRMT5 dimethylates RNA processing factors in addition to H4R3 (Deng et al., 2010; Schmitz et al., 2008; Wang et al., 2007). Moreover, histone-modifying enzymes can have essential roles in their actions on histones and/or on non-histone substrates, making it difficult to study loss-of-function mutants in these situations. On the other hand, there can be multiple histone-modifying enzymes with redundant functions, in which case, phenotypes may not be observed if all of the relevant histone-modifying enzymes are not mutated. For example, the *atxr5* and *atxr6* single mutants do not show significantly reduced levels of H3K27me1 and

consequently, they do not demonstrate the defects in chromatin organization, gene silencing, or heterochromatin amplification seen in the *atxr5/6* double mutant (Jacob et al., 2009; Jacob et al., 2010). The presence of redundant histone-modifying enzymes is especially confounding when all of the enzymes targeting a specific histone residue are not known. Therefore, the strategy of assessing the function of histone PTMs through the mutation of histone-modifying enzymes presents difficulties if there are many redundant histone-modifying enzymes, if the histone-modifying enzyme(s) have not been identified, or if the histone-modifying enzyme(s) have essential activities in the modifications of histones or additional non-histone substrates (McKay et al., 2015b).

As a solution to many of these challenges, researchers have implemented histone replacement systems, in which some or all of the endogenous copies of a histone are replaced with a histone mutated at a specific residue (or multiple residues). Histone replacement systems have been implemented in several model organisms, including the single-celled eukaryote *Saccharomyces cerevisiae*, as well as the multicellular eukaryotes *Drosophila melanogaster* and *Arabidopsis thaliana* (Dai et al., 2008; Dong et al., 2020; Govin et al., 2010a; Gunesdogan et al., 2010; Hodl and Basler, 2009, 2012; Jacob et al., 2014; Jiang et al., 2017; Nakanishi et al., 2008; Zhang et al., 2019). Rather than replacing a histone with a substitution mutant, systems have also been generated where specific histone variants are replaced (Hodl and Basler, 2012). Histone replacement systems can be partial, if some copies of the endogenous histone remain expressed, or complete, if all endogenous expression of the relevant histone is eliminated. Moreover, while not strictly histone replacement systems, histone substitution mutants have also been expressed in wild-type backgrounds (Herz et al., 2014; Lewis et al., 2013). Retaining different amounts of endogenous histone expression can provide information about the dominance of histone mutations.

Fundamental studies on histone mutations were initially performed in

Saccharomyces cerevisiae. Genetic studies on suppressors of δ insertion mutations first indicated that mutations in histones H2A and H2B alter the transcription of genes adjacent to the δ insertions (Clark-Adams et al., 1988). Moreover, mutations in histones H3 and H4 were found to partially relieve the requirement for transcription of the *HO* gene by the SWI/SNF complex (Kruger et al., 1995; Sternberg et al., 1987). These classic experiments established some of the initial links between transcription and chromatin structure using a forward genetics approach. Reverse genetics experiments further solidified the link between histones and gene expression, as well as other phenotypes including replication and cell cycle regulation (Kayne et al., 1988; Megee et al., 1990).

Modern molecular biologists often utilize a reverse genetics approach to generate a large array of targeted histone mutations, and several groups used this strategy to generate the first histone replacement systems in *Saccharomyces cerevisiae*. In one of these experiments, all residues located on the surface of the canonical nucleosome were individually mutated to alanine, with the exception of endogenous alanine residues, to conduct an unbiased screen for roles of these residues in various processes including DNA damage, replication, and transcriptional elongation (Matsubara et al., 2007). Similarly, Nakanishi et al. individually mutated all residues of the four core histones to alanine (except endogenously occurring alanine residues) to determine which histone residues are essential for viability and the methylation of H3K4 (Nakanishi et al., 2008).

Other screens focused exclusively on histones H3 and H4, rather than all of the core histones. Dai et al. generated a library of 486 histone H3 and H4 mutants, individually substituting each non-alanine residue to alanine and each alanine residue to serine (Dai et al., 2008). Additional substitutions were also made to mimic modified and

unmodified states, such as lysine to arginine (partially mimicking constitutive de-acetylation) and lysine to glutamine (partially mimicking constitutive acetylation). This large library of H3 and H4 mutants was screened for a multitude of phenotypes, including impaired response to DNA damaging agents, sensitivity to chemicals which perturb transcription and microtubules, transcriptional silencing defects, and proficiency at NHEJ. Additionally, Govin et al. generated a collection of alanine substitution mutants at each individual serine, threonine, lysine, and arginine residue to systematically assess the role of chromatin changes during gametogenesis (Govin et al., 2010a). Importantly, some of these systems eliminate all endogenous expression of the replaced histones in the mutants, as in the latter example, while others utilize partially reduced endogenous histone expression, as in the former example. Finally, the most recent system developed in *Saccharomyces cerevisiae* utilizes an efficient Clustered Regularly Interspaced Short Palindromic Repeats/ CRISPR-associated protein 9 (CRISPR/Cas9)-based histone shuffle strategy that allows for complete histone replacement at the endogenous locus (Fu et al., 2021). Complete histone replacement and histone gene replacement at the endogenous locus via gene targeting are ideal parameters for a histone replacement system, but as demonstrated by the research process in *Saccharomyces cerevisiae*, establishing a system with these parameters may take many iterations upon previous work to develop.

While histone replacement systems in yeast provide valuable information about histone function, there are likely additional roles for histone residues in multicellular eukaryotes due to their greater complexity and need for regulation across different cell types and developmental stages. For example, methylation at H3K27 is critical for development in all multicellular eukaryotes, but this PTM is absent in *Saccharomyces cerevisiae* (McKay et al., 2015a; Zhang et al., 2007a). The multicellular eukaryote *Drosophila melanogaster* has recently been used to generate histone replacement

systems that have been applied to study histone function in the context of a metazoan system. One study focused on histone variant H3.3 and found that the complete absence of H3.3 or the replacement of H3.3 with the mutant H3.3K4A, which has lysine 4 of H3.3 mutated to alanine, causes sterility but does not affect viability (Hodl and Basler, 2009). Another study expressed an inducible transgene encoding histone H1 in wild-type flies, and found that the expression of transgenic histone H1 leads to a dramatic reduction of endogenous histone H1 levels via an uncharacterized autoregulatory mechanism (Siriaco et al., 2015). Siriaco et al. utilized this partial replacement system to study the function of mutations that block H1S10 phosphorylation on chromatin compaction and chromatin structure (Siriaco et al., 2015).

In contrast to the aforementioned experiments, which focused on specific canonical histones or histone variants, Gunesdogan et al. generated an elegant histone replacement system where the entire canonical histone complement, which is clustered at a single chromosomal locus, can be replaced with modified histone transgenes (Gunesdogan et al., 2010). Animals were generated that were homozygous for a chromosomal deletion, which removes the entire histone gene cluster ($\Delta HisC$), and a transgene cassette was introduced into $\Delta HisC$ mutants providing 12 copies of the histone gene cluster (either encoding wild-type histones or particular histone mutants) (Gunesdogan et al., 2010). This system was first used to study the H3K27R mutation in a dominant context, and it was found that this mutation does not cause any dominant phenotypes in fly development, viability or fertility (Gunesdogan et al., 2010). However, a more complete replacement of wild-type histone H3 with H3K27R mutant histones caused a derepression of genes that are normally repressed by PRC2 and homeotic transformations similar to PRC2 mutants (Pengelly et al., 2013).

McKay et al. utilized a variation of this system to study the effect of individual mutations that prevent the addition of PTMs on three residues with proposed roles in

well-characterized epigenetic pathways: H4K20, H3K27, and H3K36 (McKay et al., 2015b). Experiments with H3K27A and H3K27R mutants confirmed that H3K27 is required for the repression of Polycomb target genes, and unlike the H3K27R mutation, the H3K27A mutation caused a dominant phenotype. Moreover, unlike in *Saccharomyces cerevisiae*, the H3K36R mutation caused lethality in the flies before they completed development. Finally, H4K20A replacement mutants exhibited a significant developmental delay, but were still viable and capable of DNA replication (McKay et al., 2015b).

This histone replacement system has also been utilized to replace all gene copies of the canonical histone H3.2 with the variant H3.3, or vice versa (Hodl and Basler, 2012). These experiments provided some evidence that H3.2 and H3.3 can functionally replace each other, as clones exclusively expressing either H3.2 or H3.3 differentiated into normally patterned adult tissues and displayed repression of typically silenced gene states. Cells were also generated completely replacing all endogenous copies of histones H3.3 and H3.2 with mutants that cannot be methylated at lysine 4. While these H3K4 mutants showed diminished growth, they were found to have normal expression levels for all genes assayed (Hodl and Basler, 2012). Additionally, this system was also used to generate an H3S28A replacement mutant, which was found to support normal mitosis despite the lack of H3S28ph; however, this mutant also showed derepression of Polycomb silencing and decreased H3K27 methylation (Yung et al., 2015).

Recently, Zhang et al. engineered a novel histone-mutagenesis platform in *Drosophila melanogaster* utilizing a CRISPR/Cas9-based strategy (Zhang et al., 2019). This platform provided an improvement over the former histone replacement system developed by Gunesdogan et al. in that it was less labor intensive to utilize, and thus offered a much higher throughput strategy for generating histone replacement mutants.

The system, which utilizes a CRISPR/Cas9-mediated HR pathway to knock in the replacement histone genes, was used to generate 40 alanine substitution mutations, covering all known modified residues in histones H3 and H4 in *Drosophila melanogaster* (Zhang et al., 2019). These mutants were screened for various phenotypes including viability, fertility, DNA damage sensitivity, and gene silencing defects. In sum, in their initial usages, histone replacement systems in *Drosophila melanogaster* were predominantly used to study specific residues on histone H3 that had previously been demonstrated to play important roles in the regulation of development and disease (Gunesdogan et al., 2010; McKay et al., 2015b; Muller and Verrijzer, 2009; Nichol et al., 2016; Pengelly et al., 2013; Shilatifard, 2012). However, the most recent system developed by Zhang et al. allows for more systematic screens of histone residues in this organism due to its higher efficiency (Zhang et al., 2019).

Unlike the model systems discussed previously, plant systems present additional challenges to implementing complete histone gene replacement. While all replication-dependent histone genes are clustered at a single genomic locus in *Drosophila melanogaster* (Lifton et al., 1978) and there are only two copies of each core histone gene in *Saccharomyces cerevisiae* in the haploid cell stage (Fu et al., 2021), there are 47 core histone genes in *Arabidopsis thaliana* found dispersed throughout the genome (Tenea et al., 2009). Because the histone genes are not clustered together in plants, the establishment of complex histone deletion mutants necessary for partially or completely replacing endogenous histone genes with modified histone genes via transgenesis is more challenging in plants compared to model organisms in which efficient histone replacement systems have already been established. While the earliest strategies used to implement complete histone gene replacement in both *Saccharomyces cerevisiae* and *Drosophila melanogaster* were not applicable to plants due to their reliance on either the plasmid shuffle strategy and/or site-specific recombination systems that were not

possible in plants, some aspects of the newest histone replacement strategies are applicable to plant systems. For example, recent advancements in the deployment of multiplex CRISPR/Cas9-based technologies in plants facilitate the creation of mutations in large gene families like histones. These improvements in gene editing technologies were crucial to support the development of the complete histone gene replacement system in *Arabidopsis thaliana* described in this dissertation.

Prior to the start of this dissertation project, the only histone replacement system generated in the model plant *Arabidopsis thaliana* was a partial histone replacement system for the canonical histone H3.1. To establish this system, individual H3.1 substitution mutants were expressed from transgenes in a H3.1 quadruple mutant background, which has four out of the five endogenous histone H3.1 genes mutated (with two H3.1 genes knocked-out, and two others being hypomorphic mutants) with T-DNA insertions (Jacob et al., 2014). This system was first used to study the role of alanine at position 31 in histone H3.1, as this residue demonstrates a conserved difference between H3.1 and H3.3 in plants and animals (with threonine being found at position 31 in the variant H3.3). Interestingly, H3.1A31T replacement mutants showed lower levels of H3.1K27me1 compared to wild-type plants; moreover, biochemical experiments demonstrated that the H3.1A31T mutation inhibited the ability of ATXR5 and ATXR6 to methylate H3.1K27. These results strongly support the hypothesis that the conserved difference at position 31 of histone H3 is responsible for the selective methylation of H3.1 over H3.3 by ATXR5 and ATXR6 (Jacob et al., 2014).

Members of the Jacob lab recently expanded their use of the H3.1 replacement system to conduct a systematic screen of histone H3.1, screening 81 individual H3.1 point mutants (unpublished data). The collection of point mutants covers all residues of H3.1 that can theoretically be targets for PTMs in plants (lysine, arginine, threonine, serine, and tyrosine), and thus this histone H3.1 replacement system provides a method

to systematically screen the functional roles of PTMs on H3.1 without prior knowledge of all of the relevant histone-modifying enzymes. Plant systems such as *Arabidopsis thaliana* often have multiple redundant histone-modifying enzymes, and consequently, this screen provides a more efficient method of systematically screening roles for histone PTMs compared to attempting to mutate all enzymes that modify a particular histone residue. One notable phenotype observed in this screen was that of the H3.1S28A mutant, which exhibits amplification of heterochromatic DNA, heterochromatin decondensation, and transposon derepression similar to the *atxr5/6* mutant. Moreover, biochemical experiments demonstrated that the H3.1S28A mutant histone cannot be methylated at K27 by ATXR5/6 (Dong et al., 2020).

To further assess the role of PTMs on H3.1K27 in regulating genome stability, Dong et al. also generated a mutant where they expressed H3.1K27Q from a wild-type background in order to partially mimic an acetylated lysine residue. Similar to the H3.1S28A mutant and the *atxr5/6* mutant, transcriptional derepression and upregulation of the DNA damage response element *BRCA1* were observed in the H3.1K27Q mutant, providing evidence to support the model that ectopic acetylation is added at H3.1K27 in mutants lacking monomethylation at this residue, leading to the loss of transcriptional silencing and genome instability. Biochemical experiments further supported this hypothesis by confirming that the H3.1S28A mutant histone can still be acetylated at K27 (Dong et al., 2020). Experiments were then performed to assess whether mutations on other H3.1 residues rescued the genome instability phenotypes demonstrated by the H3.1S28A mutant. Every residue in addition to K27 that is known to be acetylated by GCN5 on H3.1 (K9, K14, K18, K23, and K36) was individually mutated in combination with the H3.1S28A mutation, and mutations of K36 were found to specifically suppress the H3.1S28A mutant phenotype (Dong et al., 2020). From these results, a model was proposed whereby GCN5 acetylates lysines 27 and 36 on histone H3 in the absence of

monomethylation on lysine 27, leading to transcriptional reactivation and the subsequent genome instability phenotypes.

During the completion of this dissertation, several additional systems were described in *Arabidopsis thaliana* that approximate histone replacement. In 2017, a system using a combination of traditional crossing and artificial microRNA was developed in order to generate knockdown lines largely depleted of histone H3.1 (Jiang et al., 2017). H3.1 depletion mutants displayed developmental defects including enlarged inflorescence meristems, fasciated stems, ectopic leaflets, and reduced silique size. The developmental defects of these lines could be partially to completely rescued by the expression of microRNA-resistant H3.1. This system was used to study the effect of several H3.1 point mutants, including H3.1K4A, H3.1K9A, and H3.1K27A, on the propagation of H3K27me₃. Additionally, the authors utilized this system to assess whether the histone variant H3.3 could functionally replace canonical H3.1, and to precisely probe the impact of each of the four divergent residues between H3.1 and H3.3 on H3K27me₃ propagation.

Recently, two systems that allow the assessment of functions for histone H2A variants were also described in *Arabidopsis thaliana*. In 2019, partial H2A.Z replacement mutants were generated in order to assess the effect of H2A.Z monoubiquitination on transcriptional repression (Gomez-Zambrano et al., 2019). A double mutant containing mutations in two out of the three endogenous H2A.Z genes was generated by crossing (March-Diaz et al., 2008), and an N-terminal FLAG-tagged version of H2A.Z was then constitutively expressed under the control of cauliflower mosaic virus promoter. When wild-type H2A.Z was constitutively expressed in this system, it partially rescued the developmental phenotypes observed in the H2A.Z double mutant. The authors then utilized this system to assess the function of H2A.Z lacking monoubiquitination by expressing FLAG-tagged H2A.Z with either K129 or both K129 and K132 mutated to

arginine and showed that H2A.Z monoubiquitination plays an important role in mediating transcriptional repression. Additionally, in 2021, a system was developed that expressed H2A.W from a transgene in a background completely deprived of endogenous H2A.W expression (Schmucker et al., 2021). The *h2a.w* depletion background was generated by crossing T-DNA insertion mutants corresponding to two out of the three endogenous H2A.W genes and mutating the remaining endogenous H2A.W gene using CRISPR/Cas9 (Bourguet et al., 2021). This system was used to study the effect of differing regulatory and primary amino acid sequences of H2A.W isoforms, specifically by swapping the promoters, histone core domains, and C-terminal tails between the isoforms H2A.W.6 and H2A.W.7 and expressing these fusions in the *h2a.w* depletion background.

Although the systems described above represent useful methods to study histones, several improvements upon these systems could be implemented to facilitate the establishment of a high-throughput screen of histone function. For example, although the artificial microRNA strategy presented by Jiang et. al allows the depletion of endogenous histone H3.1, the creation of stable transgenic lines displaying complete histone H3.1 replacement would facilitate future studies and provide a more high-throughput approach, as the former strategy is relatively time-consuming and may not completely eliminate endogenous histones (Jiang et al., 2017). Similarly, the H2A.Z depletion background that was described utilized only a partial elimination of endogenous H2A.Z, from two out of the three genes coding for different H2A.Z proteins (Gomez-Zambrano et al., 2019; Lei and Berger, 2020; March-Diaz et al., 2008). Moreover, in all of the cases presented above, further work assessing the expression levels of replacement histones in these mutants is necessary to determine how the levels of transgenic histones compare to histone levels in wild-type plants.

Finally, an important feature of histone replacement systems is the rescue of

wild-type phenotypes in plants expressing unmodified transgenic histones. Although all of the systems described above displayed some degree of rescue of the histone depletion phenotype with the reintroduction of a wild-type histone transgene, this rescue was in all cases only a partial rescue, except for some individual H3.1 replacement plants, which displayed a complete rescue (Bourguet et al., 2021; Gomez-Zambrano et al., 2019; Jiang et al., 2017). Ideally, histone replacement plants expressing wild-type replacement histones would display the same phenotype, or a very similar phenotype, to the wild-type background. One contributing factor in the two systems studying H2A variants may be that only one transgenic H2A.W or H2A.Z gene was provided, even though each of the endogenous H2A genes that were depleted code for a different protein (Bourguet et al., 2021; Lei and Berger, 2020; Schmucker et al., 2021; Tenea et al., 2009). Therefore, a replacement system providing a transgenic copy of each depleted H2A variant could allow a more thorough rescue of the H2A depletion backgrounds. However, as these systems are complicated to establish and usually require the combination of multiple mutations and rounds of transgenesis, it may be unfeasible to generate a histone replacement plant with a completely wild-type phenotype given current technological limitations. Regardless, the systems described above contain numerous future applications towards studying histone function, and with further improvements and characterization, could be used as complete histone replacement systems allowing for high-throughput analysis. As demonstrated above, histone mutants provide noteworthy functionalities to study the role of histone PTMs, including their abilities to act as PTM mimics and to potentially prevent the addition of certain PTMs.

Although histone replacement systems engender numerous benefits, confounding issues do still exist with these systems, so they are often supplemented with the traditional analysis of mutations in histone-modifying enzymes. For example,

some histone residues can have multiple PTMs added to them, such as lysine mono-, di- or trimethylation or lysine acetylation. Therefore, mutating said lysine residue could eliminate all possible PTMs from being added, and thus, determining what specific PTM is causing an observed phenotype in a histone replacement mutant would be difficult without more information. Additionally, mutating a particular histone residue may not only affect PTMs added to that specific residue, but also the ability of histone-modifying enzymes to add PTMs to neighboring residues. For example, many of the phenotypes observed when expressing the H3.1S28A mutant in *A. thaliana* are likely consequences of this mutation's effect on monomethylation of the neighboring H3.1K27 residue. This feature can be a benefit of these systems, but it also introduces confounding factors that often need to be addressed. In addition, mutations of specific histone residues can also cause lethality or severe reductions in viability in the histone replacement mutants, complicating the analysis of these residues. Although lethality is a complicating factor, somewhat surprisingly, in systematic screens of core histones in *Saccharomyces cerevisiae*, only a small percentage of residues were found to be essential for viability (Govin et al., 2010a; Nakanishi et al., 2008). Therefore, despite the high level of conservation of histones, it appears that the individual mutation of most histone residues results in a viable mutant that can be studied.

Finally, mutated histone residues can have dominant gain-of-function effects that can inhibit histone-modifying enzymes. For example, individual histone H3 lysine-to-methionine substitutions of H3K9, H3K27, and H3K36 have all been shown to inhibit the enzymatic activity of PRC2 complexes by capturing these histone-modifying enzymes (Herz et al., 2014; Lewis et al., 2013). This final case represents a confounding factor, but it also presents a way to bypass some of the above obstacles presented for the traditional analysis of mutations in histone-modifying enzymes. Namely, histone gain-of-function mutations that inhibit modification pathways (such as methylation) can be used

to disrupt the addition of certain PTMs even if the relevant histone-modifying enzyme(s) have not been identified. However, if the sequestered histone-modifying enzymes have non-histone substrates and/or essential activities, these systems do not circumvent those complications. Nonetheless, the combination of evidence from a histone replacement system and evidence from mutating the relevant histone-modifying enzyme(s) often provides the clearest representation of the functional significance of a particular histone PTM given the current technological capabilities for *in vivo* work using model organisms.

Dissertation goals

Building off of previous work in this field, the goal of my dissertation was to establish a CRISPR-based histone mutagenesis platform in the plant model system *Arabidopsis thaliana* that allows for complete histone replacement. As proof-of-concept, I targeted histone H4, which is encoded by the largest number of endogenous genes (i.e., eight genes) among identical histone proteins in plants (Okada et al., 2005; Tenea et al., 2009; Wierzbicki and Jerzmanowski, 2005), for a systematic functional screen of modifiable residues on this protein. As described above, histone replacement systems provide numerous benefits toward conducting an efficient screen for roles of histone residues and PTMs, including circumventing the need to simultaneously mutate many redundant histone-modifying enzymes and even eliminating the requirement to identify all of the relevant histone-modifying enzymes.

I was specifically interested in histone H4 due to the diversity of PTMs on this histone that have been identified to have substantial roles in regulating processes such as transcription, DNA repair, and chromatin compaction (Suganuma and Workman, 2011). Unlike other core histones, which have numerous variants with distinct sequences and functions, histone H4 in contrast is the most conserved histone and most

organisms lack H4 sequence variants (Holmes et al., 2005; Kamakaka and Biggins, 2005; Tenea et al., 2009). In addition, along with histone H3, histone H4 has been proposed as a more likely carrier of long-term memory compared to the other core histones due to the slower exchange of histone H4 outside of S phase (Kimura and Cook, 2001; Probst et al., 2009). Finally, no systematic screen of histone H4 has been conducted in plants, and thus I expected to uncover many novel functions for residues on histone H4 with this approach. While a recent screen in the multicellular eukaryote *Drosophila melanogaster* generated alanine substitution mutants covering 14 residues on histone H4 (Zhang et al., 2019), I intended to cover more residues and generate additional substitution mutants in *Arabidopsis thaliana*.

In the subsequent chapters, I describe my work generating a complete histone H4 replacement system in *Arabidopsis thaliana* and assessing the phenotypes in this collection of mutants. In chapter 2, I explain the construction and initial characterizations of the histone H4 replacement system, which is the first complete histone replacement system characterized in plants allowing for a high-throughput analysis of histone function. In chapter 3, I describe the systematic screens conducted of the entire collection of histone H4 mutants to assess the role of these residues in regulating flowering time, rosette morphology, DNA replication, chromatin structure, and gene silencing. In chapter 4, I focus on one specific histone H4 residue—H4 arginine 17—and further characterize the phenotypes induced by mutations of this residue. Additionally, I pursue several alternative hypotheses to determine a molecular mechanism responsible for the observed phenotypes. Finally, in chapter 5, I discuss further applications for the histone H4 replacement system, potential improvements upon the existing system, and future directions for this work.

Chapter 2: Establishment of a histone H4 replacement system in *Arabidopsis thaliana*

Author's Note: Portions of this chapter have been submitted for publication as:

Emma Tung Corcoran¹, Chantal LeBlanc¹, Mia Arias Tsang¹, Anthony Sarkiss¹, Yuzhao Hu², Ullas V. Pedmale², and Yannick Jacob¹. 2021. Systematic histone H4 replacement in *Arabidopsis thaliana* reveals a role for H4R17 in regulating flowering time.

Affiliations:

1. Yale University, Department of Molecular, Cellular and Developmental Biology, Faculty of Arts and Sciences; 260 Whitney Avenue, New Haven, Connecticut 06511, United States.
2. Cold Spring Harbor Laboratory; Cold Spring Harbor, New York 11724, United States.

Overview

To establish a histone H4 replacement system in *Arabidopsis thaliana*, we first generated an endogenous histone H4 depletion background. This histone H4 depletion background retained expression of only one of the eight endogenous histone H4 genes, and displayed genome instability phenotypes as well as abnormal morphology. Next, we engineered an H4 replacement plasmid, which—upon transformation into the H4 depletion background—allowed for expression of a modified H4 transgene (the H4 replacement gene) as well as targeting of the remaining endogenous H4 gene for mutagenesis. Using the above strategy, we generated plants expressing the replacement H4 gene and observed a rescue of the genome instability and morphological phenotypes of the H4 depletion background due to high expression of transgenic histone H4. Additionally, we created a collection of 63 H4 point mutant replacement plasmids, covering every residue on histone H4 that can in theory be post-translationally modified, and individually expressed each of these H4 mutations in *Arabidopsis thaliana* using the described H4 replacement system. From this collection of 63 mutants, we identified six H4 mutations that cause lethality. Finally, we assessed the extent of the replacement of endogenous histone H4 in second-generation transformants generated with our histone H4 replacement system and demonstrated that we can straightforwardly identify plants displaying a complete replacement of endogenous histone H4 for future analysis.

Results

Generation of an endogenous H4 depletion background

In order to ultimately conduct a systematic screen of histone H4 for roles in regulating cellular and developmental phenotypes in a multicellular eukaryote, we utilized the plant model system *Arabidopsis thaliana* to generate a histone H4

replacement system. We aimed to use this histone H4 replacement system to engineer a large library of *Arabidopsis thaliana* plants with a complete replacement of endogenous histone H4 with histone H4 point mutants. In *Arabidopsis thaliana* (Columbia [Col] ecotype), there are eight histone H4 genes, which are found dispersed throughout four out of the five chromosomes. All eight of the endogenous histone H4 genes code for a single histone H4 variant. Histone H4 shows remarkable conservation between eukaryotic species, and *Arabidopsis thaliana* histone H4 is 100% identical to histone H4 in agriculturally important crop plants such as wheat (*Triticum aestivum*) and rice (*Oryza sativa*), and 98% identical to human histone H4 (100/102 identical aa; conservative substitutions at aa 60 and 77) (Figure 2.1). The high level of protein sequence similarity between histone H4 in *Arabidopsis thaliana* and other systems suggests that results determined from screens of histone H4 in *Arabidopsis thaliana* may not only be applicable to agriculturally relevant crop plants, but also have potential relevance to human pathology. Interestingly, *Arabidopsis thaliana* histone H4 has 91% amino acid sequence identity to *Saccharomyces cerevisiae* histone H4, for which several systematic screens have previously been conducted (Dai et al., 2008; Govin et al., 2010a; Matsubara et al., 2007; Nakanishi et al., 2008).

The first part of our strategy in establishing the histone H4 replacement system in *Arabidopsis thaliana* was to eliminate expression from all but one of the endogenous histone H4 genes by generating a histone H4 depletion background using CRISPR/Cas9. We chose the untargeted histone H4 gene (*At3g53730*) due to its relatively high and consistent expression level across different cell types so that the H4 depletion background would remain viable and fertile (Nakabayashi et al., 2005; Schmid et al., 2005; Waese et al., 2017). We reasoned that a complete loss of histone H4 or too significant of a loss of histone H4 would likely cause lethality or infertility. We designed three guide RNAs (gRNAs) to target seven out of the eight endogenous histone H4

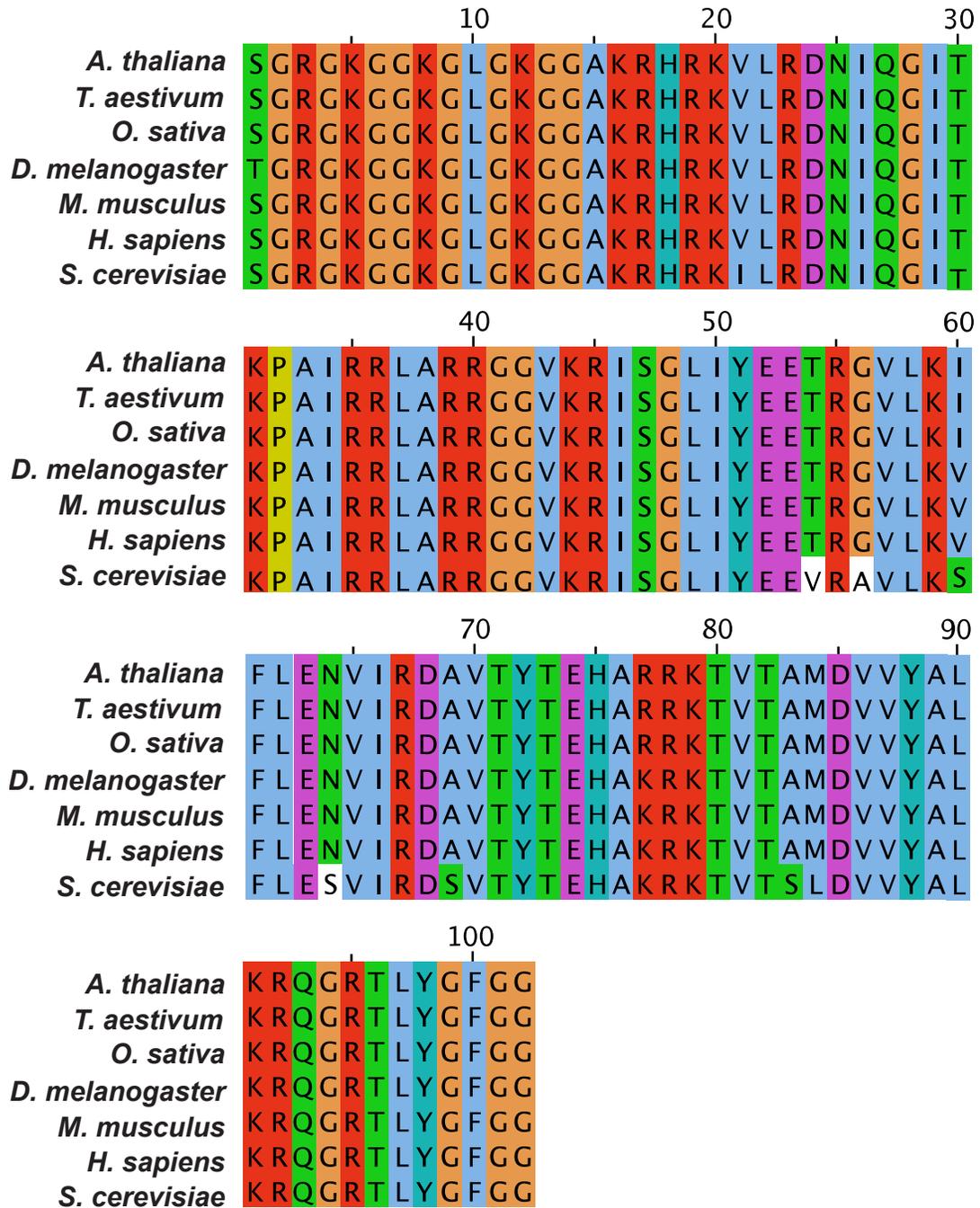


Figure 2.1 Protein sequence similarity of histone H4 across species. Multiple sequence alignment of histone H4 proteins performed with Clustal Omega. Protein sequences were obtained from UniProt and correspond to the following accession numbers: *Arabidopsis thaliana*; P59259, *Triticum aestivum*; P62785, *Oryza sativa*; Q7XUC9, *Drosophila melanogaster*; P84040, *Mus musculus*; P62806, *Homo sapiens*; P62805, and *Saccharomyces cerevisiae*; P02309. Chemical characteristics of amino acids shown with ClustalX color scheme (Larkin et al., 2007).

genes using temperature-optimized multiplex CRISPR/Cas9 (Figure 2.2). We then transformed a T-DNA coding for these three gRNAs and Cas9 into the Col background, selected first-generation transformant (T1) seeds, and exposed these T1 plants to repeated heat stress treatments at 37°C for 30h to increase the efficiency of targeted mutagenesis by Cas9 (LeBlanc et al., 2017).

We identified a mutant in the T2 generation that had homozygous loss-of-function mutations in all seven targeted H4 genes (i.e., the H4 septuple mutant) (Figure 2.2). Most of the mutations identified in the H4 septuple mutant were small insertion-deletion mutations (indels) leading to frameshift mutations and a premature termination codon (Table 2.1). Additionally, the mutation in *At1g07820* was a relatively large deletion also causing a frameshift in the coding sequence and a premature termination codon, and the mutation in *At3g45930* caused the first 79 nucleotides of the H4 coding sequence to be fused to a portion of the H4 terminator, before terminating in an ectopic stop codon (Table 2.1). Finally, the H4 septuple mutant background was homozygous for the Cas9 transgene, such that the last remaining endogenous H4 gene (*At3g53730*) in this mutant background could be targeted for mutagenesis with transformation of an additional gRNA to completely eliminate H4 production from the eight endogenous genes (see below).

Morphological and molecular characterization of the H4 septuple mutant plants showed that they were smaller than wild-type Col plants and displayed a serrated leaf phenotype (Figure 2.3A). In addition, fertility was lower in the H4 septuple mutant compared to Col plants, as demonstrated by the reduced silique size (Figure 2.3B). We found that the transcription of the remaining endogenous H4 gene (*At3g53730*) was upregulated approximately 2-fold in these mutants relative to Col, likely to compensate for the loss of function of the other seven histone H4 genes (Figure 2.4A). The H4

gRNA 1

At5g59970: 1 nucleotide (T) insertion

wild-type 59 --GGAAGGTTCTGAGAGACAACA-TCCAAGGAATCA-- 91
guide RNA 1 GTTCTGAGAGACAACA-TCCA
mutant 59 --GGAAGGTTCTGAGAGACAACATTCCTCAAGGAATCA-- 92

At3g45930: 299 nucleotide deletion

wild-type 59 --GGAAGGTTCTGAGAGACAACATCCAAGGAATCA-- 91
guide RNA 1 GTTCTGAGAGACAACATCCA
mutant 59 --GGAAGGTTCTGAGAGACAACA----- 79

At5g59690: 2 nucleotide deletion

wild-type 59 --GGAAGGTTCTGAGAGACAACATCCAAGGAATCA-- 91
guide RNA 1 GTTCTGAGAGACAACATCCA
mutant 59 --GGAAGGTTCTGAGAGACA --ATCCAAGGAATCA-- 89

At3g46320: 1 nucleotide (T) deletion

wild-type 59 --GGAAGGTTCTGAGAGACAACATCCAAGGAATCA-- 91
guide RNA 1 GTTCTGAGAGACAACATCCA
mutant 59 --GGAAGGTTCTGAGAGACAACA-CCAAGGAATCA-- 90

gRNA 2

At1g07660: 1 nucleotide (T) insertion

wild-type 45 --AGCGAAGAGGCACAGGAAGGTT-CTGAGGGATAA-- 77
guide RNA 2 AGAGGCACAGGAAGGTT-CTG
mutant 45 --AGCGAAGAGGCACAGGAAGGTTTCTGAGGGATAA-- 78

At1g07820: 56 nucleotide deletion

wild-type 45 --AGCGAAGAGGCACAGGAAGGTTCTGAGGGATAA-- 77
guide RNA 2 AGAGGCACAGGAAGGTTCTG
mutant 45 --AGCGAAGAGGCACAGGAAGGTT----- 66

gRNA 3

At2g28740: 1 nucleotide (G) insertion

wild-type 101 --CGATTCGTCGTCTTGCTCGTA-GAGGAGGTGTGA-- 133
guide RNA 3 CGTCGTCTTGCTCGTA-GAGG
mutant 101 --CGATTCGTCGTCTTGCTCGTAGGAGGAGGTGTGA-- 134

Figure 2.2 Design of guide RNAs to generate H4 septuple mutant background.

Three guide RNAs were designed to target seven endogenous histone H4 genes (*At5g59970*, *At3g45930*, *At5g59690*, *At3g46320*, *At1g07660*, *At1g07820*, *At2g28740* from top to bottom). The homozygous mutation in each of the targeted genes in the H4 septuple mutant is shown.

H4 gene Effect of mutations in H4 septuple mutant on mRNA

<i>At1g07660</i>	66 nucleotides from H4 5' CDS fused to 8 nucleotides from H4 CDS (after frameshift mutation due to 1 nucleotide insertion), terminating in premature termination codon
<i>At1g07820</i>	66 nucleotides from H4 5' CDS fused to 81 nucleotides from H4 CDS (after frameshift mutation due to 56 nt deletion), terminating in premature termination codon
<i>At2g28740</i>	121 nucleotides from H4 5' CDS fused to 79 nucleotides from H4 CDS (after frameshift mutation due to 1 nucleotide insertion), terminating in premature termination codon
<i>At3g45930</i>	79 nucleotides from H4 5' CDS fused to 14 nucleotides from H4 terminator (due to 299 nucleotide deletion), terminating in ectopic termination codon
<i>At5g59690</i>	76 nucleotides from H4 5' CDS fused to 41 nucleotides from H4 CDS (after frameshift mutation due to 2 nucleotide deletion), terminating in premature termination codon
<i>At3g46320</i>	79 nucleotides from H4 5' CDS fused to 12 nucleotides from H4 CDS (after frameshift mutation due to 1 nucleotide deletion), terminating in premature termination codon
<i>At5g59970</i>	79 nucleotides from H4 5' CDS fused to 40 nucleotides from H4 CDS (after frameshift mutation due to 1 nucleotide insertion), terminating in premature termination codon
<i>At3g53730</i>	N/A

Table 2.1 Effect of H4 septuple mutations on H4 messenger RNA (mRNA). Effect of the homozygous mutations in the H4 genes in the H4 septuple mutant background on the expression of H4 mRNA.

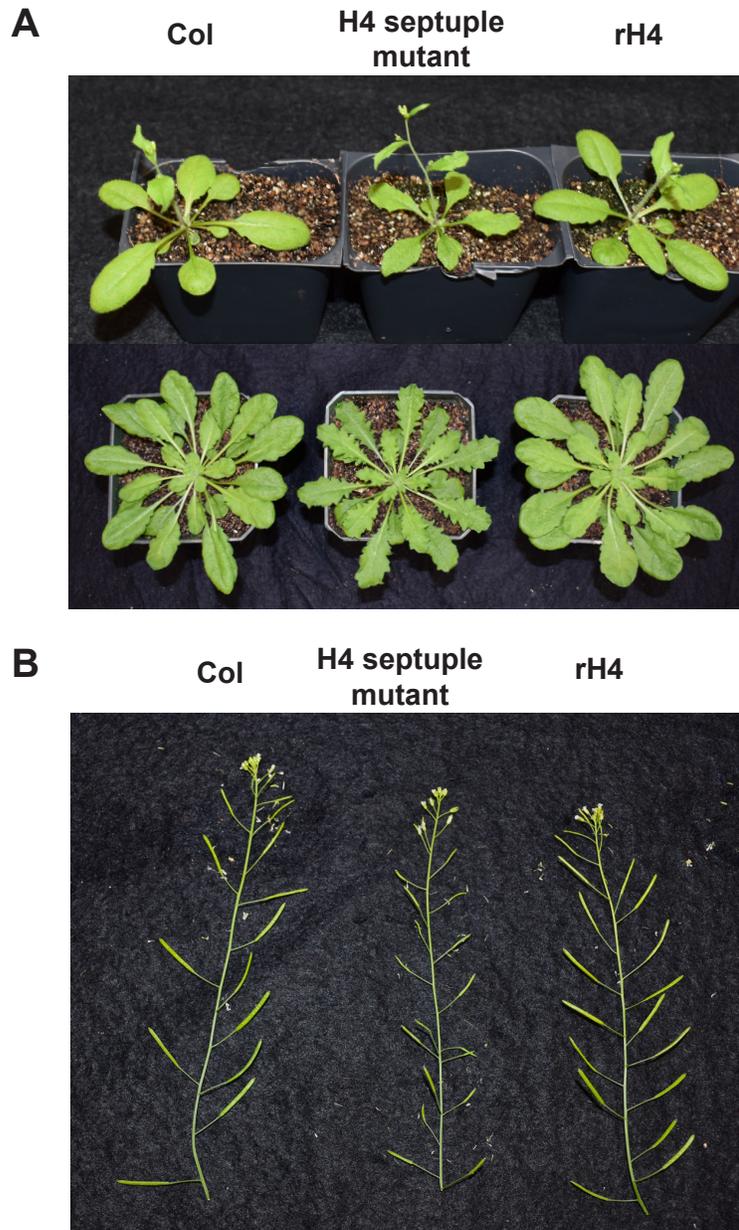


Figure 2.3 Morphological phenotypes of the H4 septuple mutant and H4 replacement plants. (A) Phenotype of Col, H4 septuple mutant, and rH4 plants grown in long-day conditions at 3.5 weeks (top) and short-day conditions at 7 weeks (bottom). (B) Siliques of Col, H4 septuple mutant, and rH4 plants grown in long-day conditions for 4 weeks.

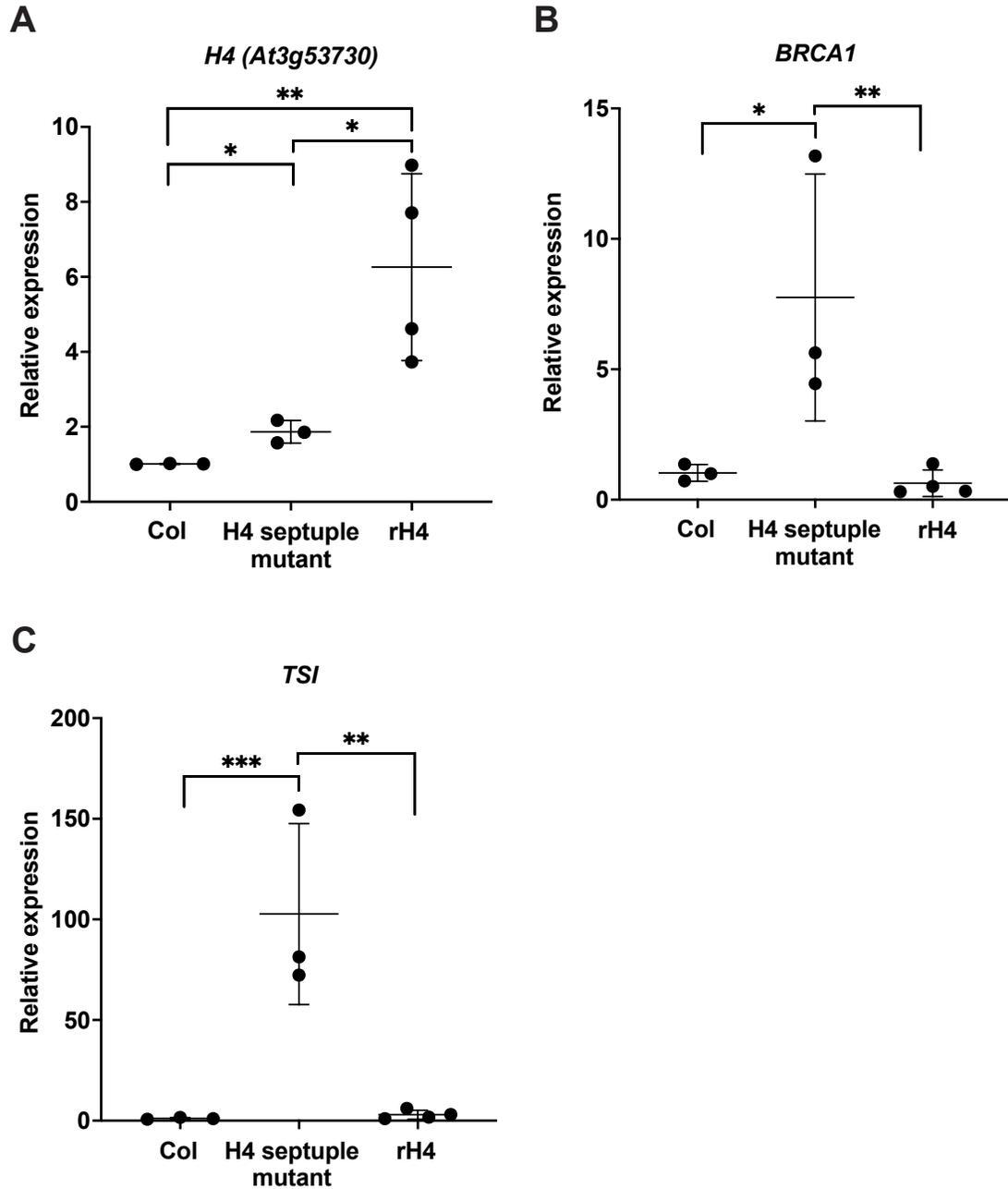


Figure 2.4 Expression levels of H4 (*At3g53730*), *BRCA1*, and *TSI* in H4 septuple mutant and rH4 lines. quantitative Reverse Transcription PCR (RT-qPCR) of (A) H4 (*At3g53730*), (B) *BRCA1* and (C) *TSI* in Col, H4 septuple, and four independent rH4 T1 lines. Three biological replicates were included for Col and H4 septuple mutant plants. Horizontal bars indicate the mean. Standard deviation denoted with error bars. *P*-value from unpaired Student's *t*-test denoted with asterisks (* $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$).

septuple mutants exhibited misregulation of several markers of genomic and epigenomic instability, including upregulation of the DNA damage response gene *BRCA1* and transcriptional de-repression of the DNA repeat *TRANSCRIPTIONALLY SILENT INFORMATION (TSI)* (Figure 2.4B-C).

In addition to the H4 septuple mutant, we also identified several transgenic lines in the T2 generation with differing numbers of mutations in endogenous H4 genes, including the H4 triple mutant (containing homozygous mutations in three out of the eight endogenous H4 genes), the H4 quadruple mutant (containing homozygous or biallelic mutations in four out of the eight endogenous H4 genes), the H4 quintuple mutant (containing homozygous or biallelic mutations in five out of the eight endogenous H4 genes), and the H4 sextuple mutant (containing homozygous or biallelic mutations in six out of the eight endogenous H4 genes) (Table 2.2). Plants with homozygous or biallelic mutations in four or fewer endogenous H4 genes were observed to display wild-type phenotypes (Table 2.2, Figure 2.5). Although H4 sextuple mutants exhibited minor developmental phenotypes including leaf serration and reduced fertility, these phenotypes were less severe than that of H4 septuple mutants (Table 2.2). Finally, we generated an additional transgenic line with homozygous mutations in seven out of the eight endogenous H4 genes, but with the removal of the Cas9 transgene (Table 2.2). The Cas9-negative H4 septuple mutant line exhibited identical phenotypes to the Cas9-positive H4 septuple mutant line (Figure 2.5), and from hereafter, H4 septuple mutant plants will refer to the H4 septuple mutant line retaining Cas9 expression.

Histone H4 replacement plasmids

To set up histone H4 replacement, we designed a new plasmid for plant transformation (i.e., H4 replacement plasmid) that contains 1) a gRNA targeting the last remaining endogenous H4 gene (*At3g53730*) and 2) a modified H4 gene allowing for

H4 gene	H4 triple mutant	H4 quadruple mutant	H4 quintuple mutant	H4 sextuple mutant	H4 septuple mutant (1)	H4 septuple mutant (2)	Expression level	guide RNA
<i>At1g07660</i>	wild-type	wild-type	wild-type	1 bp insertion	1 bp insertion	1 bp insertion	medium	2
<i>At1g07820</i>	wild-type	wild-type	1 bp insertion/ 56 bp deletion	1 bp insertion	56 bp deletion	1 bp insertion	medium	2
<i>At2g28740</i>	wild-type	wild-type	wild-type	wild-type	1 bp insertion	1 bp insertion	low	3
<i>At3g45930</i>	1 bp insertion	1 bp insertion/ 299 bp deletion	1 bp insertion/ 299 bp deletion	1 bp insertion/ 299 bp deletion	299 bp deletion	299 bp deletion	medium	1
<i>At5g59690</i>	wild-type	1 bp insertion	5 bp deletion	2 bp deletion	2 bp deletion	2 bp deletion	high	1
<i>At3g46320</i>	1 bp insertion	4 bp deletion	1 bp deletion	1 bp deletion	1 bp deletion	1 bp deletion	medium	1
<i>At5g59970</i>	1 bp insertion	1 bp insertion	7 bp deletion	1 bp insertion	1 bp insertion	1 bp insertion	high	1
<i>At3g53730</i>	wild-type	wild-type	wild-type	wild-type	wild-type	wild-type	high	N/A
Cas9	-	+	+	-	+	-	N/A	N/A
phenotype	normal	normal	minor	moderate	severe	severe	N/A	N/A

Table 2.2 Mutations in endogenous H4 genes present in H4 depletion backgrounds. Mutations in *At5g59970*, *At3g45930*, *At5g59690*, *At3g46320*, *At1g07660*, *At1g07820*, *At2g28740*, and *At3g53730* in the H4 triple mutant, H4 quadruple mutant, H4 quintuple mutant, H4 sextuple mutant, and H4 septuple mutant backgrounds are displayed. The presence or lack of the Cas9 transgene is also indicated. mRNA expression levels of H4 genes are indicated as described previously (Nakabayashi et al., 2005; Schmid et al., 2005; Waese et al., 2017). Guide RNA targeting H4 gene is displayed (Figure 2.2), and severity of reduced fertility, small size and leaf serration phenotype in each H4 depletion background is noted.

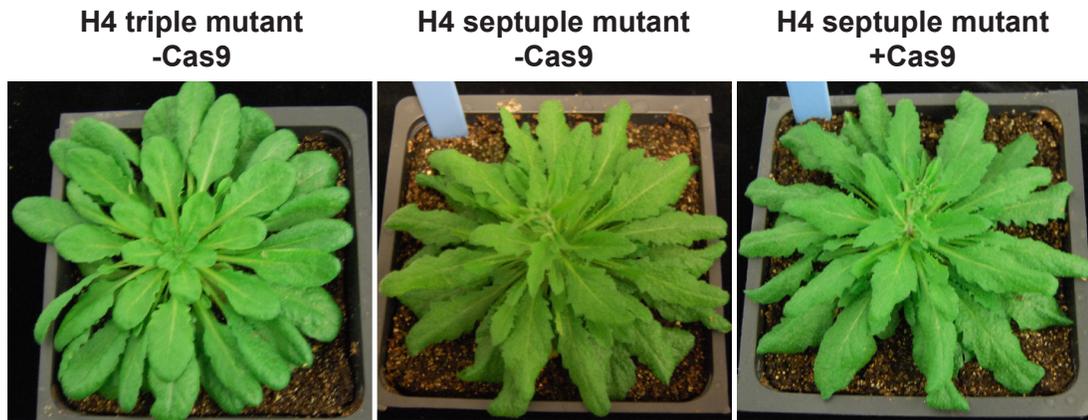


Figure 2.5 Morphological phenotypes of H4 sextuple mutant and H4 septuple mutant plants. Phenotype of H4 triple mutant and two independent transgenic H4 septuple mutant lines grown in short-day conditions at 8 weeks. -Cas9 and +Cas9 indicate presence of transgene expressing Cas9. Mutations in endogenous H4 genes are shown in Table 2.2.

expression of *At3g53730* under its native promoter (i.e., replacement H4 gene) (Figure 2.6A). Our strategy was to transform the H4 septuple mutant, which expresses Cas9, using the H4 replacement plasmid, and select T1 plants that contain mutations at the endogenous *At3g53730* gene. To prevent Cas9 from targeting the replacement H4 gene provided with the T-DNA, we introduced two silent mutations in the replacement H4 gene (*At3g53730*) that prevent recognition from the gRNA targeting the endogenous *At3g53730* (Figure 2.6B). After transformation of the H4 septuple mutant with the H4 replacement plasmid, we recovered many T1 transformants expressing the replacement H4 gene (i.e., rH4 plants), and in contrast to the H4 septuple mutant, rH4 plants were normal in size, did not exhibit serrated leaves and showed normal fertility (Figure 2.3A-B). Moreover, the RNA expression levels of *BRCA1* and *TSI* in rH4 plants were comparable to levels observed in Col (Figure 2.4A-C). The expression of the replacement H4 gene in rH4 plants was found to be upregulated approximately 4- to 9-fold relative to Col (Figure 2.4A). These results indicate that high expression levels of the replacement H4 gene are responsible for rescuing the H4 septuple mutant phenotypes in the rH4 plants.

We then used site-directed mutagenesis to create a large library of plasmid constructs carrying a point mutation in the H4 replacement gene. We generated mutations covering every amino acid in histone H4 that could theoretically be post-translationally modified (i.e., lysine, arginine, threonine, serine, and tyrosine) (Johnson et al., 2004; Moraes et al., 2015). We mutated each amino acid to a residue that cannot be post-translationally modified (i.e., alanine, valine or phenylalanine). We also mutated lysine and arginine residues to residues that are chemically similar (i.e., arginine and lysine, respectively). In total, we modified 38 residues to generate 63 different H4 replacement genes containing a specific point mutation (Figure 2.7A). We subcloned these H4 mutant genes into the H4 replacement plasmid and individually transformed

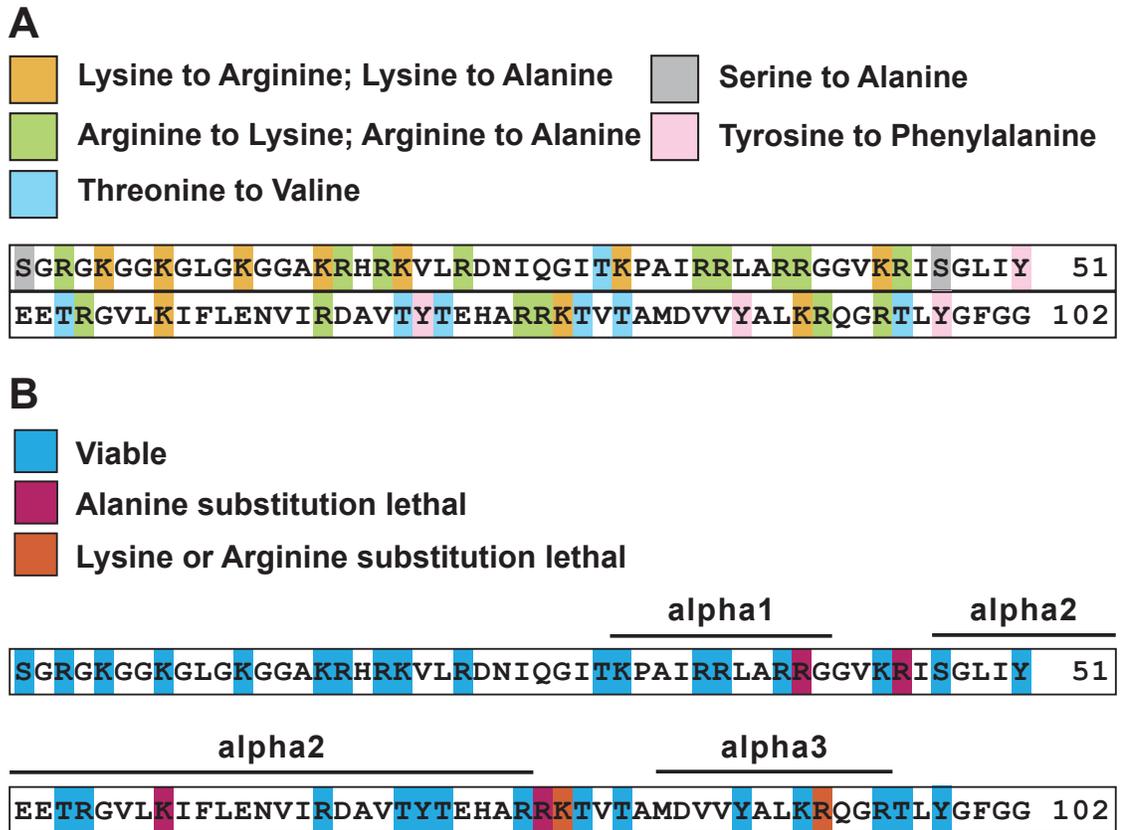


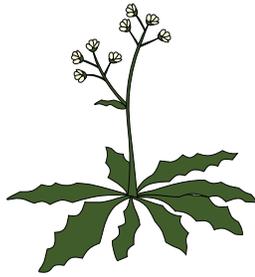
Figure 2.7 63 mutations generated for systematic screens of histone H4. (A) Schematic of point mutations in the H4 replacement plasmid library. (B) Graphic representation of viability of histone H4 substitution mutants. The histone fold is represented above the relevant amino acids (Govin et al., 2010a; White et al., 2001).

them into the H4 septuple mutant. We selected two independent transgenic lines for each H4 mutant, except for plants expressing the replacement gene H4 arginine 40 to alanine (rH4R40A), rH4R45A, rH4K59A, rH4R78A, rH4K79R and rH4R92K due to lethality induced by these mutations (Figure 2.7B). All T1 plants were exposed to heat stress treatments to maximize the efficiency of targeted mutagenesis of the remaining endogenous H4 gene by Cas9 (Figure 2.8).

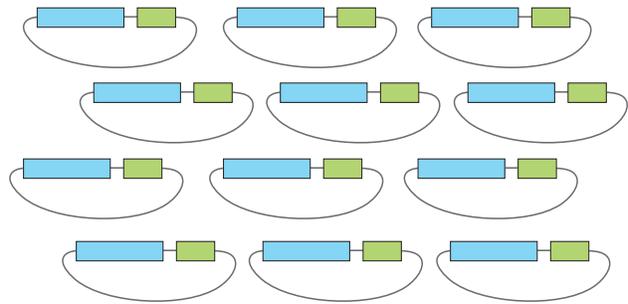
Characterization of histone H4 replacement lines

To confirm high expression of the replacement H4 gene in rH4 mutant plants, we measured H4 (*At3g53730*) expression in five rH4S1A T1 plants and demonstrated that there is a 6- to 26-fold increase in expression relative to Col (Figure 2.9). Next, to estimate the frequency of mutation of the remaining endogenous H4 gene, we genotyped three plants each from two independent rH4 lines and two independent rH4K16A lines at the T2 generation stage. The H4 mutations assayed were randomly chosen out of the entire library of H4 replacement genes. We amplified the remaining endogenous H4 gene (*At3g53730*) from these T2 plants, cloned the resulting PCR products and sequenced at least ten individual clones corresponding to each plant, and calculated the percentage of mutated alleles. Approximately half of the plants were characterized by a complete elimination of wild-type *At3g53730* alleles, while the other plants varied from approximately 50% to 75% wild-type alleles remaining (Figure 2.10). Taking into account that expression of the replacement H4 gene is either equivalent or much higher compared to the remaining endogenous H4 gene (Figure 2.4A, Figure 2.9), these results suggest that the chromatin of most T2 plants in our H4 replacement collection contains large amounts of H4 point mutants. Overall, these results show that our CRISPR-based strategy was successful in creating a large collection of *A. thaliana* plants expressing different H4 point mutants replacing wild-type H4 proteins.

Generate H4 septuple mutant



Generate 63 H4 replacement plasmids



Individually transform all H4 replacement plasmids into H4 septuple mutant background

Select T1 plants on Basta plates and subject T1 plants to short heat stress treatments during vegetative growth to optimize efficiency of CRISPR/Cas9

cold treatment (s)	T1 selection	acclimation to soil	37°C 30 h	22°C 42 h	37°C 30 h	22°C 42 h	37°C 30 h	22°C 42 h	37°C 30 h	plants recover and set seeds
3 days	5 days	4 days	10 days						remainder of plant life	

collect T2 generation seeds

Select T2 plants on Basta plates and screen for phenotypes



Figure 2.8 Workflow for histone H4 replacement system. Schematic representation of the steps and growth conditions utilized to optimize selection of histone H4 replacement plants. s: stratification at 4°C, h: hours.

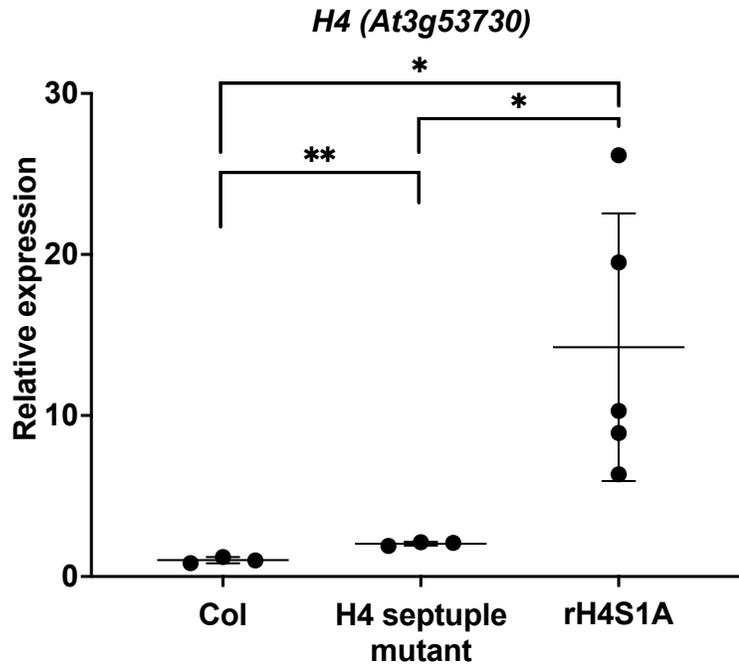


Figure 2.9 H4 expression levels in rH4S1A mutant lines. RT-qPCR of *H4* (*At3g53730*) in Col, H4 septuple, and five independent rH4S1A T1 lines. Three biological replicates were included for Col and H4 septuple mutant plants. Horizontal bars indicate the mean. Standard deviation denoted with error bars. *P*-value from unpaired Student's *t*-test denoted with asterisks (* $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$).

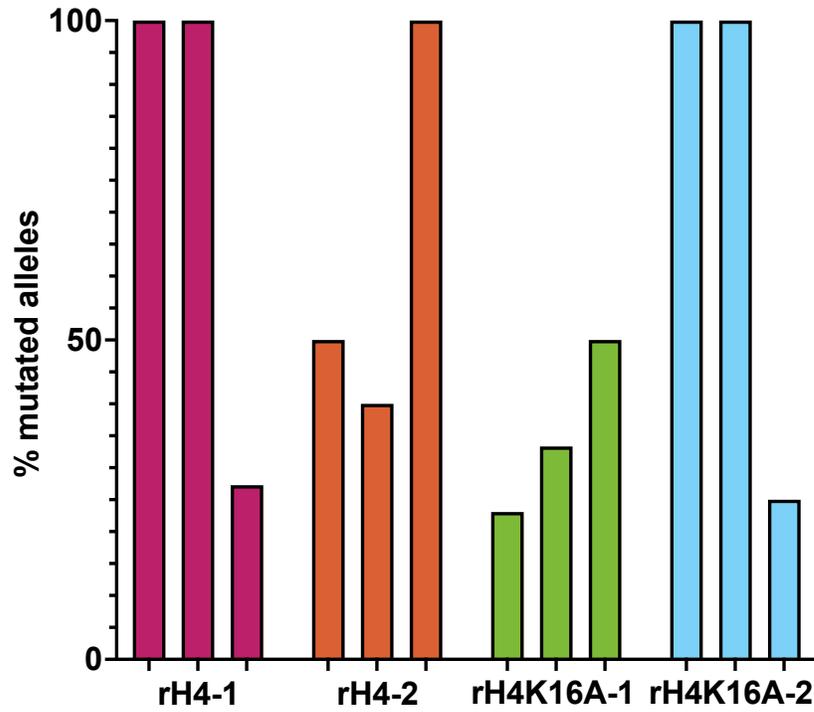


Figure 2.10 Efficiency of CRISPR/Cas9 in H4 replacement plants. Percentage of mutated alleles in six rH4 plants and six rH4K16A plants. Each plant assessed was from the T2 generation; three plants from the same T1 parent were used in this experiment (i.e., two independent T1 lines per genotype).

Discussion

A CRISPR/Cas9-based strategy for histone replacement in *Arabidopsis thaliana*

Aided by the efficiency of novel genome engineering technologies, we generated an efficient histone H4 replacement system in the model organism *Arabidopsis thaliana*. The experiments performed in this chapter outline a strategy for complete histone replacement in *Arabidopsis thaliana* that could be utilized to generate complete histone replacement systems targeting other histones. The successful generation of the H4 septuple mutant in two generations demonstrates the efficiency of temperature-optimized CRISPR/Cas9 for multiplexed gene targeting, as three gRNAs were used to simultaneously target seven histone H4 genes to create homozygous mutations (LeBlanc et al., 2017). In the initial published usage of temperature-optimized CRISPR/Cas9 by LeBlanc et al., a single gRNA targeted a single locus for mutagenesis; the generation of the H4 septuple mutant substantiates that this method can be used for the simultaneous targeting of as many as seven independent loci. Previous strategies to generate histone depletion backgrounds in plants typically relied upon traditional crossing strategies, which would take many generations to complete when a large number of genes were being targeted for inactivation. Consequently, the efficacy of temperature-optimized CRISPR/Cas9 significantly increases the ease of simultaneously mutating multiple endogenous histone genes in order to generate a complete histone replacement system.

Depletion of endogenous histone H4 causes genome instability and abnormal morphology

The H4 septuple mutant demonstrates interesting morphological and molecular phenotypes as a result of the histone H4 deficiency. One probable mechanism for the derepression of transcriptionally silenced genomic elements and the upregulation of

DNA damage response elements in the H4 septuple mutant is that the lack of sufficient histone H4 causes replication fork stalling and/or defective chromatin assembly, leading to genome instability (Aguilera and Garcia-Muse, 2013; Alabert and Groth, 2012) (Figure 2.4B-C). Extensive evidence supports that the rate of DNA synthesis and replication fork progression is tightly coupled to histone deposition and chromatin assembly (Groth et al., 2007; Gunesdogan et al., 2014; Hoek and Stillman, 2003; Mejlvang et al., 2014; Nelson et al., 2002; Zhao et al., 2004). Moreover, impeded replication fork progression and impaired chromatin assembly can cause deleterious genomic outcomes, including DNA damage (Alabert and Groth, 2012; Nabatiyan and Krude, 2004; Ye et al., 2003). Interestingly, partial depletion of histone H4 in *Saccharomyces cerevisiae* has been shown to cause an increase in HR between ectopic DNA sequences associated with an accumulation of recombinogenic DNA lesions during the S and G2/M phases (Prado and Aguilera, 2005). Therefore, an increase of DNA damage due to impaired replication caused by the decreased histone H4 supply is a likely factor responsible for the upregulation of DNA damage response elements such as *BRCA1* in the H4 septuple mutant.

The upregulation of the DNA repeat *TSI* could also be a consequence of replication stress, which can lead to the loss or gain of epigenetic information encoded on histones and DNA, as well as initiate more dramatic chromatin reorganization during fork collapse, causing transcriptional derepression of typically silenced chromatin regions (Alabert and Groth, 2012; Jasencakova et al., 2010; Sarkies et al., 2010; Zaratiegui et al., 2011). The increased expression of *TSI* could also be due to a reduction of nucleosome content caused by the depletion of histone H4. When histone H4 was depleted in *Saccharomyces cerevisiae*, approximately 15% of genes were found to have increased expression, especially within the telomeric heterochromatin (Wyrick et al., 1999). To distinguish between these two mechanisms, future experiments could

compare nucleosome occupancy and RNA levels corresponding to genomic regions such as *TSI* in the H4 septuple mutant and wild-type plants with Assay for Transposase-Accessible Chromatin using sequencing (ATAC-seq) and whole-transcriptome analysis via RNA sequencing (RNA-seq). Additionally, PTMs associated with different genomic regions could be assayed using chromatin immunoprecipitation followed by sequencing (ChIP-seq) in order to assess how the chromatin landscape is altered in the H4 septuple mutant. In summary, some combination of replication stress and impaired nucleosome assembly is likely responsible for the molecular phenotypes observed in the H4 septuple mutant, and future experiments more thoroughly assessing chromatin organization, epigenetic information, and genome-wide transcription could elucidate a mechanism for these phenotypes.

The most striking morphological phenotypes of the H4 septuple mutant (i.e., reduced growth, serrated leaves, and reduced fertility) could be a consequence of cell cycle defects caused by replication stress as a result of the lack of histone H4 (Figure 2.3A-B). Histone supply levels have been shown to regulate the length of S phase in *Drosophila melanogaster*, and a lack of sufficient histone supply causes cell cycle arrest during G2 phase (Gunesdogan et al., 2014). Additionally, defective chromatin assembly during S phase and the lack of histone expression inhibit cell cycle progression and cell proliferation in mammalian cells (Nabatiyan and Krude, 2004; Ye et al., 2003; Zhao et al., 2004). Furthermore, *Arabidopsis thaliana* mutants with reduced cell proliferation due to cell cycle defects display similar phenotypes to the H4 septuple mutant. For example, mutations in *F BOX-LIKE 17 (FBL17)* drastically alter cell cycle regulation and inhibit cell proliferation, and *fb17* mutants show decreased leaf size, leaf serration, and partial sterility (Gusti et al., 2009; Noir et al., 2015). In addition, mutant lines overproducing different KIP-RELATED PROTEINs (KRPs) exhibit impaired cell cycle progression, leaf serration, and partial sterility (De Veylder et al., 2001). Interestingly, *fb17* mutants also

display impaired DNA replication and upregulation of *BRCA1*, although these molecular phenotypes appear to be an effect of the cell cycle defects, rather than a cause of them, as proposed for the H4 septuple mutant (Noir et al., 2015). However, it is also possible that cell cycle defects contribute to the *BRCA1* upregulation observed in the H4 septuple mutant.

In summary, the lack of histone H4 expression in the H4 septuple mutant may lead to defective chromatin assembly, S phase arrest, and consequently, impaired cell proliferation, causing diminished growth, abnormal morphology, and reduced fertility. Further experiments could test this model by assaying the expression of cell cycle regulatory genes in the H4 septuple mutant using RT-qPCR and assessing cell proliferation using microscopy to examine leaf cell number and cell size. Staining experiments of primary root tips with propidium iodide followed by microscopy could also reveal cell death phenotypes of the H4 septuple mutant compared to wild-type plants.

Histone H4 depletion triggers a dosage compensation mechanism

While the lack of histone H4 supply is likely the cause of defective growth and genome instability in the H4 septuple mutant, there does appear to be a dosage compensation mechanism in this mutant partially countering the effect of eliminating expression from seven out of the eight endogenous histone H4 genes (Figure 2.4A, Figure 2.9). Elucidating the mechanism by which this histone dosage compensation occurs is one future direction for this work, and multiple questions related to the histone dosage compensation in the H4 septuple mutant remain. First, which histone genes are being upregulated in the H4 septuple mutant? For example, all of the canonical histone genes may be affected, only histone H4 genes may be targeted for upregulation, or only histone H3 and H4 genes may be affected. If the expression of all of the canonical histone genes is increased, it is possible that there may be some toxicity as a result of

overexpression of histones H2A, H2B, and H3, but there are mechanisms for the posttranscriptional regulation of excess histone proteins to combat this outcome (Eriksson et al., 2012; Gunjan and Verreault, 2003; Liang et al., 2012; Reichheld et al., 1998; Singh et al., 2010). Examining whether histone dosage compensation occurs in other histone mutants such as the H3.1 quadruple mutant, which has mutations in four out of the five endogenous histone H3.1 genes (Jacob et al., 2014), may shed light onto the specificity of this mechanism for histone H4. Moreover, examining the expression of other endogenous histone genes in the H4 septuple mutant compared to wild-type plants using RT-qPCR and Western blot would elucidate which histone genes are targeted for upregulation.

Second, how is the lack of histone H4 being sensed in the H4 septuple mutant in order to signal for the dosage compensation to occur? A prolonged insufficient supply of all of the canonical histones or the depletion of individual histone subtypes has been shown to limit replication fork progression and induce checkpoint signaling (Mejlvang et al., 2014; Nelson et al., 2002; Zhao et al., 2004; Zheng and Hayes, 2003). An activation of the S phase checkpoint due to impaired replication fork progression and resulting DNA damage is thus one major possibility for how cells are sensing the lack of histone H4 in the H4 septuple mutant. Assessing the activation of cell cycle checkpoint responses in the H4 septuple mutant could provide initial support for this hypothesis. This hypothesis could also be tested by mutating cell cycle response genes in an H4 septuple mutant background and evaluating whether the histone dosage compensation is altered. Additionally, it would be interesting to examine other histone mutants such as the H3.1 quadruple mutant to determine whether the lack of other histone subtypes induces checkpoint signaling. However, it is important to note that histone H4 is unique among histone proteins in that it lacks sequence variants. Thus, the mutation of all but one of the endogenous histone H4 genes may cause a more severe effect on replication

than the mutation of all but one of the endogenous histone H3.1 genes, as there are additional histone H3 variant genes present that may be able to compensate for the loss of histone H3.1 (Holmes et al., 2005; Yuan and Zhu, 2013). In fact, studies in *Drosophila melanogaster* have demonstrated that canonical and variant histone H3 are able to functionally replace each other in most cell types (Hodl and Basler, 2009, 2012).

Finally, what mechanisms are responsible for the upregulation of histone H4 expression in the H4 septuple mutant? The upregulation observed from RT-qPCR data in Figure 2.4A and Figure 2.9 indicates that a mechanism exists to increase histone expression at the mRNA level, but these data do not reveal whether there are additional mechanisms that regulate translation. Interestingly, histone dosage compensation has also been observed in the histone replacement systems implemented in the multicellular eukaryote *Drosophila melanogaster* (McKay et al., 2015b; Zhang et al., 2019). Although wild-type diploid flies contain approximately 200 copies of the histone repeat unit, a single transgene encoding between 12 and 24 copies of the histone repeat unit was found to not only support development of flies lacking all endogenous histone genes, but also to generate the same amount of histone protein and mRNA as found in wild-type flies (McKay et al., 2015b; Zhang et al., 2019).

While the precise mechanism for this dosage compensation requires further investigations, the clustering of the canonical histone genes at a single genomic location provides a straightforward method of upregulating the expression of all histone genes in *Drosophila melanogaster*. In this organism, *cis*-regulatory elements within the *histone3-histone4* promoter have been shown to be essential for the transcription of histones H3 and H4 through the recruitment of DNA-binding proteins, which increase chromatin accessibility and allow for the activation of the other canonical histone genes through the formation of the histone locus body (Duronio and Marzluff, 2017; Rieder et al., 2017; Salzler et al., 2013). In plant systems, the canonical histone genes are dispersed

throughout the five chromosomes rather than being clustered at a single locus, but many histone promoters similarly contain conserved sequences that act as *cis*-regulatory elements by serving as binding sites for transcription factors (Brignon and Chaubet, 1993; Chaboute et al., 1987; Chaubet et al., 1996; Minami et al., 1993; Reichheld et al., 1998; Tabata et al., 1991; Taoka et al., 1999). Some *cis*-regulatory elements have been found to be common to nearly all replication-dependent histone genes in plants, while others appear specific to particular histone subtypes (Moes et al., 2013; Taoka et al., 1998). The positive regulation of transcription factors targeting these *cis*-regulatory elements serves as one possibility for how histone dosage compensation occurs.

A second mechanism to upregulate histone genes at the transcriptional level could be related to the recently described process of transcriptional adaptation, in which mutant mRNA decay causes the upregulation of related genes (El-Brolosy et al., 2019; Serobyan et al., 2020; Sztal and Stainier, 2020). As we detected that most of the CRISPR/Cas9-induced mutations we generated in endogenous histone H4 genes resulted in a frameshift of the coding sequence leading to a premature termination codon (Table 2.1), it is thus theoretically possible that these frameshift mutations cause the nonsense-mediated decay of these mutant H4 mRNAs (He and Jacobson, 2015; Shaul, 2015; Shyu et al., 2008). However, histone H4 is a naturally intronless gene, and although it has been shown that intronless genes are capable of undergoing nonsense-mediated decay (He and Jacobson, 2015; Rajavel and Neufeld, 2001), it was also demonstrated that human H4 was incapable of undergoing nonsense-mediated decay upon introduction of a premature stop codon (Maquat and Li, 2001). Conversely, the frameshift mutations could also cause the H4 mRNA to be identified as an aberrant transcript, e.g., by generating an unusual 3' UTR (Parker and Song, 2004; Siwaszek et al., 2014; Szadeczky-Kardoss et al., 2018). In this case, mechanisms such as non-stop decay and no-go decay could also lead to mRNA degradation (Szadeczky-Kardoss et

al., 2018). To test whether mRNA degradation and transcriptional adaptation are contributing to the histone dosage compensation mechanism observed in *A. thaliana*, additional H4 depletion lines could be generated by using paired gRNAs with CRISPR/Cas9 to completely delete the coding sequence of endogenous H4 genes. These H4 depletion lines could then be compared to the H4 depletion lines containing frameshift mutations in order to assess histone H4 expression levels with RT-qPCR.

If a mechanism also exists to upregulate histone expression at the posttranscriptional level, how this process may occur in *Arabidopsis thaliana* is less clear. Replication-dependent histone mRNAs in metazoans such as *Drosophila melanogaster* are unique in that they lack polyadenylation, instead containing a stem-loop structure at their 3' ends (Davila Lopez and Samuelsson, 2008). This structure regulates the processing, translation and stability of histone mRNAs by binding stem-loop binding protein (SLBP), a protein involved in histone mRNA metabolism (Lanzotti et al., 2002; Marzluff et al., 2008; Sullivan et al., 2001; Zanier et al., 2002). Thus, the 3' stem-loop structure specific to metazoan replication-dependent histone mRNAs provides a mechanism for the post-transcriptional regulation of histone expression. However, this mechanism is not shared in plant systems such as *Arabidopsis thaliana*, as histone mRNAs in these organisms are polyadenylated, and the major way that histone expression has been shown to be positively regulated in plants is via transcriptional processes (Davila Lopez and Samuelsson, 2008; Moes et al., 2013).

Another process that may be contributing to the upregulation of histone H4 expression in the H4 septuple mutant is the regulation of S phase length. As in yeast and animals, expression of replication-dependent histone genes in plants is highly coordinated and predominantly specific to S phase (Minami et al., 2000; Moes et al., 2013; Reichheld et al., 1998; Taoka et al., 1999). In *Drosophila melanogaster*, S phase can be completed under a constant, diminished histone supply, but this diminished

histone supply has been shown to extend the length of S phase by slowing replication fork progression (Gunesdogan et al., 2014). It is possible that the reduced histone H4 supply in the H4 septuple mutant similarly extends the length of S phase and this extended S phase increases the supply of replication-dependent histones due to the coordination of histone biosynthesis with this phase of the cell cycle (Marzluff et al., 2008; Mejlvang et al., 2014). Experiments assessing histone expression in cell cycle mutants with increased S phase lengths could support this hypothesis, although regulatory mechanisms countering histone overexpression may complicate these analyses (Eriksson et al., 2012; Reichheld et al., 1998). The complete arrest of DNA replication has been shown to elicit histone gene repression or mRNA degradation in order to mediate cytotoxic effects from excess histone levels, and thus there is likely a balance between mechanisms increasing and repressing histone expression in response to impaired DNA replication (Kaygun and Marzluff, 2005b; Osley, 1991).

In summary, studies of the H4 septuple mutant have the potential to elucidate fundamental processes including the regulation of histone biosynthesis and metabolism in plants. One genetic tool in addition to the H4 septuple mutant that could be used to answer the aforementioned questions is the additional H4 depletion backgrounds generated, including the H4 triple mutant, H4 quadruple mutant, H4 quintuple mutant, and H4 sextuple mutant (Table 2.2). Additional mutants could further vary the amount of endogenous histone H4 expressed (or vary which endogenous histone H4 genes remain unmutated) in order to study the resultant phenotypes. These mutants provide a fairly analogous system in plants to the system developed in the metazoan *Drosophila melanogaster* (Gunesdogan et al., 2010). Gunesdogan et al. utilized this system to reintroduce defined numbers of transgenic histones genes into a background lacking endogenous histone expression in order to study the rate of DNA replication and S phase length relative to histone supply (Gunesdogan et al., 2014). Similar experiments

could be performed with histone H4 mutants in *Arabidopsis thaliana* to further elucidate fundamental relationships between histone supply and DNA replication in plants.

Another interesting question that these histone H4 mutants could be used to address is why *Arabidopsis thaliana* contains eight endogenous histone H4 genes that encode identical proteins, and whether any of these genes are redundant for plant fitness and development. While more detailed analyses are required to thoroughly answer this question, we observed that mutants retaining wild-type expression from four or more of the endogenous histone H4 genes were indistinguishable from wild-type plants morphologically (Table 2.2). This observation suggests that all eight of the endogenous histone H4 genes are not necessary for normal development in *Arabidopsis thaliana*. However, it is possible that while these histone H4 mutants appear morphologically wild-type, expression of the additional histone H4 genes may be advantageous under certain environmental circumstances and contribute to plant fitness. On the other hand, histone dosage compensation appears to be occurring in the absence of mRNA expression (and/or in the presence of mRNA degradation) from certain histone H4 genes, and this dosage compensation may be sufficient to ensure an identical phenotype to that of plants which retain expression from all eight endogenous histone H4 genes. With our RT-qPCR experiments, we observed a 2-fold upregulation of mRNA expression from the remaining endogenous H4 gene (*At3g53730*) in H4 septuple mutant plants (Figure 2.4A, Figure 2.9), and therefore, if a similar upregulation were applied to all unmutated H4 genes in other H4 depletion lines, such as the H4 quadruple mutant, this upregulation would likely approximate the endogenous histone H4 levels present in Col plants.

Subsequent analyses mutating different combinations and numbers of endogenous H4 genes would further elucidate the requirement in plants for different H4 genes. Interestingly, in the H4 triple mutant, H4 quadruple mutant, H4 quintuple mutant,

and H4 sextuple mutant that we identified, the loss-of-function mutations were mostly located within relatively highly expressed H4 genes (Table 2.2) (Nakabayashi et al., 2005; Schmid et al., 2005; Waese et al., 2017). For example, one of the two endogenous H4 genes that remained unmutated in the H4 sextuple mutant (*At2g28740*) displayed the lowest mRNA expression out of all of the eight endogenous H4 genes (Table 2.2) (Nakabayashi et al., 2005; Schmid et al., 2005; Waese et al., 2017). These differential rates of mutagenesis are likely due to differences in gRNA efficiency, as the H4 genes with the highest rates of mutagenesis were all targeted by the same gRNA (Figure 2.2, Table 2.2). Nonetheless, retaining wild-type expression of more highly expressed endogenous H4 genes could allow for the development of a H4 quintuple mutant or even H4 sextuple mutant that exhibits a wild-type phenotype. A different combination of gRNAs may need to be utilized in order to achieve this result, as the same gRNA targeted two of the most highly expressed H4 genes and two H4 genes with medium expression levels simultaneously with our strategy, and this gRNA demonstrated the ability to efficiently target all four of these genes for mutagenesis in the same generation (Figure 2.2, Table 2.2). However, due to the high mutagenic efficiency that we observed with our multiplex CRISPR/Cas9 strategy (LeBlanc et al., 2017), the establishment of additional H4 depletion lines remains feasible under a reasonable amount of time and resources.

Overexpression of histone H4 in H4 replacement lines may trigger dosage compensation

In an analogous manner to histone depletion, an overabundance of histone supply has also been shown to lead to genome instability in various organisms. The presence of excess histones has been linked to increased DNA damage sensitivity, cytotoxicity, and mitotic chromosome loss (Gunjan and Verreault, 2003; Meeks-Wagner

and Hartwell, 1986; Singh et al., 2009; Singh et al., 2010). Two of the major mechanisms for these deleterious phenotypes appear to be the non-specific binding of excess histones to DNA and RNA, and the saturation of some histone-modifying enzymes due to increased binding (Singh et al., 2010). In *Saccharomyces cerevisiae*, Rad53 targets excess, non-chromatin-bound histones for degradation by the ubiquitin-proteasome pathway, and a loss of Rad53 causes an accumulation of excess histones and resulting genome instability phenotypes (Gunjan and Verreault, 2003; Singh et al., 2009). In metazoans, SLBP recognizes the conserved stem-loop structure at the 3' end of histone mRNAs to target these transcripts for degradation at the end of S phase or when DNA synthesis is inhibited (Kaygun and Marzluff, 2005a, b; Sittman et al., 1983). Therefore, various mechanisms exist to combat the accumulation of excess histones, and can serve to regulate histone supply under conditions of histone overexpression or inhibition of DNA replication.

These processes to combat excess histone accumulation may be relevant for the H4 replacement lines, which display varying amounts of histone H4 mRNA expression from the H4 replacement transgene. While the mRNA levels of H4 (*At3g53730*) were found to be approximately 4- to 26-fold higher in rH4 plants compared to Col (Figure 2.4A, Figure 2.9), further RT-qPCR experiments measuring mRNA expression from all H4 genes and Western blot experiments comparing H4 protein levels in these lines would clarify whether the total histone H4 expression in rH4 plants is altered compared to wild-type plants. One expectation would be that the rH4 plants have similar histone H4 protein levels to Col, since the transformation with the wild-type H4 transgene rescued the H4 septuple mutant phenotype. As *At3g53730* is one of the most highly expressed histone H4 genes in *Arabidopsis thaliana*, a 4- to 26-fold increase in expression relative to Col likely restores wild-type histone H4 mRNA levels to these lines and at the highest end, may even cause an overexpression of histone H4. Therefore, mechanisms that

degrade excess histone proteins or transcripts would be necessary to prevent cytotoxic effects in the event of histone overexpression in these lines.

Indeed, in other organisms, the introduction of additional copies of histone genes activates dosage compensation mechanisms to maintain a constant ratio between histone supply and the rate of DNA replication. In yeast, inserting an extra H2A:H2B gene pair into the haploid genome does not increase the steady-state levels of histone H2A and H2B mRNAs (Osley and Hereford, 1981). While the transcription rate of H2A and H2B is increased in these lines, an increased turnover of the histone transcripts maintains the level of histones H2A and H2B through a post-transcriptional mechanism. Similarly, an activation of an inducible transgene encoding histone H1-GFP in *Drosophila melanogaster* leads to an approximately 10-fold increase in the total H1 mRNA level, but a post-transcriptional mechanism causes a compensatory decrease in the translation of this mRNA into protein, as the H1 protein level in these mutants is not significantly different from that of wild-type flies (Siriaco et al., 2015). In addition, expressing a transgene encoding 24 copies of the histone repeat unit in a wild-type *Drosophila melanogaster* background does not significantly alter the histone mRNA level compared to wild-type flies due to a reduction of histone expression from both endogenous and transgenic histone genes (McKay et al., 2015b). These results demonstrate that both transcriptional and post-transcriptional mechanisms for dosage compensation under conditions of histone overexpression exist.

Varying expression of histone H4 in H4 replacement lines as a consequence of T-DNA integration

Histone H4 in the H4 replacement lines was expressed under its native promoter (*At3g53730*) from T-DNA stably integrated into the genome of the H4 septuple mutant background. Despite the fact that all replacement histone H4 genes were expressed

under the same native promoter, varying H4 mRNA levels were observed in different H4 replacement lines (Figure 2.4A, Figure 2.9). A likely explanation for this observation is that random T-DNA integration causes differing levels of expression from the H4 replacement genes. When *Arabidopsis thaliana* is transformed with T-DNA from *Agrobacterium tumefaciens*, the T-DNA integrates at one or several random genomic loci (Gelvin, 2017). Additionally, the copy number of T-DNA insertions has been observed to vary from a single copy to tens of copies (De Buck et al., 2009; Gelvin, 2017). The integration of the T-DNA into more accessible, open chromatin regions could cause a higher level of expression of the H4 replacement gene. Additionally, higher T-DNA copy numbers could cause higher expression of the H4 replacement gene; however, multiple T-DNA integration events and inverted repeat T-DNA integration patterns have also been shown to decrease transgene expression via silencing mechanisms (Oltmanns et al., 2010). Therefore, there appears to be a balance regarding expression from different numbers of T-DNA insertions, as a very low copy number may result in low expression, while a very high number of T-DNA insertions can conversely cause silencing of the transgene.

Another interesting observation from H4 expression analyses of rH4 plants is that the total H4 mRNA expression from *At3g53730* always appears to be at least ~4-fold higher than in wild-type plants (Figure 2.4A, Figure 2.9). This high level of expression may seem counterintuitive given the random nature of T-DNA integration, but several explanations for this phenomenon exist. First, while T-DNA integration into the *Arabidopsis thaliana* genome occurs randomly, selection pressure (e.g., due to the use of antibiotics or herbicides to select for transgenic plants) appears to shift the recovery of T-DNA insertions into more transcriptionally active chromatin regions (Alonso et al., 2003; Brunaud et al., 2002; Kim et al., 2007; Koncz et al., 1989; Szabados et al., 2002). As transgenic plants were selected on herbicide-containing growth plates, selection bias

may partially account for the high H4 expression observed from the rH4 plants. Second, dosage compensation mechanisms acting to upregulate expression of the endogenous histone H4 gene (*At3g53730*) may also act on the histone H4 replacement gene, as the expression of both of these genes is driven by the same promoter. For both of these reasons, transgenic plants with high levels of histone H4 expression may generally be observed.

Additionally, the primers used in the RT-qPCR experiment measuring H4 mRNA expression from *At3g53730* in Figure 2.4A and Figure 2.9 amplify both the endogenous transcripts as well as transcripts expressed from the H4 replacement transgene. As the majority of CRISPR/Cas9-induced mutations that we observed were 1-2 nucleotide insertions or deletions, we would hypothesize that most CRISPR/Cas9-induced mutations in the endogenous *At3g53730* gene would not prevent the RT-qPCR primers from amplifying the endogenous transcripts, although large deletions that could remove primer binding sites were occasionally observed (Figure 2.2, Figure 2.10, Table 2.1). Additionally, while it is currently unknown whether the CRISPR/Cas9-induced H4 mutations lead to degradation of the H4 mRNAs, either through nonsense-mediated decay, non-stop decay, or no-go decay (He and Jacobson, 2015; Parker and Song, 2004; Shaul, 2015; Shyu et al., 2008; Siwaszek et al., 2014; Szadeczky-Kardoss et al., 2018), if any of these processes do indeed act upon the mutant H4 mRNAs, endogenous mutant histone H4 would not be amplified by the RT-qPCR experiment due to degradation. Additional experiments, such as RT-qPCR experiments with other endogenous H4 genes in the H4 septuple mutant background in addition to *At3g53730*, would elucidate whether degradation of histone H4 mRNA is occurring in the H4 septuple mutant.

In the absence of mRNA degradation, if the dosage compensation mechanism observed in the H4 septuple mutant also increased the expression of the H4

replacement transgene 2-fold, the lowest H4 mRNA expression levels that we recorded in the rH4 mutants (~4-fold increase compared to Col) could be explained by simply summing the expression of the endogenous *At3g53730* gene and the H4 replacement transgene. However, even the rH4 plants with the lowest levels of H4 mRNA expression rescued the H4 septuple mutant phenotype and exhibited normal morphological phenotypes, even when there was a complete elimination of expression from the remaining endogenous histone H4 gene. Thus, the amount of functional histone H4 appears to always be higher in the rH4 mutants compared to the H4 septuple mutants. Further work is required to determine the exact amount of functional histone H4 being expressed from the replacement H4 gene compared to H4 septuple mutant and Col plants. Additional RT-qPCR experiments measuring H4 mRNA levels in rH4 lines with characterized deletions in the primer binding sites located on the remaining endogenous H4 gene, as well as Western blots of histone H4, could further clarify this question. Moreover, analysis of histone mRNA expression in Col and rH4 plants through RNA-seq experiments will be discussed further in Chapter 4.

Low percentage of histone H4 mutations result in lethality

9.5% of point mutations generated in our histone H4 replacement screen resulted in lethality. Mutations were denoted as lethal when no transformants were recovered from hundreds of seeds generated from two independent transformations. Six individual mutations (H4R40A, H4R45A, H4K59A, H4R78A, H4K79R and H4R92K) out of the 63 total mutations assayed were deemed to cause lethality. Interestingly, for all six of these residues, mutants were recovered when a different amino acid substitution was made (e.g., H4R40K instead of H4R40A). In addition to these single amino acid substitution mutants, we also wanted to assess the effect of multiple amino acid substitutions on histone H4 function. As acetylation of lysines 5, 8, and 12 of histone H4 has been shown

to have an additive effect on transcription, we generated an H4 replacement plasmid in which all three of these lysines are substituted with arginine (Dion et al., 2005). We also generated H4 replacement plasmids in which four lysines (5, 8, 12, and 16) and five lysines (5, 8, 12, 16, and 20) were substituted with arginine, as all of these lysines have been implicated in transcriptional regulation (Dion et al., 2005; Kaimori et al., 2016). However, all three of these histone H4 combinatorial mutations resulted in lethality.

Screens in *Saccharomyces cerevisiae* and *Drosophila melanogaster* have also assessed the percentage of histone point mutations resulting in lethality, and differing percentages of essential residues have been calculated. In several screens in *Saccharomyces cerevisiae*, approximately 90% of residues assayed were not essential for viability, while almost 50% of residues assayed on histones H3 and H4 in *Drosophila melanogaster* were found to be essential (Dai et al., 2008; Govin et al., 2010a; Nakanishi et al., 2008; Zhang et al., 2019). Govin et al. individually mutated all endogenous serine, threonine, lysine, and arginine residues on histones H3 and H4 to alanine in *Saccharomyces cerevisiae*, essentially covering all residues assayed in our screen except for the tyrosine residues (Govin et al., 2010a). Three out of 38 (7.9%) of these H4 mutations resulted in lethality. Similar to *Arabidopsis thaliana*, both the H4R40A and H4R45A mutations caused lethality, and in addition, the H4R39A mutation was found to cause lethality in *Saccharomyces cerevisiae*.

While the H4R78A, H4K79R and H4R92K mutants were viable in *Saccharomyces cerevisiae*, H4R78A and H4K79A displayed severely reduced sporulation efficiency (0-20% of wild-type) and H4R92K displayed reduced sporulation efficiency (20-60% of wild-type). In contrast, H4K59A was viable with a sporulation efficiency comparable to that of wild-type cells. In this screen, K77-S83 on histone H4 was identified as the LOS (low sporulation) patch, which is on the edge of the nucleosome in close proximity to contacts with the DNA. DNA replication was found to

be similar to wild-type cells in LOS mutants, but the four normal meiotic products were not observed, suggesting that LOS mutations may impact meiosis. A similar role for residues on this H4 patch in plants could contribute to the lethality observed for the *Arabidopsis thaliana* rH4R78A and rH4K79R mutants.

In addition to the aforementioned screen, Dai et al. substituted each residue on the core histones with one or multiple different amino acids in two yeast strain backgrounds (Dai et al., 2008). 16.7% of histone H4 mutations (covering 16 out of 102 residues) resulted in lethality in both strain backgrounds tested. These histone H4 mutations (H4I34A, H4R39A, H4R40A, H4K44Q, H4R45A, H4R45K, H4Y51E, H4F61A, H4S69D, H4Y72A, H4T73D, H4A76S, H4T80D, H4S83D, H4L84A, H4D85N, and H4L90A) covered many residues not assayed in our screen, but H4R40A and H4R45A were again found to result in lethality. Many lethal substitutions in this screen including H4R40A and H4R45A were mapped to the nucleosome entry site on histone H4 (I34-Y51). This region may be highly sensitive to mutations due to the effect of these mutations on DNA wrapping around the nucleosome (Zhou et al., 2019).

A third screen conducted by Nakanishi et al. generated alanine mutants for all residues of the core histones, except endogenously occurring alanine residues (Nakanishi et al., 2008). Only five out of the 96 mutations assayed in this screen (5.2%) resulted in lethality. All five of these mutations (H4R39A, H4R40A, H4R45A, H4Y72A, and H4L90A) were also identified in the screen by Dai et al.; however, Dai et al. identified many additional lethal mutations compared to this screen (Dai et al., 2008). One reason for the differences in essential residues found between these three yeast screens could be the different strain backgrounds assayed. Notably, Dai et al. identified several residues where mutations only resulted in lethality in one out of the two strain backgrounds tested. Additionally, the extent of the elimination of endogenous wild-type histone expression (i.e., a complete or partial replacement of the endogenous histones)

and the amino acid substitutions generated are two other factors that could contribute to the differences in lethality observed between the yeast screens.

In addition to these screens in *Saccharomyces cerevisiae*, Zhang et al. generated individual alanine substitution mutants in *Drosophila melanogaster* covering all known modified residues on histones H3 and H4 (Zhang et al., 2019). Of the 14 histone H4 mutations assayed, they found that H4Y88A and H4K91A were embryonic lethal and H4K12A was lethal in pupae. In contrast, H4S1A, H4R3A, H4K5A, H4K8A, H4K16A, H4K20A, H4R23A, H4K31A, H4Y51A, H4K77A, and H4R92A were all viable. In total, 21.4% of histone H4 mutations were found to result in lethality. The essential histone H4 residues identified by Zhang et al. were all distinct from those identified in the yeast screens, and none of the lethal mutations identified in our screen in *Arabidopsis thaliana* were assayed by this group.

When comparing mutations assayed, the most similar experiment to our screen in *Arabidopsis thaliana* was that conducted by Govin et al. in *Saccharomyces cerevisiae* (Govin et al., 2010a). The percentage of lethal mutations in our screen was slightly higher than that of Govin et al., but our screen assessed more histone H4 residues and generated more amino acid substitutions for residues assayed. Similar to our screen, Govin et al. also generated an H4K5,8,12R mutant, in which lysines 5, 8, and 12 on histone H4 were simultaneously mutated to arginine (Govin et al., 2010a). While this mutant was viable in *Saccharomyces cerevisiae*, the H4K5,8,12R mutation caused lethality in *Arabidopsis thaliana*. Despite these differences, however, many of the same mutations were identified to either cause lethality or severely reduced fitness in these two screens.

In contrast, the percentage of lethal mutations in *Drosophila melanogaster* measured by Zhang et al. was more than 2-fold higher than that of our screen in *Arabidopsis thaliana* (Zhang et al., 2019). Given the high degree of conservation of

histone H4 (Figure 2.1), the percentage of mutations resulting in lethality was remarkably low in our screen compared to that observed in *Drosophila melanogaster*, although it was comparable to results found in yeast. Moreover, while only about 10% of mutations resulted in lethality in our screen, the 90% of viable mutations could still cause a minor to severe reduction in fitness and therefore be evolutionarily disadvantageous. Additionally, the dosage of the replacement histone appears to serve as an important factor in invoking the phenotypes observed. It has recently been demonstrated that the addition of a second copy of a mutant histone gene can rescue the lethality observed in yeast histone replacement systems expressing a single copy of an integrated mutant histone gene (Jiang et al., 2017). Therefore, the lethality observed in some of the earlier yeast histone replacement systems may have been influenced by the fact that only a single mutant histone copy was integrated. This result may have implications for our *Arabidopsis thaliana* histone H4 replacement system, as we analogously utilized a single endogenous histone H4 gene for the H4 replacement gene, rather than generating eight histone H4 replacement genes corresponding to each endogenous histone H4 gene, due to technical limitations.

From the three screens conducted in yeast, it is clear that the percentage of mutations that result in lethality can vary significantly depending on the residues covered and amino acid substitutions generated (Dai et al., 2008; Govin et al., 2010a; Nakanishi et al., 2008). One explanation for the substantially higher lethality in *Drosophila melanogaster* could be that Zhang et al. specifically modified residues in this experiment that are known to be targets for PTMs. This approach may bias the number of lethal mutations recovered compared to mutating all residues or mutating every occurrence of a certain amino acid due to a higher relative importance for residues that are confirmed targets for PTMs. Another explanation for the higher percentage of mutations resulting in lethality in *Drosophila melanogaster* is that multicellular eukaryotes have additional roles

for histone residues compared to yeast (Zhang et al., 2019). While this hypothesis is appealing and likely explains, for example, why the H3K27A mutation caused lethality in *Drosophila melanogaster* but not in *Saccharomyces cerevisiae*, the drastically different results obtained between *Drosophila melanogaster* and *Arabidopsis thaliana* would still be somewhat unexpected (Govin et al., 2010a; Zhang et al., 2019). It is possible that metazoans such as *Drosophila melanogaster* contain additional requirements for histone residues compared to plants such as *Arabidopsis thaliana*, or that plant systems have additional redundant pathways that can compensate for a mutation in an individual histone residue. However, it is important to note that different sets of residues were assayed by Zhang et al. in *Drosophila melanogaster* and our screen in *Arabidopsis thaliana*, and thus more data are needed to draw conclusions about these hypotheses. One final difference to call attention to between these screens is that, unlike in the system implemented in *Drosophila melanogaster*, all of the expression from the endogenous histone H4 genes had not been eliminated when H4 replacement mutants were selected in *Arabidopsis thaliana*. Therefore, some mutations may not have caused lethality in our screen because there was still wild-type histone H4 to compensate for a loss of function induced by these mutations. However, our results indicate that most of the expression of endogenous histone H4 had been eliminated in many of the *Arabidopsis thaliana* H4 replacement mutants by the T2 generation (Figure 2.10).

Second generation transgenic H4 replacement plants display a robust replacement of endogenous histone H4

When we assessed the percentage of endogenous histone H4 (*At3g53730*) alleles remaining in H4 replacement plants, half of the plants displayed a complete elimination of endogenous wild-type alleles, while the remaining plants varied from approximately 50% to 75% wild-type alleles remaining (Figure 2.10). In addition to our

result that our CRISPR/Cas9-based system is efficient at targeting the remaining endogenous histone H4 gene for mutagenesis in rH4 plants, we also found that the mRNA expression of the replacement H4 gene was often substantially higher than that of the remaining endogenous H4 gene (Figure 2.4A, Figure 2.9). Based on these data, it is likely that mutant histone H4 predominantly comprises the chromatin of rH4 mutants in the T2 generation. Moreover, as Cas9 remains active in all H4 replacement plants, it is highly likely that subsequent transgenic generations will display an even greater reduction in wild-type alleles compared to T2 plants, and thus the proportion of wild-type histone H4 comprising chromatin will continue to decline.

While genotyping must be performed to identify H4 replacement lines with a complete replacement of endogenous H4 with mutant H4, approximately half of the H4 replacement lines assessed in the T2 generation displayed a complete replacement of endogenous H4. Therefore, isolating complete H4 replacement lines from pools of T2 plants appears to be a simple and straightforward process. For our screens of phenotypes induced by histone H4 mutations described in the next chapter, we utilized plants in the T2 and T3 generations without genotyping the remaining endogenous histone H4 gene due to the high number of mutant lines we assessed. However, for more in-depth analyses of specific rH4 mutants described in later chapters, we performed genotyping to identify lines with a complete replacement of endogenous H4 with mutant H4, as this step is essential to remove variation between individual plants in the same transgenic line (i.e., by ensuring that no expression of wild-type histone H4 remains). Nonetheless, due to the consistently high expression from the H4 replacement transgene and the robust reduction/elimination of endogenous histone H4 expression, the variation in chromatin composition between individual plants in the same transgenic line due to wild-type histone H4 expression is likely minor.

Multiple independent transgenic lines mitigate off-target effects

Multiple studies have demonstrated that CRISPR/Cas9 is capable of generating DSBs at sites where the gRNA imperfectly matches the genomic target sequence, designating off-target effects as one of the major concerns of utilizing a CRISPR/Cas9-based system (Fu et al., 2013; Jinek et al., 2012; Wu et al., 2014). While the frequency of off-target mutagenesis by CRISPR/Cas9 has been shown to be extremely low or even undetectable in many experiments conducted in plants, temperature-optimized CRISPR/Cas9 has been demonstrated to increase the rate of off-target mutagenesis along with on-target mutagenesis in *Arabidopsis thaliana* (Feng et al., 2014c; Jia and Wang, 2014; LeBlanc et al., 2017; Li et al., 2013; Nekrasov et al., 2013; Peterson et al., 2016; Zhang et al., 2014). Concretely, one substitution at the genomic target sequence was found to prevent off-target mutagenesis in plants grown at 22°C, but not in plants subjected to short heat stress treatments at 37°C. However, two substitutions at the genomic target sequence were sufficient to prevent off-target mutagenesis in both growth conditions (LeBlanc et al., 2017).

With these parameters in mind, all four gRNAs that we utilized in the histone H4 replacement system were designed to minimize undesirable mutations, and thus the likelihood of off-target mutagenesis remains very low (LeBlanc et al., 2017; Tang et al., 2018). To further ensure that phenotypes observed in the rH4 mutant plants result as an effect of the H4 mutation, rather than a consequence of off-target mutagenesis by CRISPR/Cas9, we generated two independent transgenic lines corresponding to each H4 mutation, as well as multiple independent rH4 lines. By analyzing multiple independent transgenic lines, we minimized the likelihood of off-target mutagenesis affecting our system, as well as controlled for differential effects caused by random transgene integration (Gelvin, 2017).

Perspectives

Our histone H4 replacement system represents the first system to offer a complete replacement of endogenous histone H4 with histone H4 point mutants in plants. Although an additional step of genotyping the remaining endogenous histone H4 gene is required to establish complete H4 replacement lines, this collection of H4 mutants can easily be utilized for phenotypic screens of complex phenotypes. Comparing the phenotypes induced by H4 mutations in *Arabidopsis thaliana* to phenotypes in other histone replacement systems, such as *Saccharomyces cerevisiae* and *Drosophila melanogaster*, comprises an exciting implementation for this system. One benefit of our system over that implemented in *Saccharomyces cerevisiae* is that we are able to assess more complex developmental phenotypes in *Arabidopsis thaliana*. Additionally, our work comprises a deeper coverage of histone H4 than the systematic screen conducted in *Drosophila melanogaster*, which was only used to generate 14 H4 point mutants (Zhang et al., 2019). Comparing the results observed in *Drosophila melanogaster* to those in our system will allow us to examine the conservation of these mechanisms between plants and animals.

In sum, our histone H4 replacement system offers a useful tool that can be utilized for diverse studies of histone function in a multicellular eukaryote. Our histone H4 replacement system circumvents traditional challenges of potentially having to identify and simultaneously mutate many redundant histone-modifying enzymes and/or chromatin binding proteins. Thus, this platform provides a method for high-throughput reverse genetic screening of histone H4 function, and our strategy can be adapted to interrogate the function of other histones or histone H4 mutations in *Arabidopsis thaliana*.

Chapter 3: Phenotypic screens of histone H4 mutants

Author's Note: Portions of this chapter have been submitted for publication as:

Emma Tung Corcoran¹, Chantal LeBlanc¹, Mia Arias Tsang¹, Anthony Sarkiss¹, Yuzhao Hu², Ullas V. Pedmale², and Yannick Jacob¹. 2021. Systematic histone H4 replacement in *Arabidopsis thaliana* reveals a role for H4R17 in regulating flowering time.

Affiliations:

1. Yale University, Department of Molecular, Cellular and Developmental Biology, Faculty of Arts and Sciences; 260 Whitney Avenue, New Haven, Connecticut 06511, United States.
2. Cold Spring Harbor Laboratory; Cold Spring Harbor, New York 11724, United States.

Overview

We conducted systematic screens of our collection of histone H4 replacement mutants for phenotypes related to flowering time, rosette morphology, DNA content, chromatin structure, and gene silencing. We found that many histone H4 mutants displayed early flowering phenotypes in both short-day and long-day conditions. These early flowering mutants clustered into two distinct groups corresponding to moderately early flowering and severely early flowering phenotypes. Many histone H4 mutants also exhibited abnormal morphological phenotypes at the vegetative stage of growth. Additionally, several histone H4 mutations compromised normal endoreduplication, chromatin condensation, and/or transposon silencing. First, the rH4R35K, rH4K91R, and rH4Y98F mutants displayed increased proportions of higher ploidy nuclei compared to wild-type plants. Second, the rH4K16A, rH4K20A, rH4R35K, rH4R39A, rH4R77K, and rH4T80V mutants exhibited decondensation of heterochromatic chromocenters. Finally, the rH4R39K and rH4R39A mutations caused transcriptional derepression of the DNA repeat *TSI*. Taken together, we demonstrated that our histone H4 replacement system enables the assessment of expressing numerous histone mutations on the regulation of diverse biological processes, thus demonstrating the usefulness of the system for probing histone H4 function in plants.

Results

Systematic screens of histone H4 mutants

For all of the 57 viable histone H4 replacement mutants, we evaluated a range of cellular and developmental phenotypes, including flowering time, rosette morphology, DNA content, chromatin structure, and gene silencing. For these phenotypic screens, we utilized plants in the T2 generation or later without genotyping the remaining endogenous histone H4 gene. Therefore, individual plants utilized in these screens do

not necessarily exhibit a complete replacement of endogenous histone H4 with mutant histone H4, but our results indicate that the majority of chromatin in the H4 replacement plants is composed of histone H4 expressed from the H4 replacement gene (see Chapter 2). For all of these experiments, we first screened multiple transgenic lines corresponding to each of the 57 H4 point mutations along with Col, H4 septuple mutant, and rH4 plants as controls, with a few exceptions (Table 3.1). Initially, we were unable to recover any transgenic lines expressing the H4R45K, H4R55K, H4R55A, and H4R78K mutations, and thus these mutations were excluded from phenotypic screens discussed in these chapters. However, as multiple transgenic lines corresponding to each of these H4 mutations were generated after the completion of these screens, these histone H4 mutants remain for phenotypic assessment in future work. Nonetheless, after these initial screens broadly assessing phenotypes in our entire collection of histone H4 mutants, we performed further assays to characterize mutants of interest in greater depth (see below).

Flowering time

To demonstrate the utility of the H4 replacement collection in identifying pathways regulated by H4 in *A. thaliana*, we initiated a screen of the plants expressing H4 mutants for defects in flowering time. The transition between vegetative and reproductive development in *A. thaliana* has been shown to be sensitive to various chromatin disruptions, but most of the findings in this field have focused on the roles of post-translational modifications on histone H3 (Berry and Dean, 2015; He, 2009; He and Amasino, 2005; Srikanth and Schmid, 2011; Yaish et al., 2011).

We performed a preliminary assessment of flowering time for all of the H4 point mutants in both short-day (8 hours light, 16 hours dark) and long-day (16 hours light, 8 hours dark) conditions in order to evaluate the impact of H4 residues on flowering time

rH4 mutants	Viability^a	Flowering time^b	DNA content^c	Gene silencing defects^d
rH4S1A	+	wt	wt	wt
rH4R3K	+	wt	wt	wt
rH4R3A	+	wt	wt	wt
rH4K5R	+	wt	wt	wt
rH4K5A	+	wt	wt	wt
rH4K8R	+	wt	wt	wt
rH4K8A	+	wt	wt	wt
rH4K12R	+	wt	wt	wt
rH4K12A	+	wt	wt	wt
rH4K16R	+	wt	wt	wt
rH4K16A	+	wt	wt	wt
rH4R17K	+	both	wt	wt
rH4R17A	+	both	wt	derepression
rH4R19K	+	wt	wt	wt
rH4R19A	+	wt	wt	wt
rH4K20R	+	wt	wt	wt
rH4K20A	+	both	wt	wt
rH4R23K	+	wt	wt	wt
rH4R23A	+	wt	wt	wt
rH4T30V	+	wt	wt	derepression
rH4K31R	+	wt	wt	wt
rH4K31A	+	wt	wt	wt
rH4R35K	+	both	high	wt
rH4R35A	+	both	wt	wt
rH4R36K	+	wt	wt	wt
rH4R36A	+	both	wt	derepression
rH4R39K	+	both	wt	derepression
rH4R39A	+	both	wt	derepression
rH4R40K	+	long	wt	wt
rH4R40A	-	nt	nt	nt
rH4K44R	+	short	wt	wt
rH4K44A	+	both	wt	wt
rH4R45K	+	nt	nt	nt
rH4R45A	-	nt	nt	nt
rH4S47A	+	wt	wt	wt
rH4Y51F	+	wt	wt	wt
rH4T54V	+	wt	wt	wt
rH4R55K	+	nt	nt	nt
rH4R55A	+	nt	nt	nt
rH4K59A	-	nt	nt	nt
rH4K59R	+	wt	wt	wt
rH4R67K	+	wt	wt	wt
rH4R67A	+	wt	wt	wt
rH4T71V	+	wt	wt	wt
rH4Y72F	+	wt	wt	wt
rH4T73V	+	wt	wt	wt
rH4R77K	+	wt	wt	wt
rH4R77A	+	wt	wt	wt

rH4 mutants	Viability ^a	Flowering time ^b	DNA content ^c	Gene silencing defects ^d
rH4R78K	+	nt	nt	nt
rH4R78A	-	nt	nt	nt
rH4K79R	-	nt	nt	nt
rH4K79A	+	wt	wt	wt
rH4T80V	+	both	wt	wt
rH4T82V	+	wt	wt	wt
rH4Y88F	+	wt	wt	wt
rH4K91R	+	wt	high	wt
rH4K91A	+	wt	wt	wt
rH4R92K	-	nt	nt	nt
rH4R92A	+	wt	wt	wt
rH4R95K	+	wt	wt	wt
rH4R95A	+	wt	wt	wt
rH4T96V	+	wt	wt	wt
rH4Y98F	+	wt	high	wt

Table 3.1 Phenotypic analysis of plants with mutations in histone H4 residues.

nt: not tested

^aViability was represented by recovery of at least two independent transgenic lines.

^bFlowering time was represented by days to flower in long-day and short-day conditions for two independent transgenic lines. long: early flowering in long-day conditions, short: early flowering in short-day conditions, both: early flowering in both long-day and short-day conditions, wt: flowering time for both lines not significantly different from rH4 plants. Flowering was denoted early when the number of days to flower for both independent transgenic lines was significantly different compared to both rH4 lines ($p < 0.05$). The p -value was determined by Student's t -test (unpaired).

^cDNA content represented by ratio of 2N:4N:8N:16N:32N nuclei in mature leaves measured by flow cytometry. high: increased proportions of 16N/32N nuclei for multiple plants, wt: ploidy ratios not different from rH4 plants.

^dGene silencing defects represented by *TSI* expression measured by RT-qPCR. derepression: upregulation of *TSI* by >10-fold relative to Col for all transgenic lines initially assessed, wt: *TSI* expression less than 10-fold higher than Col.

regulation (Table 3.1, Figure 3.1). While individual plants expressing certain H4 mutations exhibited a late floral transition compared to the controls, we observed no rH4 mutants for which there was a consistent and significant late flowering time phenotype for both independent transgenic lines assessed. However, sixteen rH4 mutants were initially noted to display an early flowering phenotype, and were thus pursued for more in-depth flowering time analyses.

We proceeded to measure flowering time for the selected rH4 mutants and controls in both long-day and short-day conditions (Figure 3.2A,C). Additionally, we measured the rosette leaf number at flowering in both growth conditions (Figure 3.2B,D). Under long-day conditions, Col plants flowered with an average of 13 rosette leaves at 22 days. The H4 septuple mutant and both rH4 lines flowered slightly earlier, at 18 days, with an average of 10 rosette leaves. Under short-day conditions, Col plants flowered with an average of 50 leaves at 53 days. The rH4 lines flowered on average at 47 and 50 days, with 43 and 46 rosette leaves, respectively. The H4 septuple mutant flowered at 43 days with 34 rosette leaves. Many rH4 mutant lines exhibited early flowering in both long-day and short-day conditions, with the rH4R17A, rH4R36A, rH4R39K, rH4R39A, and rH4K44A mutants exhibiting the most consistent and drastic decrease in flowering time. The rH4R17A mutants flowered around 16 days in long-day conditions with 9 rosette leaves and 30 days in short-day conditions with 12 rosette leaves. The rH4R36A mutants flowered at 14 days in long-day conditions with 8 rosette leaves and on average 27 days in short-day conditions with 12 rosette leaves. The rH4R39K and rH4R39A mutants similarly flowered at 13 days in long-day conditions with 7 rosette leaves and between 28 and 35 days in short-day conditions with between 12 and 21 rosette leaves.

Many of the early flowering mutants also displayed morphological phenotypes (Figure 3.3, Figure 3.4). For example, the rH4R39K and rH4R39A mutants were very

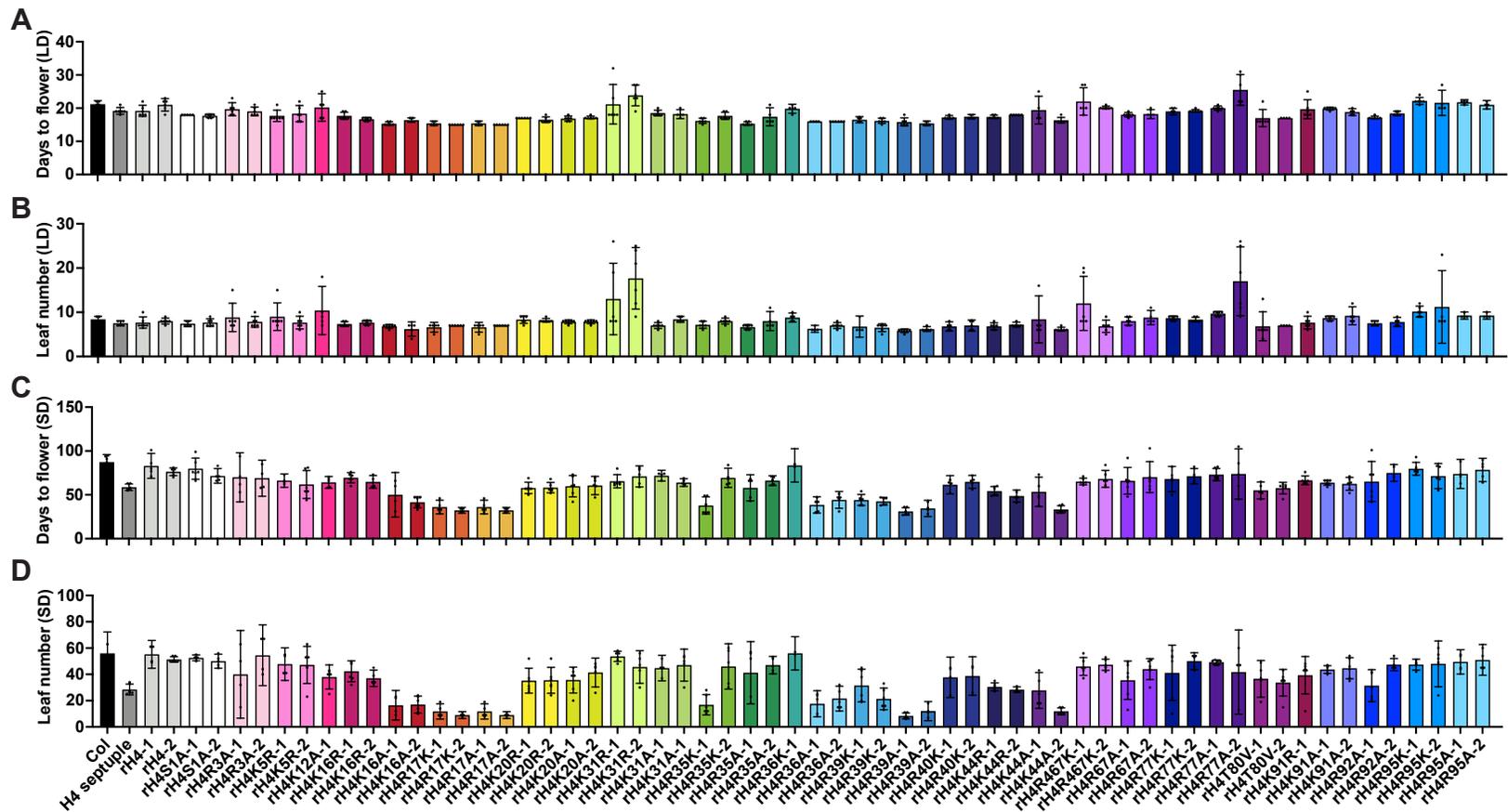


Figure 3.1 Flowering time screen of rH4 mutants. Mean (A) days to flower and (B) rosette leaf number at flowering in long-day conditions (LD) and (C) days to flower and (d) rosette leaf number at flowering in short-day conditions (SD) for Col, H4 septuple mutant, rH4, rH4S1A, rH4R3A, rH4K5R, rH4K12A (one line), rH4K16R, rH4K16A, rH4R17K, rH4R17A, rH4K20R, rH4K20A, rH4K31R, rH4K31A, rH4R35K, rH4R35A, rH4R36K, rH4R36A, rH4R39K, rH4R39A, rH4R40K, rH4K44R, rH4K44A, rH4R67K, rH4R67A, rH4R77K, rH4R77A, rH4T80V, rH4K91R (one line), rH4K91A, rH4R92A, rH4R95K, and rH4R95A plants (two independent transgenic lines each except where specified). 95% confidence intervals shown with error bars. $n \geq 4$ for long-day, $n \geq 3$ for short-day.

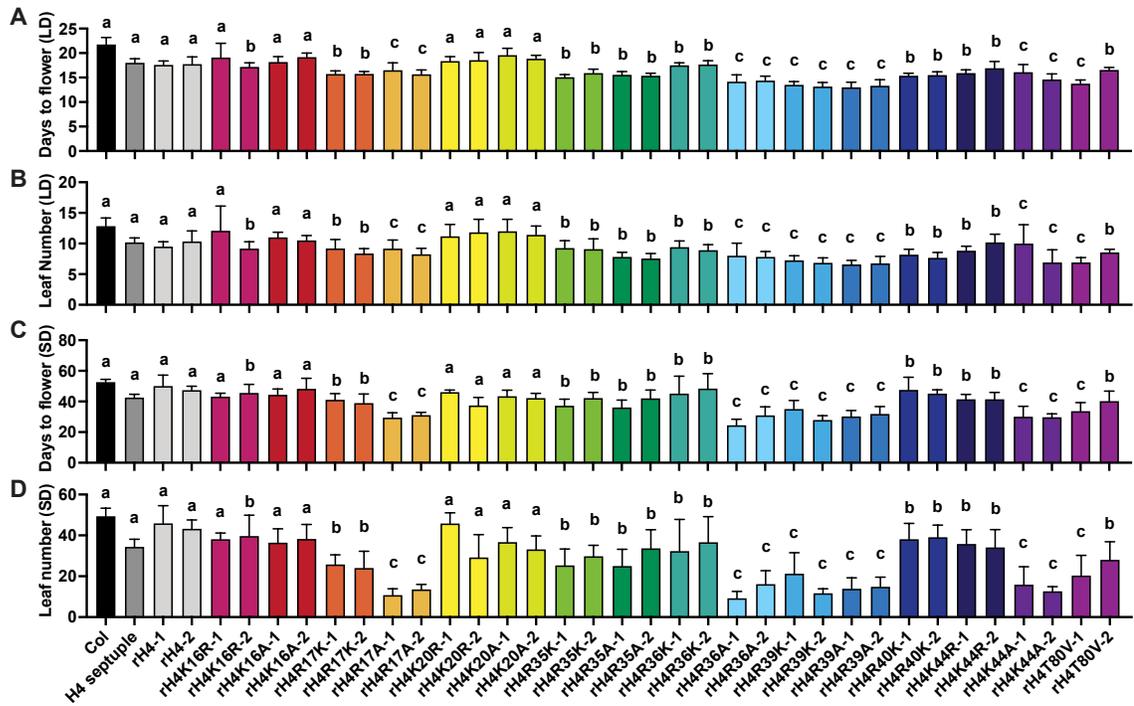


Figure 3.2 Multiple rH4 mutants display early flowering phenotypes. Mean (A) days to flower and (B) rosette leaf number at flowering in long-day conditions (LD) and (C) days to flower and (D) rosette leaf number at flowering in short-day conditions (SD) for Col, H4 septuple mutant, rH4, rH4K16R, rH4K16A, rH4R17K, rH4R17A, rH4K20R, rH4K20A, rH4R35K, rH4R35A, rH4R36K, rH4R36A, rH4R39K, rH4R39A, rH4R40K, rH4K44R, rH4K44A, and rH4T80V plants (2 independent transgenic lines each). Standard deviation shown with error bars. $n \geq 11$ for long-day, $n \geq 7$ for short-day. Letters (a,b,c) indicate cluster identified by *k*-means clustering (Figure 3.5).

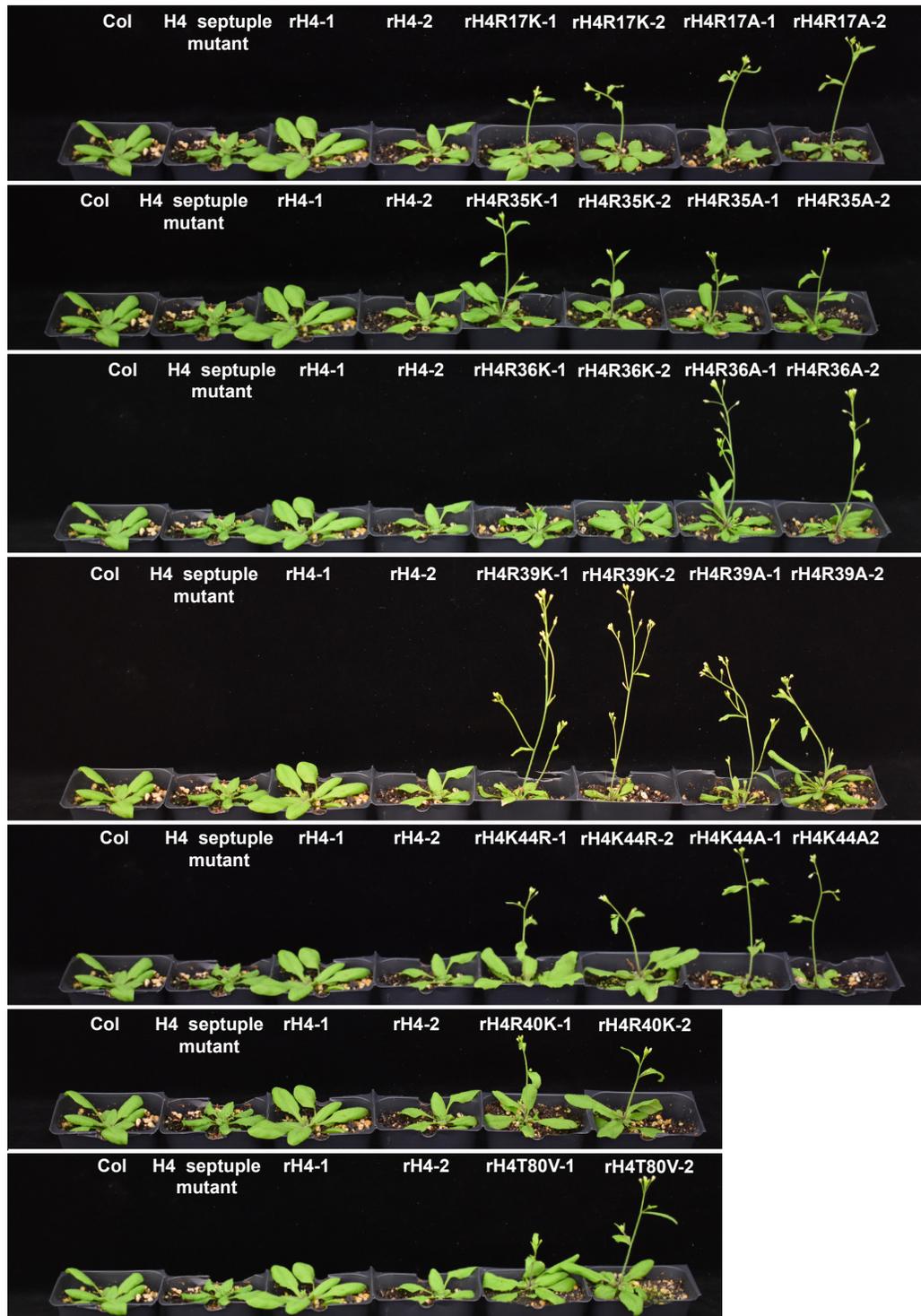


Figure 3.3 Morphological phenotypes of early flowering rH4 mutants in long-day growth conditions. Phenotypes of Col, H4 septuple mutant, and rH4 plants along with (top to bottom) rH4R17K, rH4R17A, rH4R35K, rH4R35A, rH4R36K, rH4R36A, rH4R39K, rH4R39A, rH4K44R, rH4K44A, rH4R40K, and rH4T80V mutants grown in long-day conditions at 19 days. Two independent transgenic lines were assessed per H4 replacement construct.

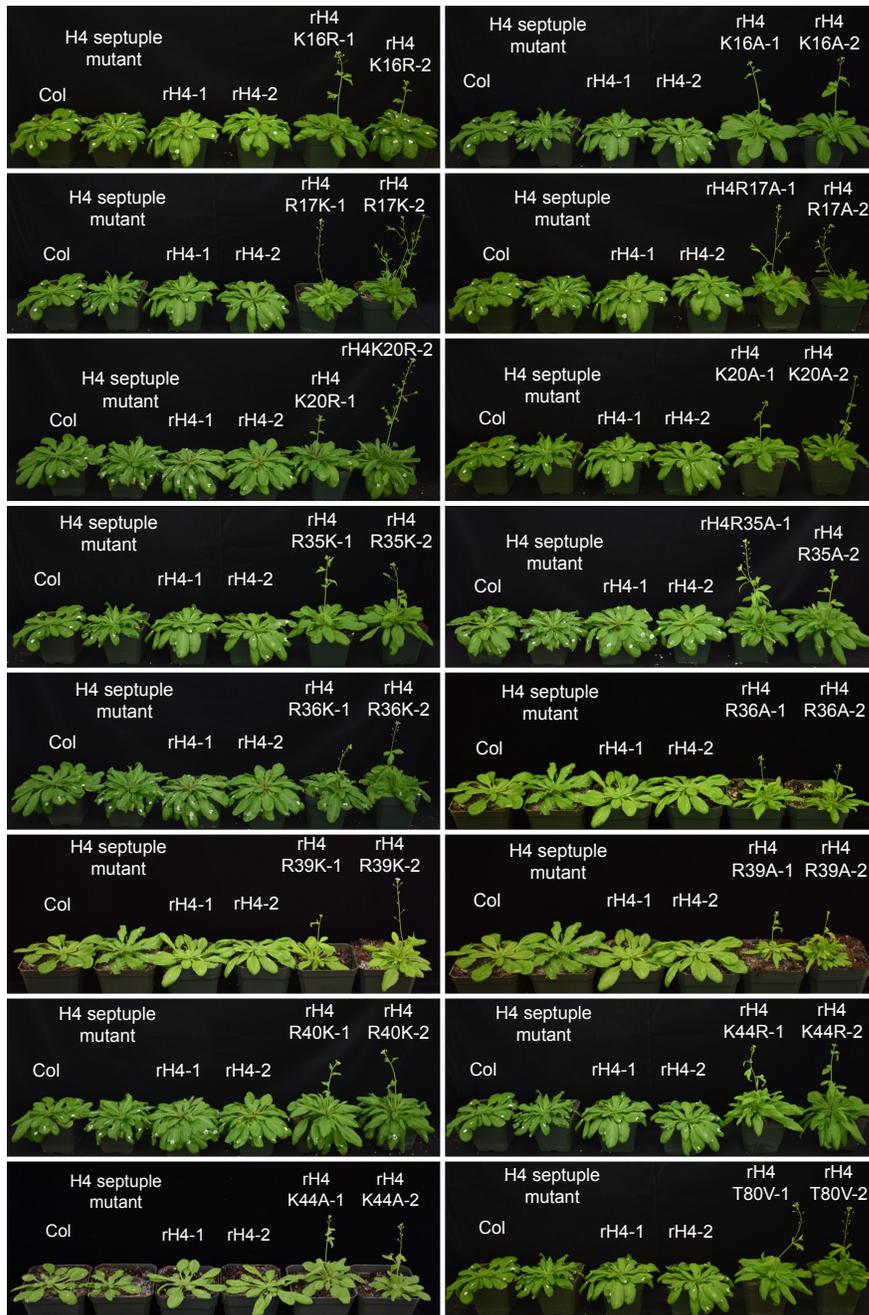


Figure 3.4 Morphological phenotypes of early flowering rH4 mutants in short-day growth conditions. Phenotype of Col, H4 septuple mutant, and rH4 plants along with (top to bottom) rH4K16R, rH4K16A, rH4R17K, rH4R17A, rH4K20R, rH4K20A, rH4R35K, rH4R35A, rH4R36K, rH4R36A, rH4R39K, rH4R39A, rH4R40K, rH4K44R, rH4K44A, and rH4T80V mutants grown in short-day conditions between 5 and 7.5 weeks. Two independent transgenic lines were assessed per H4 replacement construct. White marks present on certain rosette leaves due to leaf counting measurements.

small and displayed narrow leaves and rH4R39A mutants displayed a more severe morphological phenotype than rH4R39K mutants. The rH4R17K and rH4R17A mutants were also small and displayed an upward curling of their leaves, and the rH4R17A mutants similarly displayed a more severe morphological phenotype than the rH4R17K mutants.

In order to reduce the dimensionality of the data, we performed principal component analysis of the four variables measured corresponding to flowering time: mean day number in long-day, mean leaf number in long-day, mean day number in short-day, and mean leaf number in short-day (Figure 3.5A). We took the first two principal components (PC1 and PC2), which explained 98% of the variance (Figure 3.5B). We then performed *k*-means clustering on PC1 and PC2 to identify three clusters in the data. Cluster a, corresponding to a flowering response most similar to wild-type plants, contained Col, H4 septuple mutant, rH4, rH4K16A, rH4K20R, and rH4K20A. Cluster b, corresponding to a moderately early flowering time phenotype, contained rH4R17K, rH4R35K, rH4R35A, rH4R36K, rH4R40K, and rH4K44R. Cluster c, corresponding to a drastically early flowering time phenotype, contained rH4R17A, rH4R36A, rH4R39K, rH4R39A, and rH4K44A. The two rH4K16R lines were split between Cluster a and Cluster b, and the two rH4T80V lines were split between Cluster b and Cluster c. While the rH4K16R, rH4K16A, rH4K20R, and rH4K20A mutants appeared slightly early flowering relative to rH4 plants, all of these mutant lines except for a single rH4K16R line clustered within Cluster a through these analyses. H4R17K and H4R17A were thus the only mutations located within the H4 N-terminal tail that displayed a significantly early flowering time phenotype for both lines assessed.

We next performed RT-qPCR analyses on the key flowering time regulatory genes *FT* and *SOC1* for several rH4 mutants corresponding to each cluster and we observed upregulation of these genes consistent with the early flowering behavior of rH4

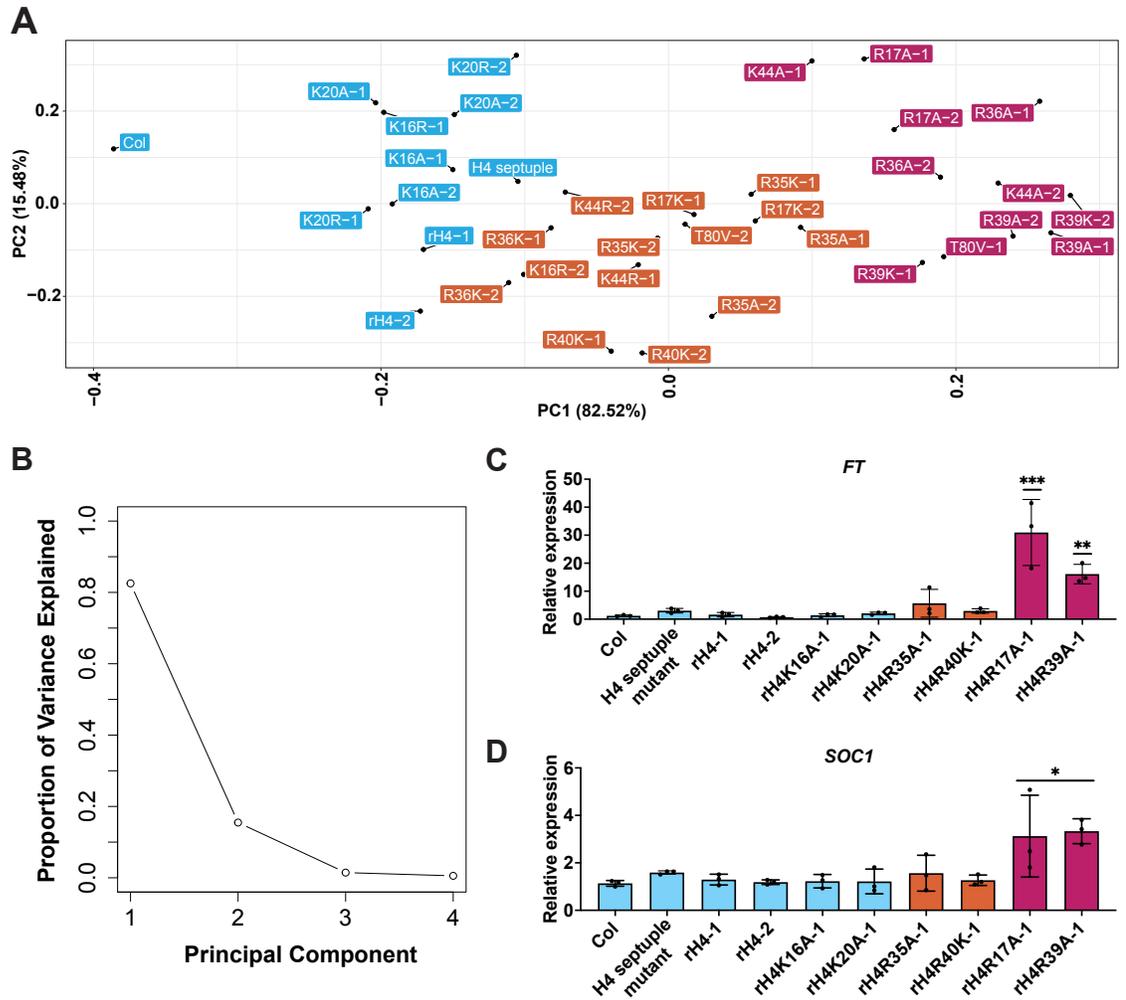


Figure 3.5 Mutations in specific residues of histone H4 generate early flowering phenotypes in *A. thaliana*. (A) Principal component plot for flowering time data along the first two principal components, PC1 and PC2. Variance explained by each principal component shown on respective axis. Three clusters produced by *k*-means clustering represented in blue (Cluster a), orange (Cluster b), and pink (Cluster c) colors. (B) Scree plot depicting the proportion of variance explained by each of the principal components. (C-D) RT-qPCR of (C) *FLC* and (D) *SOC1* in Col, H4 septuple mutant, rH4-1, rH4-2, rH4K16A-1, rH4K20A-1, rH4R35A-1, rH4R40K-1, rH4R17A-1, and rH4R39A-1 plants. Standard deviation denoted with error bars. Statistical analyses were performed using one-way ANOVA with Tukey's HSD post hoc test. *P*-value from Tukey's HSD test (genotype vs. Col) denoted with asterisks (* $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$). Bar colors represent cluster assignment from (a).

mutants from different clusters (Figure 3.5C-D). Notably, the rH4R39A mutants displayed an approximately 15-fold upregulation of *FT* and a 3-fold upregulation of *SOC1*, and the rH4R17A mutants displayed a 30-fold upregulation of *FT* and a 4-fold upregulation of *SOC1*. Thus, both the assessed mutants that clustered within the drastically early flowering cluster displayed a high increase in mRNA expression corresponding to flowering promoter genes. In contrast, rH4 mutants such as the rH4K16A and rH4K20A plants that clustered within Cluster a displayed no upregulation of these key flowering promoter genes.

While many novel functions for histone H4 residues in the regulation of flowering time were revealed in this screen, interestingly, mutations in histone H4 residues that have previously been speculated to play an important role in flowering time regulation were not observed to cause major flowering time phenotypes in these experiments. For example, point mutations in H4R3, H4K5, H4K8, H4K12, and H4K16 caused no significant change in flowering time relative to rH4 plants, apart from a minor phenotype detected for some rH4K16R mutants (Table 3.1, Figure 3.1) (Bond et al., 2009; Pei et al., 2007). However, it is important to note that some of these histone H4 residues are thought to act in combination with other H4 residues to influence flowering time regulation (e.g., tetra-acetylation of lysines 5, 8, 12, and 16 on histone H4), and thus individual mutation of these residues may be compensated by the presence of other unmutated residues on histone H4.

Rosette morphology

During our flowering time screens, we identified many morphological phenotypes at the vegetative stage of growth in T2 plants expressing the different H4 mutants (Figure 3.6), which demonstrates that our H4 replacement strategy can be used to reveal various developmental phenotypes associated with mutations on histone H4. To

further investigate the impact of H4 mutations on developmental processes, we assessed all rH4 mutants for abnormal rosette morphology phenotypes. From these investigations, we chose 15 rH4 mutants for further characterization of size, growth rate, and leaf development. The selected H4 residues were distributed over the entire histone H4 protein, covering the H4 N-terminal tail as well as the histone body. We only analyzed one mutation corresponding to each selected H4 residue, even if both substitutions at a particular residue appeared to cause morphological phenotypes in the initial screen (e.g., rH4R17K and rH4R17A). We found that many of these rH4 mutants exhibited distinct morphological phenotypes (Figure 3.7). For example, the rH4R36A and rH4R39A mutants were very small and exhibited narrow leaves compared to the controls. Additionally, rH4R17A mutants were small with upwardly curled leaves and rH4K8A mutants were small with more circular leaves than wild-type plants. Some rH4R77K plants appeared larger than the controls, while rH4K91A plants were small with minor leaf serration. Finally, rH4K44A plants displayed minor leaf serration as well as upwardly curled leaves.

We then utilized the ARADEEPOPSIS workflow, an image analysis pipeline applying a convolutional neural network, to extract 20 quantitative morphometric and color index traits from segmented rosette images corresponding to the 15 rH4 mutants described above (Huther et al., 2020) (Figure 3.8, Figure 3.9). We performed principal component analysis on these 20 features and took the first two principal components, which explained 74% of the variance in the data (Figure 3.8A-B). We then performed *k*-means clustering on PC1 and PC2 to identify three clusters in the data (Figure 3.8C). Cluster a contained Col, H4 septuple mutant, rH4, and rH4T82V plants, Cluster b contained rH4K8A, rH4K12A, rH4K16A, rH4R17A, rH4Y51F, and rH4K91A plants, and Cluster c contained rH4R3K, rH4K5R, rH4R19A, rH4R36A, and rH4R39A plants. Additionally, the two rH4K44A, rH4R77K, and rH4T80V lines were split between Cluster

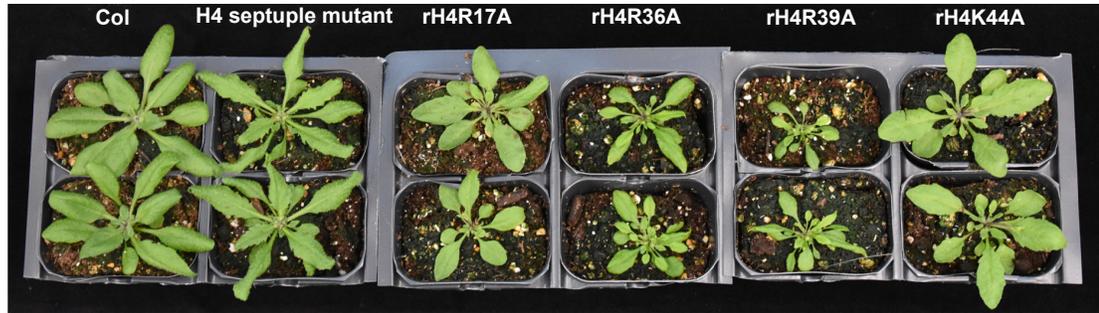


Figure 3.6 Mutations in specific residues of histone H4 generate morphological phenotypes in *A. thaliana*. Rosette phenotype of Col, H4 septuple mutant, rH4R17A, rH4R36A, rH4R39A, and rH4K44A plants grown in long-day conditions for 3 weeks. For the rH4 mutant plants, individual T2 plants (top and bottom) from independent T1 parents are shown.

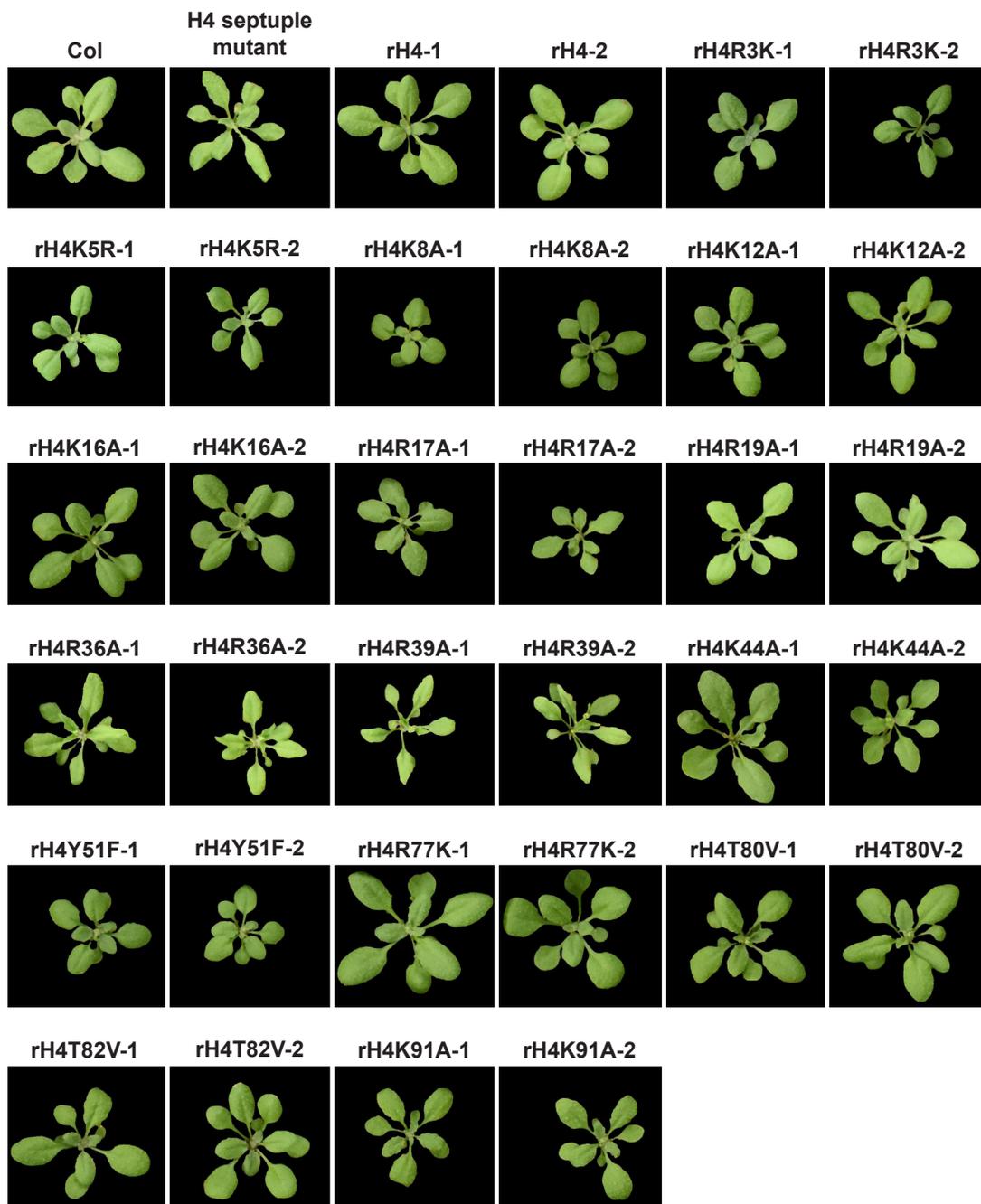


Figure 3.7 Rosette phenotypes of histone H4 mutants. Rosette phenotype of Col, H4 septuple mutant, rH4R3K, rH4K5R, rH4K8A, rH4K12A, rH4K16A, rH4R17A, rH4R19A, rH4R36A, rH4R39A, rH4K44A, rH4Y51F, rH4R77K, rH4T80V, rH4T82V, and rH4K91A plants grown in long-day conditions for 2.5 weeks. Two independent transgenic lines were assessed per H4 replacement construct.

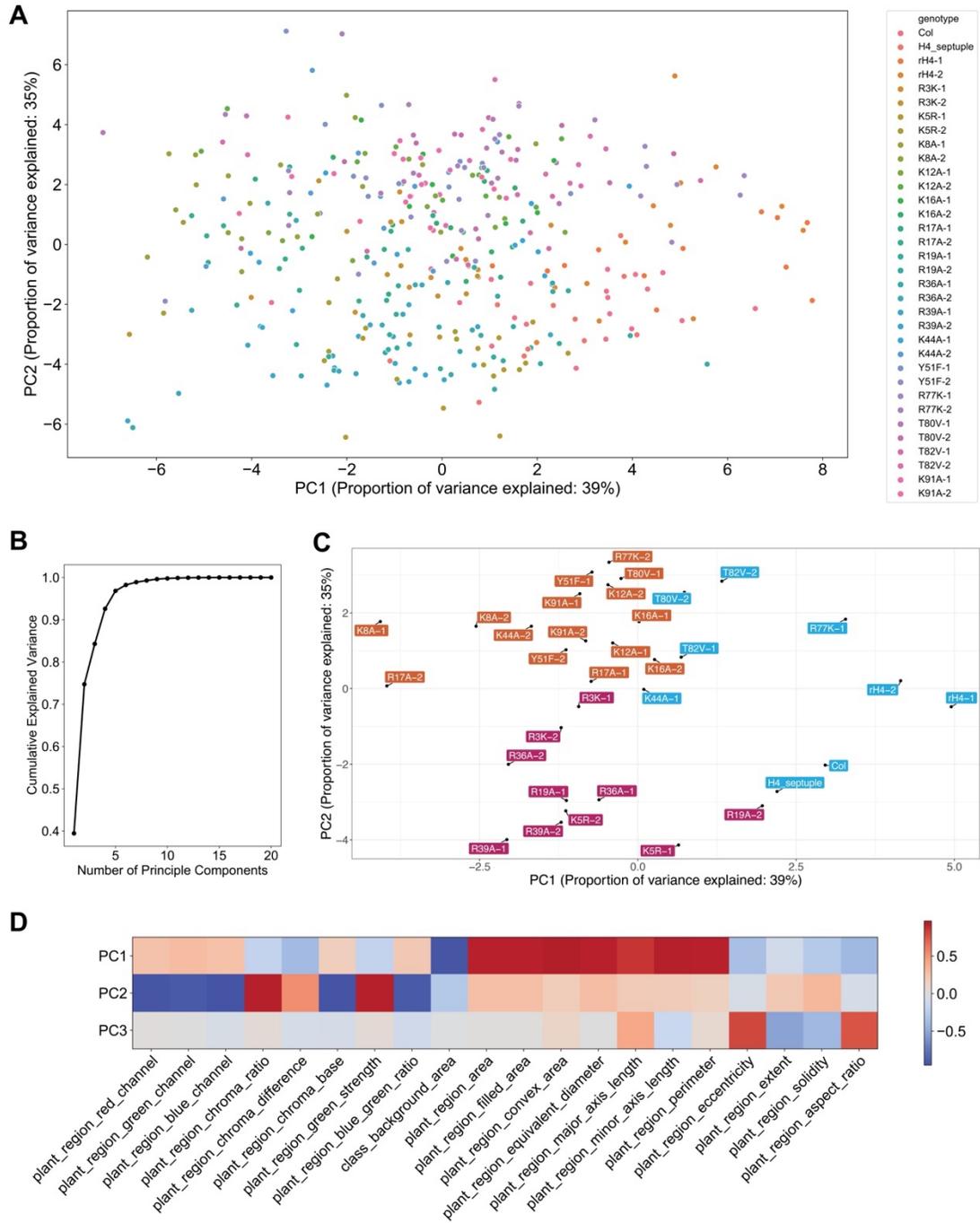


Figure 3.8 Principal component analysis of morphological features. (A) Principal component plot for rosette phenotype data along the first two principal components, PC1 and PC2. Points represent individual biological replicates. (B) Scree plot depicting the proportion of variance explained by each of the principal components. (C) *k*-means clustering of rosette phenotype data. Points represent mean of biological replicates ($n \geq 10$). Three clusters produced represented in blue (Cluster a), orange (Cluster b), and pink (Cluster c) colors. (D) Pearson's correlation matrix comparing the first three principal components (PC1 through PC3) to each rosette phenotype feature.

a and Cluster b.

We performed correlation analyses to determine which morphometric and color index traits were correlated most strongly with the first two principal components (Figure 3.8D). We observed that traits related to plant size correlated most strongly with PC1, with the convex area of the rosette showing the highest correlation (0.97, Pearson's correlation). Other morphometric traits such as the major and minor axis lengths of the rosette, the total rosette area, and the rosette perimeter displayed high Pearson's correlation coefficients between 0.94 and 0.95. In regards to PC2, we found that color index traits correlated most strongly with this principal component. For example, the green intensity and the chroma ratio of the plants both exhibited the highest Pearson's correlation coefficients of 0.95. The chroma ratio corresponds to the green intensity divided by the average of the blue and red intensities (Huther et al., 2020). A higher chroma ratio thus indicates a lower accumulation of anthocyanins—pigments that cause red, blue, and purple colors in plants (Kubo et al., 1999), while a lower chroma ratio indicates higher anthocyanin accumulation (Huther et al., 2020). From these data, we deduced that Cluster a corresponds to more phenotypically wild-type plants, Cluster b corresponds to smaller plants with a higher chroma ratio, and Cluster c corresponds to smaller plants with a lower chroma ratio.

We examined several of the traits most strongly correlated with PC1 and PC2 to compare how the total rosette area, rosette perimeter, major axis length of the rosette, and the chroma ratio of the plants compared in the different mutants assessed (Figure 3.9). Col plants showed an average total rosette area of 3.9 cm², a major axis length of 3.4 cm, a perimeter of 21 cm, and a chroma ratio of 1.56. While we confirmed our previous observation that the H4 septuple mutant plants were slightly smaller than wild-type plants, showing an average total rosette area of 3.3 cm² and a major axis length of 3.2 cm (Figure 2.3A, Figure 3.9A), the H4 septuple mutant plants clustered within

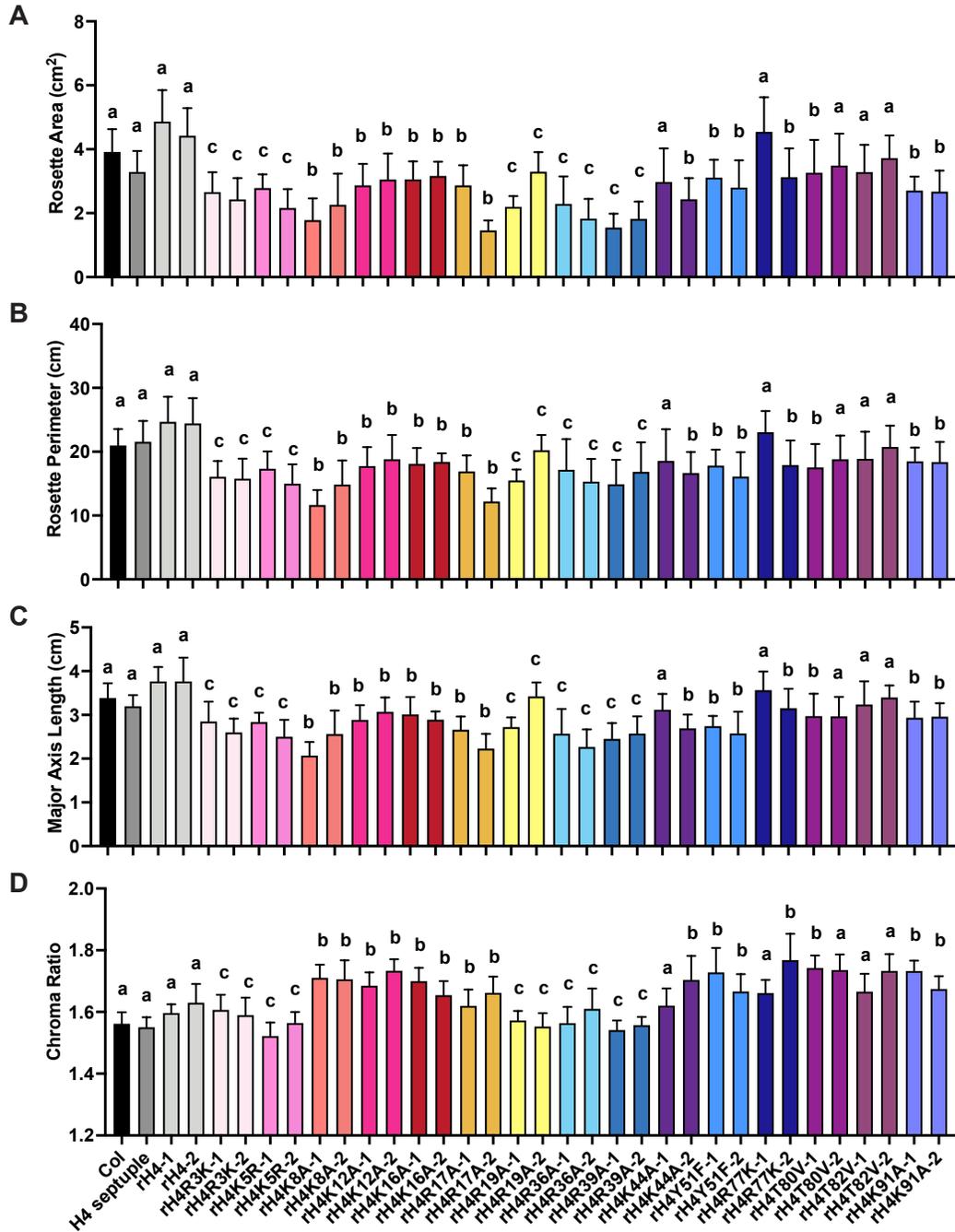


Figure 3.9 Phenotypic traits of rH4 mutant rosettes. Mean rosette (A) area, (B) perimeter, (C) major axis length, and (D) chroma ratio of Col, H4 septuple mutant (H4 septuple), rH4, rH4R3K, rH4K5R, rH4K8A, rH4K12A, rH4K16A, rH4R17A, rH4R19A, rH4R36A, rH4R39A, rH4K44A, rH4Y51F, rH4R77K, rH4T80V, rH4T82V, and rH4K91A plants (2 independent transgenic lines each). Standard deviation shown with error bars (n ≥ 10). Letters (a,b,c) indicate cluster identified by *k*-means clustering (Figure 3.8).

Cluster a due to their relatively normal chroma ratio of 1.55 and rosette perimeter of 22 cm (Figure 3.9B-D). Interestingly, the leaf serration exhibited by the H4 septuple mutant caused these mutants to display an approximately wild-type rosette perimeter despite their smaller size, and thus, the combination of this trait as well as normal colorimetric traits likely contributed to the H4 septuple mutant clustering within Cluster a with Col plants (Figure 3.8C, Figure 3.9). Moreover, rH4 plants exhibited a slightly larger average rosette size (4.9 cm² for rH4-1 plants and 4.4 cm² for rH4-2 plants) and higher chroma ratio (1.60 for rH4-1 plants and 1.63 for rH4-2 plants) than Col plants, but nonetheless clustered within Cluster a. Additionally, rH4T82V plants appeared approximately normal in size, and thus clustered within Cluster a as well.

We observed that multiple H4 mutations, including the histone N-terminal tail mutations rH4K8A, rH4K12A, rH4K16A, rH4R17A and the histone core mutations rH4Y51F and rH4K91A, caused the smaller size and higher chroma ratio (suggesting reduced anthocyanin content) observed in Cluster b. For example, rH4R17A-1 and rH4R17A-2 plants displayed an average rosette size of 2.9 cm² and 1.5 cm², and an average chroma ratio of 1.62 and 1.66, respectively. The two transgenic rH4K44A, rH4R77K, and rH4T80V lines were divided between Cluster a and Cluster b, and thus, further analyses of additional transgenic lines and endogenous H4 levels would be useful to further elucidate the phenotypes induced by these mutations. For example, we observed that rH4R77K-1 plants displayed a larger average rosette area than Col plants (4.5 cm²) (Figure 3.7, Figure 3.9). However, individual rH4R77K-1 plants displayed substantial variability, and thus, the mean rosette area of rH4R77K-1 plants did not significantly differ from Col. Additionally, rH4R77K-2 plants did not display an increase in rosette size. Finally, we found that several mutations on both the N-terminal tail and the histone core (rH4R3K, rH4K5R, rH4R19A, rH4R36A, and rH4R39A) caused the smaller rosette size and lower chroma ratio observed in Cluster c. rH4R39A mutations were

observed to produce severe morphological phenotypes, such as drastically reduced rosette area (1.6 cm² for rH4R39A-1 and 1.9 cm² for rH4R39A-2). Moreover, rH4R36A mutations also caused significantly reduced rosette area (2.3 cm² for rH4R36A-1 and 1.8 cm² for rH4R36A-2). In summary, we observed that mutations in both the H4 N-terminal tail and the globular domain caused diverse morphological phenotypes, supporting the utility of our H4 replacement system towards revealing roles for histone H4 in the regulation of plant development.

DNA content

To determine the effect of H4 residues on processes including DNA replication, endoreduplication, and cell cycle regulation, we examined the DNA content of nuclei from mature leaves from all of the rH4 mutants using flow cytometry analysis. For each H4 mutation, we assessed the phenotype of plants taken from multiple independent transgenic lines. Due to the process of endoreduplication, in which the genome is duplicated but no mitosis occurs (Galbraith et al., 1991), after staining the nuclei with propidium iodide and sorting by DNA content, the flow cytometry profile shows well-resolved populations of 2N, 4N, 8N, 16N, and 32N nuclei for wild-type plants (Figure 3.10). We assessed the width of the peaks in the flow cytometry profiles as well as the proportions of nuclei corresponding to each endoploidy level. One indication of replication/ genome stability defects in certain *Arabidopsis thaliana* mutants, such as the *atxr5/6* mutant, is that the peaks corresponding to the endoploidy levels are broader and less resolved, suggesting that the genome is undergoing unequal amplification (Jacob et al., 2010). In addition, differences in the frequencies of the different ploidies (e.g. no 16N or 32N nuclei when a specific H4 point mutant is expressed) could also indicate defects related to replication, endoreduplication, and/or the cell cycle.

We observed no rH4 mutants that showed a phenotype similar to *atxr5/6* mutants

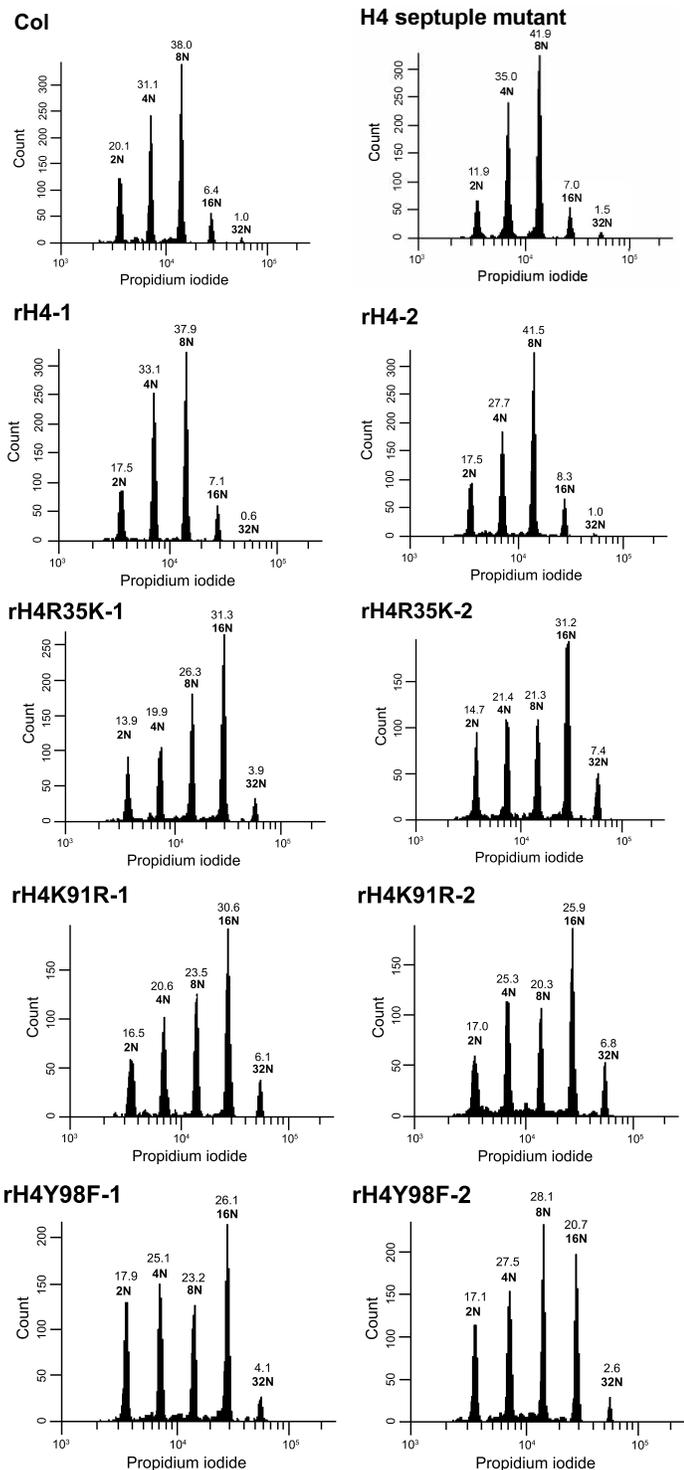


Figure 3.10 rH4R35K, rH4K91R, and rH4Y98F mutants exhibit altered DNA content. Representative flow cytometry profiles of Col, H4 septuple mutant, rH4-1, rH4-2, rH4R35K-1, rH4R35K-2, rH4K91R-1, rH4K91R-2, rH4Y98F-1 and rH4Y98F-2 plants. 1500 gated events are plotted. Labels above peaks indicate ploidy levels and percentages corresponding to outlined propidium iodide intensity.

(i.e., a widening of peaks relative to wild-type plants). However, we did observe three H4 mutations that induced altered proportions of higher ploidy nuclei relative to the controls when expressed. For these rH4 mutants, we examined the endoreduplication phenotype of additional plants to assess the proportion of plants that exhibited the increased endoploidy proportions. For Col plants, 16N nuclei typically represented between 4 and 11% of total nuclei, while 32N nuclei comprised between 0.3 and 1.6% of total nuclei (Figure 3.11). H4 septuple mutant and rH4 plants exhibited similar proportions of 16N and 32N nuclei, although some H4 septuple mutant plants exhibited up to 2.5% of 32N nuclei. In contrast, we observed increased proportions of 16N and 32N nuclei for individual plants expressing H4R35K, H4K91R, and H4Y98F mutant histones.

In our first assay with three rH4R35K plants, one rH4R35K mutant exhibited increased proportions of 16N and 32N nuclei (19.6% and 8.7%, respectively). Interestingly, the other two rH4R35K plants assessed displayed decreased proportions of the higher ploidy nuclei, presenting between 1.2 and 2.3% of 16N nuclei and 0.1% of 32N nuclei (Figure 3.11A). We proceeded to assess six additional rH4R35K plants corresponding to four independent transgenic lines. Of these plants, three out of six displayed increased proportions of both 16N and 32N nuclei (ranging between 26 and 31% of 16N nuclei and between 3.9 and 7.4% of 32N nuclei), and two out of six displayed increased proportions of 16N nuclei (ranging between 19 and 28%) (Figure 3.11B). The plants that displayed altered proportions of endoploidy levels did not correspond to specific transgenic rH4R35K lines in these experiments, but rather, they were distributed among all four transgenic lines assessed. In contrast, rH4R35A plants assessed from three independent transgenic lines exhibited approximately wild-type endoploidy levels (Table 3.1). To estimate the frequency of mutation of the remaining endogenous H4 gene in these lines, we genotyped four plants each from four independent rH4R35K lines and four independent rH4R35A lines at the T3 generation

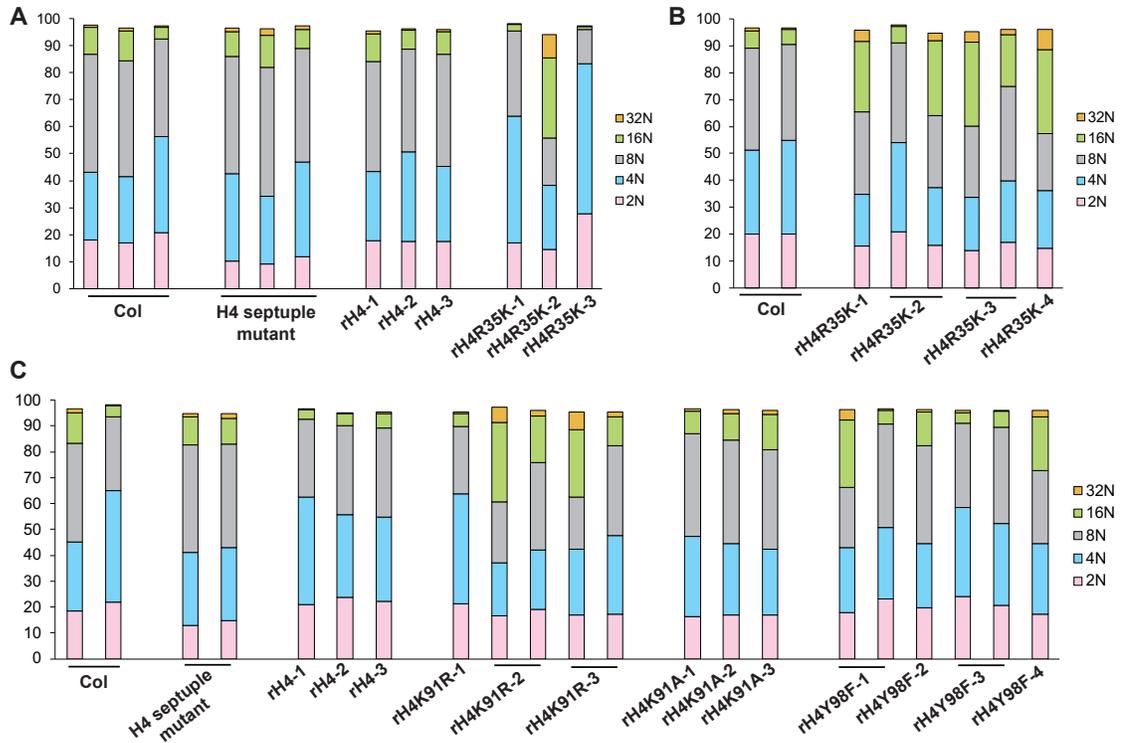


Figure 3.11 Ploidy levels of rH4R35K, rH4K91R, rH4K91A, and rH4Y98F mutants. Percentages of 2N, 4N, 8N, 16N and 32N nuclei out of 1500 gated events measured by flow cytometry for (A) Col, H4 septuple mutant, rH4-1, rH4-2, rH4-3, rH4R35K-1, rH4R35K-2, and rH4R35K-3 plants (B) Col, rH4R35K-1, rH4R35K-2, rH4R35K-3, rH4R35K-4 plants and (C) Col, H4 septuple mutant, rH4-1, rH4-2, rH4-3, rH4K91R-1, rH4K91R-2, rH4K91R-3, rH4K91A-1, rH4K91A-2, rH4K91A-3, rH4Y98F-1, rH4Y98F-2, rH4Y98F-3, and rH4Y98F-4 plants.

stage (Figure 3.12). We found that 81% of the rH4R35K plants exhibited a complete elimination of wild-type histone H4, while none of the rH4R35A plants exhibited any mutation of the remaining endogenous H4 gene. Therefore, the presence of wild-type histone H4 may restore a wild-type endoploidy phenotype in rH4R35A plants compared to rH4R35K plants.

Two out of five rH4K91R plants assessed displayed an increased proportion of 16N (between 26 and 31%) and 32N nuclei (between 6.1 and 6.8%), and one rH4K91R mutant displayed an increased proportion of 16N nuclei (18%) (Figure 3.11C). Unlike the rH4K91R mutants, all three rH4K91A mutants assessed displayed wild-type endoploidy levels. Finally, two out of six rH4Y98F mutants assessed displayed an increased proportion of 16N nuclei (between 21 and 26%) and one of these plants also exhibited an increased proportion of 32N nuclei (4.1%). Similar to the rH4R35K mutants, the individual rH4K91R and rH4Y98F plants that exhibited increased proportions of 16N and 32N nuclei were spread over multiple independent transgenic lines, rather than both corresponding to a specific transgenic line.

To further quantify variation in endoreduplication, we calculated the endoreduplication index as described previously (Sterken et al., 2012) (Figure 3.13). A higher endoreduplication index indicates increased proportions of higher ploidy nuclei within the population. Col plants exhibited endoreduplication indices between 114 and 148, and rH4 plants remained within that range. H4 septuple mutant plants displayed a significantly higher average endoreduplication index, with individual values ranging between 146 and 165. As expected from the variation in endoreduplication phenotypes between individual rH4 mutants, there was substantial variation in the endoreduplication indices calculated among plants expressing a certain H4 mutant construct. Nonetheless, rH4R35K plants in general exhibited the largest difference in the endoreduplication index compared to the controls, displaying an endoreduplication index as high as 187 and as

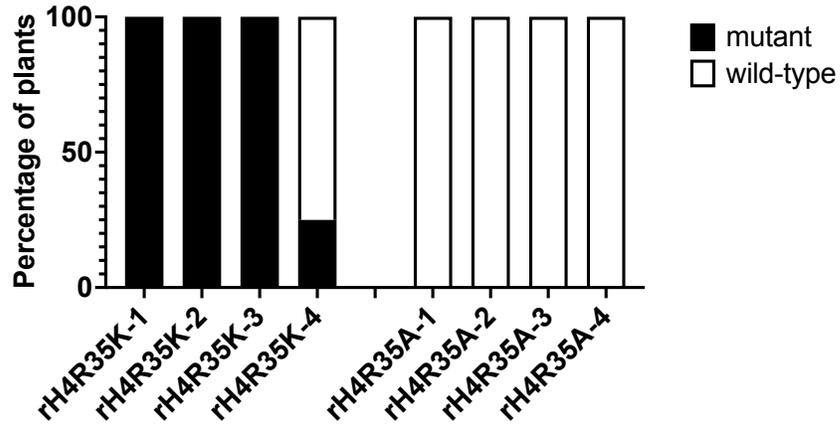


Figure 3.12 Endogenous H4 mutations in rH4R35K and rH4R35A mutant plants. Percentage of mutated endogenous H4 (*At3g53730*) alleles of sixteen rH4R35K plants and sixteen rH4R35A plants. Each plant assessed was from the T3 generation; four plants from the same initial T1 parent were used in this experiment (i.e., four independent transgenic lines per genotype).

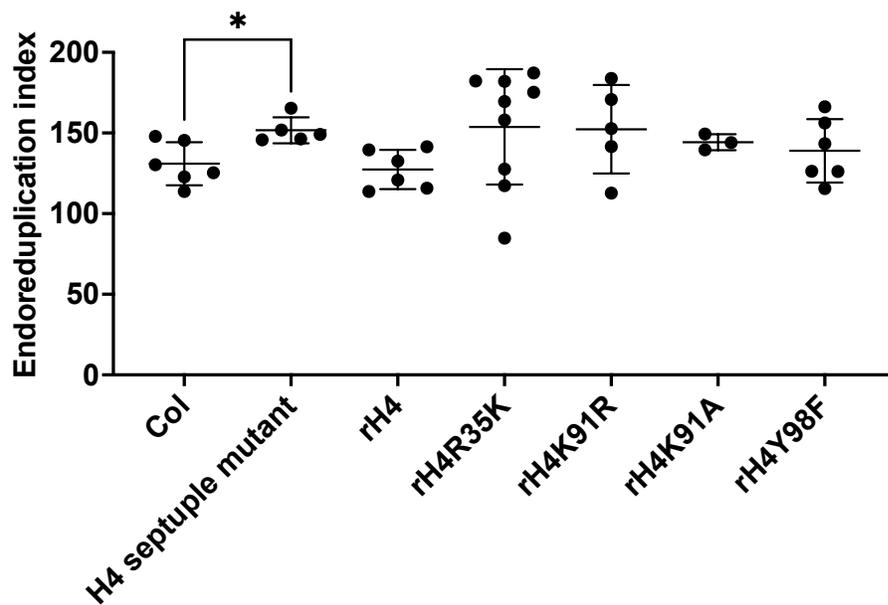


Figure 3.13 Endoreduplication indices for rH4R35K, rH4K91R, rH4K91A, and rH4Y98F mutants. Endoreduplication index calculated for Col, H4 septuple mutant, rH4, rH4R35K, rH4K91R, rH4K91A, and rH4Y98F plants. Horizontal bars indicate the mean. Standard deviation denoted with error bars. *P*-value from unpaired Student's *t*-test denoted with asterisks (* $p < 0.05$).

low as 85 (Figure 3.13). While some individual rH4R35K, rH4K91R, and rH4Y98F plants displayed substantially increased endoreduplication indices compared to wild-type plants and one individual rH4R35K plant also displayed a substantially decreased endoreduplication index, due to the high variation observed within these populations of plants, the average differences were not statistically significant compared to the controls.

Chromatin structure

To examine chromatin structure in the rH4 mutants, we first stained fixed nuclei isolated from mature leaves from one individual plant corresponding to each H4 mutation with DAPI (4',6-diamidino-2-phenylindole) and scored at least 10 nuclei images obtained from each sample (Table 3.1). Our goal was to identify roles for H4 residues in the maintenance of proper chromatin structure. Wild-type *Arabidopsis thaliana* DAPI-stained nuclei display highly condensed regions called chromocenters, corresponding to genomic regions of heterochromatin, predominantly composed of pericentromeric repeats, transposons, and ribosomal DNA (rDNA) genes (Heslop-Harrison et al., 2003; Jacob et al., 2009). Mutations in enzymes that modify or bind chromatin can cause decondensation of these heterochromatic chromocenters. For example, *atxr5/6* mutant plants (which exhibit a loss of H3.1K27me1), as well the DNA methylation mutants *met1* and *decrease in DNA methylation 1 (ddm1)*, display partial chromocenter decondensation of pericentromeric sequences, while a different DNA methylation mutant *variant in methylation 1 (vim1)* displays chromocenter decondensation corresponding to centromeric regions (Jacob et al., 2009; Soppe et al., 2002; Woo et al., 2007). Therefore, histone PTMs such as H3.1K27me1 and DNA methylation are implicated in the regulation of chromatin condensation. A loss of repressive chromatin structure around heterochromatic elements can lead to deleterious genomic outcomes, such as transposon derepression, and thus the regulation of chromatin structure is important for

the maintenance of genome stability (Henderson and Jacobsen, 2007; Hirochika et al., 2000; Jacob et al., 2009; Singer et al., 2001).

After the initial screen of all of the rH4 mutants, we then chose the rH4 mutant lines that exhibited the most drastic alterations in chromatin condensation for further work. Only one mutation corresponding to each selected H4 residue was assessed in these in-depth microscopy experiments, even if both substitutions at a particular residue were observed to cause chromatin condensation defects in the initial screen (e.g., rH4R35K and rH4R35A). We assessed nuclei from mature leaves corresponding to two independent transgenic lines for each H4 mutation selected, along with Col, H4 septuple, and rH4 plants, and obtained at least 30 nuclei images for each sample. Our results indicate that rH4 nuclei exhibited normal chromatin structure, while some H4 septuple mutant nuclei exhibited slight decondensation and loss of chromocenters (Figure 3.14). Notably, we found that rH4K16A, rH4K20A, rH4R35K, rH4R39A, rH4R77K, and rH4T80V mutant nuclei exhibited chromocenter decondensation for both transgenic lines assessed (Figure 3.14).

Interestingly, some mutations on residues that have been implicated in the regulation of chromatin condensation in plants or other systems were not identified to cause severe chromatin condensation phenotypes in our screen. For example, H4S1 phosphorylation has been demonstrated to promote nuclear compaction in *Saccharomyces cerevisiae*, but *Arabidopsis thaliana* rH4S1A mutant nuclei were not observed to display noticeable chromatin compaction defects (Table 3.1) (Govin et al., 2010b; Krishnamoorthy et al., 2006). While these results could potentially indicate a functional difference of H4S1 in plant and yeast systems, it is important to note that our screen was not exhaustive in its categorization of chromatin structural phenotypes, and thus additional experiments are needed to further validate negative results that we observed (see discussion).

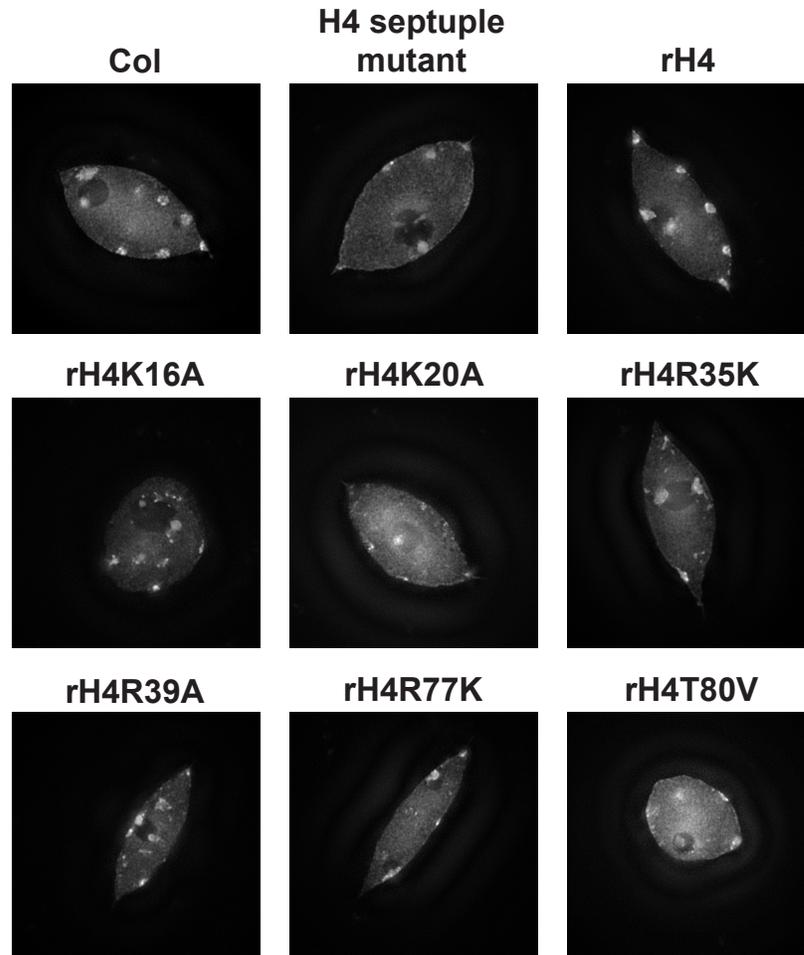


Figure 3.14 rH4 mutants exhibit altered chromatin structure. DAPI staining of Col, H4 septuple mutant, rH4, rH4K16A, rH4K20A, rH4R35K, rH4R39A, rH4R77K, and rH4T80V mutant nuclei (one representative nucleus displayed).

Transcriptional silencing

In order to assess transcription-related defects in the rH4 mutants, we used RT-qPCR to measure the expression of *TSI*, a normally silenced genomic element comprising open reading frame 1 of the *Athila* retrotransposon (Steimer et al., 2000). In Col and rH4 plants, the RNA expression of *TSI* is very low (Figure 2.4). Mutants displaying genome instability and a loss of repressive heterochromatin structure, such as the H4 septuple mutant and the *atxr5/6* mutant, exhibit a significant upregulation of *TSI* expression, often two orders of magnitude higher than wild-type plants (Figure 2.4) (Jacob et al., 2009). In addition to *TSI*, we also measured RNA expression of the DNA damage response gene *BRCA1*. Col and rH4 plants display very low RNA expression of *BRCA1*, while H4 septuple mutant and *atxr5/6* mutant plants display an upregulation of *BRCA1* approximately one order of magnitude higher than Col (Figure 2.4) (Feng et al., 2017). We performed RT-qPCR experiments to measure *TSI* and *BRCA1* RNA expression relative to Col for two independent transgenic lines corresponding to all rH4 mutants, as well as the H4 septuple mutant and two independent transgenic rH4 lines as controls (Figure 3.15, Figure 3.16).

The rH4R17A, rH4T30V, rH4R36A, rH4R39K, and rH4R39A mutants exhibited *TSI* upregulation of 10-fold or more for both transgenic lines assessed (Figure 3.15, Table 3.1). *TSI* upregulation in these rH4 mutants varied between a 21- to 78-fold increase relative to Col, and the H4 septuple mutant displayed a 56-fold increase relative to Col. We set the threshold for a 10-fold increase in *TSI* RNA expression as a stringent cutoff due to our previous observation that rH4 plants exhibited up to a 5-fold increase in *TSI* RNA expression relative to Col (Figure 2.4). Similarly, we set a threshold of a 3.5-fold increase in RNA expression relative to Col for *BRCA1* upregulation due to our previous observations of variation in the *BRCA1* expression of rH4 plants. In contrast to the multiple rH4 mutants which displayed *TSI* upregulation, no rH4 mutants exhibited

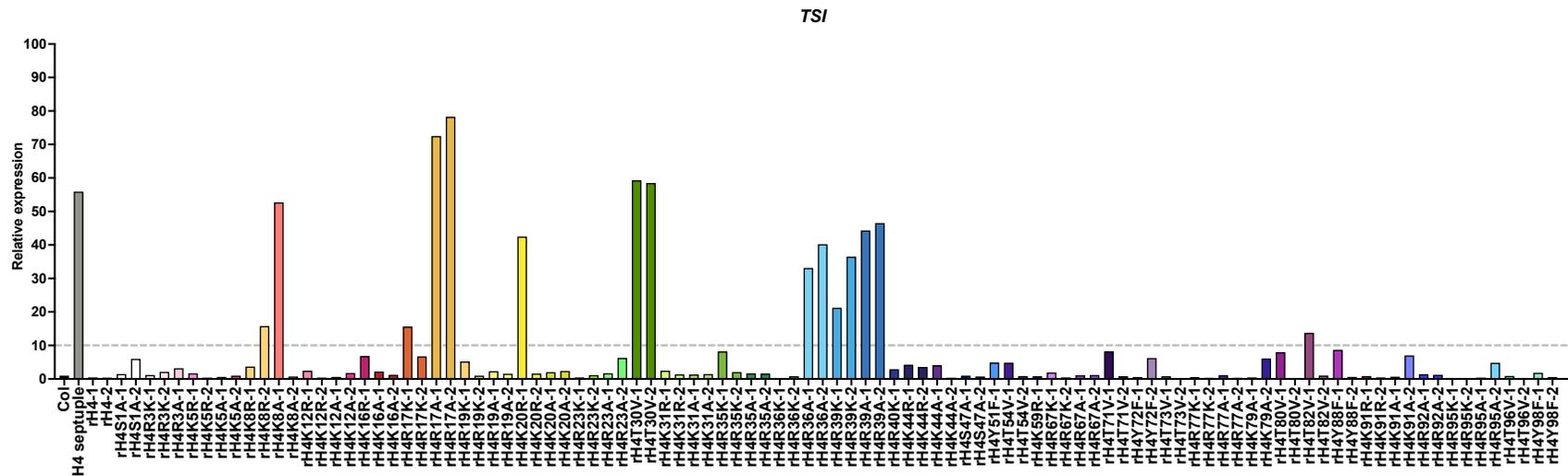


Figure 3.15 Expression levels of *TSI* in rH4 mutant lines. RNA expression of *TSI* measured by RT-qPCR in Col, H4 septuple mutant (H4 septuple), rH4, rH4S1A, rH4R3K, rH4R3A (one line), rH4K5R, rH4K5A, rH4K8R, rH4K8A, rH4K12R, rH4K12A, rH4K16R (one line), rH4K16A, rH4R17K, rH4R17A, rH4R19K, rH4R19A, rH4K20R, rH4K20A, rH4R23K, rH4R23A, rH4T30V, rH4K31R, rH4K31A, rH4R35K, rH4R35A, rH4R36K, rH4R36A, rH4R39K, rH4R39A, rH4R40K (one line), rH4K44R, rH4K44A, rH4S47A, rH4Y51F (one line), rH4T54V, rH4K59R (one line), rH4R67K, rH4R67A, rH4T71V, rH4Y72F, rH4T73V, rH4R77K, rH4R77A, rH4K79A, rH4T80V, rH4T82V, rH4Y88F, rH4K91R, rH4K91A, rH4R92A, rH4R95K, rH4R95A, rH4T96V, and rH4Y98F plants (two independent transgenic lines each except where specified). Dashed horizontal line represents threshold set for upregulation.

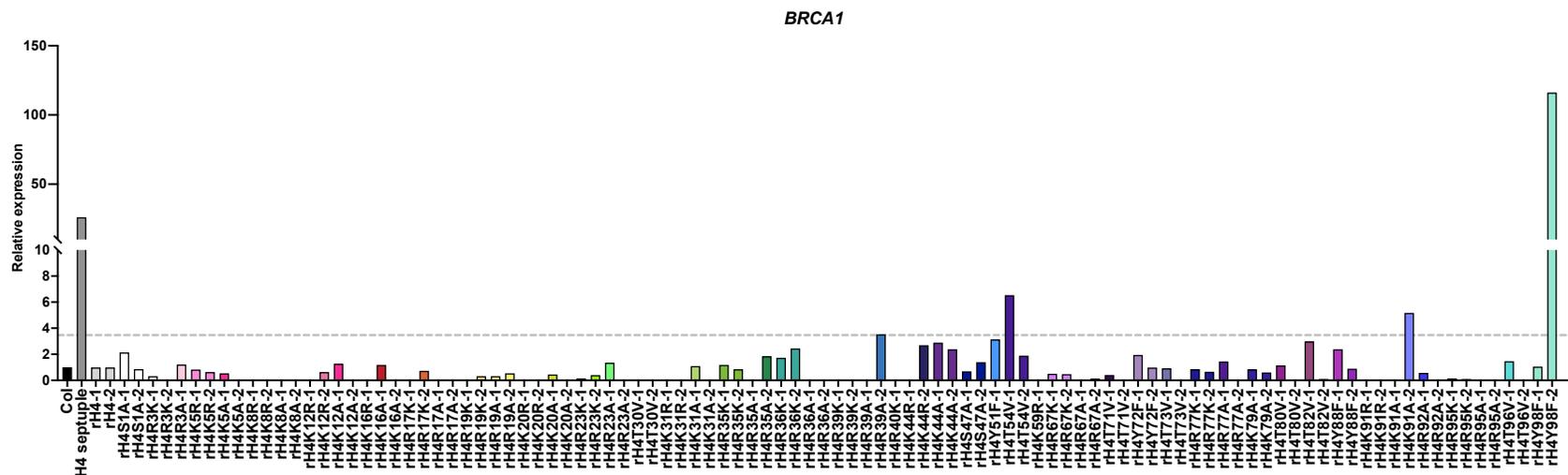


Figure 3.16 Expression levels of *BRCA1* in rH4 mutant lines. RNA expression of *BRCA1* measured by RT-qPCR in Col, H4 septuple mutant (H4 septuple), rH4, rH4S1A, rH4R3K, rH4R3A (one line), rH4K5R, rH4K5A, rH4K8R, rH4K8A, rH4K12R, rH4K12A, rH4K16R (one line), rH4K16A, rH4R17K, rH4R17A, rH4R19K, rH4R19A, rH4K20R, rH4K20A, rH4R23K, rH4R23A, rH4T30V, rH4K31R, rH4K31A, rH4R35K, rH4R35A, rH4R36K, rH4R36A, rH4R39K, rH4R39A, rH4R40K (one line), rH4K44R, rH4K44A, rH4S47A, rH4S47A (one line), rH4T54V, rH4K59R (one line), rH4R67K, rH4R67A, rH4T71V, rH4Y72F, rH4T73V, rH4R77K, rH4R77A, rH4K79A, rH4T80V, rH4T82V, rH4Y88F, rH4K91R, rH4K91A, rH4R92A, rH4R95K, rH4R95A, rH4T96V, and rH4Y98F plants (two independent transgenic lines each except where specified). Dashed horizontal line represents threshold set for upregulation.

BRCA1 upregulation of at least 3.5-fold for all transgenic lines assessed (Figure 3.16). However, one rH4Y98F mutant displayed a 116-fold upregulation in *BRCA1* expression relative to Col, compared to the 26-fold upregulation in *BRCA1* displayed by the H4 septuple mutant. As the other rH4Y98F mutant assessed displayed no upregulation in *BRCA1* expression relative to Col, we chose not to characterize the DNA damage response phenotype of this mutant in further detail. Similarly, one rH4R39A mutant, one rH4T54V mutant, and one rH4K91A mutant exhibited between a 3.5- and 6.5-fold upregulation of *BRCA1* expression compared to Col. However, the other lines assessed for each of these mutations exhibited no *BRCA1* upregulation, and thus, we did not characterize the DNA damage response phenotypes of these mutants further.

After these initial RT-qPCR screens of all of the rH4 mutants, we aimed to further investigate the *TSI* upregulation observed in the rH4R39K and rH4R39A mutants through additional RT-qPCR experiments. Out of all of the rH4 mutants that exhibited *TSI* upregulation, we chose the rH4R39K and rH4R39A mutants for further characterization due to the previously described roles for H4R39 in regulating gene silencing described in yeast (see discussion). We measured *TSI* expression in multiple plants from two independent transgenic lines expressing the rH4R39K and rH4R39A mutant constructs, along with Col, H4 septuple mutant, and rH4 plants (Figure 3.17). The rH4R39K mutants exhibited a slight upregulation in *TSI* expression (between a 9.4- and 21-fold increase relative to Col) and the rH4R39A mutants exhibited a drastic increase in *TSI* expression, displaying a 25- to 82-fold increase relative to Col. While the phenotype of the rH4R39A mutant was more severe than that of the rH4R39K mutant, the *TSI* upregulation observed for both of these mutants was much lower than that of the H4 septuple mutant, which showed an approximately 360-fold upregulation of *TSI* relative to Col. Nonetheless, the rH4R39K, rH4R39A, and H4 septuple mutant plants all showed a significant upregulation of *TSI* RNA expression compared to Col to differing

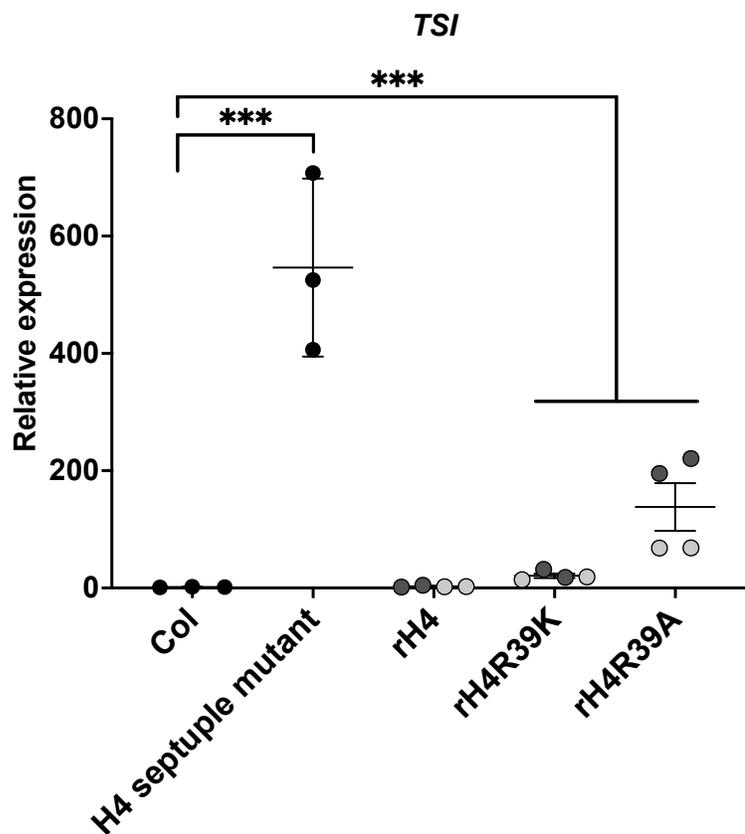


Figure 3.17 rH4R39K and rH4R39A mutants exhibit transcriptional derepression. RNA expression of *TSI* measured by RT-qPCR in Col, H4 septuple mutant, and two independent rH4, rH4R39K, and rH4R39A transgenic lines (2 plants from each line). Different shades of gray indicate independent transgenic lines. Three biological replicates were performed for Col and H4 septuple mutant plants. Horizontal bars indicate the mean. Standard deviation denoted with error bars. *P*-value from unpaired Student's *t*-test denoted with asterisks (* $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$).

extents. In sum, our screens of different molecular and cellular phenotypes have identified several H4 residues as promising candidates for future work to elucidate the ways in which histone H4 maintains genome stability in the nucleus.

Discussion

Many novel functions uncovered for histone H4 residues

With the systematic screens of histone H4 utilizing our histone H4 replacement system, we have described many novel roles for residues on histone H4. Previously, functions of histone H4 in *Arabidopsis thaliana* have mainly been studied related to known PTMs on the H4 N-terminal tail. As our histone H4 screen covered all lysine, arginine, threonine, serine and tyrosine residues on the histone H4 protein, we assessed the function of many residues on histone H4 that have previously been uncharacterized in plants. We identified novel roles for H4R35, H4R77, and H4T80 in regulating chromatin compaction, H4R35, H4K91, and H4Y98 in regulating the endocycle, and H4R39 in regulating transposon silencing in *A. thaliana*. Additionally, we identified many novel roles for histone H4 residues in the regulation of flowering time, including H4R17, H4R35, H4R36, H4R39, H4R40, H4K44, and H4T80. Finally, we determined that multiple H4 residues, including H4R3, H4K5, H4K8, H4R17, H4R19, H4R36, H4R39, H4K44, and H4K91, play a role in the regulation of rosette development. Notably, many of the above H4 residues were not only relatively uncharacterized in plants, but in all eukaryotes, demonstrating the potential of our histone H4 system to allow us to assess the impact of numerous histone H4 mutations on the regulation of diverse biological processes and reveal novel functions for residues on histone H4. While we conducted experiments to assess five major phenotypes (flowering time, rosette morphology, DNA content, chromatin structure, and transcriptional silencing) in our collection of rH4 mutants, the potential remains to examine these rH4 mutants for many additional phenotypes, such as leaf epidermal cell size and shape, root length, pollen grain viability, DNA damage sensitivity, and tolerance to various biotic and abiotic stresses.

Phenotype variability observed between transgenic lines

In our experiments assessing the phenotypes of the rH4 mutants, we occasionally observed that different transgenic lines expressing the same mutant H4 protein displayed differing phenotypes. For example, only one out of two rH4Y98F mutants assessed displayed upregulation of *BRCA1* (Figure 3.16) and only two out of three rH4K91R mutant lines displayed increased proportions of higher ploidy nuclei (Figure 3.11C). Additionally, only one rH4R77K mutant line exhibited increased rosette area (Figure 3.9). Moreover, variability within the same transgenic line was also observed. For example, individual rH4R35K mutant plants that belonged to the same transgenic line displayed three different endoploidy phenotypes: decreased proportions of higher ploidy nuclei, increased proportions of higher ploidy nuclei, and wild-type endoploidy levels, although the majority of rH4R35K mutants displayed increased levels of higher ploidy nuclei (Figure 3.11A-B).

Individual plants expressing the same mutant H4 protein may show different phenotypes due to uneven levels of expression of the H4 replacement gene, random T-DNA integration, and differing mutagenesis of the remaining endogenous histone H4 gene. To circumvent the complications that the genomic locations of the T-DNA insertions or off-target effects of CRISPR/Cas9 may cause any observed phenotypes, we assessed the phenotypes of at least two independent transgenic lines corresponding to each H4 mutation. Varying levels of expression from the remaining endogenous histone H4 gene in the rH4 mutants may also explain some of the phenotypic variability observed in the rH4 mutants. For future work on rH4 mutants of interest, identifying rH4 mutants with a complete replacement of endogenous histone H4 with mutant histone H4 would remove some of the variation between individual rH4 mutant plants. The total abundance of histone H4 could also be assessed with Western blot to ensure that total histone H4 levels are comparable between the different lines.

In the aforementioned phenotypic screens, a relatively conservative approach was taken when variation was encountered between transgenic lines. In our approach, we sought to identify the most striking and consistent phenotypes for future work. It is thus possible that some phenotypes induced by H4 mutations were overlooked due to our general requirement that two independent transgenic lines display the phenotype in order for it to be scored for analyses of flowering time, rosette morphology, DNA content, and transcriptional silencing. Similarly, in our initial experiments assessing chromatin structure, we isolated nuclei from a single plant corresponding to each H4 mutation and screened for massive defects in chromatin structure. Chromatin structure phenotypes (as well as flowering time, rosette morphology, DNA content, and transcriptional silencing phenotypes) that were more subtle were thus more likely to remain unnoticed in our approach. Assessing more plants would add further rigor to our preliminary analyses, especially for analyses of chromatin structure, for which we only assessed a single plant expressing mutant H4 gene initially. In sum, it is possible that some rH4 mutants that were marked as “wild-type” in Table 3.1 may display phenotypes related to the assessed processes that were not recorded in our results. Additionally, in these initial screens, we did not assess the phenotypes of a few of the rH4 mutants that we generated due to the low number of transgenic lines initially recovered (Table 3.1). Therefore, further work remains to evaluate the phenotypes induced by certain H4 mutations (namely, the rH4R45K, rH4R55K, rH4R55A, and rH4R78K mutations).

Flowering time analyses identify several early flowering mutants

We observed that 13 out of the 53 rH4 mutants assessed (rH4R17K, rH4R17A, rH4K20A, rH4R35K, rH4R35A, rH4R36K, rH4R36A, rH4R39K, rH4R39A, rH4R40K, rH4K44R, rH4K44A, and rH4T80V) displayed significantly early flowering compared to the control plants (Table 3.1, Figure 2.4). Certain mutants, such as the rH4K20A and the

rH4R36K mutants, differed in their assigned flowering time classification depending upon which statistical analysis was performed (i.e., using days to flower to perform a Student's *t*-test as in Table 3.1 or conducting *k*-means clustering using all four flowering time variables as in Figure 3.5). However, the vast majority of the early flowering mutants were classified as such in both of the statistical analyses performed. Notably, no rH4 mutants exhibited significantly late flowering in multiple independent transgenic lines, although multiple individual rH4K31R, rH4R77A, and rH4R95K plants were observed to display a late transition to flowering (Figure 3.1).

One explanation for why we observed that approximately 25% of the H4 mutations assessed caused significantly early flowering, while no H4 mutations assessed seemed to cause consistently late flowering, is that certain stresses have been shown to accelerate flowering (Takeno, 2016). As histone H4 is an essential and highly conserved protein in eukaryotes, the histone H4 mutations that we generated are likely to cause some level of stress in rH4 mutants. However, it is important to note that the accelerated flowering time phenotypes exhibited by the early flowering rH4 mutants are more severe than that of the H4 septuple mutant, and approximately 75% of the rH4 mutants did not exhibit early flowering (Table 3.1, Figure 3.1, Figure 3.2). Thus, the histone H4 mutations identified in our flowering time screen do appear to have a specific function in the regulation of flowering time apart from destabilizing the entire histone H4 protein.

In our flowering time analyses, several rH4 mutants that we may have expected to display flowering time phenotypes due to previously described roles for these histone H4 residues instead exhibited wild-type phenotypes. For example, while mutations in PRMT5, the enzyme that catalyzes the symmetric dimethylation of H4R3, have been shown to cause late flowering, mutations of H4R3 were not observed to cause any significant flowering time differences in our screen (Table 3.1, Figure 3.1) (Pei et al.,

2007). This result supports the hypothesis that the late flowering phenotypes in *prmt5* mutants are not caused by a loss of methylation on H4R3, but rather they are mainly caused by the action of PRMT5 on non-histone substrates such as RNA processing factors, leading to splicing defects in transcripts of flowering time regulators such as *FLK* (Deng et al., 2010; Lim et al., 2004; Mockler et al., 2004). Our data contradict the other proposed mechanism for late flowering in *prmt5* mutants by which decreased H4R3me₂s levels on *FLC* reduce the repression of this gene, causing late flowering (Schmitz et al., 2008; Wang et al., 2007). However, it is important to note that H4R3 methylation may not be completely eliminated in the rH4R3K and rH4R3A mutants assessed due to the potential expression from the remaining endogenous histone H4 gene. Therefore, generating complete H4 replacement lines for the rH4R3K and rH4R3A mutants and assessing the resultant flowering time phenotypes would garner additional support for this hypothesis.

Tetra-acetylation of histone H4 lysines 5, 8, 12, and 16 has also been shown to play an important role in the regulation of flowering time, e.g., in their effect on the induction of *VIN3* in response to vernalization (Bond et al., 2009). Mutations in components of the Nucleosome Acetyltransferase of Histone H4 (NuA4) complex, a complex which catalyzes acetylation of multiple histone H4 residues, also cause abnormal flowering time phenotypes (Espinosa-Cores et al., 2020). Moreover, depleting the expression of the putative catalytic subunits of NuA4, HISTONE ACETYLTRANSFERASE OF THE MYST FAMILY 1/2 (*HAM1/2*), causes early flowering and a reduction of H4 acetylation at flowering time regulatory genes such as *FLC* (Xiao et al., 2013). However, the action of NuA4 is not specific to histone H4 as NuA4 has also been shown to act on histones H2A and H2A.Z, and thus mutations in NuA4 components do not exclusively affect acetylation on histone H4 (Boudreault et al., 2003; Millar et al., 2006).

In our screen, we identified no significant early flowering time phenotypes for rH4K5R, rH4K5A, rH4K8R, rH4K8A, rH4K12R, and rH4K12A mutants (Table 3.1, Figure 3.1). We did observe a moderate early flowering phenotype for one rH4K16R transgenic line, but significantly altered flowering time was not consistently demonstrated by either of the rH4K16A lines (Table 3.1, Figure 3.2). Although links between the acetylation of these four lysine residues and flowering time regulation have previously been established, these individual H4 point mutants may not have displayed significant flowering time phenotypes because of the cumulative effect of acetylation on these H4 residues (Dion et al., 2005; Kaimori et al., 2016). Consequently, mutation of individual lysine residues on the H4 N-terminal tail may be compensated by the continued presence of the other N-terminal lysine residues. As combinatorial mutants of several lysines on the H4 N-terminal tail caused lethality (see Chapter 2), we were unable to assess the effect of abolition of acetylation on several or all of these lysine residues on the floral transition.

H4 mutations cause diverse effects on rosette development

We found that mutations of H4R3, H4K5, H4K8, H4K12, H4K16, H4R17, H4R19, H4R36, H4R39, H4K44, H4Y51, H4R77, H4T80 and H4K91 led to a variety of morphological phenotypes, with many plants exhibiting a reduced total rosette area and altered chroma ratios compared to wild-type plants (Figure 3.6, Figure 3.7, Figure 3.8, Figure 3.9). Many future directions remain to characterize the developmental phenotypes effected by these mutations in greater detail. For example, assessments of the rH4 mutants with altered chroma ratios to measure anthocyanin content via spectrophotometry and to perform gene expression analyses of anthocyanin biosynthesis genes would elucidate a mechanism for the altered color profiles observed in these mutants. Additionally, quantifying the expression levels of genes related to cell

growth and using microscopy to examine leaf cell number and cell size would provide further information about what developmental processes are altered in the rH4 mutants exhibiting abnormal rosette area. Our screen for phenotypes related to rosette development serves as a proof-of-concept that the generated H4 mutant library functions as an optimal resource for a high-throughput screen utilizing a software such as ARADEEPOPSIS for trait extraction (Huther et al., 2020). Additional experiments characterizing other morphological phenotypes could further elucidate how the identified H4 mutations regulate plant development. For example, developing a pipeline for quantifying additional morphological traits such as the leaf serration observed in H4 septuple mutants and rH4K44A plants, to a lesser extent, would provide additional data for the large-scale assessment of H4 mutations on plant morphology. Moreover, assessing rosette morphology in short-day conditions could reveal useful information about the ways in which histone H4 residues regulate development in other photoperiodic conditions.

H4 septuple mutant plants exhibit early flowering time

We observed that H4 septuple mutant plants flowered earlier than Col plants in both long-day and short-day conditions (Figure 3.1, Figure 3.2). Under long-day conditions, H4 septuple mutant plants on average flowered 4 days earlier than Col with 3 fewer rosette leaves, and under short-day conditions, H4 septuple mutant plants flowered 10 days earlier than Col with 16 fewer rosette leaves. Several factors may be contributing to the early flowering response displayed by H4 septuple mutant plants. First, the genome instability caused by the low dosage of histone H4 in the H4 septuple mutants is likely producing stress in these mutant plants, which has been demonstrated to cause early flowering in certain instances (Takeno, 2016). Second, it is possible that low histone H4 expression may globally diminish the nucleosome content in the

chromatin of H4 septuple mutant plants, potentially leading to the derepression of certain flowering time regulatory genes. Notably, many H4 septuple mutant nuclei were also observed to exhibit decondensation of chromocenters and/or a loss of chromocenters (Figure 3.14), suggesting that the chromatin in H4 septuple mutant nuclei is in general more open and accessible. To test this hypothesis, ATAC-seq experiments could be performed to assess genome-wide nucleosome occupancy and chromatin accessibility. RNA-seq or additional RT-qPCR experiments measuring the expression of other key flowering time regulatory genes would also serve to elucidate the early flowering response observed in these mutants.

rH4 plants exhibit minor phenotypes related to flowering time and rosette morphology

While the wild-type replacement H4 transgene was previously noted to rescue the genome instability and morphology phenotypes of the H4 septuple mutant (i.e., *BRCA1* and *TSI* upregulation, serrated leaves, small plant size, and low fertility), we observed that the wild-type H4 transgene only partially rescued the early flowering time phenotypes displayed by the H4 septuple mutants in some experiments (Figure 3.1, Figure 3.2). However, in other experiments, rH4 plants were not observed to display a significantly different flowering time from Col plants (Figure 3.1, Figure 3.2). Additionally, rH4 plants exhibited a slightly larger average rosette size than Col plants, although this result does constitute a rescue from the H4 septuple mutant phenotype, which displayed a smaller average rosette size than Col (Figure 3.8, Figure 3.9). These results indicate the importance of using rH4 plants as controls in all experiments assessing phenotypes in the rH4 mutants, as the rH4 plants are not completely identical to Col plants.

Since both independent transgenic rH4 lines exhibited the early flowering and increased rosette area phenotypes, it is unlikely that the genomic location of the T-DNA

insertion encoding the H4 replacement gene is responsible for these phenotypes. Likewise, off-target CRISPR/Cas9 activity independently occurring in both rH4 lines is similarly unlikely to cause an early floral transition in both cases. One possibility is that the genomic location of the original T-DNA insertion that generated the H4 septuple mutant, or off-target CRISPR/Cas9 activity from the generation of the H4 septuple mutant, causes the early flowering response exhibited by the H4 septuple mutant and the rH4 lines, as both rH4 lines were generated from the same H4 septuple mutant background. However, a different H4 septuple mutant line that we previously generated also exhibited the early flowering response (Figure 2.5). Therefore, we deem that it is unlikely that unintended consequences of the generation of the H4 septuple mutant background are completely responsible for the observed flowering time phenotypes of H4 septuple mutant and rH4 plants. To further support this hypothesis, we could generate additional rH4 lines using an independent transgenic H4 septuple mutant background and assess the resultant phenotypes. Additionally, we could perform whole genome-sequencing of the original H4 septuple mutant background.

Two other hypotheses may rationalize the phenotypes displayed by the rH4 plants. First, it is possible that the total histone H4 protein expression in the rH4 plants is different than that of Col plants, and this alteration in histone H4 levels leads to the different flowering time responses and rosette sizes. Western blot experiments quantifying total histone H4 expression in rH4 plants compared to Col plants would expound on this hypothesis. Second, while the seven endogenous histone H4 genes that were eliminated in the H4 septuple mutant background were observed to be redundant for viability and certain phenotypes, such as normal fertility, the possibility remains that one or multiple of these endogenous histone H4 genes play distinct roles in certain developmental stages, such as the floral transition. While all of the endogenous histone H4 genes are identical in protein sequence, they may show differential

transcriptional regulation due to their distinct regulatory sequences. Indeed, the mRNA expression levels of the eight endogenous histone H4 genes across different cell types and developmental stages are not identical (Nakabayashi et al., 2005; Schmid et al., 2005; Waese et al., 2017). Therefore, a significant alteration in histone H4 mRNA expression in certain cells prior to the floral transition, for example, may induce an early transition to flowering in rH4 plants.

No H4 mutants exhibit consistent *BRCA1* upregulation

While multiple yeast and *Drosophila melanogaster* histone H4 replacement mutants have previously been demonstrated to exhibit an impaired DNA damage response/ increased DNA damage sensitivity (Dai et al., 2008; Matsubara et al., 2007; Zhang et al., 2019), no *Arabidopsis thaliana* rH4 mutants assessed in our screen exhibited a consistent upregulation of *BRCA1* in both transgenic lines assessed (Figure 3.16). Concretely, *Drosophila melanogaster* H4K16A and H4R23A replacement mutants exhibited hypersensitivity to X-ray radiation (Zhang et al., 2019), yeast H4Y51A and H4R78A replacement mutants exhibited hypersensitivity to methyl methanesulfonate (MMS), a DNA alkylating agent, and yeast H4R36A and H4Y98A replacement mutants exhibited hypersensitivity to both MMS and hydroxyurea (HU), which impairs deoxyribonucleotide synthesis (Matsubara et al., 2007). It is important to note that we did not systematically assess sensitivity to these different DNA damaging agents in our screens, and thus, even though the *Arabidopsis thaliana* rH4 mutants did not exhibit an upregulation of the DNA damage response factor *BRCA1* when grown in normal growth conditions, they may display hypersensitivity to radiation, MMS, or HU and/or an abnormal *BRCA1* response when treated with these different agents.

Another contributing factor to the lack of rH4 mutants that we determined to exhibit an upregulation of *BRCA1* could be the stringent cutoff threshold we set at 3.5-

fold upregulation relative to Col. Moreover, we specified that both independent transgenic lines assessed must display *BRCA1* upregulation in order for this phenotype to be scored. Therefore, while one of the rH4Y98F mutants exhibited a 116-fold *BRCA1* upregulation relative to Col, the other rH4Y98F mutant that we examined displayed no increase in *BRCA1* expression (1.1-fold relative to Col) (Figure 3.16). While it is promising that mutations of H4Y98 also cause DNA damage hypersensitivity in yeast, further experiments are needed to determine whether the *BRCA1* upregulation in the rH4Y98F mutant is a consistent phenotype (see below). Moreover, one of each of the rH4R39A, rH4T54V, and rH4K91A mutant lines displayed a low upregulation of *BRCA1* between 3.5- and 6.5-fold relative to Col. However, the other mutant lines expressing these constructs displayed no upregulation of *BRCA1*, and thus additional experiments are required to investigate this phenotype in further detail. Nonetheless, while H4T54 is relatively uncharacterized in all eukaryotes, potential roles for H4R39 and H4K91 in the regulation of genome stability are discussed below.

Assays of H4K16 and H4K20 mutants support previously described roles for these residues

H4K16 and H4K20 are two of the most well-characterized residues on histone H4 due to their important regulatory roles as part of the H4 N-terminal tail. H4K16 acetylation has been demonstrated to play vital roles in mediating chromatin compaction and transcriptional activation (Oppikofer et al., 2011; Shogren-Knaak et al., 2006; Zhang et al., 2017). Concretely, H4K16 contacts the H2A/H2B acidic patch on the adjacent nucleosome and acetylation of this lysine residue neutralizes its positive charge to inhibit this inter-nucleosomal interaction and cause a decompaction of the chromatin structure (Shogren-Knaak et al., 2006; Zhang et al., 2017). H4K16 acetylation is thus believed to serve a principal role in activating transcription through its role in regulating chromatin

compaction; H4K16 yeast mutants exhibit broad, transcriptional activation and H4K16ac is found at transcriptionally active chromatin (Dion et al., 2005; Kimura et al., 2002; Soppe et al., 2002; Suka et al., 2002). Moreover, deacetylated H4K16 allows the creation of silent chromatin domains through the binding of silent information regulator (Sir) proteins in yeast (Armache et al., 2011; Hecht et al., 1996).

In contrast to H4K16 acetylation, H4K20 acetylation is associated with gene repression (Kaimori et al., 2016) and H4K20 methylation has been shown to maintain chromatin compaction in human cells, potentially related to the interaction of H4 residues K16-I26 with the neighboring H2A/H2B acidic patch and/or the interaction of this residue with other proteins involved in mediating nucleosome interactions (Kalashnikova et al., 2013; Kan et al., 2009; Shoaib et al., 2018; Trojer et al., 2007). As such, H4K20 mono-, di-, and trimethylation are often found to be associated with heterochromatin and silenced genomic regions in animals, and H4K20me1 associates with heterochromatin in plants (Ebert et al., 2004; Edwards et al., 2011; Naumann et al., 2005; Schotta et al., 2004). Similarly, H4K20 acetylation is enriched around TSSs of minimally expressed genes in human cells (Kaimori et al., 2016). However, the role of H4K20 methylation in regulating transcription appears to be more complex than simply always serving to repress transcriptional activation, as H4K20 methylation has been linked to transcriptional activation as well as repression in animals (Congdon et al., 2010; Li et al., 2011a; Li et al., 2011b; van Nuland and Gozani, 2016). In fact, in certain experiments, such as profiling of genome-wide H4K20 methylation in the human genome, H4K20me1 was found to be enriched in active gene bodies (Barski et al., 2007).

Additionally, H4K20 methylation has mainly been deduced to play an activating role in transcriptional regulation in plants, as H4K20me2 and H4K20me3 are associated with actively transcribed euchromatin in *Arabidopsis thaliana* (de la Paz Sanchez and Gutierrez, 2009; Fischer et al., 2006). Due to the differences observed in the genome-

wide localizations of H4K20 mono-, di-, and trimethylation, as well as differential effects on the regulation of transcription and replication, the functions of H4K20 methylation are thought to present significant differences between plants and animals (Balakrishnan and Milavetz, 2010; de la Paz Sanchez and Gutierrez, 2009; Fischer et al., 2006; Shoaib et al., 2018). In fact, plants even lack the Set8 (also named Pr-Set7) enzyme responsible for the monomethylation of H4K20 in metazoans (Nishioka et al., 2002; Pontvianne et al., 2010). Interestingly, H4K20A replacement mutants display severely reduced viability in *Drosophila melanogaster*, and thus the effect of this mutation on gene silencing in *Drosophila melanogaster* has not been assessed (McKay et al., 2015b; Zhang et al., 2019). In contrast, we observed *Arabidopsis thaliana* rH4K20R and rH4K20A mutants to be viable and fertile (Figure 3.4).

In our experiments, we observed that DAPI-stained rH4K16A and rH4K20A mutant nuclei both exhibited chromocenter decondensation (Figure 3.13). These results are consistent with previously published data on these residues, as the neutralization of the positive charge on lysine 16 in the H4K16A mutant would be expected to lead to a decompaction of the chromatin structure. While distinct methylation marks (mono-, di-, and tri-) on H4K20 are differentially associated with compacted or more open chromatin in *Arabidopsis thaliana*, complete ablation of H4K20 methylation in the rH4K20A mutant is expected to impact chromatin structure, and we observed a general decompaction of chromocenters (de la Paz Sanchez and Gutierrez, 2009; Fischer et al., 2006; Naumann et al., 2005). For both of these mutants, as well as other rH4 mutants exhibiting chromatin condensation defects discussed below, future work could measure chromocenter size, quantify the percentage of nuclei exhibiting chromatin condensation defects, and assess which elements of constitutive heterochromatin correspond to the decondensed chromatin observed in the DAPI-stained nuclei using fluorescent *in situ* hybridization (FISH) (Soppe et al., 2002). The chromatin structure of the rH4K16R and

rH4K20R mutants could also be assessed in more detail.

Given the substantial quantity of previously described roles for H4K16 and H4K20 in transcriptional regulation, it was somewhat unexpected that the rH4K16R, rH4K16A, rH4K20R, and rH4K20A mutants did not exhibit any significant transcriptional changes in our RT-qPCR experiments (Figure 3.15, Figure 3.16). One rH4K20R mutant assessed displayed a 43-fold *TSI* upregulation relative to Col, but all other rH4K16R, rH4K16A, rH4K20R, and rH4K20A mutant plants exhibited no increase in *TSI* RNA expression over 2.5-fold relative to Col. Further experiments measuring genome-wide RNA expression with RNA-seq, or measuring the RNA expression of additional transposable elements such as *COPIA 28* with RT-qPCR (Jing et al., 2016), would elucidate the effect that H4K16 and H4K20 mutations have on transcriptional regulation, as only one repressed genomic element was assayed in our RT-qPCR experiments. However, we also note that *Drosophila melanogaster* H4K16A replacement mutants were observed to exhibit no gene silencing defects (measured by RNA expression of the *copja* retrotransposon), consistent with our results (Zhang et al., 2019).

Finally, unlike in animal systems, where the depletion of H4K20 methylation appears to severely impact viability and development (Huen et al., 2008; Karachentsev et al., 2005; McKay et al., 2015b; Oda et al., 2009; Zhang et al., 2019), *A. thaliana* rH4K20R and rH4K20A mutants did not display any major phenotypes (e.g., fertility impairment or small size) or a reduction in viability (Figure 3.4). While H4K20 methylation has been demonstrated to regulate the DNA damage response in yeast (Du et al., 2006; Sanders et al., 2004), yeast H4K20 mutants were observed to lack major defects in viability and sporulation efficiency, similar to plants (Govin et al., 2010a). Additionally, H4K20 methylation has recently been shown to play a vital role in DNA replication licensing and the prevention of excess DNA damage in animals (Botuyan et al., 2006; Kuo et al., 2012; Shoaib et al., 2018; Tardat et al., 2010; van Nuland and

Gozani, 2016; Wang et al., 2009b). In the absence of the SET8 enzyme or the H4K20 residue in human cell lines, excessive ORC loading occurs due to a global loss of chromatin compaction, leading to genome instability (Shoaib et al., 2018). In addition, unmethylated H4K20 residues have been shown to recruit the BRCA1/BRCA1-associated RING domain protein 1 (BARD1) and MMS22-Like (MMS22L)/Tonsoku-like (TONSL) complexes which mediate HR and enable the homology-directed error-free repair of DNA breaks (Nakamura et al., 2019; Saredi et al., 2016). In contrast, *Arabidopsis thaliana* rH4K20R and rH4K20A mutants displayed no apparent defects in replication when DNA content was measured by flow cytometry (Table 3.1), and moreover, rH4K20R and rH4K20A mutants exhibited no upregulation of *BRCA1* expression (Figure 3.16). These results indicate that the regulatory functions of H4K20 methylation in plants and animals may have diverged, as H4K20 methylation appears to serve a more important role in the regulation of development, replication initiation and DNA damage in animals.

In contrast, we observed that *A. thaliana* rH4K16A mutants displayed altered morphology, most notably in their reduced rosette area compared to Col plants (Figure 3.7, Figure 3.8, Figure 3.9). Additionally, we demonstrated that rH4K5R, rH4K8A, and rH4K12A mutants also exhibited moderately to severely reduced rosette area (Figure 3.7, Figure 3.8, Figure 3.9). As discussed previously, these four lysine residues are all targets for acetylation by histone acetyltransferases, with the acetylation on lysine residues 5, 8, and 12 demonstrating a cumulative effect on transcriptional activation (Dion et al., 2005; Govin et al., 2010a). It is thus interesting that the rH4K5R and rH4K8A mutants exhibited a more severe reduction in rosette size than the rH4K16A mutants, as acetylation on H4K16 has been demonstrated to regulate transcription in a more independent fashion than acetylation on lysine residues 5, 8, and 12 (Oppikofer et al., 2011; Shogren-Knaak et al., 2006; Zhang et al., 2017). However, future experiments

assessing the remaining expression of endogenous wild-type histone H4 in these different rH4 lines would be important to determine whether there has been a complete replacement of endogenous histone H4 with mutant H4. Nonetheless, previous experiments with rH4K16A plants revealed that 33% of these plants exhibited a complete replacement of endogenous histone H4 (Figure 2.10), demonstrating that many of the rH4K16A mutants no longer contain any wild-type H4 protein.

Mutations in acetyltransferases targeting histone H4 have been demonstrated to cause pleiotropic developmental phenotypes, including extremely reduced size, as well as other phenotypes not observed in our H4 N-terminal tail lysine replacement mutants (e.g., sterility) (Bertrand et al., 2003; Bond et al., 2009; Vlachonasios et al., 2003). However, as these histone acetyltransferases have multiple targeting specificities on not only histone H4, but other histones as well, mutations in these chromatin-modifying enzymes would be expected to cause more severe developmental phenotypes than individual mutations in histone residues (Benhamed et al., 2006; Earley et al., 2007). Notably, we observed that combinatorial mutations of multiple lysines on the H4 N-terminal tail caused lethality (see Chapter 2), thus suggesting that these lysines do indeed regulate plant development in a cumulative manner.

H4R17 mutants demonstrate unique phenotypes relative to H4 N-terminal tail residues

In contrast to H4K16 and H4K20, the role of H4R17 in the regulation of development and transcription has been characterized in much less depth. While H4R17 has been shown to be a target for methylation in mammals and yeast (Crespo et al., 2020; Tweedie-Cullen et al., 2012), methylation on H4R17 has not been identified in *Arabidopsis thaliana* or *Drosophila melanogaster* (Brabencova et al., 2017; Plazas-Mayorca et al., 2009; Zhang et al., 2007a; Zhang et al., 2019). In mammals,

monomethylation of H4R17 is thought to be performed by the enzyme PROTEIN ARGININE METHYLTRANSFERASE 7 (PRMT7) (Feng et al., 2014b; Feng et al., 2013; Jain and Clarke, 2019). Developmental defects are observed in mammals lacking PRMT7, but the mechanism for these developmental changes remains to be elucidated (Jain and Clarke, 2019). The unmodified form of H4R17 has also been shown to play an important role in regulating various genomic processes along with two other residues on the H4 N-terminal tail basic patch (H18 and R19). Residues R17 to R19 on the H4 basic patch are required for H3K79 methylation and telomere silencing in yeast (Dai et al., 2008; Fingerman et al., 2007), and moreover, these three residues have been demonstrated to regulate the nucleosome remodeling activity by ISWI family remodeling complexes in yeast and animals (Clapier et al., 2001; Clapier et al., 2002; Dann et al., 2017; Fazio et al., 2005; Hamiche et al., 2001; Yan et al., 2016).

In our screens, rH4R17K and rH4R17A mutants were the only mutants assessed in the H4 N-terminal tail that displayed a consistent and significant early flowering phenotype relative to the controls (Figure 3.1, Table 3.1). The H4 N-terminal tail comprises residues S1-T30 and residues S1-R19 are not well localized by crystallography (Iwasaki et al., 2013; Luger et al., 1997). The impact of the rH4R17K and rH4R17A mutations on flowering time represent a novel role for H4R17 in the regulation of this developmentally important process. Additionally, rH4R17A mutants also displayed an upregulation of *TSI* (Figure 3.15), but RNA levels of *BRCA1* were normal (Figure 3.16). Finally, rH4R17A mutants exhibited a small rosette size and high chroma ratio, suggesting reduced anthocyanin content (Figure 3.6, Figure 3.7, Figure 3.8, Figure 3.9). Potential mechanisms for the regulation of these processes by H4R17 will be discussed in detail in Chapter 4.

Several mutations of the nucleosome entry site (H4R35, H4R36, H4R39, H4R40, and H4K44) cause early flowering

Mutations of H4R35 induced early flowering (Figure 3.1, Figure 3.2) and chromocenter decondensation (Figure 3.14). Moreover, rH4R35K mutants also exhibited altered proportions of 16N and 32N nuclei in mature leaves (Figure 3.11). These phenotypes suggest roles for H4R35 in the regulation of chromatin structure, the cell cycle and/or endoreduplication, and the floral transition. rH4R36K and rH4R36A mutants also displayed early flowering (Figure 3.1, Figure 3.2), although they did not display chromocenter decondensation or altered endoploidy levels (Table 3.1). rH4R36K mutants displayed no transposon derepression, while rH4R36A mutants displayed increased RNA expression of *TSI* (Figure 3.15). Mutations of H4R39 similarly induced early flowering (Figure 3.1, Figure 3.2) and chromocenter decondensation (Figure 3.14) as well as transposon derepression (Figure 3.15, Figure 3.17). Additionally, one rH4R39A mutant line displayed upregulation of *BRCA1* (Figure 3.16), but further assessments are required to determine whether this phenotype is consistently observed in plants expressing this mutant histone. Finally, rH4R36A and rH4R39A mutants both exhibited a reduced rosette size and low chroma ratio, indicating increased anthocyanin content (Figure 3.6, Figure 3.7, Figure 3.8, Figure 3.9).

The transcriptional derepression in rH4R39 K and rH4R39A mutants could indicate greater genome instability and further work could assess the RNA expression of other transposable elements in these mutant lines. It is possible that the chromatin decondensation exhibited by rH4R39K and rH4R39A mutants leads to transcriptional derepression of heterochromatic elements, as observed for the transposon *TSI*, through the generation of a more accessible chromatin environment for the transcriptional machinery to bind. Interestingly, the H4R39A mutation has been shown to cause lethality in *Saccharomyces cerevisiae*, but the role of this residue is relatively uncharacterized in

multicellular eukaryotes (Dai et al., 2008; Govin et al., 2010a; Nakanishi et al., 2008). H4R35 and H4R36 are similarly uncharacterized, and there have been no post-translational modifications identified on any of these three residues.

One potential effect of the H4R35, H4R36 and H4R39 mutations could be the disruption of the nucleosome entry-exit site on histone H4 (mapped to I34-Y51), affecting the wrapping of DNA around the nucleosome (Dai et al., 2008; Zhou et al., 2019). The regulation of the unwrapping rate of the DNA from the nucleosome has been demonstrated to control DNA accessibility and approximately 30% of *Saccharomyces cerevisiae* transcription factor binding sites are located on the nucleosome entry-exit region (North et al., 2012). Histone PTMs residing on the nucleosome entry-exit region, such as H3K56 acetylation, can enhance DNA accessibility by promoting partial DNA unwrapping (Neumann et al., 2009). This partial unwrapping is thought to promote transcription by allowing transcription factor binding without altering the nucleosome structure (North et al., 2012; Shimko et al., 2011). Substituting the H4R35, H4R36 and H4R39 residues with alanine or lysine residues could similarly modulate DNA unwrapping and affect transcription, leading to the observed phenotypes in these mutants. Interestingly, H4R39K mutations were identified in a yeast screen to suppress mutations in the LRS (loss of rDNA silencing) patch on the nucleosome (Norris et al., 2008). In contrast to our results in *Arabidopsis thaliana*, *Saccharomyces cerevisiae* H4R39K mutations were found to enhance silencing and lead to the spreading of silent regions (Xu et al., 2005). The mechanism for this silencing enhancement in H4R39K yeast mutants remains unknown.

Nonetheless, the expression levels of additional genomic regions would need to be assessed to draw a broader conclusion about the effect that the H4R35, H4R36 and H4R39 mutations have on global transcription. It is also possible that H4R39 mutations have a stronger effect on nucleosome wrapping than the H4R35 and H4R36 mutations

due to their location on the nucleosome entry-exit site, which is why more severe transcriptional changes are observed in the rH4R39K and rH4R39A mutants. The side chain of H4R36 makes direct hydrogen bonds with the DNA, while the interaction of H4R39 with DNA phosphate groups is assisted through a water molecule (Davey et al., 2002). From this structural data, it would be expected that the H4R36 mutations have a major effect on the interaction of the DNA strand with the nucleosome, but further work is necessary to test this hypothesis. To assess the effects of these histone H4 mutations on nucleosome wrapping, fluorescence resonance energy transfer (FRET) could be performed to measure nucleosome unwrapping kinetics (North et al., 2012).

Another potential mechanism for the H4R35, H4R36 and H4R39 mutant phenotypes could be the effect of these mutations on the interaction of histone H4 and the CAF1 complex. Structural studies of the human CAF1 subunit RbAp48 found that the section encompassing residues 34 to 40 on histone H4 is particularly important for the recognition of histone H4 by RbAp48 (Murzina et al., 2008). Site-directed mutagenesis of multiple residues on this section of histone H4 (including H4R35 and H4R39) disrupts the binding of histone H4 to RbAp48, although single substitutions of H4R35, H4R36 and H4R39 were not assessed. Studies of the *Drosophila melanogaster* homolog of RbAp48, p55, similarly found that the interactions of histone H4 residues 34 to 40 with the p55 binding pocket were a major contributor to the binding of p55 to histone H4 (Song et al., 2008). The H4R39 residue was deemed to play a critical role in this recognition and the H4R39A mutation disrupted binding of p55 to histone H4. While structural studies of the *Arabidopsis thaliana* RbAp48 homolog, MULTICOPY SUPPRESSOR OF IRA 1 (MSI1), to determine the basis of histone H4 recognition have not been performed, mutations in genes encoding CAF-1 subunits such as *FASCIATA 1* (*FAS1*) lead to increased proportions of higher ploidy nuclei in *A. thaliana*, similar to rH4R35K mutants (Hisanaga et al., 2013; Ramirez-Parra and Gutierrez, 2007).

One interesting structural feature of the interacting surface of histone H4 is that H4 residues 33 to 40 form alternating pairs of hydrophobic and basic amino acids (AI-RR-LA-RR), and thus one side of the alpha helix is hydrophobic and the other side is positively charged (Song et al., 2008). With this structure in mind, alanine substitutions of H4R35, H4R36 and H4R39 would be expected to cause a more severe phenotype than lysine substitution mutants due to the elimination of the positive charge. We did observe that the rH4R39A and the rH4R36A mutants displayed a more severe phenotype than that of the rH4R39K and rH4R36K mutants, respectively (Figure 3.15, Figure 3.17), but the rH4R35A mutants did not substantially differ from the rH4R35K mutants in flowering time or chromatin structure (Figure 3.1, Figure 3.2). In addition, only the rH4R35K mutants displayed altered endoploidy levels (Figure 3.11), while the rH4R35A, rH4R36K, rH4R36A, rH4R39K, and rH4R39A mutants displayed no alterations in endoploidy levels (Table 3.1).

Therefore, while the H4R35, H4R36 and H4R39 mutations may cause the observed phenotypes due to the impaired interaction of histone H4 with CAF1 and subsequent disruption of chromatin assembly, further work is required to elucidate why mutations of these three residues lead to differing phenotypes, including why some H4R36 mutants lack indications of impaired chromatin structure or genome instability. Interestingly, H4R36 has been demonstrated to play a key role in promoting the proper localization and degradation of the centromeric histone H3 variant CENP-A in yeast, and H4R36A mutant cells display CENP-A enrichment in euchromatin along with decreased nucleosome occupancy at transcribed genes and transcriptional defects, while H4R36K mutant cells lack these defects (Deyter et al., 2017). Therefore, the specific role of H4R36 in regulating the localization of the centromeric histone H3 variant may also partially result in the distinct phenotype observed for rH4R36K and rH4R36A mutants.

Additionally, it is unclear why only H4R35K mutations but not H4R35A mutations

cause the observed changes in endoreduplication, as alanine substitutions may be expected to cause more severe effects than lysine substitutions at this residue. However, we did find that the endogenous histone H4 gene (*At3g53730*) in rH4R35A mutants remained wild-type in all lines assessed, indicating that no mutagenesis by CRISPR/Cas9 had occurred, while the endogenous H4 gene in rH4R35K plants displayed homozygous or biallelic mutations for the vast majority of plants (Figure 3.12). It is thus possible that additional phenotypes were observed in rH4R35K plants due to the complete elimination of wild-type histone H4 expression in these lines.

A third hypothesis for the mechanism causing rH4R35K and rH4R35A mutant phenotypes relates to the role of this residue in translocation by chromatin remodelers such as ISWI. Namely, recent cryo-EM data on “canonical,” “distorted,” and “translocated” nucleosome structures (i.e., three differently organized nucleosome structures that occur before and during translocation) demonstrate that H4R35 makes the main contact of the DNA with histone H4 in the “translocated” class (Bilokapic et al., 2018). Thus, H4R35 mutations may affect the translocation of nucleosomes by chromatin remodelers and the resultant alteration of nucleosome spacing could cause various effects on gene expression and chromatin structure, producing the observed phenotypes.

In addition to the aforementioned rH4 mutants, the rH4R40K, rH4K44R, and rH4K44A mutants also displayed early flowering (Figure 3.1, Figure 3.2) and rH4K44A mutants exhibited developmental phenotypes including minor leaf serration and upwardly curled leaves (Figure 3.6). Along with the rH4R17A, rH4R36A, rH4R39K, and rH4R39A lines, the rH4K44A mutants exhibited the most drastic early flowering response (Figure 3.5), and similar to H4R35, H4R36 and H4R39, the H4R40 and H4K44 residues reside on the nucleosome entry-exit site. The early flowering observed in all of these mutants may thus be related to transcriptional changes of flowering time

regulatory genes caused by partial unwrapping of the DNA, leading to a more accessible chromatin structure (North et al., 2012; Shimko et al., 2011). Mutations of H4R40 may also impact the interaction of histone H4 with the CAF1 complex (Murzina et al., 2008; Song et al., 2008), and it is possible that the lethality induced by the rH4R40A mutation may be related to this interaction (Figure 2.7).

In contrast to H4R40 mutants, which complicate analyses due to the lethality induced by the H4R40A mutation, H4K44 mutants have previously been characterized in more detail. Concretely, *Saccharomyces cerevisiae* H4K44A replacement mutants exhibit chromosome missegregation, poor growth and aneuploidy (Ng et al., 2013) as well as intragenic cryptic transcription due to a loss of H3K36 methylation (Du and Briggs, 2010). H4K44 acetylation has also been shown to be enriched in yeast recombination hotspots during meiosis and H4K44 mutants exhibit defects in sporulation and increased nucleosome content around recombination hotspots (Govin et al., 2010a; Hu et al., 2015). H4K44 acetylation is thus believed to promote chromatin accessibility for normal DSB formation during meiotic recombination in yeast (Fenley et al., 2018; Hu et al., 2015). While H4K44 acetylation has not been identified in plants, it has recently been identified in mammals (Li et al., 2019). Experiments assessing the levels of H3K36 methylation in H4K44 mutants and the fertility of H4K44 mutants may begin to elucidate whether a similar role for H4K44 exists in plants as in yeast. Further mass spectrometry experiments to determine whether H4K44 is acetylated in *Arabidopsis thaliana* would also clarify the function of this residue.

Mutations corresponding to the yeast H4 low sporulation patch (H4R77 and H4T80) impact chromatin structure

Nuclei isolated from mature leaves of rH4R77K and rH4T80V mutant plants displayed chromocenter decondensation for both transgenic lines assessed (Figure

3.14). In addition, rH4T80V replacement mutants exhibited early flowering (Figure 3.1, Figure 3.2). With our clustering analyses, we found that rH4T80V-1 was very early flowering and rH4T80V-2 was moderately early flowering (Figure 3.5). In contrast, while some rH4R77K and rH4R77A mutant plants were noted to exhibit late flowering in our initial screen (Figure 3.1), rH4R77K and rH4R77A mutants failed to exhibit any consistent and significant alteration in flowering time compared to rH4 plants (Table 3.1). rH4R77K, rH4R77A, and rH4T80V mutants also displayed no significant changes in *TSI* or *BRCA1* RNA expression (Figure 3.15, Figure 3.16). As only one amino acid substitution corresponding to each H4 residue of interest was included for detailed analyses of chromatin structure, the rH4R77A mutant was not assessed in-depth for chromatin condensation defects. However, rH4R77A mutant nuclei were noted to exhibit chromocenter decondensation in our initial screens for chromatin condensation phenotypes (data not shown). Finally, one rH4R77K mutant line exhibited increased rosette area, while the other rH4R77K mutant line and both rH4T80V lines exhibited increased chroma ratios (Figure 3.9).

Our results indicate novel functions for H4R77 and H4T80 in the regulation of chromatin structure, as well as a novel role for H4T80 in flowering time regulation. While further analyses of the developmental phenotypes induced by the H4R77 and H4T80 mutations are required to determine the precise effect of these mutations on rosette morphology, preliminary results also suggest a novel role for these residues on the regulation of plant development. Interestingly, both of these residues lie within a region of histone H4 denoted the LOS (low sporulation) patch. The LOS patch was identified in a screen for sporulation efficiency in yeast histone replacement mutants and corresponds to H4 residues 77–83 (Govin et al., 2010a). This patch is located on the edge of the nucleosome and lies in close proximity to contacts with the DNA strand as it wraps around the nucleosome (Govin et al., 2010a). The LOS patch almost completely

coincides with the L2 loop of histone H4 (residues 77-82), which composes part of the lateral surface of the core histone octamer in direct contact with the DNA and the L1 loop of histone H3 (Luger et al., 1997; Millan-Zambrano et al., 2018). Concretely, H4K77 in *Saccharomyces cerevisiae* can interact with the DNA strand via a water molecule, while the side chain of H4T80 makes direct hydrogen bonds with the DNA (Davey et al., 2002). Analysis of yeast H4 mutants corresponding to the LOS patch determined that while DNA replication and the silencing of sporulation-repressing genes appeared normal, DAPI staining indicated that these mutants lacked the four normal meiotic products, and thus the LOS mutations may impact meiosis (Govin et al., 2010a).

Similarly, we did not observe any changes in DNA content in the rH4 mutants covering residues 77 to 83, supporting the idea that DNA replication is unaffected in these rH4 mutants (Figure 3.10). One explanation for the chromatin condensation defects in the *Arabidopsis thaliana* rH4R77 and rH4T80 mutants is that these mutations impact the interaction of the nucleosome with the DNA, facilitating a less inaccessible chromatin structure (Davey et al., 2002). This hypothesis could also explain why LOS mutants lack the four normal meiotic products in yeast (Govin et al., 2010a). Further work characterizing fertility and pollen grain formation in these *Arabidopsis thaliana* rH4 mutants could serve to elucidate whether meiosis and gametogenesis are affected in histone H4 L2 loop mutants in plants as well.

Several residues on the LOS patch are shared by the LRS (loss of rDNA silencing) patch, covering H4 residues 78 to 81 as well as residues on histone H3 (Park et al., 2002). Certain LRS mutations were demonstrated to cause gene silencing defects via the disruption of Sir3 recruitment (Norris et al., 2008). Some of these mutations may also affect the interaction of the nucleosome with the DNA strand to modulate chromatin accessibility (Hyland et al., 2005). Notably, yeast H4K79A mutations cause a decrease in both telomeric and rDNA silencing, while H4K79R mutations do not affect silencing of

these heterochromatic regions (Hyland et al., 2005). In contrast, yeast H4K77A mutations lead to an increase in rDNA silencing and a decrease in telomeric silencing, while H4K77R mutations do not affect rDNA silencing but increase telomeric silencing. Consequently, H4K77A mutations were deemed to confer an *irs* (increased rDNA silencing), rather than an *Irs*, phenotype (Hyland et al., 2005). The effects of H4R78, H4T80, and H4T82 mutations on gene silencing in yeast were not assessed.

In our screen using *Arabidopsis thaliana*, we did not observe any gene silencing defects in the H4R77, H4K79, or H4T80 mutants (Figure 3.15). One rH4T82V mutant line exhibited a 14-fold upregulation of *TSI*, but the other rH4T82V mutant line exhibited no increase in *TSI* RNA expression (Figure 3.15). We did not assess the phenotypes of rH4R78K mutants due to the low number of rH4R78K transgenic lines that were initially recovered, and similarly, we were unable to assess the gene silencing and chromatin structure phenotypes of the rH4R78A and rH4K79R mutants due to lethality induced by these mutations (Table 3.1). Our RT-qPCR results may indicate a functional difference for H4 residues 77 and 79 in regulating gene silencing in yeast and plants. Residue 77 represents a conserved amino acid substitution at histone H4, as it is an arginine residue in plants and a lysine residue in yeast and animals (Figure 2.1). This difference in protein sequence may lead to a divergence in the function of residue 77 on histone H4, as H4K77 is acetylated in animals (Li et al., 2019; Zhang et al., 2003), but is unable to be acetylated in plants. H4K79 has also been shown to be a target for acetylation in animals (Zhang et al., 2003).

Acetylation of H4K79 in yeast was speculated to decrease gene silencing in *Saccharomyces cerevisiae* through the neutralization of the positive charge on H4K79, as H4K79A mutants exhibited decreased silencing while H4K79R mutants exhibited no change in silencing (Hyland et al., 2005). In contrast, the role of acetylation on H4K77 in regulating global gene silencing is less clear, as both the H4K77R and H4K77A

mutations differentially affected silencing of telomeric and rDNA regions (Hyland et al., 2005). However, further experiments are required to determine whether lysine acetylation is indeed responsible for the observed gene silencing defects in these H4 mutants. While it is unknown whether H4K79 is acetylated in plants, H4K79 acetylation has thus far not been identified in *Arabidopsis thaliana*, and moreover, H4R77 is unable to be acetylated (Zhang et al., 2007a). This difference in acetylation status may explain the different gene silencing phenotypes observed in *Arabidopsis thaliana* and *Saccharomyces cerevisiae*. However, we also note that we did not systematically assess silencing of multiple genomic regions in our gene silencing screen in *Arabidopsis thaliana*, and concretely, we only assayed a single genomic region *TSI*, located within the pericentromeric heterochromatin (Simon et al., 2015; Steimer et al., 2000). Thus, it is possible that mutations such as the H4K79A mutation affect rDNA and telomeric silencing in plants, as they do in yeast mutants (Hyland et al., 2005).

Additionally, H4T80 phosphorylation has recently been demonstrated to regulate DNA damage checkpoint recovery in yeast and *S. cerevisiae* H4T80A mutant cells exhibit hypersensitivity to DNA damaging agents (Millan-Zambrano et al., 2018). While *Arabidopsis thaliana* rH4T80V mutants exhibited no upregulation of *BRCA1* (Figure 3.16) and H4T80 phosphorylation has not been identified in plants, it would be interesting to assess the DNA damage response in the rH4T80V mutant (e.g., through examining MMS hypersensitivity) as a future direction for this work.

C-terminal H4 mutants (H4K91 and H4Y98) demonstrate altered DNA content

rH4K91R and rH4Y98F mutants exhibited increased proportions of higher ploidy nuclei by flow cytometry, while rH4K91A mutants displayed wild-type ploidy levels (Figure 3.5, Figure 3.11, Figure 3.13). Additionally, one rH4Y98F mutant line exhibited substantial *BRCA1* upregulation, while the other rH4Y98F mutant demonstrated no

increase in the RNA expression of *BRCA1* (Figure 3.16). As this RT-qPCR experiment was a preliminary assessment of *BRCA1* RNA expression in these mutant lines, further analyses are required to determine whether the rH4Y98F mutants display a consistent upregulation of *BRCA1*. Similarly, one rH4K91A mutant displayed low upregulation of *BRCA1*, while the other transgenic line assessed did not (Figure 3.16). As opposed to the rH4K91A and rH4Y98F mutants, neither of the rH4K91R mutant lines displayed any upregulation of *BRCA1* (Figure 3.16). rH4K91A and rH4K91R mutants were noted to exhibit serrated leaves and a small size compared to wild-type plants (Figure 3.7, Figure 3.8, Figure 3.9). rH4Y98F mutants, in contrast, displayed large leaves (data not shown), potentially due to the increased proportion of higher ploidy cells (Melaragno et al., 1993; Robinson et al., 2018).

Future experiments measuring leaf epidermal cell size in rH4R35K, rH4K91R and rH4Y98F mutants would elucidate the relationship between ploidy and leaf size in these mutants. rH4Y98F mutants are the only rH4 mutants that displayed both an increase in higher ploidy levels and an increase in leaf size, while the rH4R35K mutants appeared slightly small in size (Figure 3.3, Figure 3.4) and both the rH4K91R and rH4K91A mutants were very small (Figure 3.7, Figure 3.8, Figure 3.9). Previous studies have demonstrated direct linear proportionality between cell size and ploidy in many cell types in *Arabidopsis thaliana*, but in contrast, the increase in organ size due to increased ploidy is often counteracted by a decrease in cell number so that organ size increases more gradually (Melaragno et al., 1993; Robinson et al., 2018). A decrease in cell number may explain why the rH4R35K and rH4K91R mutants are small despite their increased proportions of higher ploidy cells, but the mechanism for this decrease remains to be elucidated. While cell cycle arrest and/or cell death may contribute to the small size of the rH4K91R and rH4K91A mutants, as speculated for the H4 septuple mutant (see Chapter 2 discussion), a potential mechanism for rH4R35K mutant plants is

less clear.

With our screen, we have identified novel roles for H4K91 and H4Y98 in the regulation of endoreduplication in plants. In contrast to many of the aforementioned H4 core residues, which lie on the lateral surface of the nucleosome, H4K91 is buried within the nucleosome core and has been shown to be acetylated in animals (Hyland et al., 2005; Li et al., 2019; Yang et al., 2011; Zhang et al., 2003). H4K91 is located at the interface between the H3-H4 tetramer and the H2A-H2B dimer, and it is hypothesized that the acetylation of this residue regulates nucleosome stability through the modulation of histone-histone interactions (Fenley et al., 2018). Concretely, H4K91ac is believed to influence chromatin assembly through its role in regulating the formation of histone octamers and H4K91ac has been proposed to increase global chromatin accessibility (Fenley et al., 2018; Ye et al., 2005).

To support this hypothesis, H4K91A yeast mutants exhibit transcriptional derepression, alteration of silent chromatin structure and hypersensitivity to DNA damaging agents (Ye et al., 2005). Moreover, yeast strains containing mutations in chromatin assembly factors similarly exhibit hypersensitivity to DNA damaging agents and defective silent chromatin structure, and genetic evidence indicates that H4K91A mutations are involved in the same pathway as histone chaperones that assemble chromatin during DNA repair (Adams and Kamakaka, 1999; Tyler et al., 1999; Ye et al., 2005; Yu et al., 2011b). Finally, depletion of the histone acetyltransferase that targets H4K91 impairs nucleosome assembly, inhibits cell proliferation, causes hypersensitivity to DNA damage, and provokes apoptosis in animal cells (Yang et al., 2011). H4K91 has also been demonstrated to be a target for monoubiquitination in animals and H4K91ub has recently been shown to play an important role in the regulation of genome stability, as the loss of H4K91ub impairs the DNA damage response (Tessadori et al., 2017; Yan et al., 2009). H4K91R mutant histones expressed in zebrafish also lead to abnormal cell

cycle progression and apoptosis (Tessadori et al., 2017). While *A. thaliana* rH4K91A and rH4K91R mutants did not exhibit defects in gene silencing, they did display morphological phenotypes that could indicate impaired cell proliferation (i.e., serrated leaves and small size). rH4K91R mutants also demonstrated increased DNA content, suggesting that cell cycle defects may be present in these plants. Finally, one rH4K91A mutant line demonstrated a 5-fold increase in *BRCA1* expression, but additional analyses are required to determine whether this DNA damage response phenotype is consistently displayed.

Further assessing the DNA damage response and cell proliferation in the rH4K91R and rH4K91A mutants would elucidate the mechanisms responsible for the phenotypes observed. Additionally, further mass spectrometry experiments are required to determine whether acetylation and monoubiquitination of H4K91 are present in *A. thaliana*. It is interesting that the rH4K91R and rH4K91A mutants exhibit different phenotypes, most notably in their DNA content. One explanation for this observation could be that the H4K91A mutation neutralizes the positive charge on lysine 91, partially mimicking an acetylated residue, while the H4K91R mutation does not. These differences may impact the chromatin assembly and other genomic processes in these mutants in distinct manners, with subsequent effects on the cell cycle.

Similar to H4K91, H4Y98 has also previously been demonstrated to play roles in regulating genome stability in yeast; however, H4Y98 has not been identified as a target for any PTMs. H4Y98 yeast mutants exhibit defects in chromosome segregation, poor growth, and a rapid onset of polyploidy/aneuploidy, likely due to an impact on kinetochore function (Ng et al., 2013; Yu et al., 2011c). Additionally, H4Y98A yeast mutants exhibit hypersensitivity to DNA damaging agents (Matsubara et al., 2007). Further work assessing *BRCA1* RNA expression and DNA damage sensitivity in *Arabidopsis thaliana* rH4Y98F mutants would elucidate whether this role for H4Y98 is

conserved in yeast and plants. While the mechanism responsible for these phenotypes in the yeast H4Y98 mutants is unknown, H4Y98 does appear to play an important role in regulating nucleosome stability, and in fact, H4Y98H mutations destabilize the histone octamer, potentially through their impact on the interaction of H4Y98 with the H2A-H2B dimer (Kitevski-LeBlanc et al., 2018; Xu et al., 2005). Therefore, the effect of both the H4K91 and H4Y98 mutations in *Arabidopsis thaliana* may lead to the observed cell cycle/endoreduplication defects due to their impact on chromatin assembly and nucleosome stability. As noted above, mutations in genes encoding CAF-1 subunits lead to increased proportions of higher ploidy nuclei in *Arabidopsis thaliana* and thus, there appears to be a relationship between chromatin assembly and endoreduplication in plants.

Perspectives

As we identified many novel functions for histone H4 residues with our systematic screen of histone H4 in *Arabidopsis thaliana*, substantial opportunities remain for future work characterizing the mechanisms responsible for the phenotypes observed (see above). In the next chapter, we will discuss our experiments to further characterize the rH4R17A mutant plants. These mutants were chosen for further work due to the relatively uncharacterized role for H4R17 in plants, as well as the variety of interesting morphological and molecular phenotypes displayed, including a novel role for H4R17 in the regulation of flowering time. Importantly, H4R17 is a conserved residue in plants, animals, and yeast located within the N-terminal tail (Figure 2.1) and mutations of H4R17 are linked to several types of human cancers (Bennett et al., 2019; Nacev et al., 2019). While we focus on the H4R17 residue for the remainder of this dissertation, following up with other residues identified in our screens remains an exciting prospective direction for this research. In summary, our histone H4 replacement system has allowed

the rapid assessment of the impact of numerous histone mutations on important developmental and genomic processes to open up many avenues for future research on the function of histone H4 in plants.

Chapter 4: Regulation of ISWI chromatin remodeling by histone H4 arginine 17

Author's Note: Portions of this chapter have been submitted for publication as:

Emma Tung Corcoran¹, Chantal LeBlanc¹, Mia Arias Tsang¹, Anthony Sarkiss¹, Yuzhao Hu², Ullas V. Pedmale², and Yannick Jacob¹. 2021. Systematic histone H4 replacement in *Arabidopsis thaliana* reveals a role for H4R17 in regulating flowering time.

Affiliations:

1. Yale University, Department of Molecular, Cellular and Developmental Biology, Faculty of Arts and Sciences; 260 Whitney Avenue, New Haven, Connecticut 06511, United States.
2. Cold Spring Harbor Laboratory; Cold Spring Harbor, New York 11724, United States.

Overview

After our comprehensive screens of phenotypes induced by histone H4 mutations, we chose to further characterize the phenotype of the rH4R17A mutant in order to propose a model responsible for its early flowering phenotype. We first generated rH4 and rH4R17A lines for which there was a complete replacement of endogenous H4 with replacement H4 and assessed the dominance of the H4R17A mutation to demonstrate that the H4R17A mutation shows incomplete dominance. We then investigated two alternative hypotheses to explain the rH4R17A mutant phenotype: (1) that loss of methylation on H4R17 induces the observed phenotypes and (2) that impaired regulation of ISWI chromatin remodelers by the H4R17A mutant histone induces the observed phenotypes. Through phenotypic analysis, flowering time assessment, and assays of global gene expression and nucleosome positioning, we found evidence of a functional relationship between ISWI components and H4R17, supporting the second model. Our work is the first to establish a regulatory relationship between H4R17 and ISWI chromatin remodeling in plants to support a conserved role for H4R17 in the regulation of ISWI chromatin remodeling complexes across eukaryotic species.

Results

Generation of complete H4R17A replacement backgrounds

Among the five H4 mutations (H4R17A, H4R36A, H4R39K, H4R39A, and H4K44A) identified in our screen to cause the strongest accelerating effect on flowering time, only one of them (H4R17A) is present in the N-terminal tail of H4 where most of the histone PTMs are made (Figure 3.6). Mutations in the unstructured N-terminal tail of H4 are less likely to affect flowering time by disrupting histone H4 folding and/or nucleosome structure than mutations in the histone-fold domain. Therefore, we decided

to focus our subsequent analyses on elucidating the mechanism by which H4R17A affects the timing of the transition to reproductive development. For this work, we used H4 replacement plants (rH4 and rH4R17A) for which there was a complete replacement of endogenous histone H4 with H4 expressed from the replacement H4 transgene. Consequently, we genotyped leaf tissue from several transgenic lines to identify two independent transgenic lines (rH4-1, rH4-2, rH4R17A-1, and rH4R17A-2) with homozygous or biallelic mutations in the remaining endogenous H4 gene, and thus, a complete replacement of endogenous H4 (Figure 4.1A, Table 4.1). We also genotyped leaf tissue from the progeny of these plants to confirm the presence of the same mutations in the next generation. As observed previously, the complete rH4R17A replacement lines both exhibited multiple developmental phenotypes including smaller and upwardly curled leaves, reduced fertility, and a significantly early floral transition (Figure 4.1B-D, Figure 4.2). Moreover, we found that the rH4R17A-2 mutant line displayed a more severe early flowering phenotype and more severe developmental phenotypes compared to the rH4R17A-1 mutant line (Figure 4.1B-D, Figure 4.2B-C).

Dominance of H4R17A mutation

In order to gain more insight into the mechanism of the rH4R17A mutant phenotype, we assessed the dominance of the H4R17A mutation. We generated mutants individually expressing the H4R17A mutation under the native H4 promoter in a wild-type background and assessed the resultant phenotypes of first-generation transgenic (i.e., *pH4::H4R17A* T1) plants compared to plants with a complete replacement of endogenous histone H4 with the H4R17A mutant histone (i.e., rH4R17A-1 and rH4R17A-2 plants). We also generated plants individually expressing wild-type H4 under the native H4 promoter in a wild-type background (i.e., *pH4::H4* T1) to use as a control for the experiment. We observed that the *pH4::H4* T1 plants displayed a wild-

A

rH4-1: homozygous 1 nucleotide (A) insertion

```

At3g53730 wild-type 24 --AAAAGGATTAGGAAAGGGAGGAGCCAAGAGA--CATCGGAAAGTACTC-- 69
                        guide RNA  AGGGAGGAGCCAAGAGA--CAT
At3g53730 mutant 24 --AAAAGGATTAGGAAAGGGAGGAGCCAAGAGACATCGGAAAGTACTC-- 70

```

rH4-2: biallelic 1 nucleotide (A) insertion/ 1 nucleotide (C) deletion

```

At3g53730 wild-type 24 --AAAAGGATTAGGAAAGGGAGGAGCCAAGAGA--CATCGGAAAGTACTC-- 69
                        guide RNA  AGGGAGGAGCCAAGAGA--CAT
At3g53730 mutant allele 1 24 --AAAAGGATTAGGAAAGGGAGGAGCCAAGAGACATCGGAAAGTACTC-- 70
At3g53730 mutant allele 2 24 --AAAAGGATTAGGAAAGGGAGGAGCCAAGAGAGAAA---ATCGGAAAGTACTC-- 68

```

rH4R17A-1: homozygous 20 nucleotide deletion

```

At3g53730 wild-type 24 --AAAAGGATTAGGAAAGGGAGGAGCCAAGAGACATCGGAAAGTACTC-- 69
                        guide RNA  AGGGAGGAGCCAAGAGACAT
At3g53730 mutant 24 --AAAAGGATTAGGAAAGGGAGGAGCCAAGAGACA----- 56

```

rH4R17A-2: homozygous 20 nucleotide deletion

```

At3g53730 wild-type 24 --AAAAGGATTAGGAAAGGGAGGAGCCAAGAGACATCGGAAAGTACTC-- 69
                        guide RNA  AGGGAGGAGCCAAGAGACAT
At3g53730 mutant 24 --AAAAGGATTAGGAAAGGGAGGAGCCAAGAGACA----- 56

```

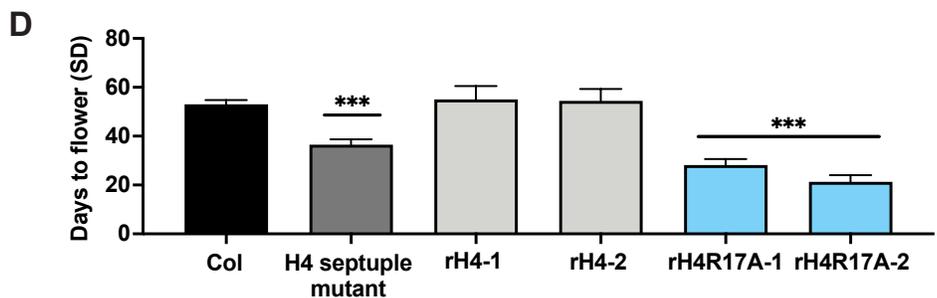
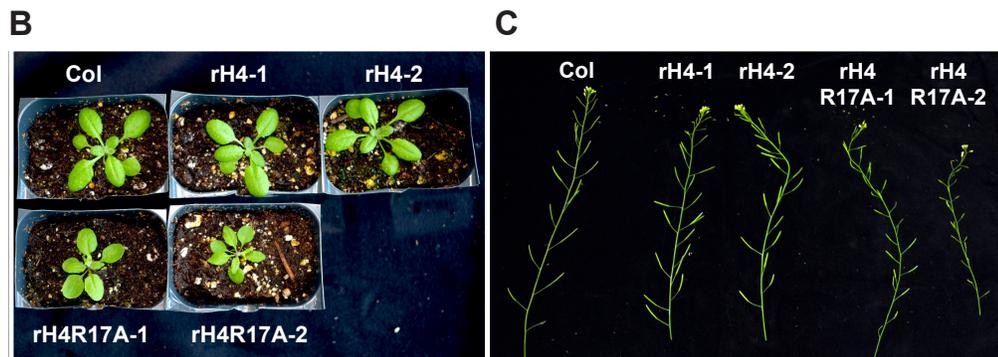


Figure 4.1 Generation of complete H4 replacement lines. (A) Homozygous or biallelic mutation in the remaining endogenous histone H4 gene (*At3g53730*) in rH4-1, rH4-2, rH4R17A-1 and rH4R17A-2 plants. (B) Rosette phenotype of Col, rH4-1, rH4-2, rH4R17A-1 and rH4R17A-2 plants grown in long-day conditions at 3 weeks. (C) Siliques of Col, rH4-1, rH4-2, rH4R17A-1 and rH4R17A-2 plants grown in long-day conditions at 4 weeks. (d) Mean days to flower in short-day conditions (SD) for Col, H4 septuple mutant, rH4-1, rH4-2, rH4R17A-1, and rH4R17A-2 plants. Standard deviation shown with error bars. Statistical analyses were performed using one-way ANOVA with Tukey's HSD post hoc test. *P*-value from Tukey's HSD test (genotype vs. Col) denoted with asterisks (**p*<0.05, ***p*<0.005, ****p*<0.0005). *n*≥6.

Genotype	Effect of mutation in endogenous <i>At3g53730</i> on H4 mRNA
rH4-1	54 nucleotides from H4 5' CDS fused to 101 nucleotides from H4 CDS (after frameshift mutation due to 1 nucleotide insertion), terminating in premature termination codon
rH4-2	<u>Allele 1</u> : 54 nucleotides from H4 5' CDS fused to 101 nucleotides from H4 CDS (after frameshift mutation due to 1 nucleotide insertion), terminating in premature termination codon <u>Allele 2</u> : 54 nucleotides from H4 5' CDS fused to 75 nucleotides from H4 CDS (after frameshift mutation due to 1 nucleotide deletion), terminating in premature termination codon
rH4R17A-1	56 nucleotides from H4 5' CDS fused to 79 nucleotides from H4 CDS (after frameshift mutation due to 20 nucleotide deletion), terminating in premature termination codon
rH4R17A-2	56 nucleotides from H4 5' CDS fused to 79 nucleotides from H4 CDS (after frameshift mutation due to 20 nucleotide deletion), terminating in premature termination codon

Table 4.1 Effect of H4 (*At3g53730*) mutations on mRNA.

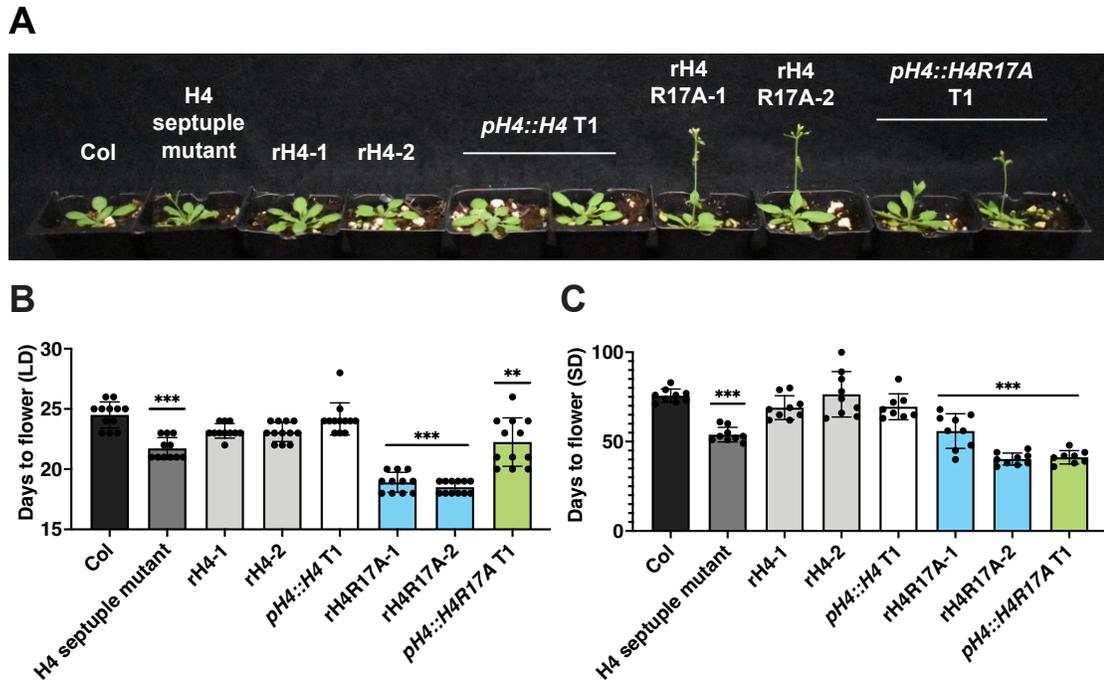


Figure 4.2 Phenotypic analyses of H4R17A mutants. (A) Phenotype of Col, H4 septuple mutant, rH4-1, rH4-2, *pH4::H4* T1 (2 individual plants shown), rH4R17A-1, rH4R17A-2, and *pH4::H4R17A* T1 (2 individual plants shown) plants grown in long-day conditions for 21 days. (B-C) Mean days to flower in (B) long-day conditions (LD) and (C) short-day conditions (SD) for Col, H4 septuple mutant, rH4-1, rH4-2, *pH4::H4* T1, rH4R17A-1, rH4R17A-2, and *pH4::H4R17A* T1 plants. Standard deviation shown with error bars. Statistical analyses were performed using one-way ANOVA with Tukey's HSD post hoc test. *P*-value from Tukey's HSD test (genotype vs. Col) denoted with asterisks (**p*<0.01, ***p*<0.001, ****p*<0.0001). *n*≥11 for long-day, *n*≥5 for short-day.

type flowering response and normal morphological phenotypes (Figure 4.2).

Additionally, we found that *pH4::H4R17A* T1 plants displayed a similar early flowering response to *rH4R17A* plants when grown in short-day conditions (Figure 4.2C). In contrast, when grown in long-day conditions, some individual *pH4::H4R17A* T1 plants displayed a wild-type floral transition, while other individual *pH4::H4R17A* T1 plants displayed an early flowering response, though less severe than that of most of the *rH4R17A* plants (Figure 4.2B). Additionally, *pH4::H4R17A* T1 plants displayed a similar rosette phenotype (smaller and upwardly curled leaves) to *rH4R17A* mutant plants (Figure 4.2A). From these results, the *H4R17A* mutation appears to function via incomplete dominance.

No relationship between PRMT7 and H4R17 in plants

One hypothesis regarding the mechanism by which the *H4R17A* mutation causes early flowering is that it prevents deposition of a post-translational modification on the *H4R17* residue. PROTEIN ARGININE METHYLTRANSFERASE 7 (*PRMT7*) is the only known modifier of *H4R17* in eukaryotes, as it has been shown to monomethylate R17 on histone H4 in mammals (Feng et al., 2014b; Feng et al., 2013; Jain and Clarke, 2019). However, no PTMs on *H4R17* have been identified in plants (Brabencova et al., 2017; Zhang et al., 2007a).

The *A. thaliana* genome contains a single orthologous gene for *PRMT7* (*At4g16570*), which has never been functionally characterized. To assess a potential role for *PRMT7* in regulating flowering time via methylation of *H4R17*, we measured flowering time in *prmt7* mutants obtained from the Salk T-DNA collection (SALK_028160 and SALK_039529). Moreover, we also generated plants overexpressing the *PRMT7* gene under the cauliflower mosaic virus promoter (i.e., *35S::PRMT7*) in order to assess flowering time and developmental phenotypes (Fromm et al., 1985). We confirmed by

RT-qPCR that both T-DNA alleles used in these experiments prevent the expression of a full-length *PRMT7* transcript and that *PRMT7* was overexpressed in the 35S::*PRMT7* T1 plants that we generated (Figure 4.3). Our analyses of flowering time caused by modulation of the *PRMT7* gene in plants showed that neither *prmt7* mutants nor *PRMT7* overexpressing plants affect flowering time in either long-day conditions or short-day conditions (Figure 4.4). In addition, none of the other vegetative or reproductive phenotypes observed in rH4R17A plants were found in plants lacking or overexpressing *PRMT7*. These results strongly suggest that replacement of H4 with H4R17A does not affect development in *A. thaliana* by interfering with PRMT7 activity on histone H4.

Functional relationship between H4R17 and ISWI in the regulation of flowering

In addition to affecting the deposition of histone PTMs, mutation of histone residues can prevent binding of proteins to chromatin. Therefore, we next investigated the possibility that replacement of histone H4 with H4R17A affects plant development by negatively impacting the function of plant ISWI chromatin remodeling complexes. In yeast and animals, R17 of H4 has been shown to directly interact with ISWI to regulate nucleosome remodeling activity *in vitro* and *in vivo* (Clapier et al., 2001; Clapier et al., 2002; Dann et al., 2017; Fazzio et al., 2005; Hamiche et al., 2001; Ludwigsen et al., 2017; Mueller-Planitz et al., 2013; Racki et al., 2014; Yan et al., 2016). Notably, mutations in genes coding for different *A. thaliana* ISWI subunits (CHR11, CHR17, RLT1, RLT2 and ARID5) result in plants showing similar phenotypes to rH4R17A mutants, including early flowering, upwardly curled leaves, reduced fertility, and a small size relative to wild-type plants (Figure 4.5, Figure 4.6) (Li et al., 2012). Defects in the timing of the floral transition and other developmental aspects are more similar between the rH4R17A mutants and the ISWI accessory subunit single mutant *arid5* and double mutant *rlt1 rlt2* (*rlt1/2*; *RLT1* and *RLT2* were shown to act redundantly (Li et al., 2012))

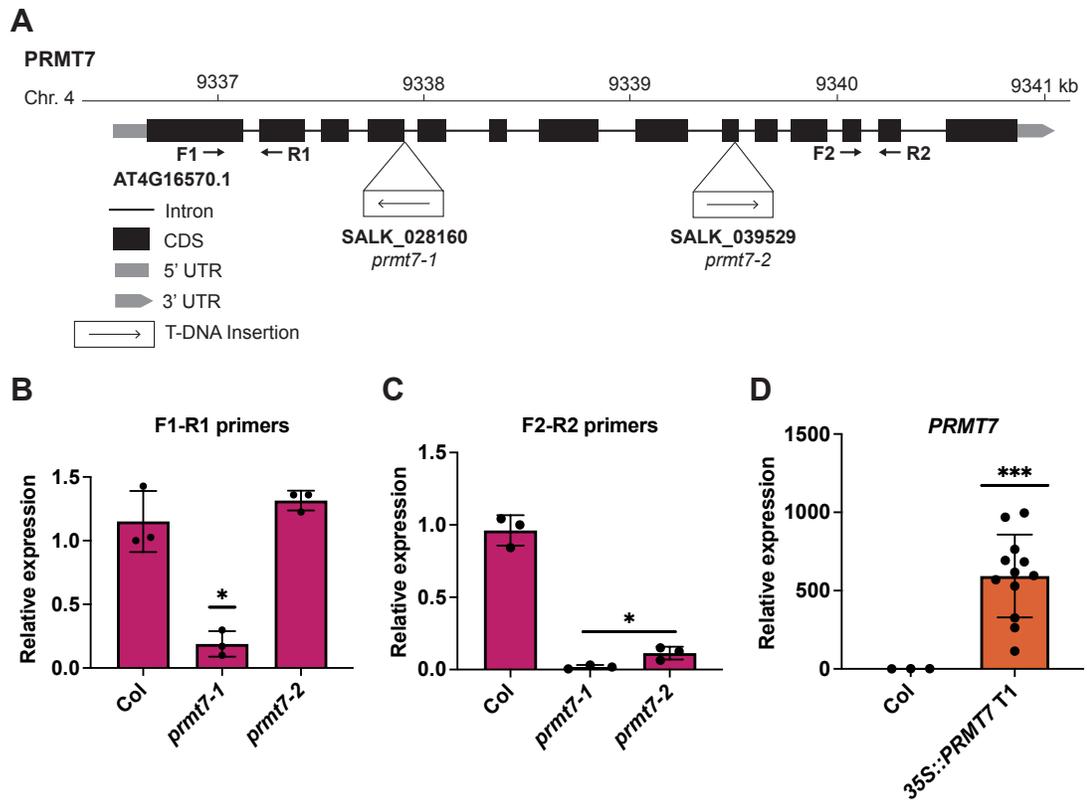


Figure 4.3 Validation of *PRMT7* mutations. (A) Gene structure of *PRMT7*. The location of the T-DNA insertions and the primers (F1-R1 and F2-R2) used for gene expression analyses are shown. (B-D) RT-qPCR showing *PRMT7* expression in (B-C) Col, *prmt7-1*, and *prmt7-2* plants and (D) Col and 35S::*PRMT7* T1 plants. The average of three biological replicates and standard deviation are shown for Col and *prmt7* mutants. For the 35S::*PRMT7* plants, individual data points represent independent T1 plants. *P*-value from unpaired Student's *t*-test (sample vs. Col) denoted with asterisks (* $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$).

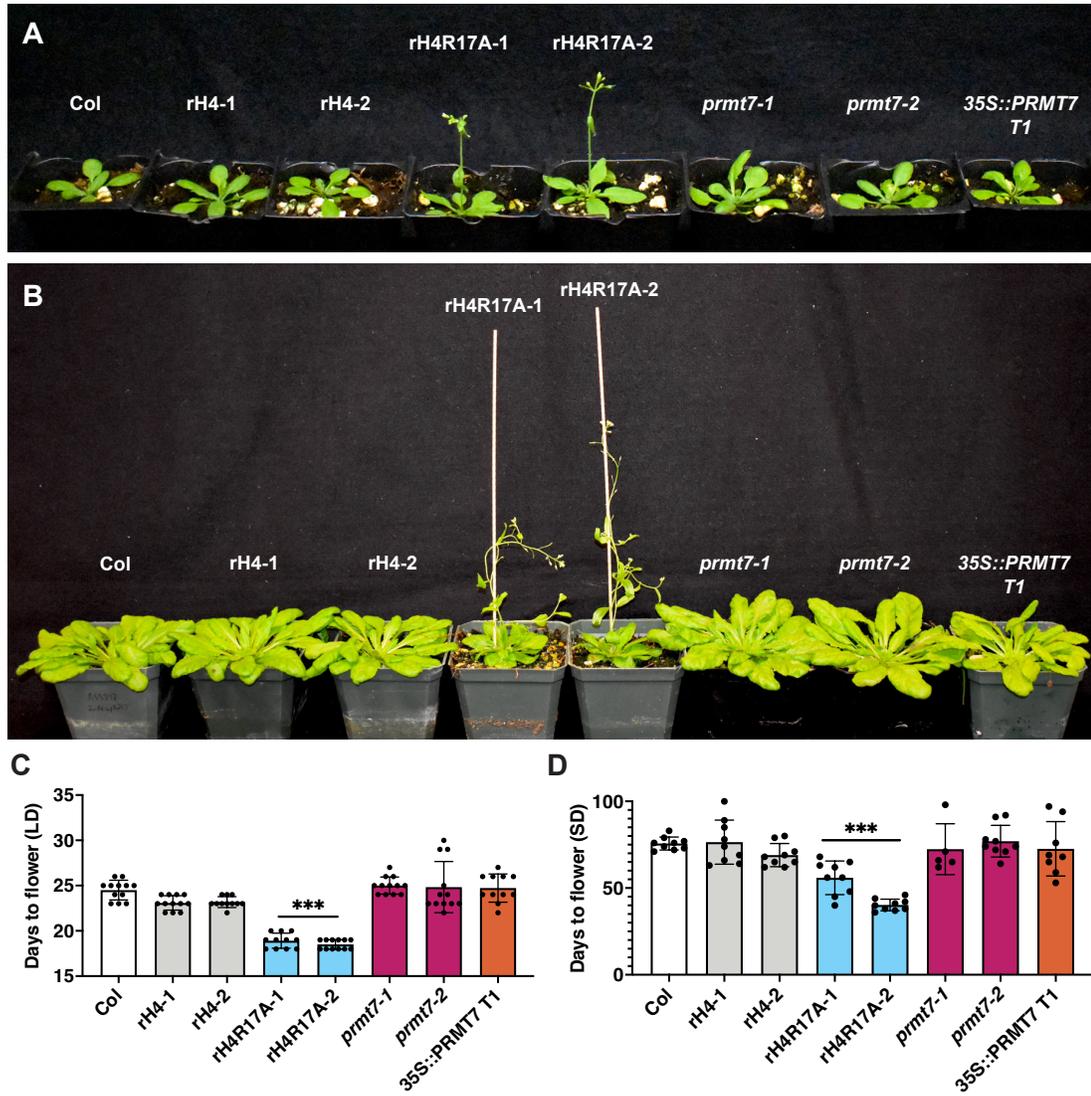


Figure 4.4 The effect of *PRMT7* mutations on the floral transition. (A-B) Phenotype of Col, rH4-1, rH4-2, rH4R17A-1, rH4R17A-2, *prmt7-1*, *prmt7-2*, and 35S::*PRMT7* T1 plants grown in (A) long-day conditions at 3 weeks and (B) short-day conditions at 8 weeks. (C-D) Mean days to flower in (C) long-day conditions (LD) and (D) short-day conditions (SD) for Col, rH4-1, rH4-2, rH4R17A-1, rH4R17A-2, *prmt7-1*, *prmt7-2*, and 35S::*PRMT7* T1 plants. Standard deviation denoted with error bars. Statistical analyses were performed using one-way ANOVA with Tukey's HSD post hoc test. *P*-value from Tukey's HSD test (genotype vs. Col) denoted with asterisks (**p*<0.05, ***p*<0.005, ****p*<0.0005). *n*≥11 for long-day, *n*≥5 for short-day.

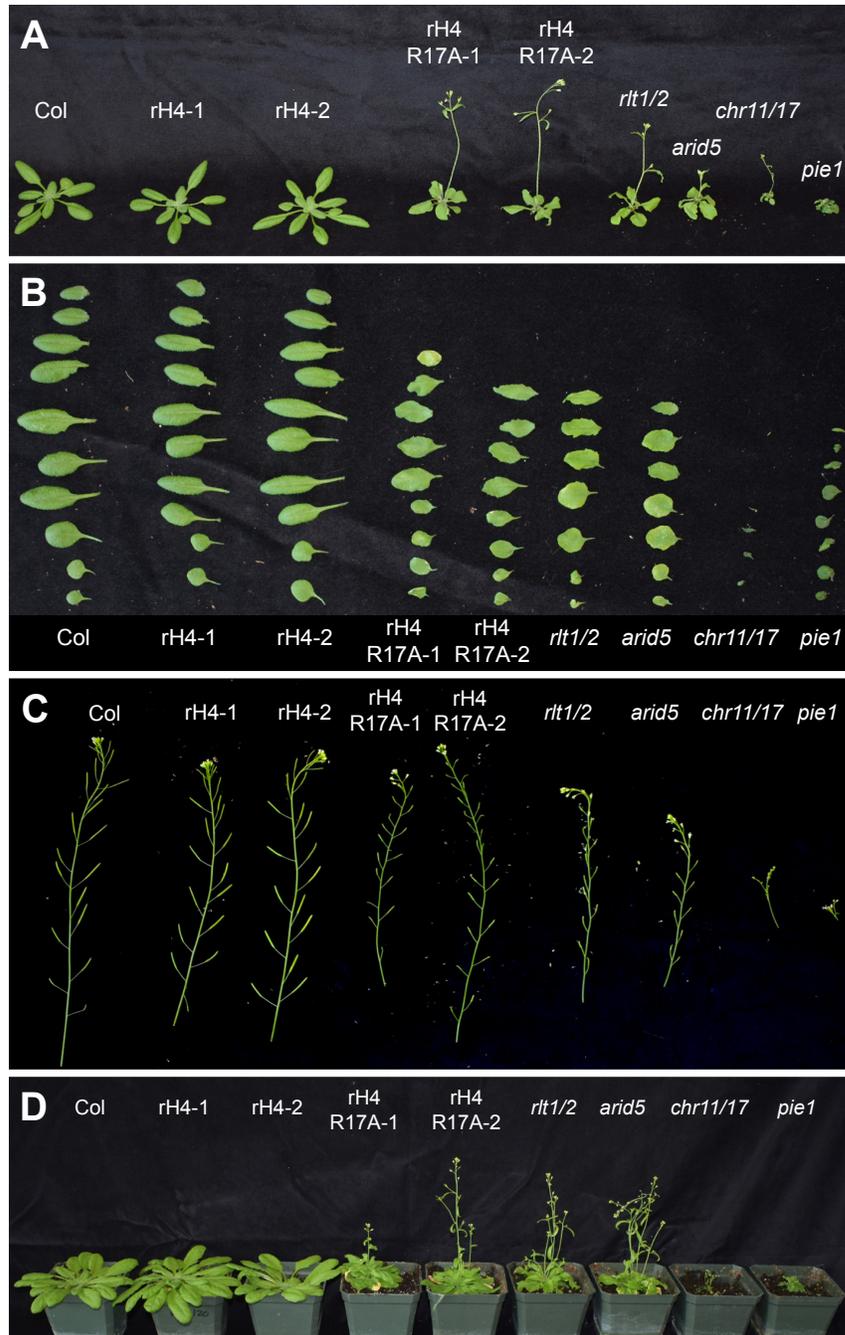


Figure 4.5 The effect of ISWI and rH4R17A mutations on the floral transition and development. (A-C) (A) 21-day-old morphological phenotype, (B) rosette leaf phenotype, and (C) silique phenotype of Col, rH4-1, rH4-2, rH4R17A-1, rH4R17A-2, *rlt1/2*, *arid5*, *chr11/17*, and *pie1* plants grown in long-day conditions. Rosette leaves were cut from plants shortly after bolting. (D) Morphological phenotypes of Col, rH4-1, rH4-2, rH4R17A-1, rH4R17A-2, *rlt1/2*, *arid5*, *chr11/17*, and *pie1* plants grown in short-day conditions at 7 weeks.

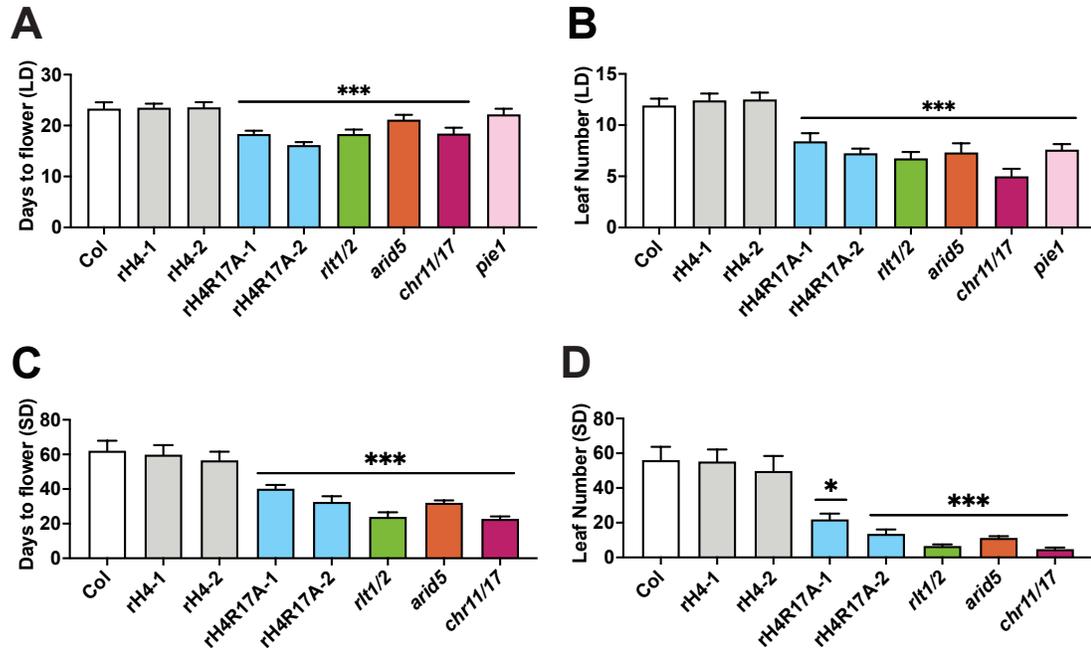


Figure 4.6 ISWI and rH4R17A mutations cause early flowering. (A-B) Mean (A) days to flower and (B) rosette leaf number at flowering in long-day conditions (LD) for Col, rH4-1, rH4-2, rH4R17A-1, rH4R17A-2, *rlt1/2*, *arid5*, *chr11/17*, and *pie1* plants. (C-D) Mean (C) days to flower and (D) rosette leaf number at flowering in short-day conditions (SD) for Col, rH4-1, rH4-2, rH4R17A-1, rH4R17A-2, *rlt1/2*, *arid5*, and *chr11/17* plants. Standard deviation shown with error bars. Statistical analyses were performed using one-way ANOVA with Tukey's HSD post hoc test. *P*-value from Tukey's HSD test (genotype vs. Col) denoted with asterisks (* $p < 0.01$, ** $p < 0.001$, *** $p < 0.0001$). $n = 12$ for all lines except *pie1* ($n \geq 5$).

compared to mutations in the ISWI catalytic subunits *CHR11* and *CHR17* (*CHR11/17*; also shown to act redundantly (Li et al., 2012)), which cause more severe developmental phenotypes (Figure 4.5, Figure 4.6) (Li et al., 2012; Tan et al., 2020). The chromatin remodelers *CHR11/17* are core components of ISWI in plants (Knizewski et al., 2008; Li et al., 2014; Li et al., 2012), but are also present in the plant SWR1 complex responsible for exchanging H2A-H2B dimers with H2A.Z-H2B (Luo et al., 2020). Therefore, the increased severity of the phenotypes displayed by the *chr11/17* double mutant may be caused by the joint disruption of the ISWI and SWR1 chromatin remodeling complexes, which both contain *CHR11/17* (Luo et al., 2020). In contrast, *ARID5* and *RLT1/2* are present in ISWI, but not in SWR1. In addition, *RLT1* and *RLT2* are only two of 12 DDT-domain proteins in *A. thaliana* and different ISWI complexes were found to associate with different DDT-domain proteins *in vivo*, including DDT-PHD PROTEIN 1 (DDP1), DDP2, DDP3, DDT-RELATED PROTEIN 1 (DDR1), DDR3, DDR4, DDR5, and DDT-WAC PROTEIN 1 (DDW1) (Dong et al., 2013; Tan et al., 2020).

To further investigate the interplay in plants between *H4R17* and ISWI, we performed RNA-seq on the *rH4R17A*, *arid5*, *rtl1/2*, *chr11/17* and *pie1* (catalytic subunit of the SWR1 complex) mutants grown in short-day conditions. We confirmed RNA quality through analysis of Agilent Bioanalyzer 2100 electropherograms (Figure 4.7). Furthermore, we utilized two biological replicates corresponding to each genotype and confirmed consistency between biological replicates through Spearman correlation coefficient analysis (Figure 4.8). With our RNA-seq data, we first investigated histone expression levels in our replacement H4 lines. We observed an approximately 3- to 5-fold increase in H4 (*At3g53730*) RNA expression when comparing *rH4-1*, *rH4-2*, *rH4R17A-1* and *rH4R17A-2* lines to Col plants (Figure 4.9A). For the remaining seven endogenous H4 genes, we observed that *rH4-1* and *rH4-2* plants displayed approximately wild-type RNA levels of *At3g46320*, while the other six H4 genes

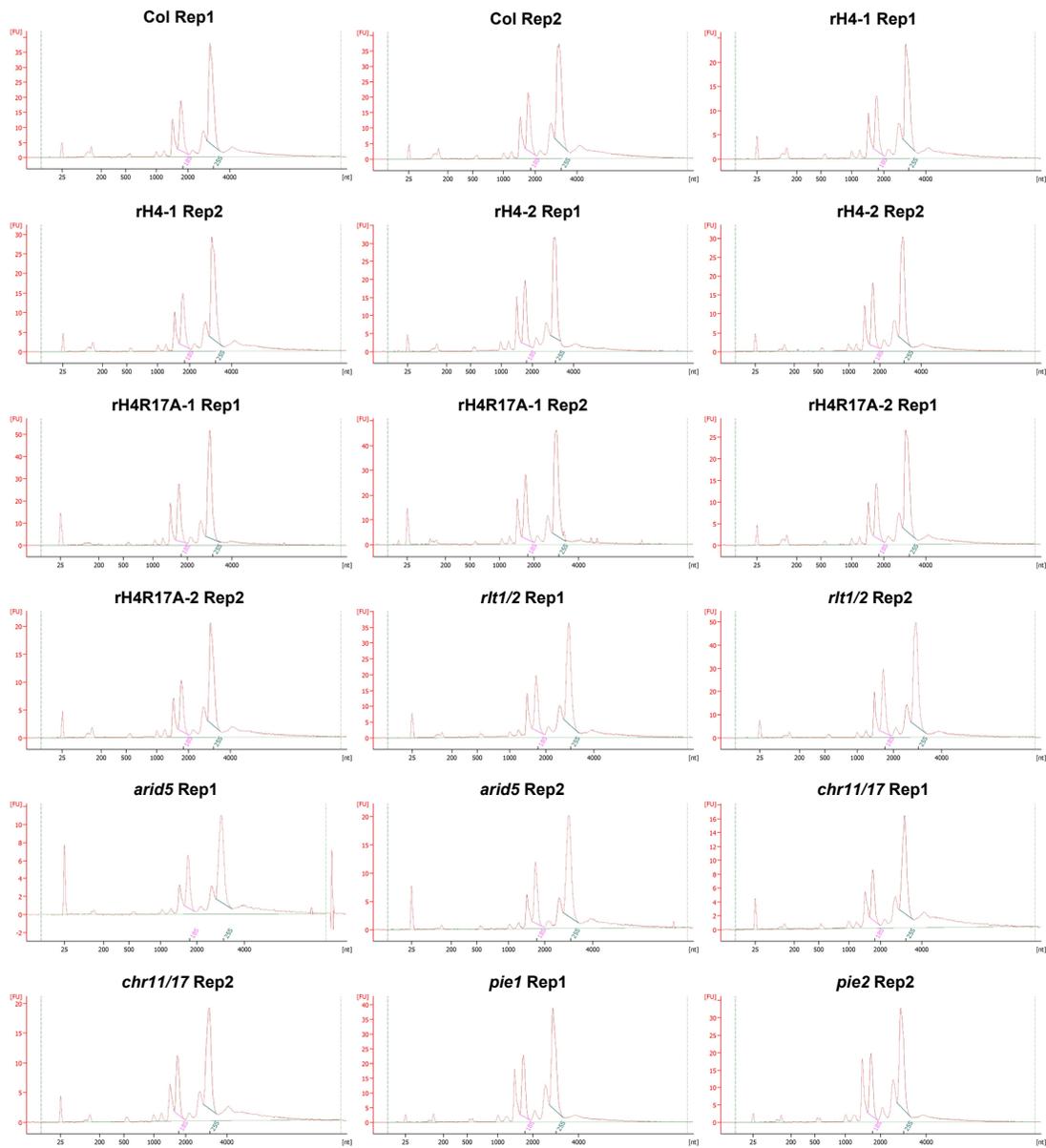


Figure 4.7 Bioanalyzer electropherograms of RNA-seq replicates. Agilent Bioanalyzer 2100 electropherograms for RNA-seq replicates of Col, rH4-1, rH4-2, rH4R17A-1, rH4R17A-2, *rtt1/2*, *arid5*, *chr11/17*, and *pie1*.

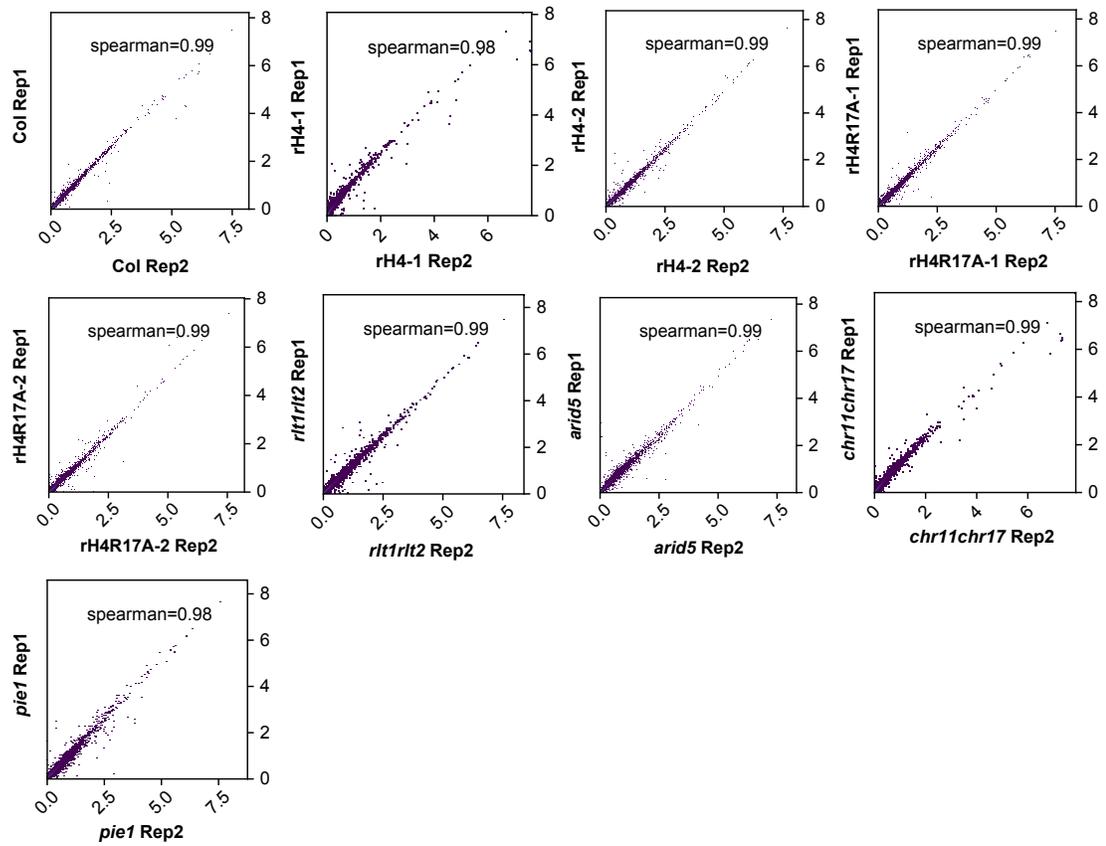


Figure 4.8 Spearman correlation of RNA-seq replicates. Spearman correlation coefficient analysis for RNA-seq replicates of Col, rH4-1, rH4-2, rH4R17A-1, rH4R17A-2, *rtt1/2*, *arid5*, *chr11/17*, and *pie1*.

displayed an approximately 25 to 50% reduction in RNA expression compared to Col (Figure 4.9A). We also investigated the expression of H2A, H2B, and H3 genes in replacement H4 plants and found that rH4-1 and rH4-2 plants displayed approximately wild-type levels of H2A, H2B, and H3 RNA (Figure 4.10). In contrast, rH4R17A-1 and rH4R17A-2 plants showed an upregulation of certain histone genes, with H3 genes showing the most consistent upregulation among rH4R17A mutants.

We then calculated the percentage of reads aligning to the *At3g53730* locus that were expressed from the replacement H4 gene versus the endogenous H4 gene. We selected the gRNA-targeting region of the *At3g53730* locus (37 bp after start codon to 59 bp after start codon) for this analysis due to the single nucleotide polymorphisms present in this region that could be used to distinguish endogenous and replacement H4 RNA. We found that 47 to 76% of the observed reads were expressed from the replacement H4 gene (Figure 4.9B) and thus, approximately 24 to 53% of the *At3g53730* RNA expression observed in these lines corresponds to non-functional H4 mRNA. In order to estimate the amount of functional H4 mRNA in Col plants compared to replacement H4 plants, we summed the reads mapping to all eight endogenous H4 genes in Col and multiplied the reads mapping to *At3g53730* in replacement H4 lines by the percentage of replacement H4 observed (Figure 4.9C). We determined that the estimated functional H4 mRNA levels in replacement H4 lines ranged from around 34 to 54% of the total sum of H4 RNA reads observed in Col.

Next, we proceeded to compare differential gene expression in rH4R17A and ISWI subunit mutants. Our results from the RNA-seq analyses showed that there were 3045 downregulated genes and 2476 upregulated genes in *chr11/17* double mutants (5521 differently expressed genes [DEGs] in total), while there were only 877 downregulated genes and 598 upregulated genes in the rH4R17A-1 mutant (1475 DEGs), and 724 downregulated genes and 685 upregulated genes in the rH4R17A-2

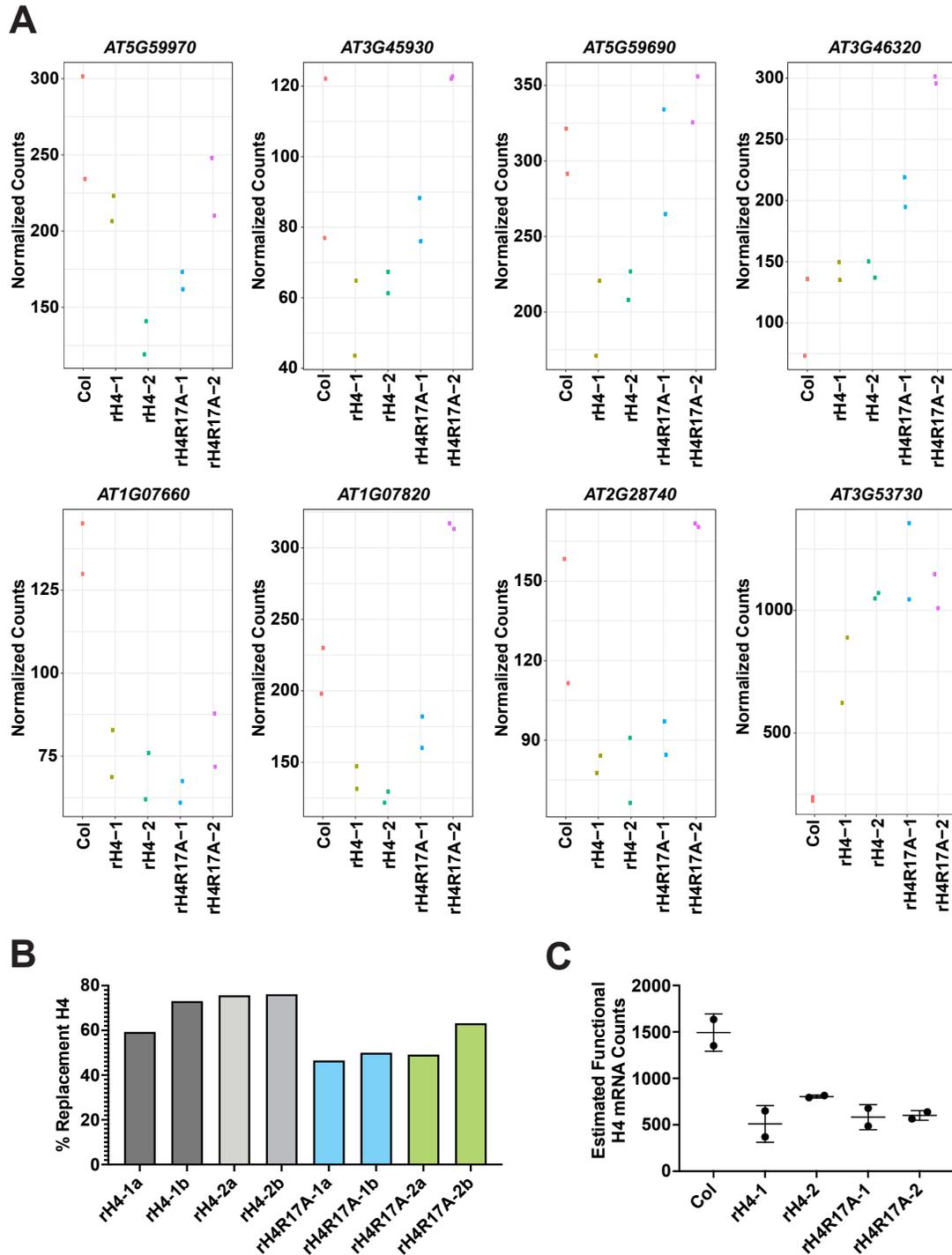


Figure 4.9 H4 expression levels in rH4 lines. (A) Normalized read counts at endogenous H4 genes (*AT5G59970*, *AT3G45930*, *AT5G59690*, *AT3G46320*, *AT1G07660*, *AT1G07820*, *AT2G28740*, *AT3G53730*) in Col, rH4-1, rH4-2, rH4R17A-1, and rH4R17A-2 plants. (B) Percentage replacement H4 vs. endogenous H4 (*At3g53730*) mRNA expressed in biological replicates of rH4-1, rH4-2, rH4R17A-1, and rH4R17A-2 plants. (C) Estimated functional H4 mRNA read counts in Col, rH4-1, rH4-2, rH4R17A-1, and rH4R17A-2 plants.

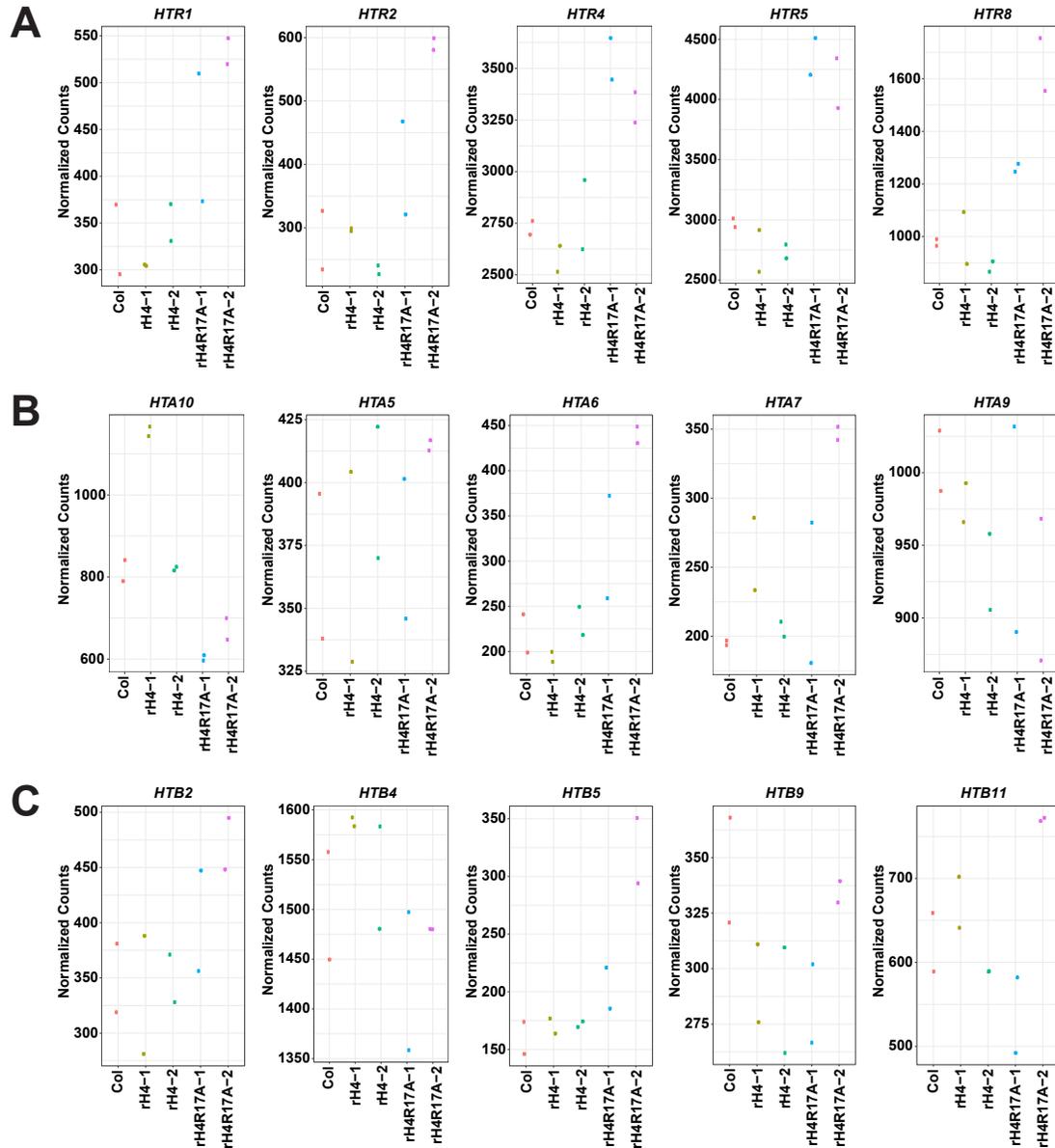


Figure 4.10 Expression levels of most highly expressed H2A, H2B, and H3 genes in rH4 lines. Normalized read counts at (A) *AT5G65360* (*HTR1*, H3.1), *AT1G09200* (*HTR2*, H3.1), *AT4G40030* (*HTR4*, H3.3), *AT4G40040* (*HTR5*, H3.3), *AT5G10980* (*HTR8*, H3.3), (B) *AT1G51060* (*HTA10*, H2A), *AT1G08880* (*HTA5*, H2A.X), *AT5G59870* (*HTA6*, H2A.W), *AT5G27670* (*HTA7*, H2A.Z), *AT1G52740* (*HTA9*, H2A.Z), and (C) *AT5G22880* (*HTB2*, H2B), *AT5G59910* (*HTB4*, H2B), *AT2G37470* (*HTB5*, H2B), *AT3G45980* (*HTB9*, H2B), and *AT3G46030* (*HTB11*, H2B) in Col, rH4-1, rH4-2, rH4R17A-1, and rH4R17A-2 plants.

mutant (1409 DEGs) (Figure 4.11). In spite of the large differences in the total amounts of DEGs between *chr11/17* mutants and the rH4R17A plants, we observed a high overlap between the DEGs in the *chr11/17* and rH4R17A-1 mutants (43.5%, 641/1475), as well as the DEGs in the *chr11/17* and rH4R17A-2 mutants (53.7%, 757/1409) (Figure 4.11). In the rH4R17A plants and *chr11/17* mutants, we detected a highly similar pattern of RNA expression not shared by rH4 or Col plants (Figure 4.12). Moreover, correlation analyses revealed a high correlation between the *chr11/17* and rH4R17A-1 mutants (0.69, Pearson's correlation) and the *chr11/17* and rH4R17A-2 mutants (0.74, Pearson's correlation) that was not displayed when comparing the *chr11/17* and rH4 backgrounds (Figure 4.12).

Additionally, we observed a high overlap of DEGs (average 49.5% rH4R17A vs. *arid5*; average 52.9% rH4R17A vs. *rtt1/2*), a similar pattern of RNA expression, and a high correlation (average 0.62 rH4R17A vs. *arid5*; average 0.66 rH4R17A vs. *rtt1/2*) when comparing the rH4R17A lines with the ISWI subunit mutants *arid5* and *rtt1/2* (Figure 4.11, Figure 4.12). In contrast, we did not observe substantial overlap (average 25.7%) or correlation (average 0.17, Pearson's correlation) between the DEGs identified in *pie1* mutants and the DEGs of rH4R17A mutants (Figure 4.11, Figure 4.12). We performed principal component analysis of the 200 most variable differentially expressed genes identified in the rH4R17A mutants and took the first two principal components, which explained 60% of the variance in the data (Figure 4.13). We found that Col, rH4-1, and rH4-2 plants clustered closely within the same region of the principal component plot and the remaining genotypes were separated from this cluster. Additionally, the *arid5*, *rtt1/2*, *chr11/17*, rH4R17A-1 and rH4R17A-2 mutants clustered within the same general region of the embedding (with rH4R17A-1 and rH4R17A-2 clustering closely together and *arid5* and *rtt1/2* clustering closely together), while the *pie1* mutants were separated from these five genotypes. From these results, we concluded that ISWI subunit mutants

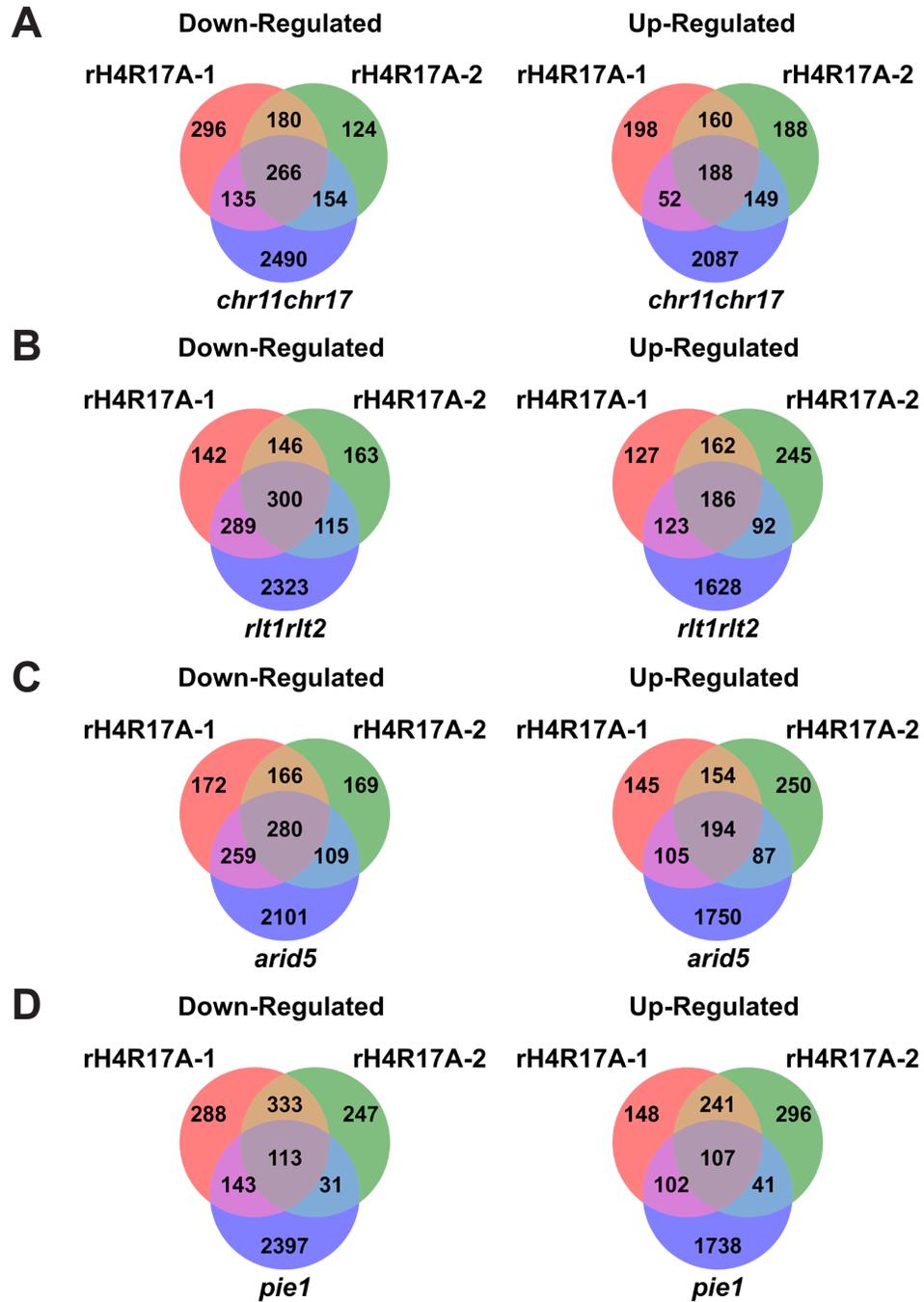


Figure 4.11 Overlap of differentially expressed genes in rH4R17A, ISWI and SWR1 mutants. Venn diagrams showing DEGs (relative to Col) identified by RNA-seq in (A) rH4R17A and *chr11/17* mutants, (B) rH4R17A and *rlt1/2* mutants, (C) rH4R17A and *arid5* mutants, and (D) rH4R17A and *pie1* mutants.

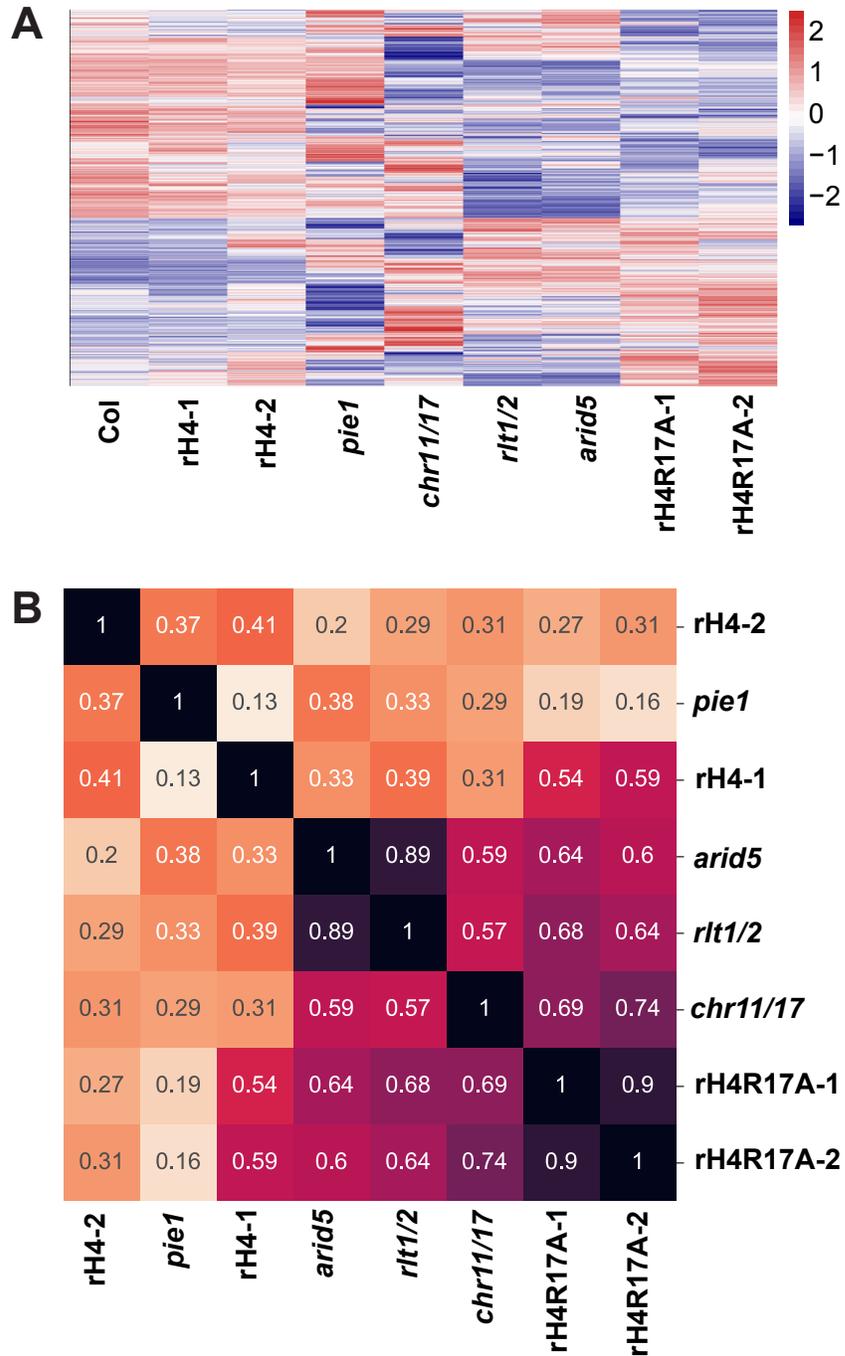


Figure 4.12 Patterns of differential gene expression in rH4R17A, ISWI and SWR1 mutants. (A) Heatmap of relative expression patterns of DEGs identified in the rH4R17A mutant. Red and blue represent up- and downregulated genes, respectively. (B) Pearson's correlation matrix of log₂fold expression changes (relative to Col) for DEGs identified in the rH4R17A mutant.

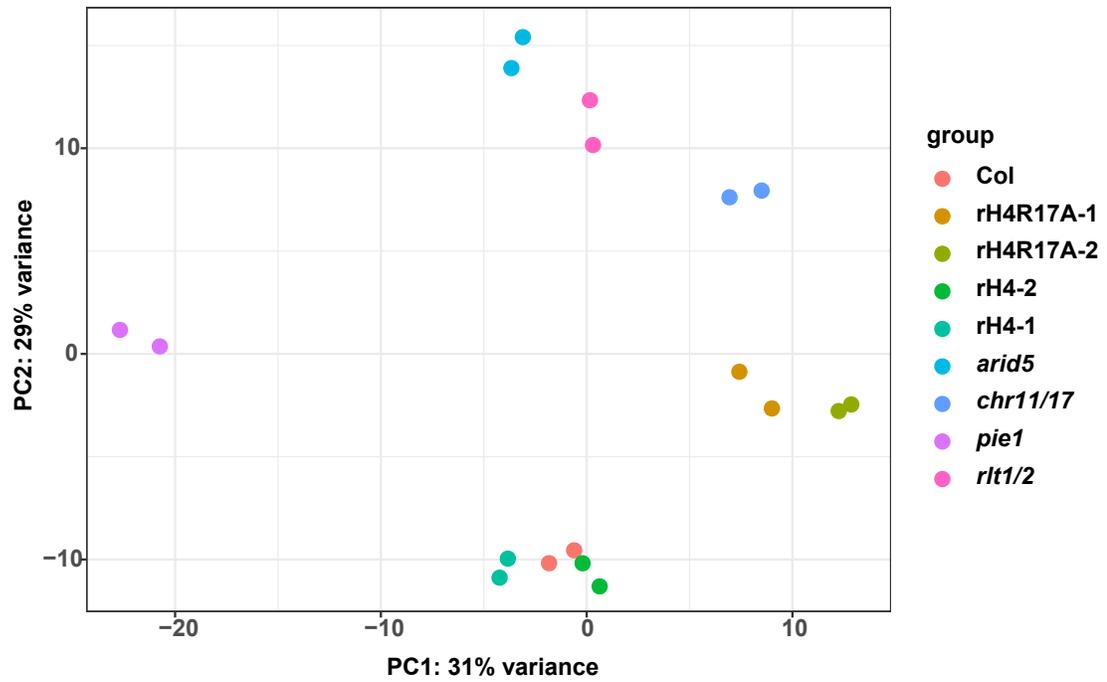


Figure 4.13 Principal component analysis of differentially expressed genes. Principal component plot for 200 most variable differentially expressed genes in rH4R17A mutants along the first two principal components, PC1 and PC2. Variance explained by each principal component shown on respective axis.

(*arid5*, *rtl1/2*, and *chr11/17*) exhibit similar transcriptional defects to rH4R17A mutants, but *pie1* mutants exhibit distinct global transcriptional changes.

To examine how these different mutations affect specific regulatory pathways, we first assessed the impact of these mutations on flowering regulatory genes due to the early flowering phenotypes observed for rH4R17A and ISWI subunit mutants. We investigated the expression of 306 flowering time regulatory genes identified in a previous study (Bouche et al., 2016) and found a high correlation between the effects of the rH4R17A mutation and the *chr11/17*, *arid5*, and *rtl1/2* mutations on the expression of flowering time genes (Figure 4.14). For example, the flowering promoter genes *FUL*, *SOC1*, *FT*, *CURLY LEAF (CLF)*, *SUMO-TARGETED UBIQUITIN E3 LIGASE 4 (STUbL4)*, *PROTEIN ARGININE METHYLTRANSFERASE 10 (PRMT10)*, *ARABIDOPSIS TRITHORAX 1 (ATX1)* and *ATX2* were all co-upregulated in the rH4R17A, *rtl1/2*, *arid5*, and *chr11/17* mutants (Figure 4.15, Figure 4.16). Additionally, the flowering repressor genes *MADS AFFECTING FLOWERING 2 (MAF2)*, *MAF3*, *EARLY FLOWERING MYB PROTEIN (EFM)*, and *ALTERED PHLOEM DEVELOPMENT (APL)* were all co-downregulated in the rH4R17A, *rtl1/2*, *arid5*, and *chr11/17* mutants (Figure 4.15). Importantly, these patterns of co-expression were not observed when comparing *rtl1/2*, *arid5*, and *chr11/17* mutants to rH4 plants (Figure 4.12, Figure 4.13, Figure 4.14, Figure 4.15, Figure 4.16).

Interestingly, several of these flowering regulatory genes encode histone-modifying enzymes, as PRMT10 catalyzes H4R3me2s (Niu et al., 2007), ATX1 catalyzes H3K4me3 (Alvarez-Venegas et al., 2003), and ATX2 catalyzes H3K4me2 (Saleh et al., 2008). Moreover, CLF, which controls leaf morphogenesis and results in small rosettes and curled leaves when mutated (Goodrich et al., 1997; Kim et al., 1998), also serves as the catalytic subunit of the PRC2 complex, which performs H3K27me3 (Schubert et al., 2006). In addition to CLF, other identified genes have also previously

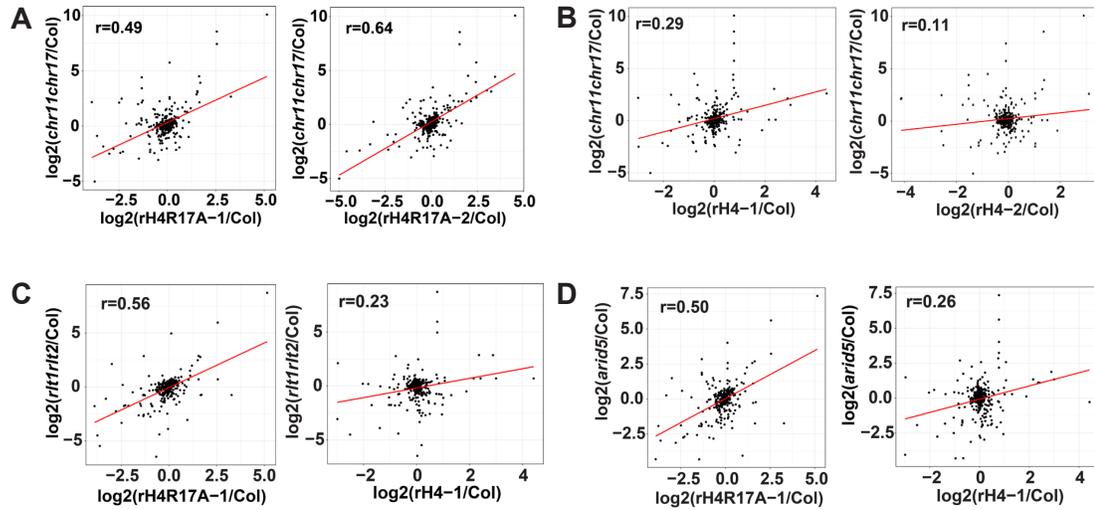


Figure 4.14 Correlation of flowering regulatory gene expression in rH4R17A, ISWI and SWR1 mutants. Scatterplots showing the correlation of the log₂fold expression changes (relative to Col) of flowering time regulatory genes between (A) rH4R17A mutants and *chr11/17* plants, (B) rH4 and *chr11/17* plants, (C) *rit1/2*, rH4R17A, and rH4 plants, and (D) *arid5*, rH4R17A, and rH4 plants.

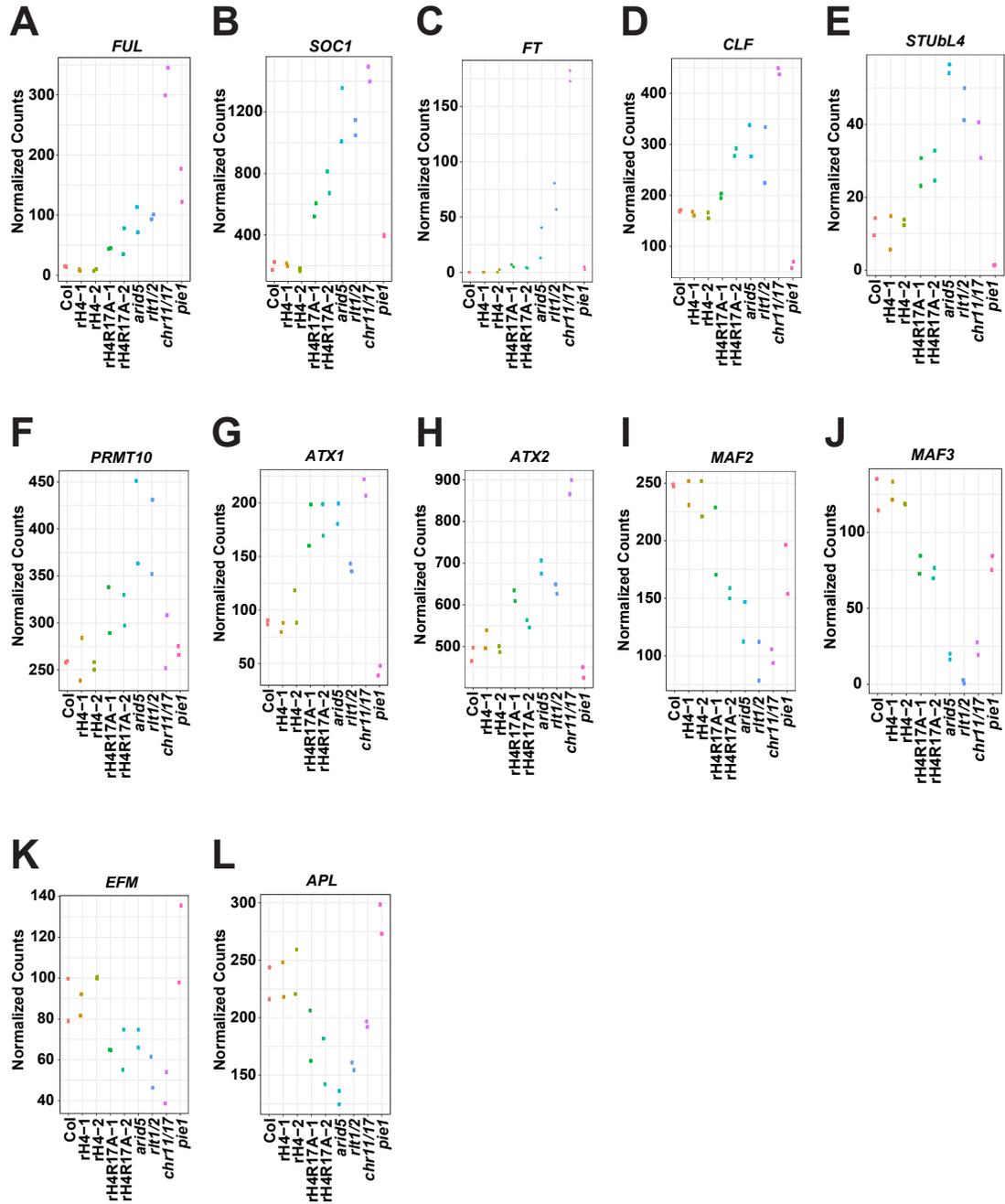


Figure 4.15 Expression of flowering regulatory genes in rh4R17A, ISWI and SWR1 mutants. Normalized read counts at (A) *FUL*, (B) *SOC1*, (C) *FT*, (D) *CLF*, (E) *STUbl4*, (F) *PRMT10*, (G) *ATX1*, (H) *ATX2*, (I) *MAF2*, (J) *MAF3*, (K) *EFM*, and (L) *APL* in Col, rh4-1, rh4-2, rh4R17A-1, rh4R17A-2, *arid5*, *rtt1/2*, *chr11/17*, and *pie1* plants.

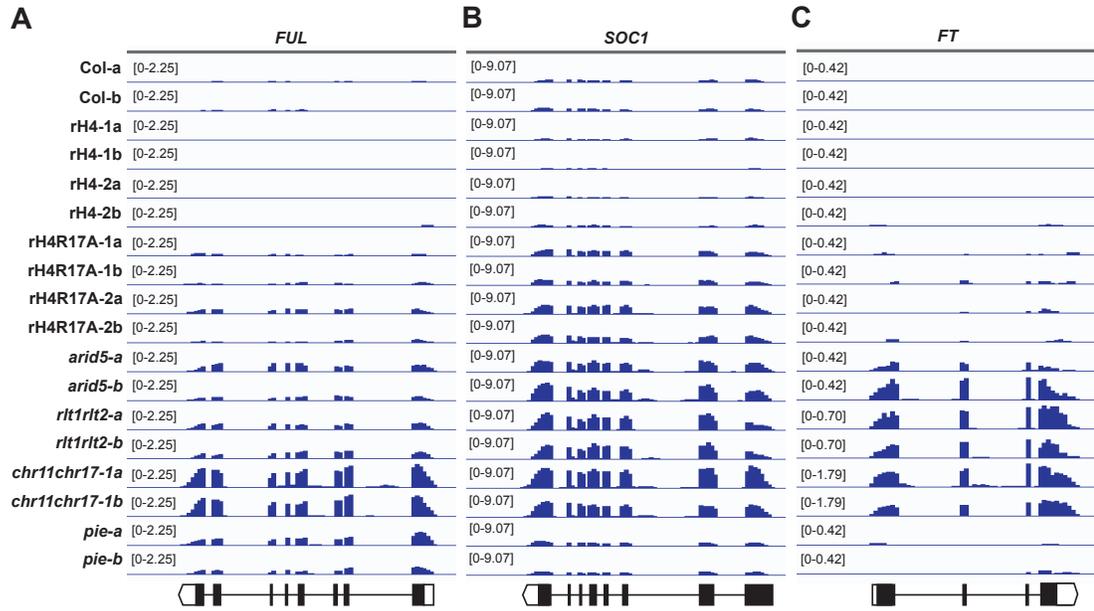


Figure 4.16 Genome browser views of key flowering regulatory genes in rH4R17A, ISWI and SWR1 mutants. Genome browser view of RNA-seq signals at (A) *FUL*, (B) *SOC1*, and (C) *FT* in biological replicates for Col, rH4-1, rH4-2, rH4R17A-1, rH4R17A-2, *arid5*, *rit1/2*, *chr11/17*, and *pie1* plants. Diagrams of genes shown at the bottom, with white boxes, black boxes, and black lines representing untranslated regions, exons, and introns, respectively.

been demonstrated to control aspects of plant development apart from flowering, such as *APL*, which regulates root development (Bonke et al., 2003). The altered expression of the above genes therefore helps explain not only the accelerated flowering response of rH4R17A and ISWI subunit mutants, but several morphological phenotypes observed as well. Experiments characterizing additional morphological phenotypes such as root development would be of interest for these mutants given their altered RNA expression levels of genes involved in diverse developmental pathways.

After these analyses of flowering regulatory genes, we then performed Gene Ontology (GO) term enrichment analysis and identified several additional pathways that are co-regulated by H4R17 and ISWI (Figure 4.17, Supplemental Table 1, Supplemental Table 2, Supplemental Table 3, Supplemental Table 4, Supplemental Table 5). For example, flavonoid metabolism and biosynthesis were impacted in rH4R17A-1, rH4R17A-2, *arid5*, *rtl1/2*, and *chr11/17* mutants, while processes related to pattern specification, the specification of symmetry, and morphogenesis were also affected in these genotypes. Additionally, our GO term enrichment analysis indicated that the rH4R17A-1, rH4R17A-2, *arid5*, *rtl1/2*, and *chr11/17* mutations also all impacted the UV response pathway and the regulation of cell death. We next analyzed RNA levels corresponding to individual genes related to the above pathways to identify several co-regulated genes of interest (Figure 4.18). We observed co-upregulation of genes involved in the DNA damage response and DNA repair such as *ARGONAUTE 9* (*AGO9*), *BRCA1*, and *RADIATION SENSITIVE 17* (*RAD17*) (Figure 4.18A) (Bilichak et al., 2014; Lafarge and Montane, 2003; Liu et al., 2003; Nisa et al., 2019; Oliver et al., 2014; Yoshiyama et al., 2013). Additionally, we observed co-upregulation of cell cycle genes including *DP-E2F-LIKE PROTEIN 3* (*DEL3*), *RETINOBLASTOMA-RELATED 1* (*RBR*), and *CYCLIN D3;2* (*CYCD3;2*) (Figure 4.18B) (Dewitte et al., 2007; Ebel et al., 2004; Vandepoele et al., 2002; Vlieghe et al., 2005). Finally, we observed the co-

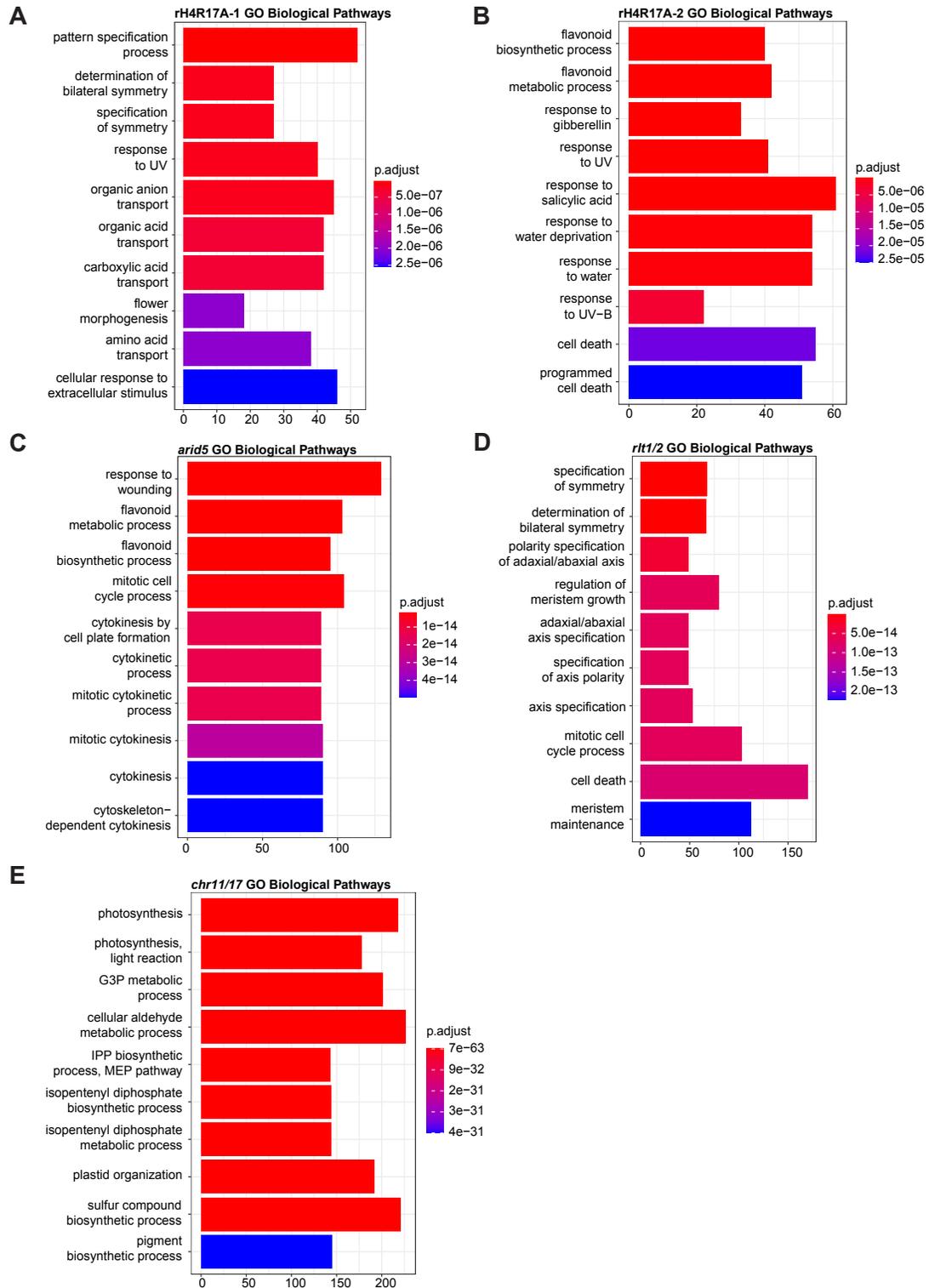


Figure 4.17 GO term enrichment analysis for rH4R17A and ISWI mutants. GO term enrichment analysis for (A) rH4R17A-1, (B) rH4R17A-2, (C) *arid5*, (D) *rtt1/2*, and (E) *chr11/17* mutants. Top 10 enriched biological pathways are shown.

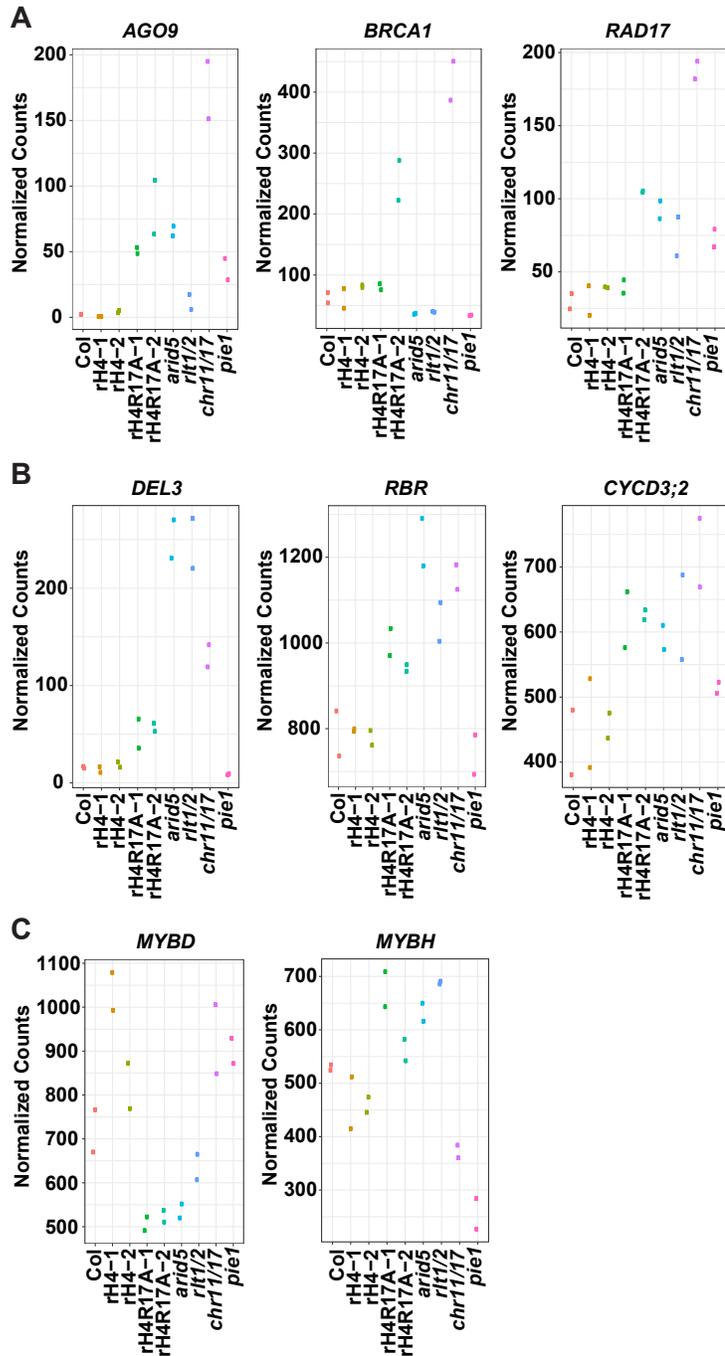


Figure 4.18 Expression of DNA damage response, cell cycle, and anthocyanin pathway regulatory genes in rH4R17A, ISWI and SWR1 mutants. Normalized read counts at (A) *AGO9*, *BRCA1*, *RAD17*, (B) *DEL3*, *RBR*, *CYCD3;2*, (C) *MYBD* and *MYBH* in Col, rH4-1, rH4-2, rH4R17A-1, rH4R17A-2, *arid5*, *rlt1/2*, *chr11/17*, and *pie1* plants.

downregulation of *MYB-LIKE DOMAIN TRANSCRIPTION FACTOR (MYBD)*, which plays a positive role in anthocyanin accumulation, and the co-upregulation of *MYB HYPOCOTYL ELONGATION-RELATED (MYBH)*, which plays a negative role in anthocyanin accumulation (Nguyen et al., 2015) (Figure 4.18C). *CYCD3;2* and *MYBH* also play important roles in regulating leaf morphogenesis, as *CYCD3;2* controls leaf cell number to regulate leaf flatness (Baekelandt et al., 2018) and *MYBH* regulates cell expansion during leaf development to result in leaf curling when overexpressed (Lu et al., 2014).

Due to the altered expression of genes involved in the maintenance of genome stability, we also investigated transposable element (TE) expression in these different lines (Figure 4.19). We detected a very low level of transcriptional reactivation of heterochromatic regions in *rH4R17A* mutants, as indicated by the increased expression of several TEs. We also observed transposon reactivation in *rtl1/2*, *arid5*, and *chr11/17* mutants, but not in *rH4* plants. This TE derepression was very low compared to other previously characterized mutants demonstrating genome instability, such as the *atxr5/6* mutant, which exhibits a low level of TE derepression in which several hundred TEs show increased expression (Dong et al., 2021), compared to the tens of TEs that demonstrate increased expression in *rH4R17A* and/or *ISWI* subunit mutants. We previously observed derepression of the DNA repeat *TSI* by RT-qPCR in *rH4R17A* mutants (Figure 3.15), supporting our observations of TE derepression in these mutants by genome-wide sequencing. Additional experimental work assessing the expression levels of specific TEs using RT-qPCR could further clarify transcriptional derepression phenotypes observed in *rH4R17A* and *ISWI* subunit mutants, as genome-wide sequencing has previously been shown to have limitations in detecting low levels of TE derepression (Dong et al., 2021). Nonetheless, the shared developmental phenotypes and transcriptional profiles of the *rH4R17A*, *rtl1/2*, *arid5*, and *chr11/17* mutants suggest

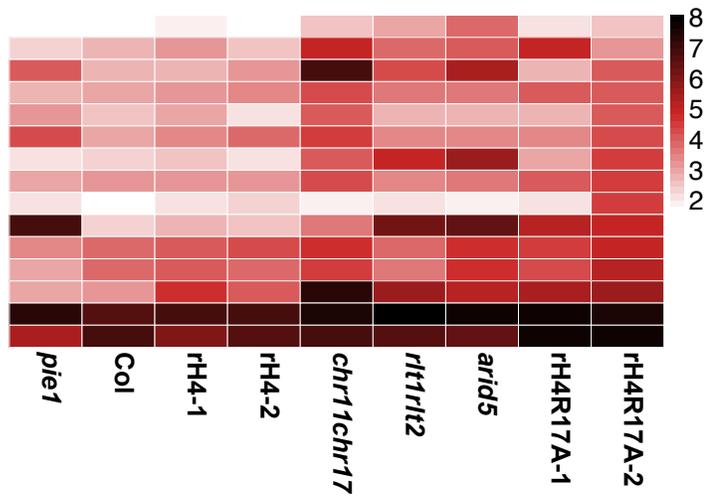


Figure 4.19 TE expression in rH4R17A, ISWI, and SWR1 mutants. Heatmap showing normalized expression (log₂ scale) of TEs found to be upregulated in rH4R17A mutants.

that H4R17 plays an important role in plants as in other eukaryotes in regulating the activity of ISWI on chromatin.

Impact of H4R17A mutation on global nucleosome positioning

ISWI functions as a chromatin remodeling complex that properly organizes nucleosome spacing at transcriptionally active genes in eukaryotes (Clapier and Cairns, 2009; Gkikopoulos et al., 2011; Li et al., 2014; Yadon and Tsukiyama, 2011). Due to the similarity in the phenotypes and transcriptional profiles between rH4R17A plants and mutants in the *Arabidopsis* ISWI complex, we hypothesized that expression of H4R17A interferes with nucleosome spacing in plants.

To address this hypothesis, we assessed global nucleosome positioning in rH4R17A mutants using micrococcal nuclease digestion followed by deep sequencing (MNase-seq). We evaluated two biological replicates corresponding to each genotype and confirmed consistency between biological replicates through Spearman correlation coefficient analysis (Figure 4.20). Consistent with previous results, a relatively lower nucleosome density was found in the 1-kb region upstream of the TSS of protein-coding genes, while a relatively high density, evenly spaced nucleosome distribution was found in the 1-kb region downstream of the TSS for Col plants (Figure 4.21A) (Li et al., 2014). Moreover, active genes were generally observed to display more highly phased nucleosome arrays in the gene body and a sharper peak of nucleosome-free DNA in the promoter when compared to inactive genes, in line with previous studies (Figure 4.21B-C) (Li et al., 2014; Zhang et al., 2015).

In terms of the different genotypes analyzed, rH4 plants displayed highly similar nucleosome positioning patterns to Col as expected. In contrast, while rH4R17A, *arid5*, and *rt1/2* mutants displayed the same general pattern of lower nucleosome density upstream of the TSS and high nucleosome density downstream of the TSS, these

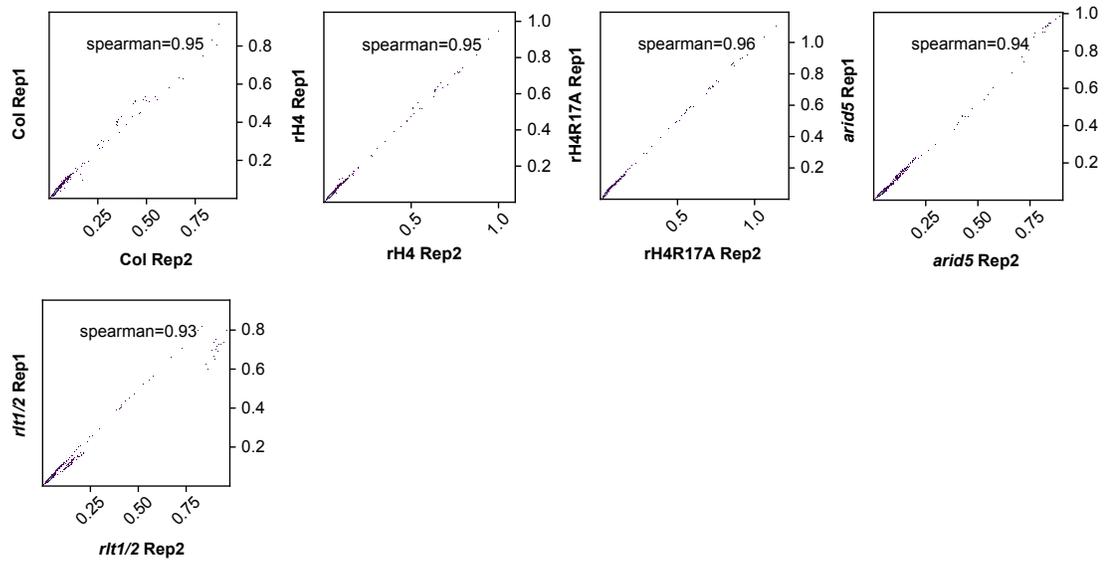


Figure 4.20 Spearman correlation of MNase-seq replicates. Spearman correlation coefficient analysis for MNase-seq replicates of Col, rH4, rH4R17A, *arid5*, and *rlt1/2*.

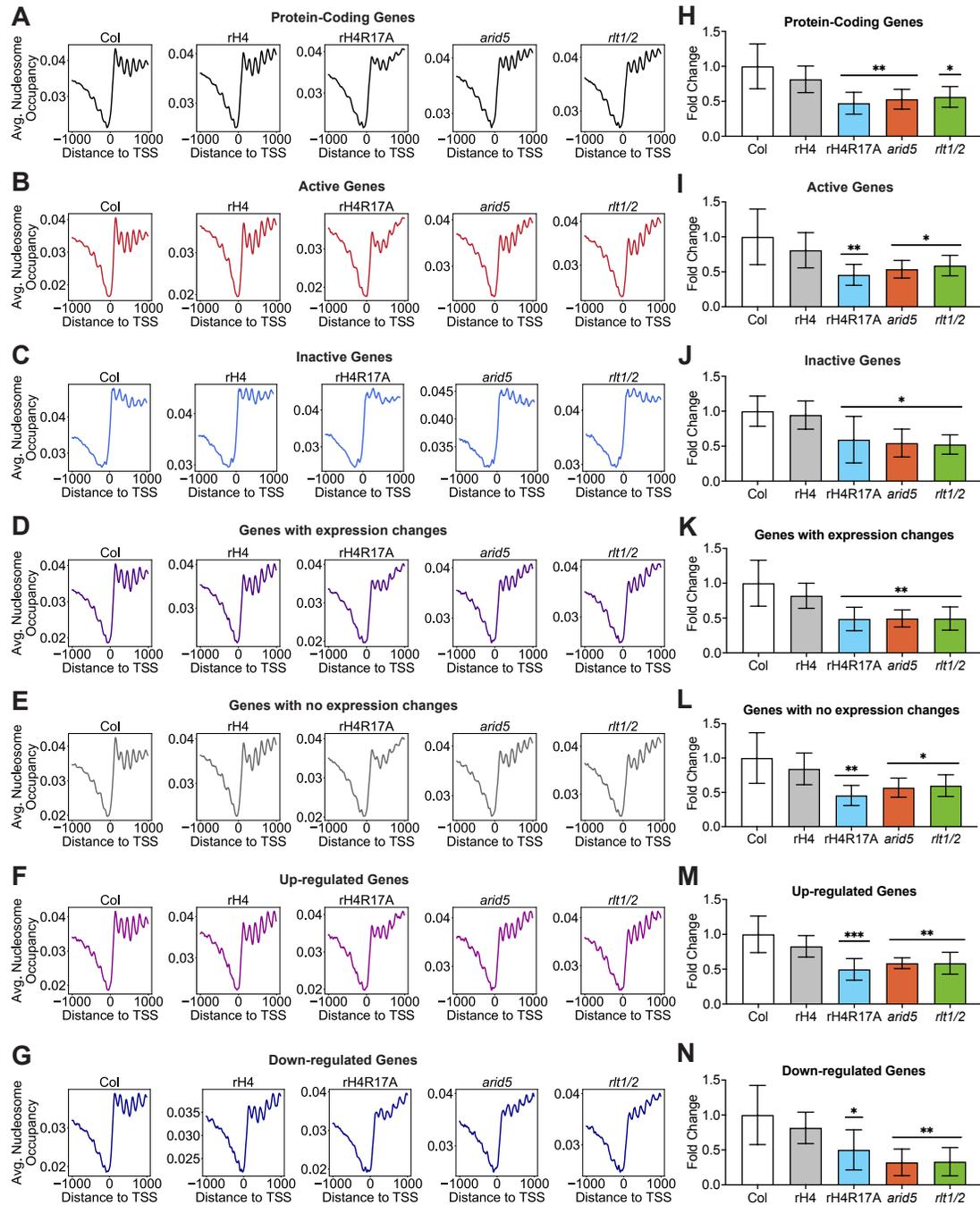


Figure 4.21 Determination of the function of H4R17 on regulating nucleosome positioning. (A-G) Average nucleosome occupancy relative to the TSS (in bp) of (A) all protein-coding genes, (B) active protein-coding genes, (C) inactive protein-coding genes, (D) genes with expression changes, (E) genes with no expression changes, (F) upregulated genes, and (G) downregulated genes in rH4R17A, *arid5*, *rit1/2*, and/or *chr11/17* mutants. The MNase-seq results were generated from two independent biological replicates and RNA-seq data were obtained from the same tissues used for MNase-seq. Cutoffs were defined as follows: Active ≥ 0.5 TPM (Transcripts Per Million); Inactive < 0.5 TPM. Genes with expression changes were defined as $> \pm 1.5$ -fold for

rH4R17A, *arid5*, *rlt1/2*, and/or *chr11/17* vs. Col and genes with no expression changes were defined as $<\pm 1.1$ -fold for rH4R17A, *arid5*, *rlt1/2*, and/or *chr11/17* vs. Col. (H-N) Average fold change in Δ Nucleosome Occupancy of all nucleosome peaks 1 kb downstream of the TSS relative to Col corresponding to (H) all protein-coding genes, (I) active protein-coding genes, (J) inactive protein-coding genes, (K) genes with expression changes, (L) genes with no expression changes, (M) upregulated genes, and (N) downregulated genes in rH4R17A, *arid5*, *rlt1/2*, and/or *chr11/17* mutants. Standard deviation denoted with error bars. Statistical analyses were performed using one-way ANOVA with Tukey's HSD post hoc test. *P*-value from Tukey's HSD test (genotype vs. Col) denoted with asterisks (* $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$).

genotypes all exhibited a loss of evenly spaced nucleosome distributions in the gene body (Figure 4.21A), similar to the pattern reported for the *chr11/17* mutant (Li et al., 2014). Additionally, we analyzed the nucleosome distribution patterns at genes with expression changes in rH4R17A, *chr11/17*, *rlt1/2*, and/or *arid5* mutants as well as genes without expression changes in these mutants. We found that the nucleosome distribution patterns at DEGs and non-DEGs were both affected by the rH4R17A, *arid5*, and *rlt1/2* mutations (Figure 4.21D-E), in line with previously published MNase-seq results for the *chr11/17* mutant (Li et al., 2014). Additionally, nucleosome distribution patterns at DEGs were affected by the rH4R17A, *arid5*, and *rlt1/2* mutations regardless of whether the expression of these genes was up- or downregulated (Figure 4.21F-G). To provide a more quantitative assessment of nucleosome spacing in our assays, we calculated the average change in nucleosome occupancy at the +2 through +6 nucleosome peaks as a measure of nucleosome phasing. This analysis confirmed that the rH4R17A, *arid5*, and *rlt1/2* mutations caused a significant reduction in regular nucleosome phasing in gene bodies (Figure 4.21H-N).

While we could observe some differences in nucleosome positioning at individual genes that we previously identified to be either co-upregulated or co-downregulated in rH4R17A and ISWI subunit mutants, we found that analyzing nucleosome positioning profiles corresponding to individual genes without complex signal processing techniques was complicated by the noisiness inherent in MNase-seq datasets (Figure 4.22). In order to further assess how nucleosome positioning patterns at individual genes are affected by rH4R17A and ISWI mutations, we performed t-distributed stochastic neighbor embedding (t-SNE) on the nucleosome occupancy data corresponding to the protein-coding genes. We utilized 10-bp bins encompassing the 1-kb region upstream and the 1-kb region downstream of the TSS, generating an array of 200 features in total for each gene. We first analyzed all 27,443 total protein-coding genes and found that there was a

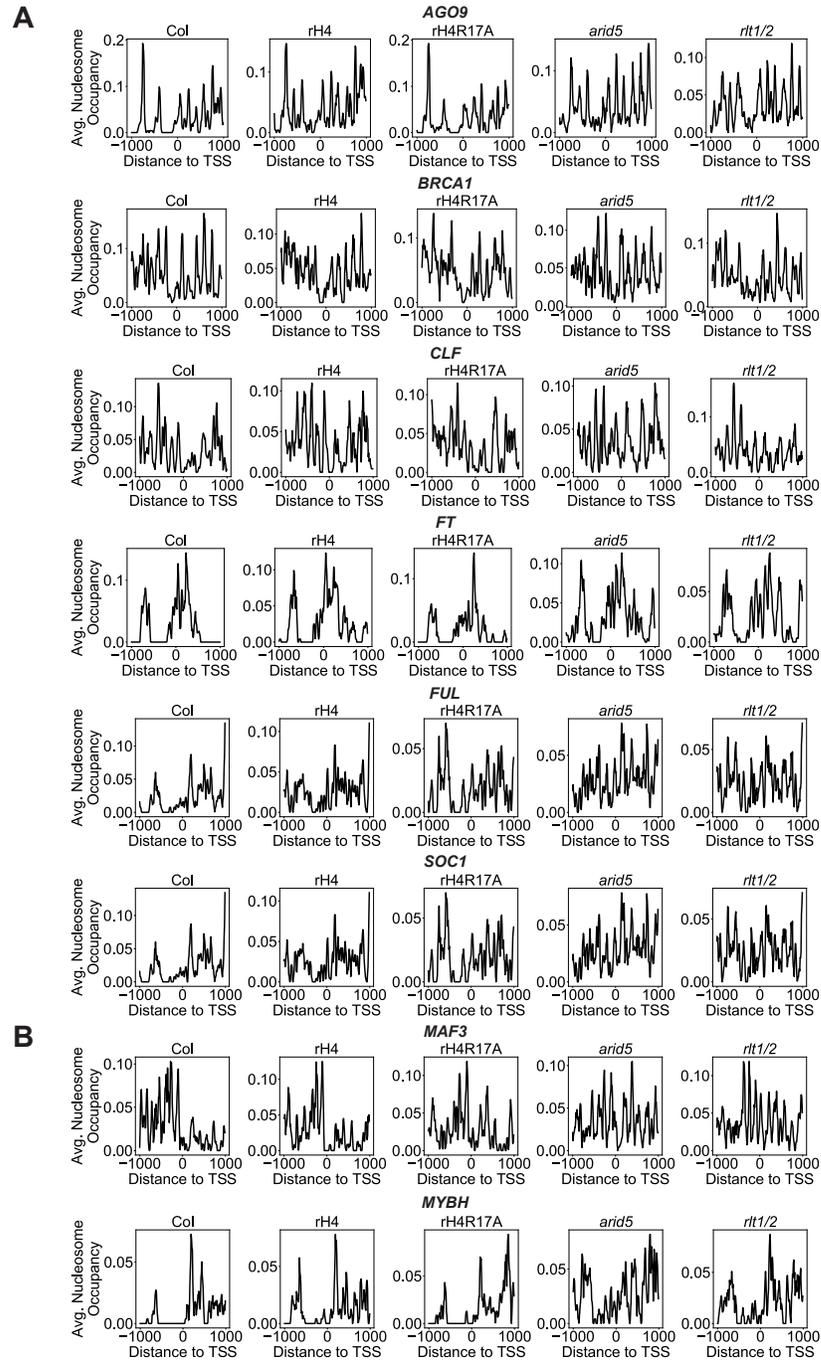


Figure 4.22 Nucleosome occupancy profiles at individual genes. Average nucleosome occupancy relative to the TSS (in bp) of individual (A) co-upregulated genes (*AGO9*, *BRCA1*, *CLF*, *FT*, *FUL*, and *SOC1*) and (B) co-downregulated genes (*MAF3* and *MYBH*) identified by RNA-seq in rH4R17A and ISWI subunit mutants.

general tendency of more highly expressed genes to cluster together, while more lowly expressed genes tended to cluster towards a separate region of the embedding (Figure 4.23A-B). While the separation between highly and lowly expressed genes was not completely discrete, this result confirms our observation that more highly expressed genes generally show distinct profiles from more lowly expressed genes, for example, seen in the metaprofiles as higher phasing and a more severe depletion of nucleosome occupancy in the 5' NFR (Figure 4.21).

We then utilized *k*-means clustering to identify groups of genes with similar nucleosome occupancy profiles (Figure 4.23C). We discarded clusters containing fewer than 10 genes and subsequently identified four main clusters in the data. Two clusters (Clusters 1 and 2) mainly contained active genes and two clusters (Clusters 3 and 4) mainly contained inactive genes (Figure 4.23D-G). We defined active genes as having an expression level greater than or equal to 0.5 transcripts per million (TPM) (containing “high” and “medium” expression categories) and inactive genes as having an expression level less than 0.5 TPM (containing “low” and “none” expression categories). We examined nucleosome occupancy metaprofiles corresponding to the genes in each of the identified clusters and found that both of the clusters mainly containing inactive genes showed a pattern of relatively highly phased and evenly sized nucleosome peaks, while the clusters mainly containing active genes showed a strong enrichment at the +1 nucleosome and a successive reduction in phasing at each subsequent nucleosome (Figure 4.23D). Comparing the t-SNE embeddings corresponding to the different genotypes analyzed, we observed little apparent change in Col vs. rH4, rH4R17A, *arid5*, or *rlt1/2* mutants (Figure 4.23A-B). Moreover, we observed that clustering patterns were maintained in the different genotypes assessed (Figure 4.23C). In contrast, we upheld our previous observation that rH4R17A, *arid5*, and *rlt1/2* mutants displayed a reduction of regularly spaced nucleosome peaks in gene bodies in groups of genes demonstrating

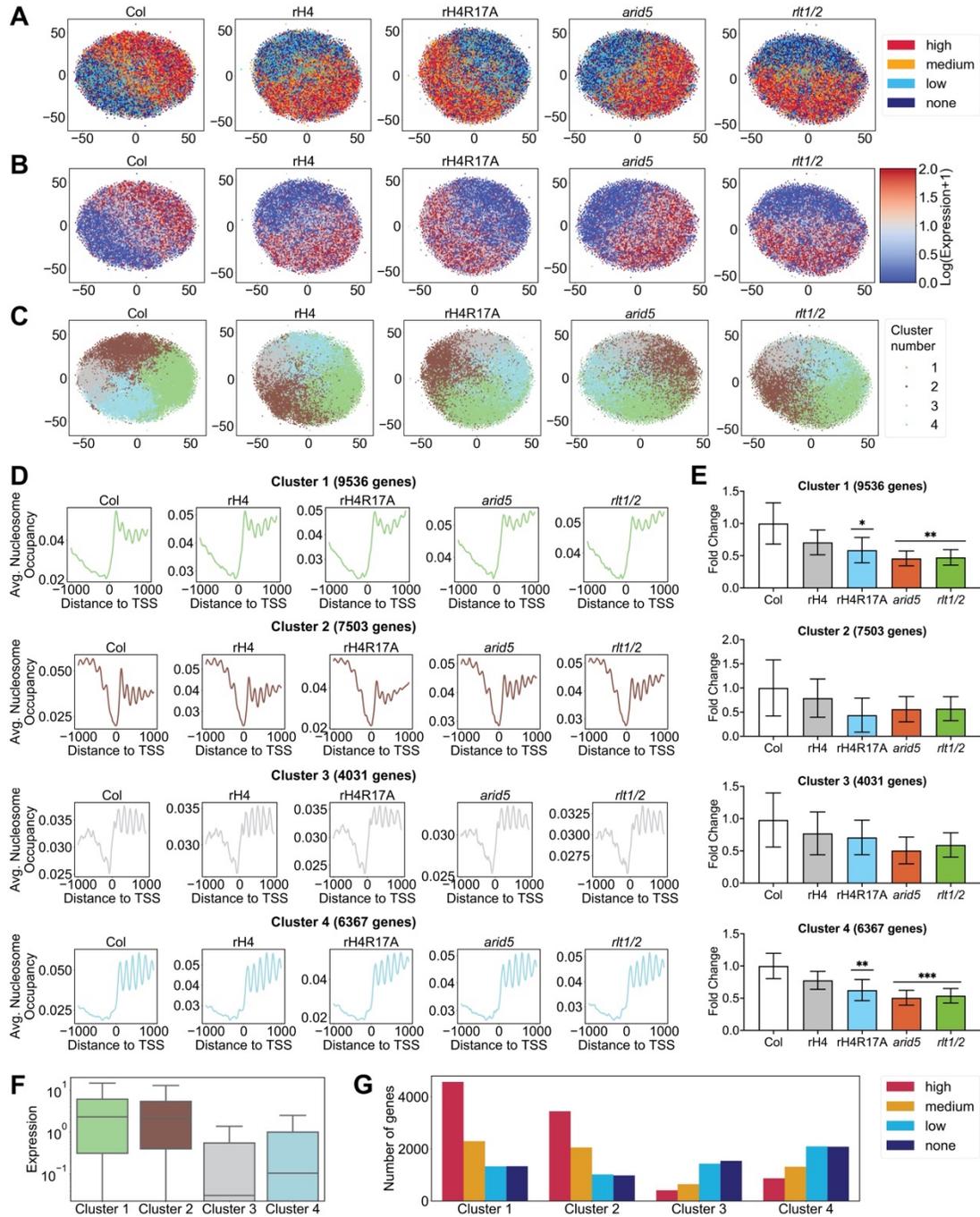


Figure 4.23 Dimensionality reduction and clustering analysis of nucleosome occupancy profiles at all protein-coding genes. (A-C) t-SNE embeddings of nucleosome occupancy data at all protein-coding genes colored by (A) expression level, (B) $\log(\text{expression}+1)$, and (C) k -means cluster. Maximum cutoff for (B) was set at 2 for ease of viewing. Expression level cutoffs were defined as follows: High ≥ 2.5 TPM; Medium < 2.5 TPM and ≥ 0.5 TPM; Low < 0.5 TPM and > 0 TPM; None = 0 TPM. (D-E) (D) Average nucleosome occupancy relative to the TSS (in bp) and (E) average fold change in Δ Nucleosome Occupancy of all nucleosome peaks 1 kb downstream of the TSS for each k -means cluster identified in (C). Standard deviation denoted with error bars.

Statistical analyses were performed using one-way ANOVA with Tukey's HSD post hoc test. *P*-value from Tukey's HSD test (genotype vs. Col) denoted with asterisks (* $p < 0.05$, ** $p < 0.005$, *** $p < 0.0005$). (F-G) (F) Boxplot showing expression levels (TPM) of genes and (G) number of genes in each of the expression categories in the clusters identified in (C).

varying expression levels (Figure 4.23D-G). We observed a significant reduction of nucleosome phasing in Clusters 1 and 4 for rH4R17A and ISWI subunit mutants, and moreover, while the mean reduction of peak size in Cluster 2 was not statistically significant due to the large variance observed between peaks in this cluster, we also observed a notable reduction of phasing in Cluster 2 through metaprofile analysis, particularly for the rH4R17A mutant (Figure 4.23D-E).

Surprisingly, we observed that the highest level of nucleosome phasing was exhibited by one of the clusters mainly containing inactive genes (Cluster 4). While this observation may seem to conflict with our previous results assessing metaprofiles of all active versus inactive genes (Figure 4.21B-C), examination of the position of each nucleosome in the nucleosomal arrays corresponding to these different clusters revealed that the position of the +1 nucleosome was altered in these different groups of genes. Notably, the +1 nucleosomes at Clusters 1 and 3 are shifted in the 3' direction compared to the +1 nucleosomes at Clusters 2 and 4, and therefore, six nucleosome peaks are visible in the Cluster 2 and 4 arrays, while only five nucleosome peaks are visible in the Cluster 1 and 3 arrays (Figure 4.23D). Consequently, taking the average of the two more lowly expressed clusters or the two more highly expressed clusters would lead to interference due to the different phasing patterns of the nucleosome arrays, and thus, the size of the resultant average peaks would depend on the degree to which the two arrays are offset from each other, their relative amplitudes, and the number of genes belonging to each cluster. These results suggest that the translational phasing of the nucleosome arrays is one of the major factors separating the different clusters. Other distinguishing features between the four clusters were the height of the +1 nucleosome peak and the total nucleosome occupancy in the 1-kb region upstream of the TSS, which could indicate a neighboring gene in the case of high occupancy. The general maintenance of these features distinguishing the four clusters in rH4R17A and ISWI

subunit mutants could explain why severe disruptions to the t-SNE embeddings are not observed in these mutants in spite of the loss of nucleosome phasing. Importantly, our results suggest that individual groups of more lowly expressed genes can display higher phasing than individual groups of more highly expressed genes, and thus, analyzing subsets of genes within expression categories can reveal further information about the relationship between nucleosome positioning and gene expression that can be obfuscated by analyzing larger groups of genes.

To further investigate nucleosome positioning patterns at groups of inactive genes compared to groups of active genes, we individually performed *k*-means clustering on all active genes (15,617 genes) and all inactive genes (11,826 genes) in Col (Figure 4.24). We discarded clusters with less than 10 genes and subsequently identified seven clusters among active genes and seven clusters among inactive genes. Among active genes, four clusters contained arrays of very highly phased nucleosomes, whereas two clusters contained lowly to moderately phased nucleosomes, and one cluster contained a disordered nucleosome array (Figure 4.24A). Among inactive genes, five clusters contained very highly phased nucleosome arrays and two clusters contained relatively disordered nucleosome arrays (Figure 4.24B). Thus, we observed that groups of genes with equally high phasing can be observed among inactive genes and active genes. Additionally, certain clusters of genes appeared to demonstrate a more severe reduction of nucleosome phasing compared to other clusters in the rH4R17A mutant (e.g., Cluster 3 of active genes), although further analyses are required to quantitatively compare the reduction in phasing phenotypes across different clusters. We then used *k*-means clustering to identify 20 clusters in all active genes (Figure 4.25) and all inactive genes (Figure 4.26) and observed that the majority of the clusters in both active and inactive genes again corresponded to highly phased nucleosome arrays that oscillated in different phases relative to the TSS.

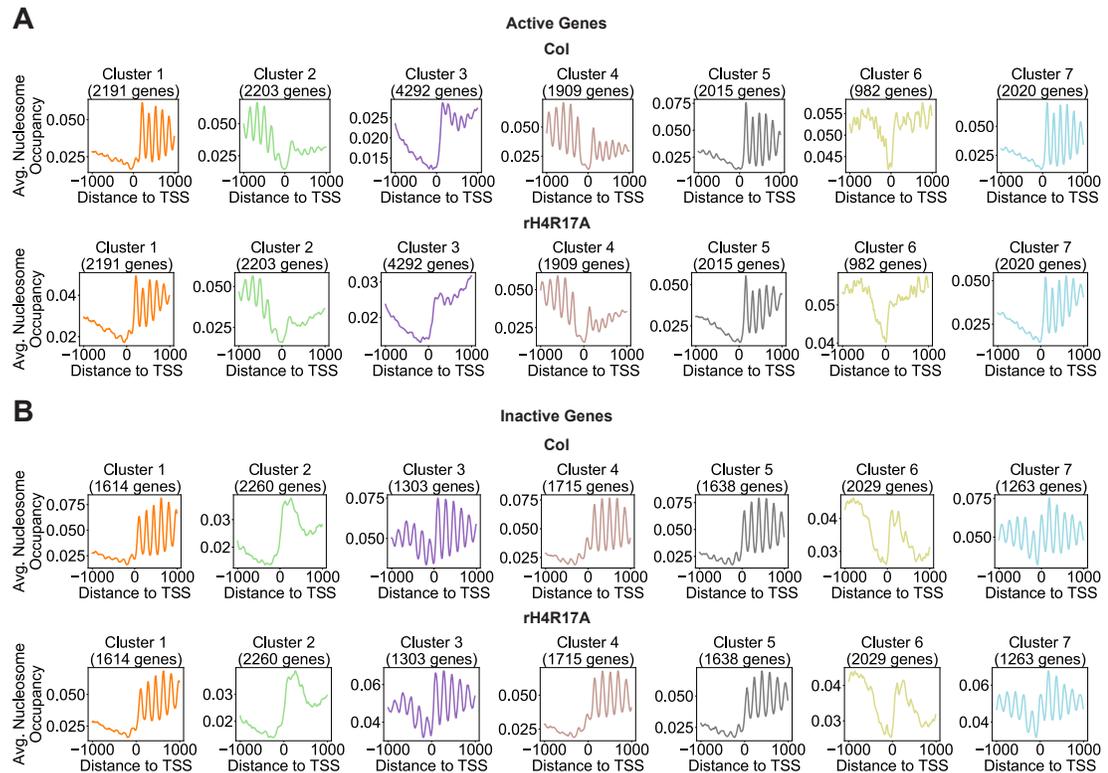


Figure 4.24 Clustering analysis of nucleosome occupancy profiles at active genes and inactive genes in Col and rH4R17A plants. Average nucleosome occupancy relative to the TSS (in bp) of each *k*-means cluster identified in (A) active protein-coding genes and (B) inactive protein-coding genes in Col and rH4R17A plants. Cutoffs were defined as follows: Active ≥ 0.5 TPM; Inactive < 0.5 TPM.

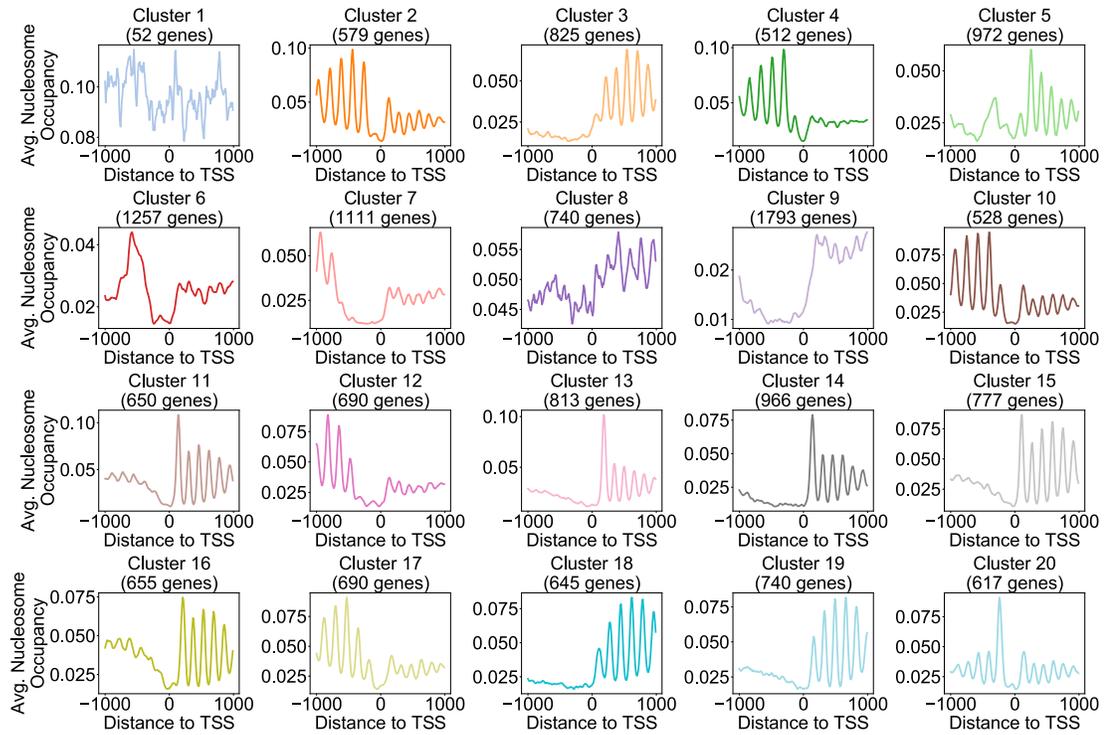


Figure 4.25 Clustering analysis of nucleosome occupancy profiles at active genes. Average nucleosome occupancy relative to the TSS (in bp) of each *k*-means cluster identified in active protein-coding genes in *Col* plants. Cutoffs were defined as follows: Active ≥ 0.5 TPM.

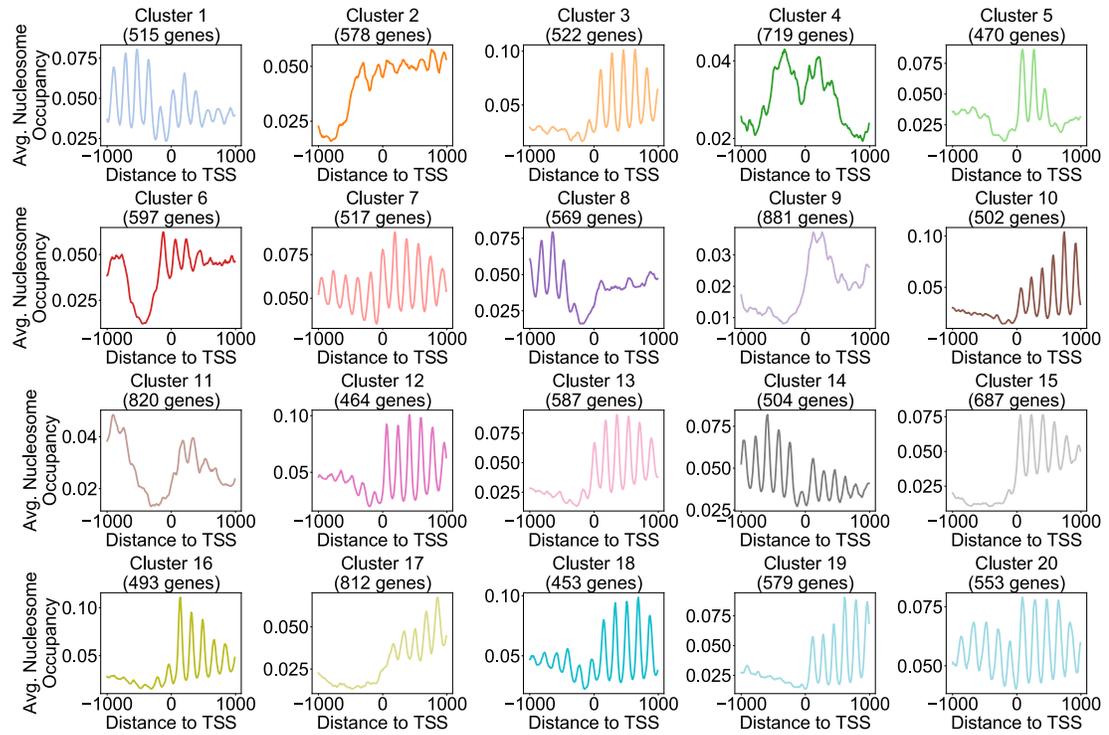


Figure 4.26 Clustering analysis of nucleosome occupancy profiles at inactive genes. Average nucleosome occupancy relative to the TSS (in bp) of each *k*-means cluster identified in inactive protein-coding genes in *Col* plants. Cutoffs were defined as follows: Inactive <0.5 TPM.

Finally, to specifically assess how nucleosome positioning patterns at genes with transcriptional alterations were affected by the rH4R17A, *arid5*, and *rtt1/2* mutations, we performed dimensionality reduction using t-SNE on groups of genes with expression changes, groups of genes with no expression changes, upregulated genes, and downregulated genes in rH4R17A and/or ISWI subunit mutants (Figure 4.27). In all four of these groups of genes, we observed that the separation between more highly and more lowly expressed genes was maintained in Col plants. Moreover, rH4R17A, *arid5*, and *rtt1/2* mutants similarly maintained separation between more highly and more lowly expressed genes. While we failed to observe major changes in the t-SNE embeddings caused by the mutations analyzed, we nonetheless observed a loss of regular nucleosome positioning in gene bodies in rH4R17A, *arid5*, and *rtt1/2* mutants in metaprofiles produced from our clustering analyses (Figure 4.23D-E). The maintenance of other major features of nucleosome occupancy profiles corresponding to more highly or more lowly expressed genes (e.g., the height of the +1 nucleosome peak compared to downstream nucleosome peaks) may be responsible for the preservation of t-SNE embedding patterns in rH4R17A and ISWI subunit mutants. Additionally, given that rH4R17A and ISWI subunit mutations affect nucleosome phasing at both active and inactive genes, we would not necessarily expect the reduction of nucleosome phasing caused by these mutations to differentially affect the t-SNE embeddings of more highly expressed genes compared to more lowly expressed genes. Nonetheless, our finding that the rH4R17A and ISWI subunit mutations caused a significant reduction of nucleosome phasing in gene bodies remained consistent across the majority of the groups of genes that we assessed (Figure 4.21, Figure 4.23, Figure 4.24). Taken together, these results indicate that H4R17 positively regulates the action of the ISWI complex to establish nucleosome arrays in protein-coding genes.

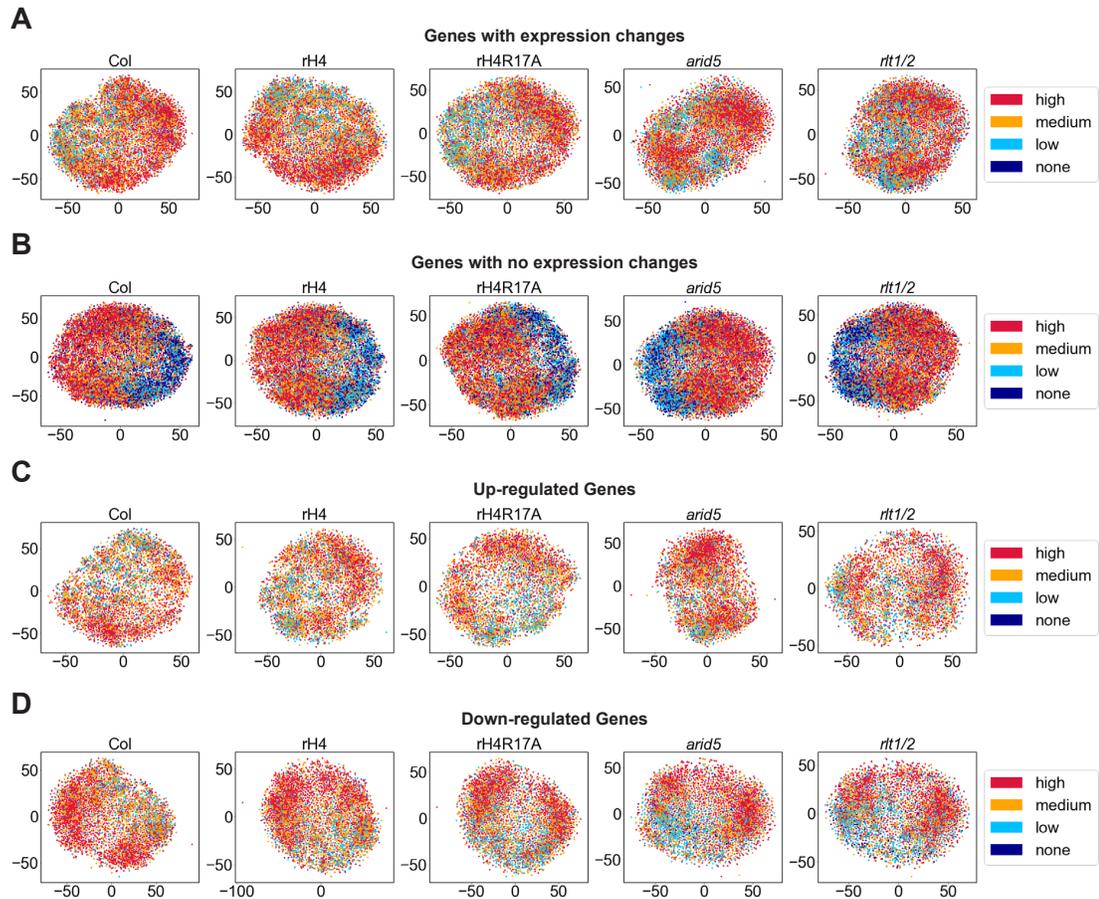


Figure 4.27 Dimensionality reduction of nucleosome positioning profiles at differentially expressed genes. t-SNE embeddings of nucleosome occupancy data at (A) genes with expression changes, (B) genes with no expression changes, (C) upregulated genes, and (D) downregulated genes in *rH4R17A*, *arid5*, *rtt1/2*, and/or *chr11/17* mutants colored by expression level. Expression level cutoffs were defined as follows: High ≥ 2.5 TPM; Medium < 2.5 TPM and ≥ 0.5 TPM; Low < 0.5 TPM and > 0 TPM; None = 0 TPM. Genes with expression changes were defined as $> \pm 1.5$ -fold for *rH4R17A*, *arid5*, *rtt1/2*, and/or *chr11/17* vs. Col. Genes with no expression changes were defined as $< \pm 1.1$ -fold for *rH4R17A*, *arid5*, *rtt1/2*, and/or *chr11/17* vs. Col.

Proposed mechanism for action of H4R17 on the regulation of flowering time and development

With our results, we have identified a novel role for H4R17 in regulating multiple development processes in *A. thaliana* including leaf development, fruit development, and flowering. Our findings suggest that this role for H4R17 is not mediated via post-translational modification of this residue. However, comparative analysis of the protein sequence of the ISWI catalytic subunits in *A. thaliana* (CHR11 and CHR17) reveals strict conservation of the amino acids involved in making contacts with histone H4 arginine 17 in the ISWI orthologs from other species (Figure 4.28) (Yan et al., 2016; Yan et al., 2019), which supports the hypothesis that the H4R17A mutation impacts ISWI chromatin remodeling. In both *Myceliophthora thermophila* and *Saccharomyces cerevisiae*, two conserved residues (Asp524 and Glu474 in *M. thermophila* and Asp536 and Glu486 in *S. cerevisiae*) have been demonstrated to contact H4R17 through structural studies (Yan et al., 2016; Yan et al., 2019). These residues are identical in *Arabidopsis thaliana* and high conservation is also observed throughout the core2 domain, which contains these two residues (Figure 4.28). We built a homology model for *A. thaliana* CHR11 (a.a. 176-706) based on the X-ray crystallography structure from *M. thermophila* ISWI (5jxr.1.a; a.a. 173-718) and observed 60.49% identity and a QMEANisCo score of 0.68, indicating high model quality (Figure 4.29) (Biasini et al., 2014). With this homology model, we detected structural conservation of the H4R17-binding region in plant ISWI proteins (Figure 4.29E-F). Moreover, we calculated a high confidence score for the conserved Asp524 and Glu474 residues (0.78 and 0.65, respectively). These results indicate that it is highly likely that the H4R17-binding function of *A. thaliana* ISWI proteins is conserved.

Based on these results, we propose a model similar to that of animal systems where H4R17 regulates developmental processes in plants through its regulation of the

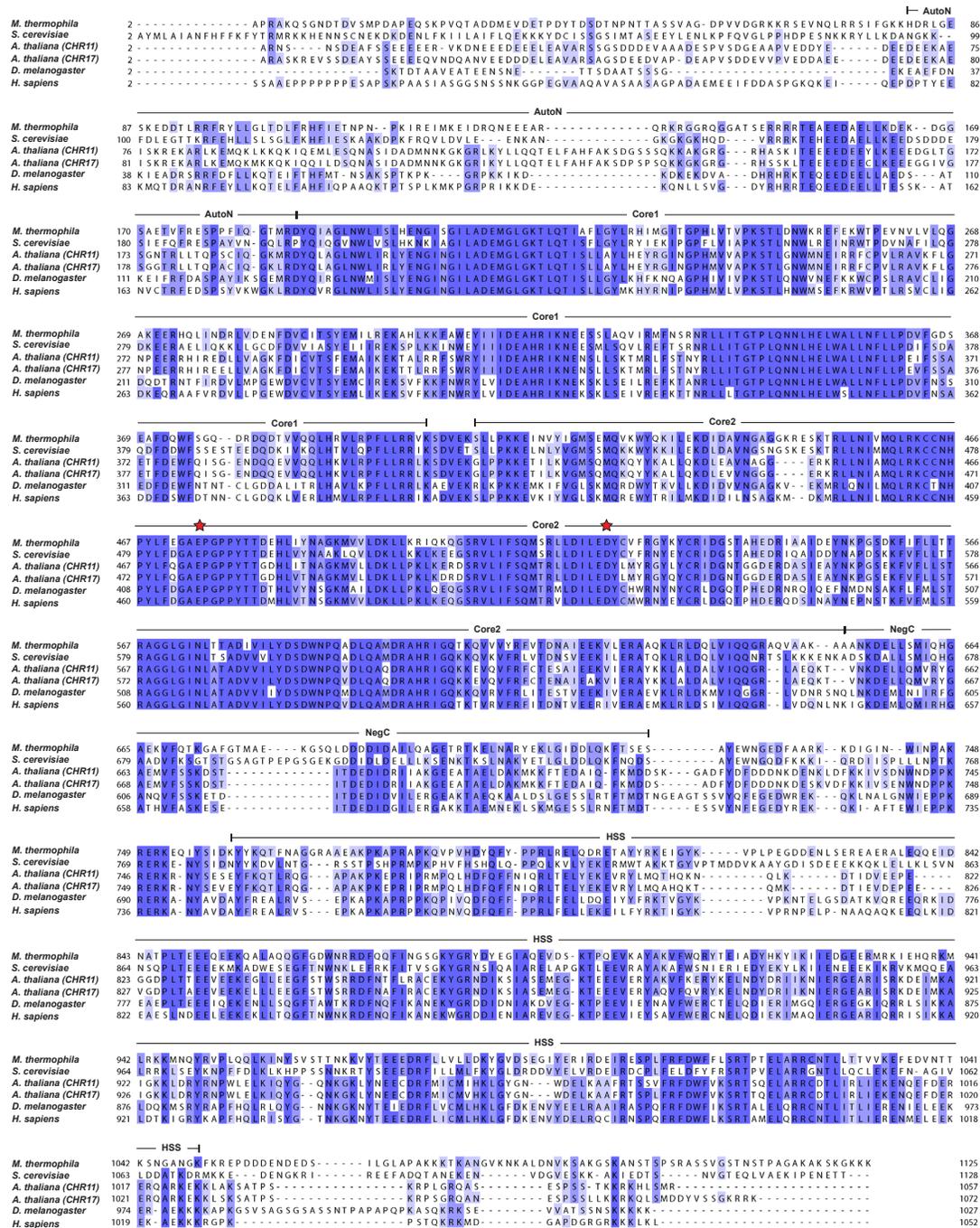


Figure 4.28 Conservation of ISWI proteins. Multiple sequence alignment of ISWI proteins performed with Clustal Omega. Protein sequences were obtained from UniProt and correspond to the following accession numbers: *Myceliophthora thermophila*: G2QFM3, *Saccharomyces cerevisiae*: P38144, *Arabidopsis thaliana* (CHR11): F4JAV9, *Arabidopsis thaliana* (CHR17): F4JY25, *Drosophila melanogaster*: Q24368, *Homo sapiens* (SNF2H): O60264. Darker shading indicates higher similarity between residues. Red stars above a.a. indicate the residues implicated in binding H4R17 on the second RecA-like ATPase core domain (core2) identified in *Myceliophthora thermophila* (Yan et al., 2016) and *Saccharomyces cerevisiae* (Yan et al., 2019). Protein domains are

assigned as reported in a previous study (Yan et al., 2016). HSS: HAND-SAND-SLIDE, core1: first RecA-like domain, core2: second RecA-like domain.

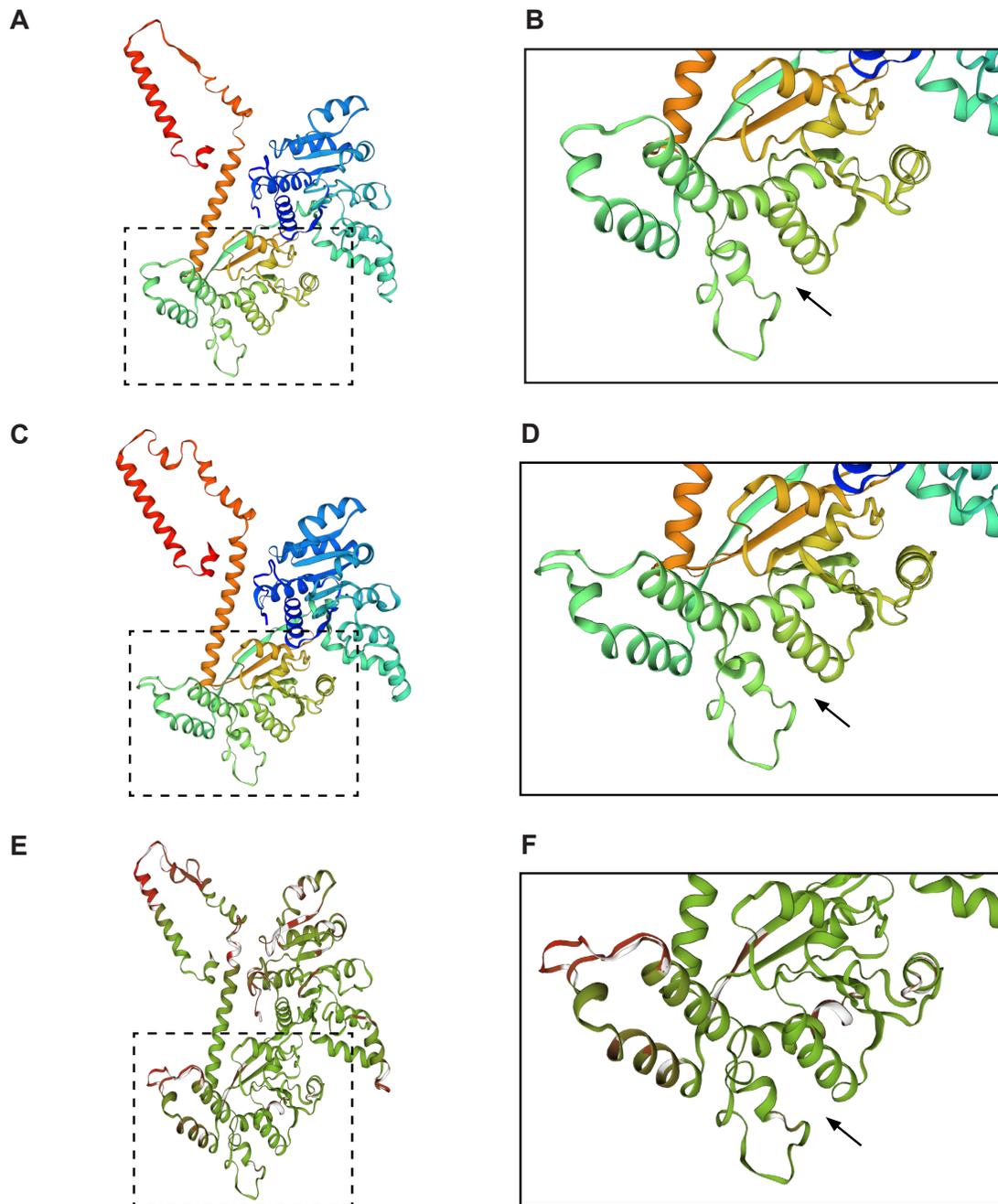


Figure 4.29 Homology model of *Arabidopsis thaliana* CHR11. (A-B) Homology model of *Arabidopsis thaliana* CHR11 a.a. 176-706. (C-D) Reference structure of *Myceliophthora thermophila* ISWI (5JXR) a.a. 173-718. (E-F) Superposition of *Arabidopsis thaliana* CHR11 and *Myceliophthora thermophila* ISWI structures with consistency color scheme (green indicates more consistent and red indicates less consistent). Black arrow denotes the predicted (*A. thaliana*) or validated (*M. thermophila*) binding pocket of histone H4 arginine 17 (Yan et al., 2016). The boxed regions are enlarged for further examinations in (B), (D), and (F).

ISWI complex (Clapier et al., 2001; Clapier et al., 2002; Dann et al., 2017; Fazio et al., 2005; Hamiche et al., 2001). In wild-type plants, H4R17 positively regulates the ISWI complex to slide nucleosomes and properly establish the nucleosome positioning patterns in the gene bodies of protein-coding genes (Figure 4.30). In rH4R17A mutant plants however, the positive regulation of ISWI by histone H4 is impaired so that evenly-spaced nucleosome distributions are no longer observed in gene bodies. The altered nucleosome positioning patterns in gene bodies and the large-scale transcriptional changes in turn cause the observed pleiotropic developmental phenotypes.

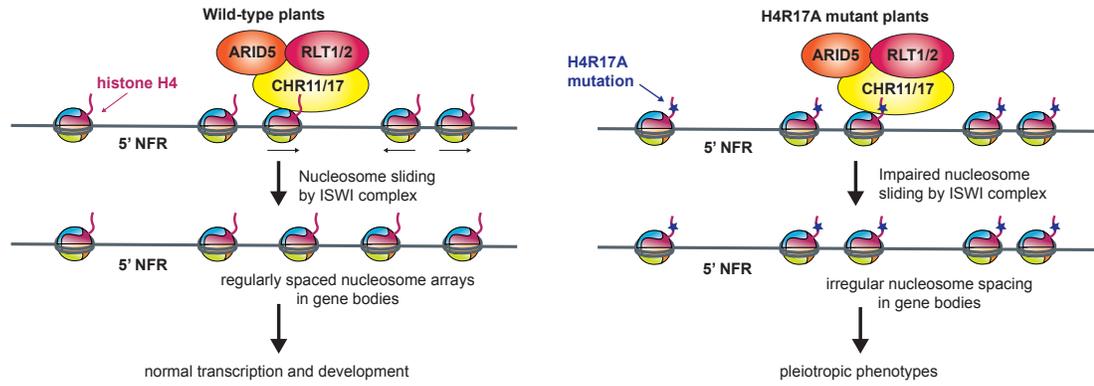


Figure 4.30 Model for the regulation of flowering and plant development by H4R17. Proposed model for the function of histone H4 arginine 17 in the regulation of ISWI chromatin remodeling in plants. 5' NFR: 5' Nucleosome-Free Region.

Discussion

Utility of histone H4 replacement system for assessing function of H4 residues

Our analyses of the rH4R17A mutant demonstrates the utility of our histone H4 replacement system towards the assessment of H4 function in plants. While H4R17A replacement mutants have previously been generated in *Saccharomyces cerevisiae* (Dai et al., 2008; Govin et al., 2010a; Nakanishi et al., 2008), our *Arabidopsis thaliana* rH4R17A mutants represent the first mutants that allow us to study the effect of H4R17 mutations in a multicellular eukaryote. With these experiments, we have identified several novel roles for H4R17 in plants, including the regulation of developmental processes such as flowering and fertility (Figure 4.1). Our analyses of global gene expression also revealed the abnormal expression of genes involved in the regulation of genome stability, such as gene silencing and the DNA damage response, in the rH4R17A mutant (Figure 4.17, Figure 4.18, Figure 4.19). Interestingly, in previous studies in budding yeast, H4R17A mutants were not observed to display negatively impacted sporulation (Govin et al., 2010a). Whether H4R17 regulates gametogenesis and other developmental processes in animals remains to be elucidated; however, mammals lacking PRMT7, the enzyme proposed to monomethylate H4R17, have been observed to display developmental defects (Akawi et al., 2015; Blanc et al., 2016; Jeong et al., 2016; Ying et al., 2015). Additionally, the H4R17 residue was found to be frequently mutated in human cancers (Bennett et al., 2019; Nacev et al., 2019). In summary, the *Arabidopsis thaliana* rH4R17A mutant provides us with a useful tool to study the effect of the H4R17A mutation *in vivo* in a multicellular organism. Moreover, we have generated a multitude of other rH4 mutants, which remain a resource for further characterization.

H4R17A mutation shows incomplete dominance

Our results indicate that the H4R17A mutation demonstrated incomplete dominance over wild-type H4 (Figure 4.1, Figure 4.2). We also observed that the rH4R17A-2 mutant line displayed a more severe phenotype than the rH4R17A-1 mutant line (Figure 4.1, Figure 4.2). This differential phenotype was likely not attributable to different amounts of wild-type histone H4 remaining present in the plants, as both mutant lines displayed loss-of-function mutations in all endogenous histone H4 genes (Figure 2.2, Figure 4.1). Previous work has shown that CRISPR/Cas9-mediated frameshift mutations can, in certain instances, lead to aberrant splicing that results in shorter or altered proteins that may retain their functionality (Kapahnke et al., 2016; Lalonde et al., 2017). Therefore, it is possible that some potentially functional endogenous histone H4 protein may continue to be translated, despite the frameshift mutations in the endogenous histone H4 genes, and differences in the amount of this functional histone H4 produced may result in a less drastic phenotype displayed by one of the rH4R17A lines. However, it is important to note that histone H4 is a relatively small, yet incredibly conserved protein (Figure 2.1) and we have found that plants can fail to tolerate some single amino acid substitutions in histone H4 (Figure 2.7B). Therefore, we propose that it would be unlikely that a truncated version of histone H4 would retain functionality, as even minimal changes to histone H4 can cause lethality in plants. Another factor that may cause differences in the phenotypes between the two rH4R17A mutant lines is random T-DNA integration and copy number variation of the T-DNA insertions (De Buck et al., 2009; Gelvin, 2017). These differences in T-DNA integration could cause different amounts of total histone H4 to be expressed from the transgene (e.g., by differential epigenetic regulation), which may also explain minor differences in the phenotypes observed. Analyses of histone abundance in replacement H4 lines are discussed in more detail in the next section.

Nevertheless, a mechanism of incomplete dominance aligns with our proposed model of action for the rH4R17A mutation. When some rH4R17A mutant histones are present in the chromatin of plants, ISWI function would be inhibited at some, but not all, nucleosomes. In contrast, when all nucleosomes contain rH4R17A mutant histones, ISWI function would be more severely inhibited. Therefore, compared to the situation where rH4R17A mutant histones comprise every nucleosome, we would expect to observe an intermediate phenotype when only some histone H4 proteins contain the R17A mutation while other histone H4 proteins are wild-type.

Histone dosage compensation observed in rH4 plants

Supporting the results from our RT-qPCR experiments in Chapter 2, we observed significantly increased expression of H4 (*At3g53730*) in replacement H4 plants compared to Col, although the increase in H4 expression observed by RNA-seq corresponded to a lower set of values than those observed by RT-qPCR (Figure 4.9). With the RNA-seq data, we were further able to approximate the percentage of *At3g53730* RNA expressed from the endogenous locus versus the replacement H4 transgene. While we found that there was a 3- to 5-fold increase in *At3g53730* RNA expression in replacement H4 plants compared to Col, approximately 25 to 50% of this increase could be attributed to non-functional endogenous *At3g53730* RNA. With these results in mind, the amount of functional *At3g53730* RNA in rH4 plants was closer to a 2- to 4-fold increase in expression compared to Col.

We also assessed the RNA levels corresponding to the other core histone genes (H2A, H2B, and H3) in replacement H4 plants. We found that the RNA expression levels of H2A, H2B, and H3 were not increased in rH4-1 or rH4-2 plants, while the expression of several histone genes was upregulated in rH4R17A-1 and rH4R17A-2 plants (Figure 4.10). The increase in expression of H2A, H2B, and H3 genes in rH4R17A plants may

be due to global transcriptional changes observed in these mutants due to the misregulation of ISWI-dependent nucleosome positioning. We also observed that the RNA expression for six out of the seven endogenous H4 genes targeted in the H4 septuple mutant background was decreased in the rH4-1 and rH4-2 lines compared to Col. While two of these genes contain large deletions in the replacement H4 backgrounds (Figure 2.2), potentially causing a reduction in the number of read counts aligning to these genes, the other genes contain small indels that would not majorly impact read alignment and therefore, these mutations would not be expected to affect the calculated read counts. Therefore, this observation suggests that RNA degradation is occurring on the mutant H4 transcripts expressed from the endogenous H4 genes in the replacement H4 lines.

The wild-type expression levels of H2A, H2B, and H3 observed in rH4-1 and rH4-2 plants and the reduced expression of many H4 genes in these lines conflict with the hypothesis of histone dosage compensation occurring by transcription factors targeting all histone genes, all replication-dependent histone genes, or all histone H4 genes (see Chapter 2). Further work assessing transcription factor binding sites in histone promoters could clarify which specific histone genes would be expected to be co-regulated by transcription factors in this hypothesis. It also remains possible that the postulated degradation of H4 RNA is occurring due to compensatory mechanisms to prevent an excess of histone supply due to increased expression of certain histone genes, in which case, the above hypothesis could be correct. However, the current results seem to more strongly support the hypothesis that degradation of mutant H4 mRNA causes increased H4 expression due to transcriptional adaptation (see Chapter 2).

We further used these data to estimate the amount of functional H4 mRNA in replacement H4 plants compared to Col plants and calculated that replacement H4

plants contained significantly lower levels of functional H4 mRNA compared to Col. While these values are merely an estimate of the total levels of mRNA that could be utilized to produce functional H4 proteins, they do suggest that dosage compensation may also be occurring at the level of translation to ensure that wild-type levels of histone H4 proteins are being produced in replacement H4 plants to restore an approximately wild-type phenotype. Conversely, it may be possible that the levels of H4 mRNA observed in replacement H4 plants are sufficient for normal growth and development. Additionally, it is important to note that the mRNAs expressed from the different endogenous H4 genes are not necessarily translated at the same rates, and thus, the estimated functional H4 mRNA counts may not correlate precisely to the levels of H4 protein observed in the plants.

Finally, our results from these experiments remain somewhat puzzling given the fact that we did not observe the same degree of histone H4 dosage compensation by RNA-seq as we observed by RT-qPCR. We estimated that at the lowest end, some replacement H4 lines showed only a 2-fold increase in functional H4 (*At3g53730*) mRNA compared to Col plants, which is the same degree of increase that we calculated in the H4 septuple mutant with RT-qPCR (Figure 2.4A, Figure 2.9). However, the rescue of the H4 septuple mutant phenotype by the replacement H4 transgene strongly indicates that the amount of functional H4 in replacement H4 lines is higher than in the H4 septuple mutant background. As amplification efficiency of RT-qPCR primers is not exactly 100%, and moreover, the 2-fold increase calculated from the RNA-seq data was simply an estimation of functional H4 mRNA, these values are not entirely comparable. Additionally, one difference between the plants utilized in our RT-qPCR experiments and our RNA-seq experiments was that the remaining endogenous H4 gene had homozygous or biallelic mutations in our lines utilized in the RNA-seq experiments, while this gene was not necessarily uniformly mutated in the lines used in the RT-qPCR

experiments. Additional independent transgenic lines were also utilized in the RT-qPCR experiments compared to the RNA-seq experiments. Therefore, rH4 lines would not be expected to behave in exactly the same manner in the RT-qPCR and RNA-seq experiments. Further experiments assessing H4 protein levels in these different lines and assessing H4 mRNA and protein levels in H4 depletion backgrounds (such as the H4 septuple mutant) would further elucidate requirements for H4 expression and histone dosage compensation occurring in plants.

Genome-wide transcriptional analyses uncover multiple biological processes affected by H4R17A mutation

With our analyses of global transcriptional changes in rH4R17A mutant plants, we discovered several genomic and developmental processes regulated by H4R17, including flavonoid metabolism and biosynthesis, specification of symmetry and cell death, the cell cycle, the UV response/ the DNA damage response, and gene silencing (Figure 4.17, Figure 4.18, Figure 4.19). Interestingly, we previously observed that rH4R17A mutants displayed a higher chroma ratio than wild-type plants, indicating reduced anthocyanin content in these mutants (Figure 3.9D), which supports our RNA-seq results. Specifically, we observed downregulation of genes playing a positive role in anthocyanin accumulation and upregulation of genes playing a negative role in anthocyanin accumulation (Figure 4.18C) (Nguyen et al., 2015), supporting our hypothesis that rH4R17A mutations result in reduced anthocyanin content. As flavonoids contribute to UV protection (Li et al., 1993), the abnormal UV response pathways observed in rH4R17A mutants could be an effect of altered flavonoid content. In fact, genes related to the process of anthocyanin accumulation in response to UV light were identified by our GO term enrichment analysis in rH4R17A-1 and rH4R17A-2 mutants. Further analysis of flavonoid content in rH4R17A mutants would further clarify how the

biosynthesis of secondary metabolites is regulated by H4R17.

The low level of transposon reactivation and the upregulation of elements involved in the DNA damage response in rH4R17A mutants suggest that H4R17 plays a role in regulating genome stability. While we observed an approximately 3-fold increase in *BRCA1* expression for rH4R17A-2 plants by RNA-seq, we did not observe this same increase for rH4R17A-1 plants (Figure 4.18A). Additionally, we did not observe *BRCA1* upregulation by RT-qPCR in our initial screens of rH4 mutants (Figure 3.16). However, we did previously observe transcriptional derepression of the typically silenced genomic element *TSI* in both the rH4R17A-1 and rH4R17A-2 mutants in our initial screens (Figure 3.15). Interestingly, we also observed the upregulation of *AGO9*, which plays a role in TE silencing, with our RNA-seq analyses of rH4R17A mutants (Figure 4.18A) (Duran-Figueroa and Vielle-Calzada, 2010). Further analysis of DNA damage in rH4R17A plants through experiments such as the comet assay, as well as quantification of transposon derepression through RT-qPCR of individual TEs, would further elucidate genome instability phenotypes observed in rH4R17A mutant plants.

Altered PRMT7 function not responsible for H4R17 mutant phenotypes

While no modifications on H4R17 have been identified in plants (Brabencova et al., 2017; Zhang et al., 2007a), monomethylation of H4R17 has been identified in yeast (Crespo et al., 2020), and mono- and dimethylation of H4R17 have been identified in mammals (Luense et al., 2016; Tweedie-Cullen et al., 2012). *In vitro* studies have implicated PRMT7 as the enzyme responsible for monomethylation of H4R17 in mammals, but further studies are required to confirm this interaction *in vivo* (Feng et al., 2014b; Feng et al., 2013). PRMT7 has also been demonstrated to monomethylate H4R19 *in vitro*, although its preference was for H4R17 (Feng et al., 2013), and monomethylation of H4R19 has not been identified in mammals (Luense et al., 2016).

Due to PRMT7's low activity compared to other methyltransferases, such as PRMT1 and PRMT5, and the technical limitations associated with using mass spectrometry for detecting rare modifications or modifications at highly basic regions, such as the basic patch of histone H4, complications exist towards assessing PRMT7 substrates *in vivo* (Feng et al., 2014b; Jain and Clarke, 2019; Jain et al., 2017). Nonetheless, PRMT7 appears to have important functions in regulating development in mammals (Akawi et al., 2015; Blanc et al., 2016; Jeong et al., 2016; Ying et al., 2015), although PRMT7 has been shown to target non-histone substrates in addition to histone H4 (Feng et al., 2013; Gonsalvez et al., 2007; Jung et al., 2011). Intriguingly, monomethylation of H4R17 has also been demonstrated to significantly increase PRMT5-mediated methylation of H4R3 in mammalian cells by allosterically activating PRMT5 into a high-activity conformation (Feng et al., 2014b), although further *in vivo* analyses of the relationship between PRMT7, PRMT5, and histone H4 methylation are required.

We found that modulation of *PRMT7* expression in *Arabidopsis thaliana* failed to result in phenotypes similar to rH4R17A mutants (Figure 4.3, Figure 4.4). Additionally, neither *Arabidopsis thaliana* H4R19 mutants nor H4R3 mutants display early flowering time phenotypes or other phenotypic similarities to H4R17 mutants (Table 3.1). Because PRMT7 is the only known modifier of H4R17 in any eukaryote (Feng et al., 2014b; Feng et al., 2013; Jain and Clarke, 2019), we therefore found that it is unlikely that the rH4R17A mutant phenotypes are caused by the interference of PTM deposition on H4R17. Further analyses could confirm whether there are any PTMs on H4R17 in *Arabidopsis thaliana*, although complicating factors outlined above regarding the difficulty of mass spectrometric analyses towards this application remain. The high level of conservation of H4 across eukaryotes should allow for commercial antibodies raised against H4R17me to be used in *Arabidopsis thaliana* as an alternative method to probe for the existence of this chromatin mark in plants. However, commercial H4R17me

antibodies are currently rare and poorly validated.

H4R17 regulates nucleosome remodeling in plants

Our results support a model where H4R17 regulates various developmental processes in plants through positive regulation of the ISWI chromatin remodeling complex. Precisely how the rH4R17A and ISWI subunit mutations cause the differential gene expression that leads to the observed developmental phenotypes is one area for future investigations. Because the rH4R17A and ISWI subunit mutations cause a reduction of regular nucleosome positioning in gene bodies, and moreover, the nucleosome metaprofile corresponding to all inactive genes showed lower phasing in gene bodies than the nucleosome metaprofile corresponding to all active genes (Figure 4.21), one hypothesis could be that impaired chromatin remodeling by the ISWI complex may cause highly expressed genes to behave similarly to lowly expressed genes due to a loss of nucleosome phasing. However, our results do not support this hypothesis for several reasons. While transcription and nucleosome positioning have been shown to be highly interconnected processes (Hughes et al., 2012; Jiang and Pugh, 2009; Struhl and Segal, 2013; Workman and Kingston, 1998), we and others have observed that mutations in H4R17 and plant ISWI complex subunits affect nucleosome positioning patterns in differentially and non-differentially expressed genes alike (Li et al., 2014). This result is consistent with the idea many factors on top of nucleosome positioning in gene bodies affect the transcription level of a gene, and thus in some cases, altered genic nucleosome positioning appears to majorly impact transcription, while in others, little change is observed (Bai and Morozov, 2010; Jiang and Pugh, 2009). Moreover, processes related to genetic robustness may also serve to counteract transcriptional fluctuations due to perturbations of nucleosome positioning (Masel and Siegal, 2009). Additionally, we observed that more highly and more lowly expressed genes maintained

separation in their t-SNE embeddings in rH4R17A and ISWI subunit mutants, whereas we may expect more highly expressed genes to lose separation from more lowly expressed genes if the above hypothesis were correct (Figure 4.23, Figure 4.25). Finally, we found a surprising result from our clustering analyses that groups mainly containing inactive genes can actually show higher phasing than groups mainly containing active genes (Figure 4.23). This result diverges from previously published data as well as our own initial analyses demonstrating that the average over all inactive (or lowly expressed) genes shows decreased nucleosome phasing compared to the average over all active (or highly expressed) genes (Figure 4.21) (Li et al., 2014). We hypothesized that the low phasing previously seen in our metaprofiles of all inactive genes may be due to the averaging of several highly phased, oscillating nucleosome array patterns, rather than being an indication of broadly disorganized nucleosome arrays.

To test this hypothesis, we performed clustering analyses on all active genes and all inactive genes and found that the majority of the clusters that we identified in both groups corresponded to highly phased nucleosome arrays. Therefore, when the nucleosome occupancy over all active or all inactive genes is calculated, the averaging of the different oscillating phases of these nucleosome arrays likely leads to a degradation of the phasing pattern. As we observed that the metaprofile over all active genes displayed higher phasing than the metaprofile over all inactive genes, this result suggests that the oscillating phases of the nucleosome arrays in inactive genes are more discordant with each other compared to active genes, although further analysis is required to assess this hypothesis. Methods to quantify the phase shifts, amplitudes, and periods of the nucleosome occupancy metaprofiles corresponding to different groups of genes would help to elucidate differences in nucleosome positioning patterns between active and inactive protein-coding genes. Additionally, investigating the genes in these individual clusters in more detail could allow us to determine why specific phase shifts

are observed for different groups of genes.

Our results conflict with one of the predominant models proposing why nucleosome metaprofiles averaging all inactive genes show lower phasing than nucleosome metaprofiles averaging all active genes in multicellular eukaryotes. In this model, “barrier complexes” (e.g., transcription factors) occupy the 5' NFR of active genes and are flanked by highly phased -1 and +1 nucleosomes on either side (Chereji et al., 2019). In contrast, inactive genes lack barrier complexes upstream of the TSS due to the lack of transcription and consequently, nucleosome positioning is mainly determined by weak preference for DNA sequences. However, we observed that inactive genes and active genes both contained a 5' NFR in plants (Figure 4.21) (Li et al., 2014; Zhang et al., 2015), unlike what is observed in animals (Chereji et al., 2019; Voong et al., 2016; West et al., 2014), opposing this model. Additionally, we found that a large percentage of both active and inactive genes contained highly phased +1 nucleosomes resulting in regularly spaced nucleosome arrays in the gene body, and while we observed oscillating arrays due to phase shifts in different groups of genes, these oscillations were observed in both active and inactive genes.

As we collected the genomic DNA for these MNase-seq experiments from leaf tissue, we assessed a heterogeneous population of cells pooling different cell types and cell cycle stages in these experiments. Experiments assessing a homogenous population of cells through tissue culture could theoretically further reduce the noise of our nucleosome positioning data, although epigenetic aberrations can be induced by the process of tissue culture (Bar and Benvenisty, 2019; Nestor et al., 2015; Wang et al., 2013; Weissbein et al., 2017), complicating these analyses. Most MNase-seq experiments examining nucleosome phasing in animals have been performed in cultured cells (Chereji et al., 2019; Voong et al., 2016; West et al., 2014), which could partially explain differences that have been observed between plants and animals (e.g., in the

depletion/ lack of depletion of nucleosome occupancy around the TSS of inactive genes). Clustering analyses of animal nucleosome positioning data could elucidate whether active and inactive genes show similar nucleosome positioning patterns to plants. Studies in *S. cerevisiae* have demonstrated that transcription disrupts nucleosome arrays, leading to a reduction in phasing in the most highly expressed genes compared to less highly expressed genes (Chereji and Clark, 2018; Singh et al., 2021; Voong et al., 2016), although there are not large groups of inactive genes in yeast with which to perform a comparative analysis. These results in yeast support our finding that more lowly expressed genes can demonstrate higher phasing than more highly expressed genes (Figure 4.23D), although extensive further experiments are required to determine precisely what factors determine nucleosome positioning patterns in different groups of genes at a global scale.

Factors including DNA sequence (e.g., nucleosome destabilizing sequences in the 5' NFR and sequences with high nucleosome affinity) and chromatin remodeling complexes play important roles in the establishment of proper nucleosome positioning at protein-coding genes (Singh et al., 2021). Notably, chromatin remodeling complexes such as the ISWI complex are proposed to be necessary for restoring proper nucleosome positioning in the case of disruption by the transcriptional machinery. In the absence of regularly phased nucleosome arrays in yeast, markers of genome instability such as cryptic transcription, DNA damage, ectopic recombination, and transposon integration within the gene accumulate (Singh et al., 2021), and thus, regularly phased nucleosome arrays appear to play a protective role against genome instability. With our results, we found that *A. thaliana* rH4R17A and ISWI subunit mutants exhibited several markers of genome instability, including the alteration of cell death pathways, DNA damage response pathways, and gene silencing pathways, in line with results described in yeast.

Further analysis of the ways in which rH4R17A and ISWI subunit mutations affect nucleosome positioning at the individual gene and global level would clarify how impaired ISWI chromatin remodeling leads to the transcriptional changes observed. For example, analyzing the genes comprising the different clusters that we identified (e.g., through GO term enrichment analysis, motif analysis, and analysis of neighboring genes) could allow us to identify patterns or similarities between genes in the same cluster, elucidating targeting of the ISWI complex and differential activity of the ISWI complex on diverse sets of genes. Additional analysis methods to comparatively evaluate the severity of the reduction of nucleosome phasing in the different clusters, as high variance between nucleosome peaks was observed to complicate analysis of mean peak height, could also allow us to determine which clusters of genes are most severely affected by rH4R17A or ISWI mutations and thus, elucidate how different groups of genes are affected by ISWI chromatin remodeling.

ISWI chromatin remodeling complexes contain between two and four subunits in eukaryotes, including a conserved ATPase subunit and at least one accessory subunit (Aydin et al., 2014; Clapier and Cairns, 2009; Corona and Tamkun, 2004). Multiple types of ISWI complexes have been identified in animals, and the different accessory subunits in these complexes have been proposed to modulate the activity of the shared catalytic subunit as well as the specificity and target recognition of the complex (Aydin et al., 2014; Lusser et al., 2005; Toto et al., 2014). In plants, there are three types of ISWI complexes that have been identified: the plant-specific CHR11/17-RLT1/2-ARID5 (CRA)-type complex, the CHR11/17-DDP1/2/3-MSI3 (CDM)-type complex, and the CHR11/17-DDR1/3/4/5-DDW1 (CDD)-type complex (Tan et al., 2020). The shared ISWI catalytic subunits CHR11 and CHR17 were also recently demonstrated to act as accessory subunits of the SWR1 chromatin remodeling complex in plants (Luo et al., 2020). Given that there are multiple types of ISWI complexes in plants, we demonstrated that

mutations in the CRA-type complex demonstrate a less severe impact on nucleosome positioning at all protein-coding genes compared to rH4R17A mutations, which is in line with the model that H4R17 regulates all three types of ISWI complexes through its interaction with the catalytic subunits CHR11 and CHR17 (Clapier et al., 2001; Clapier et al., 2002; Dann et al., 2017; Fazio et al., 2005; Hamiche et al., 2001). Further characterization of the different ISWI complexes in plants, including their different targeting specificities to chromatin loci and the impact of the other identified CDM-type and CDD-type complexes on the regulation of both global transcription and nucleosome positioning, will contribute to elucidating their specific impacts on chromatin regulation.

We observed that rH4R17A mutants displayed less severe transcriptional changes than *arid5*, *rlt1/2*, and *chr11/17* mutants, with *chr11/17* mutants displaying the most drastic transcriptional defects. Although we were unable to perform MNase-seq assays with *chr11/17* mutants due to their severely reduced size and viability in our growth settings, previous MNase-seq experiments with *chr11/17* mutants also appear to show more drastic nucleosome positioning defects in this background compared to what we observed in rH4R17A mutants (Li et al., 2014); however, assays performed on plants grown concurrently would be needed to confirm this observation. The rH4R17A mutant could display a less severe transcriptional and nucleosome positioning phenotype than *chr11/17* mutants because the H4R17A mutation may not cause a complete loss of ISWI function, while abolition of the catalytic subunits of ISWI may be expected to cause a more severe impact on ISWI chromatin remodeling (and potentially an impact on SWR1 chromatin remodeling as well). While *in vitro* experiments in *M. thermophila* showed that the H4R17A mutation significantly reduced the binding of ISWI with the H4 N-terminal peptide and the ATPase activity of ISWI (Yan et al., 2016), similar ITC measurements and ATPase assays evaluating how the H4R17A mutation affects the binding and activity of plant ISWI catalytic subunits would be informative. Additionally, biochemical

experiments using the human ISWI-family enzyme ACF demonstrated that the H4K16A mutation and deletion of H4 residues 1-19 both increased the duration of the pausing phase of ISWI remodeling without affecting the translocation phase (Hwang et al., 2014) and thus, while ISWI activity is affected by H4 N-terminal tail mutations, this complex appears to retain the ability to translocate nucleosomes in the presence of H4 mutations. Moreover, we would not expect the rH4R17A mutation to completely abolish the binding of the ISWI complex to chromatin, as the ISWI catalytic subunit likely binds additional histone residues (see below) as well as the DNA itself (Yan et al., 2016; Yan et al., 2019). Accessory ISWI subunits also have chromatin-binding domains, such as ARID5, which recognizes H3K4me3 and AT-rich DNA (Tan et al., 2020). In addition, a human ISWI complex was demonstrated to associate with H4K4me3 and H4K16ac through its accessory subunit BPTF (Li et al., 2006; Ruthenburg et al., 2011; Wysocka et al., 2006). Therefore, it is likely that the ISWI complex retains some activity in the presence of the H4R17A mutation, which is supported by our results. Expressing CHR11 and CHR17 with point mutations in the two conserved H4R17-binding residues in the *chr11/17* mutant background could be a useful future experiment to test our model.

Experiments with yeast ISWI enzymes also revealed that neighboring residues on the H4 N-terminal tail (i.e., H4K16 and H4R19) likely play a role in ISWI regulation in addition to H4R17. For example, H4R19A and H4K16ac peptides were found to reduce the binding and the ATPase activity of *M. thermophila* ISWI, although to a lesser extent than the H4R17A mutation (Yan et al., 2016). Additionally, structural studies indicated that H4K16 may interact with two *M. thermophila* ISWI residues (Glu523 and Asp524), although not as strongly as H4R17 interacts with Glu474 and Asp524 (Yan et al., 2016). We observed in our experiments that the rH4K16R, rH4K16A, rH4R19K, and rH4R19A mutants failed to exhibit similar morphological or early flowering phenotypes to rH4R17A or rH4R17K mutants, although some H4K16 mutants displayed a slightly early floral

transition and/or modestly reduced rosette size (Table 3.1, Figure 3.1, Figure 3.2, Figure 3.9). Further experiments assessing global gene expression and nucleosome positioning in H4K16 and H4R19 mutants, or combinatorial mutants at H4K16, H4R17 and H4R19, could reveal whether all of these H4 residues participate in the regulation of ISWI in plants. *In vitro* experiments measuring binding and ATPase activity of plant ISWI enzymes in the presence of these different mutant histone peptides would also be informative. Additionally, residues on histone H3 (Gln76 and Asp81 in *S. cerevisiae*) have also been implicated in ISWI binding via structural studies (Yan et al., 2019) and biochemical experiments have demonstrated that residues and PTMs on and neighboring the H2A/H2B acidic patch play an important role in the regulation of ISWI chromatin remodeling activity (Dann et al., 2017; Gamarra et al., 2018). Therefore, future experiments assessing these additional ISWI-interacting residues individually or in combination with the H4R17 mutation would elucidate how different histone residues interact to regulate ISWI chromatin remodeling.

In the same thread, a screen of combinatorial mutations of H4 residues (or other histone residues) in combination with the rH4R17A mutation could comprehensively reveal whether mutations on other histone residues enhance or suppress the rH4R17A mutant phenotype. As we found that multiple chromatin remodelers displayed altered expression levels in our RNA-seq experiments, assessing the interplay of H4R17 and ISWI with other histone residues/ chromatin modifications would be an interesting future direction for this work. For example, we found that PRMT10 (which catalyzes H4R3me2s), ATX1 (which catalyzes H3K4me3), ATX2 (which catalyzes H3K4me2), and CLF (which catalyzes H3K27me3) all displayed altered expression in rH4R17A and ISWI subunit mutant backgrounds (Alvarez-Venegas et al., 2003; Niu et al., 2007; Saleh et al., 2008; Schubert et al., 2006). Therefore, H4R3, H3K4, and H3K27 would be interesting candidates to investigate further, although we did not observe evident phenotypes for

individual H4R3 mutants except for smaller rosette size (Table 3.1, Figure 3.9).

However, it is possible that combinatorial mutations of H4R3 and H4R17 could still enhance and/or suppress the rH4R17A mutant phenotype.

Interestingly, the *arid5* and *rtt1/2* mutants displayed less severe nucleosome positioning defects over all protein-coding genes but more severe transcriptional defects than rH4R17A mutants. Additional experiments characterizing how these different mutations affect the nucleosome sliding activity of the ISWI complex and how precisely the ISWI complex regulates transcription would be necessary to determine why we observe these results. However, one possibility is that the rH4R17A mutation affects the action of all types of ISWI complexes in plants, but it does not completely abolish their function. In contrast, the *arid5* and *rtt1/2* mutations could have a more severe impact on only one specific type of ISWI complex (i.e., the CRA-type complex). For example, as ARID5 recognizes H3K4me3 (Tan et al., 2020), it may play a role in directing the ISWI complex to a specific subset of genes, and thus mutation of ARID5 could cause the mislocalization of the ISWI complex. These different mechanisms of action of the rH4R17A and ISWI accessory subunit mutations could induce the differing gene expression and nucleosome positioning phenotypes that we observed, although further work assessing the interplay between ISWI chromatin remodeling and gene expression is required.

Perspectives

Our studies with the rH4R17A mutant exemplify the utility of our histone H4 replacement system for assessing histone function. Currently, the *A. thaliana* rH4R17A mutant is the only genetic system allowing researchers to study the H4R17A mutation in a complete replacement setting in multicellular eukaryotes. While this mutant would theoretically be possible to generate in *Drosophila melanogaster* using existing histone

replacement strategies, replacement H4R17A mutant flies have not yet been assessed because PTMs on H4R17 have not been identified in this organism (Zhang et al., 2019). Our results gathered with *A. thaliana* rH4R17A mutants point to the benefit of our systematic screen, where we targeted every residue that can in theory be a target for PTMs on histone H4, rather than every residue that is a confirmed target for PTMs. This result also underscores how residues that are not targets for PTMs can play important roles in the regulation of chromatin and thus, including non-modifiable residues in future screens for histone function would be informative. An additional benefit of our system in *A. thaliana* is that mutant seed stocks are relatively easy to maintain in comparison to *D. melanogaster* mutant lines, facilitating our analyses of a large number of H4 mutations. In sum, the *A. thaliana* rH4R17A mutant has provided us with a novel system to study the role of H4R17 *in vivo* in multicellular eukaryotes and using this system, we identified a conserved regulatory relationship between H4R17 and ISWI chromatin remodeling. Future work assessing molecular mechanisms for other histone H4 replacement mutant phenotypes could therefore reveal important results about the ways in which histone residues regulate chromatin.

Chapter 5: Conclusions

Author's Note: Portions of this chapter have been submitted for publication as:

Emma Tung Corcoran¹, Chantal LeBlanc¹, Mia Arias Tsang¹, Anthony Sarkiss¹, Yuzhao Hu², Ullas V. Pedmale², and Yannick Jacob¹. 2021. Systematic histone H4 replacement in *Arabidopsis thaliana* reveals a role for H4R17 in regulating flowering time.

Affiliations:

1. Yale University, Department of Molecular, Cellular and Developmental Biology, Faculty of Arts and Sciences; 260 Whitney Avenue, New Haven, Connecticut 06511, United States.

2. Cold Spring Harbor Laboratory; Cold Spring Harbor, New York 11724, United States.

Summary

Histone replacement systems represent valuable tools to directly assess the ways in which histone residues regulate different processes occurring on chromatin. However, multiple challenges have encumbered the development of histone replacement systems in plants. One major challenge is that plants contain large gene families with many copies of the core histone genes that are dispersed throughout the genome (Tenea et al., 2009). Additionally, gene targeting technologies in plants have very low efficiency compared to yeast and animal systems, mainly due to the fact that multicellular plants predominantly use the NHEJ pathway over the HR pathway for DNA repair (Fauser et al., 2012; Huang and Puchta, 2019; Puchta, 2005). Gene targeting also requires an exogenous DNA “template”, which is difficult to introduce into plant cells (Fauser et al., 2012; Huang and Puchta, 2019). However, recent improvements in gene editing technologies (i.e., CRISPR/Cas9) have facilitated the establishment of the complex histone deletion mutants that are necessary for the generation of histone replacement systems in plants. Moreover, as additional advancements are made, continued improvements and adaptations of existing histone replacement systems have become possible.

Here, we have presented the establishment of a CRISPR/Cas9-based histone mutagenesis platform in the plant model system *Arabidopsis thaliana*. As proof-of-concept, we targeted histone H4 for a systematic screen of modifiable residues on this protein. In the first generation after transformation, the histone H4 replacement plants contain loss-of-function mutations in seven out of the eight endogenous H4 genes, a gRNA targeting Cas9 to the remaining endogenous H4 gene, and a replacement H4 gene expressing wild-type or modified histone H4 that is resistant to mutagenesis by Cas9. In subsequent generations, H4 replacement plants with a complete replacement of endogenous histone H4 with replacement histone H4 can be identified by genotyping.

We found that H4 septuple mutant plants, which have loss-of-function mutations in seven out of the eight endogenous H4 genes, exhibit drastic morphological phenotypes and increased genome instability due to the lack of sufficient histone H4 dosage. However, we demonstrated that H4 replacement plants expressing the wild-type H4 replacement gene in the H4 septuple mutant background are phenotypically similar to wild-type plants and moreover, fail to exhibit any indications of increased genome instability.

After this *in vivo* validation of our platform, we directly assessed the function of 38 histone H4 residues, covering every residue on histone H4 that can in theory be a target for PTMs, to generate 63 H4 point mutants in total. Using our system, we identified many novel roles for residues on histone H4 in regulating diverse processes including flowering time, rosette morphology, endoreduplication, chromatin condensation, and gene silencing. One phenotype of interest that we identified was a previously uncharacterized role for H4R17 in regulating the floral transition. We then focused on determining a mechanism for the impact of H4R17 on the regulation of flowering time. Through phenotypic analyses and assessment of global nucleosome positioning and gene expression, we demonstrated the functional relationship between H4R17 and the ISWI chromatin remodeling complex in plants. We subsequently proposed a model where H4R17 positively regulates the catalytic subunits of the ISWI complex to establish regularly spaced nucleosome arrays in gene bodies in plants, representing a conserved mechanism in regard to the function of H4R17 in yeast and animals. Together, these studies demonstrate the utility of directly assessing the function of histone residues to reveal novel roles for histones in plants and more broadly, in all multicellular eukaryotes.

A novel system for studying histone function in plants

In this dissertation, we have presented a new histone replacement system that facilitates the analysis of histone H4 functions on diverse processes in plants. Our results serve as proof-of-concept that complete histone replacement systems can be rapidly established in *A. thaliana*. In the future, this approach may be applied to generate similar systems to probe the functions of different histones or histone variants. The histone replacement system developed in this study for histone H4 will supplement already existing systems in *S. cerevisiae* and *D. melanogaster* to offer new biological insights into the roles of H4 in plants. In its current iteration, our system already provides the most extensive coverage of H4 mutants in a multicellular eukaryote, as the histone replacement system generated in *D. melanogaster* has only been used to generate 14 H4 point mutants, compared to the 63 H4 point mutants generated with our system (Zhang et al., 2019). The incredibly high level of conservation of histone H4 between *Arabidopsis thaliana* and humans (Figure 2.1) provides one future direction for our system in its ability to be used to study the effect of histone H4 mutations implicated in human disease. Mutations in histones (“oncohistones”) have been linked to human cancers and several mutations in histone H4 have recently been identified with relatively high prevalence in tumor samples (Nacev et al., 2019). While we did generate substitution mutants at some of the implicated histone H4 residues in the experiments presented in this work, we did not generate the exact amino acid substitutions identified in the above studies. Studying the effect of these precise mutations on chromatin and development is thus a promising future direction for our histone H4 replacement system.

Another future use of the histone replacement system is to study the effects of multiple mutations on histone H4, for example, to inactivate more than one pathway. While we attempted to generate combinatorial mutants of multiple lysines on the H4 N-terminal tail in our screens, these mutations caused lethality and thus, we were unable to

perform detailed experiments assessing their impact on the regulation of chromatin. However, we identified many additional residues in our work that may be involved in the co-regulation of different biological processes and therefore would be interesting candidates for combinatorial mutations. Due to the relative ease of expressing new histone mutations in plants with our system, the generation of mutants for the above future directions should be straightforward and facilitate the rapid assessment of phenotypes caused by these mutations.

Our CRISPR-based strategy to replace endogenous histones offers several advantages over other methods that can potentially be used to achieve complete histone replacement. For example, the rapid successful generation of the H4 septuple mutant in this work demonstrates that multiplex CRISPR/Cas9 can be used to efficiently inactivate a large number of histone genes in plants (LeBlanc et al., 2017). Using CRISPR/Cas9 greatly reduces the amount of time and resources required to generate a histone depletion background, especially when compared to crossing individual histone mutants. The presence of tandem duplicated copies of histone genes (e.g., the H3.1 genes *At5g10390* and *At5g10400*) can also preclude using traditional crossing schemes to generate backgrounds lacking a specific histone or histone variant. In addition, deploying multiplex CRISPR/Cas9 to inactivate endogenous histones will allow researchers to rapidly re-establish histone replacement systems in a specific mutant background, for example, to screen for point mutations in histones that enhance or suppress a phenotype of interest. Another advantage of our histone H4 replacement strategy is that we consistently observed high expression of the H4 replacement gene, which rescues the morphological phenotype of the H4 septuple background in all of our T1 lines (Fig. 1A-C). While further work is required to determine why we observe high levels of expression of replacement H4, several factors including selection pressure during the recovery of transformants and histone dosage compensation could be contributing to

this phenomenon. Due to the above reasons, rH4 plants exclusively expressing mutant histones can predictably be obtained using our strategy.

Several changes could be implemented to improve future histone replacement systems in *A. thaliana*. To control for differential effects caused by random T-DNA integration (Gelvin, 2017), we characterized in this work two independent transgenic lines expressing each H4 replacement construct. Ideally, gene targeting would be utilized to introduce the H4 mutations directly at an endogenous histone H4 locus. While gene targeting technologies in plants are currently inefficient (Fauser et al., 2012; Huang and Puchta, 2019; Puchta, 2005), as additional improvements in gene targeting are developed, *in situ* histone replacement systems in plants analogous to platforms already existing in yeast and *Drosophila melanogaster* may also become feasible. Additionally, more precise control over the dosage of the replacement histone could also serve to improve this system. It was recently shown that while yeast histone replacement systems utilizing single-copy integrated histone genes expressing certain mutant histones cannot survive, the addition of a second copy of the mutant histone gene rescues this lethality (Jiang et al., 2017). For our system, we utilized a single endogenous histone H4 gene for the H4 replacement gene, rather than generating eight histone H4 replacement genes corresponding to each endogenous histone H4 gene present in the *A. thaliana* genome. While we observed that rH4 plants appear morphologically wild-type due to high H4 expression, it may be important to study the function of the other endogenous H4 genes, or the requirement for *A. thaliana* to have eight copies of the H4 genes in its genome. Although labor-intensive, future strategies simultaneously using multiple endogenous H4 genes as H4 replacement genes could therefore be more reflective of the H4 supply available to wild-type plants.

Compared to simply expressing mutant histones in a wild-type background, there are several advantages to completely replacing expression of endogenous histones with

mutant histones. First, complete histone replacement reveals all phenotypes caused by the histone mutation, instead of only dominant phenotypes. In certain cases, researchers may want to examine histone mutations in a dominant context, such as when studying some histone mutations that are known to cause diseases in humans but which comprise only a portion of chromatin. For example, H3K27M mutant histones were overexpressed in flies in a wild-type background as a model to study the role of this dominant gain-of-function mutation in pediatric brain cancers (Herz et al., 2014). However, there are many cases where researchers may want to perform broader screens that can reveal roles for histone residues that are more subtle. While we observed that the H4R17A mutation displayed incomplete dominance in plants, expression of wild-type histone H4 may completely restore a normal phenotype in the presence of other mutant H4 proteins assessed in our screens and thus, we may miss mutant phenotypes by studying histone mutations in a wild-type or partial replacement background. As such, researchers may want to study genetic systems where the mutated histone is the only protein pertaining to that histone type that is present in chromatin to eliminate extenuating factors in the experiment. On the other hand, histone mutations are more likely to cause lethality or sterility in a complete replacement background and consequently can be more difficult or impossible to study. In addition, it is possible that expressing mutant histones in *Col* may lead to cytotoxicity due to the excess accumulation of histone supply (Gunjan and Verreault, 2003; Meeks-Wagner and Hartwell, 1986; Singh et al., 2009; Singh et al., 2010). However, various mechanisms exist to combat the accumulation of excess histones and can serve to regulate histone supply under conditions of histone overexpression (Gunjan and Verreault, 2003; Kaygun and Marzluff, 2005a, b; Singh et al., 2009; Sittman et al., 1983). Finally, one disadvantage of histone replacement systems is that they are more labor-intensive to establish. Nonetheless, once these mutant lines are generated, they can serve as a

valuable resource for answering questions in chromatin biology. For example, with the existing library of rH4 mutants described in this work, many other screens can be performed to assess the impact of H4 residues on regulating various biological processes, such as stress responses and developmental processes. Consequently, this histone replacement system serves as a valuable resource for the plant research and epigenetics research communities, as it represents the largest collection of H4 point mutants in any multicellular organism.

Materials and Methods

Plant materials

All *Arabidopsis* plants were derived from the Columbia ecotype and grown in Pro-Mix BX Mycorrhizae soil under cool-white fluorescent lights (approximately $100 \mu\text{mol m}^{-2} \text{s}^{-1}$). Seeds were surface-sterilized with a 70% ethanol, 0.1% Triton solution for 5 minutes, and then with 95% ethanol for one minute. Seeds were spread on sterilized paper and plated on 0.5% Murashige-Skoog (MS) plates. Seeds were stratified in the dark at 4°C for 2 to 4 days, transferred to the growth chamber for 5 days, and then transplanted to soil. Plants grown in long-day conditions were grown for 16 hours light/ 8 hours dark and plants grown in short-day conditions were grown for 8 hours light/ 16 hours dark.

The *chr11* (GK-424F01)/ *chr17* (GK-424F04) double mutant was described previously (Li et al., 2012). The *arid5* (SALK_111627), *prmt7-1* (SALK_028160), and *prmt7-2* (SALK_039529) T-DNA insertion mutants were obtained from the Arabidopsis Biological Resource Center. The *pie1* T-DNA insertion mutants were initially obtained from the Arabidopsis Biological Resource Center (SALK_096434) and the *pie1* mutants used in this work were seeds collected from homozygous *pie1* plants. The *rlt1* (SALK_099250)/ *rlt2* (SALK_132828) double mutants were generated by crossing. Due to severely reduced fertility, *chr11/17* and *arid5* mutants were maintained in a heterozygous state.

Generation of transgenic *Arabidopsis* plants

Binary vectors were transformed into *Agrobacterium tumefaciens* (strain GV3101) using heat shock and plants were transformed with these constructs using the floral dip procedure as described previously (Clough and Bent, 1998). Transgenic plants for the generation of the H4 septuple mutant were selected on 0.5 MS plates containing

1% sucrose, carbenicillin (200 $\mu\text{g ml}^{-1}$) and kanamycin (100 $\mu\text{g ml}^{-1}$). Transgenic H4 replacement plants were selected on 0.5 MS plates containing 1% sucrose, carbenicillin (200 $\mu\text{g ml}^{-1}$) and glufosinate ammonium (25 $\mu\text{g ml}^{-1}$). Plants were subjected to heat stress treatments as described previously (LeBlanc et al., 2017). The plants were grown continuously at 22°C from that point on.

Flowering time

Days to flower was measured when a 1 cm bolting stem was visible. Rosette leaves were counted at day of flowering.

Rosette measurements

Rosette area was measured using the ARADEEPOPSIS workflow (Huther et al., 2020). The chroma ratio was calculated as follows: $2 \times \text{mean green intensity} / (\text{mean blue intensity} + \text{mean red intensity})$

Plasmid construction

CRISPR constructs used to generate the H4 septuple mutant were inserted into the *pYAO-Cas9-SK* vector as described previously (Yan et al., 2015).

The H4 replacement plasmid was made by amplifying the promoter (967 bp upstream of start codon), gene body, and terminator (503 bp downstream of stop codon) of H4 (*At3g53730*) into pENTR/D (ThermoFisher Scientific, Waltham, MA, USA). Site-directed mutagenesis of *pH4::H4* in pENTR/D using QuikChange II XL (Agilent Technologies, Santa Clara, CA, USA) was first performed to create plasmids with 10 silent mutations in the H4 coding sequence. These silent mutations were engineered to test the resistance of the H4 replacement gene against multiple gRNAs. Additional site-

directed mutagenesis of this vector was performed to generate a library of 63 H4 point mutant genes.

Each *pH4::H4* sequence was then transferred into the binary vector pB7WG, containing the H4 gRNA, using Gateway Technology. The binary vector pB7WG containing the H4 gRNA was generated as follows: The AtU6-26-gRNA vector containing the gRNA targeting H4 (*At3g53730*) was first digested with the restriction enzymes *SpeI* and *NheI*, and the digestion products were run on a 1% agarose gel. Then, the band containing the H4 gRNA was cut out and ligated into the binary vector pB7WG, which had been digested with the restriction enzyme *SpeI*.

The PRMT7 overexpression construct was created by cloning the genomic *PRMT7* gene (from ATG to stop codon, including introns) into pDONR207, and then subcloning the gene into pMDC32 (Curtis and Grossniklaus, 2003).

DNA extraction, PCR and sequencing analyses

Genomic DNA was extracted from *Arabidopsis* plants by grinding one leaf in 500 μ l of Extraction Buffer (200 mM of Tris-HCl pH8.0, 250 mM NaCl, 25 mM ethylenediaminetetraacetic acid (EDTA) and 1% SDS). Phenol/chloroform (50 μ l) was added and tubes were vortexed, followed by centrifugation for 10 minutes at 3220g. The supernatant was transferred to a new tube and 70 μ l of isopropanol was added, followed by centrifugation for 10 minutes at 3200g. The supernatant was removed and the DNA pellets were resuspended in 100 μ l of water.

PCR products were sequenced and analyzed using Sequencher 5.4.6 (Gene Codes Corporation, Ann Arbor, MI, United States) to identify CRISPR-induced mutations. To assess the rate of mutation of the remaining endogenous H4 gene (*At3g53730*) in the H4 septuple mutant by the gRNA in the H4 replacement plasmid, endogenous H4 PCR products were cloned into TOPO TA cloning vectors (Invitrogen,

Carlsbad, CA, United States). Ten to sixteen individual clones corresponding to each plant were sequenced.

RT-qPCR

RNA was extracted from 4-week-old leaf tissue with TRIzol (Invitrogen) and DNase treated using RQ1 RNase-Free DNase (Promega, Madison, WI, USA). Three biological replicates were assessed. SuperScript II Reverse Transcriptase (Invitrogen) was used to produce cDNA. Reverse transcription was initiated using random hexamers (Applied Biosystems, Foster City, CA, United States). Quantification of cDNA was done by real-time PCR using a CFX96 Real-Time PCR Detection System (Bio-Rad, Hercules, CA, USA) with KAPA SYBR FAST qPCR Master Mix (2x) Kit (Kapa Biosystems, Wilmington, MA, USA). Relative quantities were determined by using a comparative C_t method as follows: $\text{Relative quantity} = 2^{-(C_t \text{ GOI unknown} - C_t \text{ normalizer unknown}) - (C_t \text{ GOI calibrator} - C_t \text{ normalizer calibrator})}$, where GOI is the gene of interest (Livak and Schmittgen, 2001). *ACTIN7* (*AT5G09810*) was used as the normalizer.

Flow cytometry

Rosette leaves from four-week-old plants were finely chopped in 0.5 ml Galbraith buffer (45 mM MgCl_2 , 20 mM MOPS, 30 mM sodium citrate, 0.1% Triton X-100, 40 $\mu\text{g}/\mu\text{l}$ RNase A) using a razor blade. The lysate was filtered through a 30 μm mesh (Sysmex Partec, Gorlitz, Germany). Propidium iodide (Sigma, St. Louis, MO) was added to each sample to a concentration of 20 $\mu\text{g}/\text{ml}$ and vortexed for 3 seconds. Each sample was analyzed using a BD FACS LSR Fortessa X20 (Becton Dickinson, Franklin Lakes, NJ). Quantification (nuclei counts and robust CV values) was performed using Flowjo 10.0.6 (Tree Star, Ashland, Oregon).

Immunostaining

Leaves from 4-week-old plants were fixed in 3.7% formaldehyde in cold Tris buffer (10 mM Tris-HCl pH 7.5, 10 mM NaEDTA, 100 mM NaCl) for 20 minutes. Formaldehyde solution was removed and leaves were washed twice for 10 minutes in Tris buffer. The leaves were then finely chopped with razor blade in 500 μ l LB01 buffer (15 mM Tris-HCl pH7.5, 2 mM NaEDTA, 0.5 mM spermine-4HCl, 80 mM KCl, 20 mM NaCl and 0.1% Triton X-100). The lysate was filtered through a 30 μ m mesh (Sysmex Partec, Gorlitz, Germany). 5 μ l of lysate was added to 10 μ l of sorting buffer (100 mM Tris-HCl pH 7.5, 50 mM KCl, 2mM MgCl₂, 0.05% Tween-20 and 5% sucrose) and spread onto a coverslip until dried. Cold methanol was added onto each coverslip for 3 minutes, then rehydrated with TBS-Tx (20 mM Tris, pH 7.5, 100 mM NaCl, 0.1% Triton X-100) for 5 minutes. The coverslips were mounted onto slides with Vectashield mounting medium DAPI (Vector Laboratories, Burlingame, CA).

Nuclei were imaged on a Nikon Eclipse Ni-E microscope with a 100X CFI PlanApo Lamda objective (Nikon). Digital images were obtained using an Andor Clara camera. Z-series optical sections of each nucleus were obtained at 0.3 μ m steps. Images were deconvolved by ImageJ using the deconvolution plugin.

Dimensionality reduction and clustering

For flowering time data, principal component analysis of four variables (day number in long-day, leaf number in long-day, day number in short-day, and leaf number in short-day) was performed. For rosette morphology, principal component analysis of 20 variables was performed. We centered variables at mean 0 and set the standard deviation to 1. We took the first two principal components, which explained 98% of the variance for flowering time data and 74% of the variance for rosette morphology data. *k*-means clustering was performed 40 times with random initializations on the first two

principal components to identify three clusters. Analyses were performed in RStudio with R version 3.6.1 (Team, 2018).

For nucleosome positioning data, t-SNE of 200 variables corresponding to 10-bp bins centered around the TSS was performed. *k*-means clustering was performed on Col nucleosome positioning data that was centered at mean 0 with standard deviation set to 1. Clusters with fewer than 10 genes were discarded. Analyses were implemented using Scikit-learn (Pedregosa et al., 2011) in JupyterLab (Python version 3.8.8).

Endoreduplication index

To quantify variation in endoreduplication, we calculated the endoreduplication index as described previously (Sterken et al., 2012). The endoreduplication index was calculated as the weighted number of endoreduplication cycles per nucleus: $(0 \times \%2C) + (1 \times \%4C) + (2 \times \%8C) + (3 \times \%16C) + (4 \times \%32C)$

Next-generation sequencing library preparation

RNA-seq and MNase-seq libraries were prepared at the Yale Center for Genome Analysis (YCGA). Leaves of 4-week-old plants grown in short-day conditions were frozen in liquid nitrogen, ground with a mortar and pestle, and then RNA was extracted using the RNeasy Plant Mini Kit (Qiagen, Hilden, Germany). RNA quality was confirmed through analysis of Agilent Bioanalyzer 2100 electropherograms. Library preparation was performed using Illumina's TruSeq Stranded Total RNA with Ribo-Zero Plant in which samples were normalized with a total RNA input of 1 µg and library amplification with 8 PCR cycles. MNase-digested DNA was collected as described previously (Pajoro et al., 2018) with the following modifications: 2 g of leaf tissue from 4-week-old plants grown in short-day conditions was ground in liquid nitrogen and resuspended in 20 ml of lysis buffer for 15 minutes at 4°C. The resulting slurry was filtered through a 40 µm cell

strainer into a 50 ml tube. Samples were centrifuged for 20 minutes at 3200g. The resulting pellets were resuspended in 10 ml of HBB buffer and centrifuged for 10 minutes at 1500g. Pellets were successively washed in 5 ml wash buffer and 5 ml reaction buffer. MNase-seq library preparation was performed using the KAPA Hyper Library Preparation kit (KAPA Biosystems, Part#KK8504). For each biological replicate, pooled leaf tissue from three plants was used. Libraries were validated using Agilent Bioanalyzer 2100 Hisense DNA assay and quantified using the KAPA Library Quantification Kit for Illumina® Platforms kit. Sequencing was done on an Illumina NovaSeq 6000 using the S4 XP workflow.

RNA-seq processing and analysis

Two independent biological replicates for Col, rH4-1, rH4-2, rH4R17A-1, rH4R17A-2, *arid5*, *rlt1/2*, *chr11/17*, and *pie1* were sequenced. Paired-end reads were filtered and trimmed using Trim Galore! (version 0.5.0) with default options for quality (<https://github.com/FelixKrueger/TrimGalore>). The resulting data sets were aligned to the Araport11 genome (Cheng et al., 2017) using STAR (version 2.7.2a) allowing 2 mismatches (--outFilterMismatchNmax 2) (Dobin et al., 2013). Protein-coding genes were defined as described in the Araport11 genome annotation (Cheng et al., 2017). The program featureCounts (version 1.6.4) (Liao et al., 2014) was used to count the paired-end fragments overlapping with the annotated protein-coding genes. Differential expression analysis of protein-coding genes was performed using DESeq2 version 1.26 (Love et al., 2014) on raw read counts to obtain normalized fold changes and *Padj*-values for each gene. Genes were considered to be differentially expressed if they showed $>\pm 1.5$ -fold-change and *Padj*-value < 0.05 . TPM (transcripts per million) values were calculated for TEs. TEs were considered to be upregulated if they showed > 2 -fold upregulation compared to Col and had a TPM value > 5 in both biological replicates.

Heatmaps, Venn diagrams, correlation plots and correlation matrices were plotted using RStudio with R version 3.6.1 (Team, 2018). GO term enrichment analysis was performed in R with clusterProfiler using `pvalueCutoff = 0.05` and `qvalueCutoff = 0.10` (version 3.0.4) (Yu et al., 2012). Consistency between biological replicates was confirmed by Spearman correlation using deepTools2 (version 2.7.15) (Ramirez et al., 2016). deepTools2 was used to generate bam coverage profiles for visualization with Integrative Genomics Viewer version 2.8.9 (Robinson et al., 2011).

MNase-seq processing and analysis

Two independent biological replicates for Col, rH4-2, rH4R17A-1, *arid5*, and *rtt1/2* were sequenced. Paired-end reads were filtered and trimmed using Trim Galore! (version 0.5.0) with default options for quality (<https://github.com/FelixKrueger/TrimGalore>). Bowtie2 version 2.4.2 (Langmead and Salzberg, 2012) was used to align the reads to the Araport11 genome (Cheng et al., 2017) with the `--very-sensitive` parameter. Protein-coding genes were defined as described in the Araport11 genome annotation (Cheng et al., 2017). Duplicate reads were removed using Picard toolkit version 2.9.0 (toolkit., 2019) (`MarkDuplicates` with `REMOVE_DUPLICATES=true`) and the insertion size was filtered from 140 bp to 160 bp using SAMtools version 1.11 (Li et al., 2009). The average nucleosome occupancy corresponding to the regions 1 kb upstream and downstream of the TSS of all protein-coding genes was calculated using the `bamCoverage` (`--MNase` parameter specified) and `computeMatrix` functions of deepTools2 version 2.7.15 (Ramirez et al., 2016). Normalization was performed by scaling with the effective library size calculated by the `calcNormFactors` function using edgeR version 3.28.1 (Robinson et al., 2010). Consistency between biological replicates was confirmed by Spearman correlation using deepTools2. Fold change in Δ Nucleosome Occupancy at each nucleosome peak relative

to Col was calculated with a custom Python script (https://github.com/etc27/MNase-seq-workflow/analysis/peak_height) as follows: $\Delta\text{Nucleosome Occupancy} = \text{peak maximum} - (\text{5' peak minimum} + \text{3' peak minimum})/2$. The 3' peak minimum was used as the minimum for the +1 nucleosome.

Model building

The homology model for *Arabidopsis thaliana* CHR11 (a.a. 176-706) was built with Swiss-Model against the *Myceliophthora thermophila* ISWI reference structure (5JXR) (Biasini et al., 2014; Yan et al., 2016).

Primers

All primers used in this study are listed in Supplemental Table 6.

Statistical analyses

All measurements with statistical analyses were taken from multiple independent biological replicates. Means and error bars are indicated in the figure legends and are measured from the indicated sample sizes. Appropriate statistical tests were applied and statistical significance was defined as specified in figure legends ($p < 0.05$ or $p < 0.01$).

Data availability statement

Raw RNA-seq and MNase-seq data have been deposited in the Gene Expression Omnibus database with the accession code GSE190317. Statistics for mapping and coverage of the NGS data are provided in Supplemental Table 7.

Accession numbers

Accession numbers of genes reported in this study include: AT3G53730 (*H4*), AT1G07660 (*H4*), AT1G07820 (*H4*), AT2G28740 (*H4*), AT3G45930 (*H4*), AT5G59690 (*H4*), AT3G46320 (*H4*), AT5G59970 (*H4*), AT5G65360 (*HTR1*, H3.1), AT1G09200 (*HTR2*, H3.1), AT4G40030 (*HTR4*, H3.3), AT4G40040 (*HTR5*, H3.3), AT5G10980 (*HTR8*, H3.3), AT1G51060 (*HTA10*, H2A), AT1G08880 (*HTA5*, H2A.X), AT5G59870 (*HTA6*, H2A.W), AT5G27670 (*HTA7*, H2A.Z), AT1G52740 (*HTA9*, H2A.Z), AT5G22880 (*HTB2*, H2B), AT5G59910 (*HTB4*, H2B), AT2G37470 (*HTB5*, H2B), AT3G45980 (*HTB9*, H2B), AT3G46030 (*HTB11*, H2B), AT4G21070 (*BRCA1*), AT1G65480 (*FT*), AT2G45660 (*SOC1*), AT4G16570 (*PRMT7*), AT3G06400 (*CHR11*), AT5G18620 (*CHR17*), AT3G43240 (*ARID5*), AT1G28420 (*RLT1*), AT5G44180 (*RLT2*), AT3G12810 (*PIE1*), AT5G60910 (*FUL*), AT2G23380 (*CLF*), AT1G66650 (*STU_bL4*), AT1G04870 (*PRMT10*), AT2G31650 (*ATX1*), AT1G05830 (*ATX2*), AT5G65050 (*MAF2*), AT5G65060 (*MAF3*), AT2G03500 (*EFM*), AT1G79430 (*APL*), AT5G21150 (*AGO9*), AT5G66130 (*RAD17*), AT3G01330 (*DEL3*), AT3G12280 (*RBR*), AT5G67260 (*CYCD3;2*), AT1G70000 (*MYBD*), AT5G47390 (*MYBH*), and AT5G09810 (*ACTIN7*).

Supplemental Tables

Biological Processes Enriched in rH4R17A-1 mutants	p.adjust	Count
pattern specification process	3.17E-08	52
determination of bilateral symmetry	1.56E-07	27
specification of symmetry	1.56E-07	27
response to UV	1.58E-07	40
organic anion transport	1.71E-07	45
organic acid transport	3.70E-07	42
carboxylic acid transport	3.70E-07	42
flower morphogenesis	2.03E-06	18
amino acid transport	2.03E-06	38
pigmentation	2.54E-06	23
pigment accumulation	2.54E-06	23
pigment accumulation in response to UV light	2.54E-06	23
pigment accumulation in tissues in response to UV light	2.54E-06	23
pigment accumulation in tissues	2.54E-06	23
anthocyanin accumulation in tissues in response to UV light	2.54E-06	23
xylem and phloem pattern formation	5.08E-06	21
flavonoid metabolic process	8.91E-06	35
flavonoid biosynthetic process	1.54E-05	32
meristem maintenance	4.01E-05	37
shoot system morphogenesis	4.10E-05	52
meristem initiation	0.0001887	25
response to gibberellin	0.00032577	24
phenylpropanoid metabolic process	0.00020403	26
xylem development	0.00107481	15
anthocyanin-containing compound metabolic process	0.00121694	15
programmed cell death	0.00127349	45
cell death	0.00159375	48
anthocyanin-containing compound biosynthetic process	0.00165102	12
meristem structural organization	0.00168174	30
leaf morphogenesis	0.00193551	26
root hair elongation	0.00387671	23
root hair cell development	0.00468721	24
axis specification	0.00589279	14
regulation of meristem growth	0.01278699	20
polarity specification of adaxial/abaxial axis	0.01389042	12
regulation of meristem development	0.01439659	24

Biological Processes Enriched in rH4R17A-1 mutants	p.adjust	Count
pigment metabolic process	0.01483195	33
positive regulation of flavonoid biosynthetic process	0.01635689	14
adaxial/abaxial axis specification	0.01643902	12
specification of axis polarity	0.01643902	12
phenylpropanoid biosynthetic process	0.0169028	17
cellular response to red light	0.0169028	4
gibberellic acid mediated signaling pathway	0.01784692	12
gibberellin mediated signaling pathway	0.01951612	12
meristem growth	0.01992772	20
response to hydrogen peroxide	0.02221529	21
adaxial/abaxial pattern specification	0.0279999	12
cotyledon morphogenesis	0.02840673	5
regulation of programmed cell death	0.02933355	35
response to light intensity	0.0297311	27
regulation of cell size	0.03141515	9
leaf development	0.0332286	37
regulation of cell death	0.03746393	35
pigment biosynthetic process	0.04537603	25
root cap development	0.04537603	4

Supplemental Table 1 Selected biological pathways identified with GO term enrichment analysis in rH4R17A-1 mutants.

Biological Processes Enriched in rH4R17A-2 mutants	p.adjust	Count
flavonoid biosynthetic process	4.41E-09	40
flavonoid metabolic process	6.56E-09	42
response to gibberellin	1.42E-08	33
response to UV	1.89E-08	41
response to salicylic acid	2.30E-08	61
response to water deprivation	2.21E-07	54
response to water	3.77E-07	54
response to UV-B	3.12E-06	22
cell death	2.21E-05	55
programmed cell death	2.50E-05	51
anthocyanin-containing compound biosynthetic process	3.99E-05	15
anthocyanin-containing compound metabolic process	3.99E-05	18
regulation of flavonoid biosynthetic process	0.00013345	21
regulation of programmed cell death	0.00029788	43
regulation of cell death	0.00041307	43
phenylpropanoid metabolic process	0.00047144	25
positive regulation of flavonoid biosynthetic process	0.00087651	17
circadian rhythm	0.00103196	23
gibberellic acid mediated signaling pathway	0.00201712	14
gibberellin mediated signaling pathway	0.0022728	14
cellular response to red or far red light	0.00245134	10
regulation of hydrogen peroxide metabolic process	0.00260193	23
red or far-red light signaling pathway	0.00717385	9
cuticle development	0.01326372	9
response to ionizing radiation	0.01633143	13
secondary metabolite biosynthetic process	0.01768475	30
pigment biosynthetic process	0.02426568	26
regulation of anthocyanin metabolic process	0.02524047	7
pigmentation	0.0281034	14
pigment accumulation	0.0281034	14
pigment accumulation in response to UV light	0.0281034	14
pigment accumulation in tissues in response to UV light	0.0281034	14
pigment accumulation in tissues	0.0281034	14
anthocyanin accumulation in tissues in response to UV light	0.0281034	14
phenylpropanoid biosynthetic process	0.03051204	16
cellular response to radiation	0.03106486	11
pattern specification process	0.03215051	31
double-strand break repair	0.03696349	13

Biological Processes Enriched in rH4R17A-2 mutants	p.adjust	Count
regulation of seed germination	0.03729301	10
regulation of seedling development	0.04820239	10

Supplemental Table 2 Selected biological pathways identified with GO term enrichment analysis in rH4R17A-2 mutants.

Biological Processes Enriched in <i>arid5</i> mutants	p.adjust	Count
response to wounding	2.07E-15	129
flavonoid metabolic process	2.07E-15	103
flavonoid biosynthetic process	2.07E-15	95
mitotic cell cycle process	2.47E-15	104
cytokinesis by cell plate formation	1.31E-14	89
cytokinetic process	1.31E-14	89
mitotic cytokinetic process	1.31E-14	89
mitotic cytokinesis	3.01E-14	90
cytokinesis	4.93E-14	90
cytoskeleton-dependent cytokinesis	4.93E-14	90
regulation of DNA replication	2.48E-13	65
microtubule cytoskeleton organization	1.08E-12	96
meristem maintenance	1.10E-11	105
phenylpropanoid metabolic process	1.71E-10	74
regulation of cell cycle phase transition	2.72E-10	38
regulation of mitotic cell cycle phase transition	2.72E-10	38
cell cycle phase transition	4.82E-10	39
mitotic cell cycle phase transition	4.82E-10	39
mitotic cell cycle	4.88E-10	130
DNA replication	5.14E-10	105
histone lysine methylation	5.80E-10	87
histone phosphorylation	9.44E-10	35
cell division	9.44E-10	113
pattern specification process	9.44E-10	115
response to UV	1.20E-09	89
positive regulation of flavonoid biosynthetic process	1.78E-09	48
secondary metabolite biosynthetic process	2.21E-09	106
regulation of G2/M transition of mitotic cell cycle	2.56E-09	35
regulation of cell cycle G2/M phase transition	2.56E-09	35
histone modification	3.10E-09	116
cell death	3.93E-09	149
regulation of flavonoid biosynthetic process	4.13E-09	53
G2/M transition of mitotic cell cycle	4.13E-09	35
cell cycle G2/M phase transition	4.13E-09	35
histone H3-K9 methylation	5.59E-09	70
histone H3-K9 modification	5.59E-09	70
gene silencing	7.37E-09	122
regulation of cell cycle process	1.12E-08	57

Biological Processes Enriched in <i>arid5</i> mutants	p.adjust	Count
chromatin silencing	1.26E-08	85
DNA-dependent DNA replication	1.64E-08	87
negative regulation of gene expression, epigenetic	2.34E-08	85
regulation of growth	2.62E-08	100
programmed cell death	3.35E-08	134
cell proliferation	3.69E-08	89
polarity specification of adaxial/abaxial axis	3.71E-08	38
regulation of cell cycle	3.71E-08	96
specification of symmetry	4.62E-08	51
regulation of meristem growth	6.37E-08	64
adaxial/abaxial axis specification	7.82E-08	38
specification of axis polarity	7.82E-08	38
DNA methylation or demethylation	8.08E-08	68
determination of bilateral symmetry	9.57E-08	50
histone methylation	1.02E-07	94
axis specification	1.15E-07	41
regulation of cell death	1.43E-07	121
regulation of programmed cell death	1.51E-07	119
regulation of developmental growth	1.64E-07	90
cytoskeleton organization	2.11E-07	140
plant organ formation	2.19E-07	79
adaxial/abaxial pattern specification	2.62E-07	39
meristem growth	3.53E-07	64
regulation of mitotic cell cycle	4.24E-07	43
phenylpropanoid biosynthetic process	4.55E-07	53
regulation of meristem development	5.52E-07	75
anatomical structure formation involved in morphogenesis	5.97E-07	104
gene silencing by RNA	1.27E-06	95
chromatin silencing by small RNA	1.57E-06	46
sepal development	3.34E-06	33
leaf development	6.72E-06	120
methylation-dependent chromatin silencing	7.17E-06	45
circadian rhythm	9.44E-06	58
anthocyanin-containing compound metabolic process	9.44E-06	35
xylem and phloem pattern formation	1.16E-05	39
sepal morphogenesis	1.18E-05	31
petal formation	1.18E-05	31
sepal formation	1.18E-05	31

Biological Processes Enriched in <i>arid5</i> mutants	p.adjust	Count
regulation of flower development	1.91E-05	102
floral organ formation	3.06E-05	57
regulation of shoot system development	3.32E-05	109
DNA replication initiation	3.70E-05	28
shoot system morphogenesis	3.76E-05	126
pigmentation	4.89E-05	41
pigment accumulation	4.89E-05	41
pigment accumulation in response to UV light	4.89E-05	41
pigment accumulation in tissues in response to UV light	4.89E-05	41
pigment accumulation in tissues	4.89E-05	41
anthocyanin accumulation in tissues in response to UV light	4.89E-05	41
anthocyanin-containing compound biosynthetic process	6.73E-05	26
multidimensional cell growth	7.54E-05	40
meristem initiation	9.28E-05	55
regulation of post-embryonic development	0.000100319	127
response to reactive oxygen species	0.000234537	88
cell fate specification	0.000241367	24
flower morphogenesis	0.000276594	27
DNA endoreduplication	0.000277117	41
response to gibberellin	0.000309926	52
petal morphogenesis	0.000316519	32
cell cycle DNA replication	0.000331525	41
xylem development	0.000376107	31
petal development	0.000578042	38
corolla development	0.000578042	38
leaf morphogenesis	0.001130627	61
epithelium development	0.001342668	16
post-embryonic plant organ morphogenesis	0.001412554	76
embryo sac development	0.001459465	68
leaf vascular tissue pattern formation	0.001659463	13
floral organ morphogenesis	0.00197777	64
response to UV-B	0.002094214	34
epidermal cell fate specification	0.002520677	9
regulation of nuclear division	0.002520677	9
reactive oxygen species metabolic process	0.002522379	95
response to hydrogen peroxide	0.002596781	55
nuclear division	0.002922132	67

Biological Processes Enriched in <i>arid5</i> mutants	p.adjust	Count
root hair elongation	0.003626285	53
response to red or far red light	0.003628088	102
regulation of cell size	0.004345922	21
epidermis development	0.004488477	14
epidermal cell differentiation	0.004488477	14
epithelial cell differentiation	0.004488477	14
meristem structural organization	0.006053699	68
gamete generation	0.006088215	56
positive regulation of growth	0.006206977	7
aging	0.006695179	44
root hair cell development	0.011097312	54
regulation of hydrogen peroxide metabolic process	0.012398879	50
leaf senescence	0.017564026	23
plant organ senescence	0.017564026	23
cellular response to oxidative stress	0.017564026	18
de-etiolation	0.017869998	6
posttranscriptional gene silencing	0.023246647	51
regulation of cell growth	0.023265198	32
regulation of DNA endoreduplication	0.028522596	10
removal of superoxide radicals	0.030480108	6
cellular response to oxygen radical	0.030480108	6
cellular response to superoxide	0.030480108	6
cellular oxidant detoxification	0.030480108	6
cellular detoxification	0.030480108	6
female gamete generation	0.031979532	39
RNA interference	0.032538698	46
hydrogen peroxide metabolic process	0.034244294	79
maintenance of meristem identity	0.036639274	18
inflorescence development	0.037308444	10
regulation of DNA-dependent DNA replication	0.037308444	10

Supplemental Table 3 Selected biological pathways identified with GO term enrichment analysis in *arid5* mutants.

Biological Processes Enriched in <i>rlt1rlt2</i> mutants	p.adjust	Count
specification of symmetry	6.27E-16	68
determination of bilateral symmetry	9.20E-16	67
polarity specification of adaxial/abaxial axis	2.97E-14	49
regulation of meristem growth	6.30E-14	80
adaxial/abaxial axis specification	6.44E-14	49
specification of axis polarity	6.44E-14	49
axis specification	6.58E-14	53
mitotic cell cycle process	6.59E-14	103
cell death	8.48E-14	170
meristem maintenance	2.20E-13	112
pattern specification process	2.43E-13	129
cytokinesis by cell plate formation	2.70E-13	88
cytokinetic process	2.70E-13	88
mitotic cytokinetic process	2.70E-13	88
cytokinesis	3.53E-13	90
cytoskeleton-dependent cytokinesis	3.53E-13	90
shoot system morphogenesis	3.53E-13	161
meristem growth	3.53E-13	80
adaxial/abaxial pattern specification	3.62E-13	50
flavonoid metabolic process	4.42E-13	99
mitotic cytokinesis	4.55E-13	89
flavonoid biosynthetic process	4.69E-13	91
programmed cell death	6.12E-13	154
regulation of cell death	1.01E-11	139
regulation of cellular response to stress	1.01E-11	134
regulation of programmed cell death	1.82E-11	136
meristem initiation	4.90E-11	73
regulation of meristem development	6.78E-11	88
regulation of growth	2.29E-10	109
regulation of developmental growth	1.59E-09	99
phenylpropanoid metabolic process	1.75E-09	73
positive regulation of flavonoid biosynthetic process	1.75E-09	49
secondary metabolite biosynthetic process	1.87E-09	109
regulation of DNA replication	3.55E-09	58
regulation of flavonoid biosynthetic process	5.47E-09	54
cytoskeleton organization	1.10E-08	150
histone phosphorylation	1.13E-08	34
regulation of cell cycle phase transition	1.21E-08	36

Biological Processes Enriched in <i>rlt1rlt2</i> mutants	p.adjust	Count
regulation of mitotic cell cycle phase transition	1.21E-08	36
cell cycle phase transition	1.94E-08	37
mitotic cell cycle phase transition	1.94E-08	37
xylem and phloem pattern formation	1.99E-08	46
regulation of G2/M transition of mitotic cell cycle	2.66E-08	34
regulation of cell cycle G2/M phase transition	2.66E-08	34
leaf morphogenesis	3.08E-08	78
G2/M transition of mitotic cell cycle	4.38E-08	34
cell cycle G2/M phase transition	4.38E-08	34
mitotic cell cycle	5.30E-08	126
cell division	5.42E-08	110
cell proliferation	9.61E-08	90
circadian rhythm	1.07E-07	65
leaf development	1.16E-07	131
histone lysine methylation	1.61E-07	82
flower morphogenesis	1.62E-07	34
histone H3-K9 methylation	3.29E-07	67
histone H3-K9 modification	3.29E-07	67
meristem structural organization	1.74E-06	85
anatomical structure arrangement	1.74E-06	86
reactive oxygen species metabolic process	1.93E-06	113
histone modification	2.06E-06	109
regulation of DNA metabolic process	2.19E-06	70
DNA replication	3.24E-06	95
regulation of hydrogen peroxide metabolic process	3.27E-06	65
covalent chromatin modification	3.62E-06	120
organelle assembly	4.13E-06	27
xylem development	6.34E-06	36
plant organ formation	1.03E-05	76
negative regulation of transcription, DNA-templated	1.03E-05	114
negative regulation of RNA biosynthetic process	1.03E-05	114
negative regulation of nucleic acid-templated transcription	1.03E-05	114
chromatin silencing	1.03E-05	78
phenylpropanoid biosynthetic process	1.03E-05	51
regulation of cell cycle process	1.03E-05	51
hydrogen peroxide metabolic process	1.23E-05	100
anatomical structure formation involved in morphogenesis	1.34E-05	102
negative regulation of gene expression, epigenetic	1.62E-05	78

Biological Processes Enriched in <i>rlt1rlt2</i> mutants	p.adjust	Count
regulation of mitotic cell cycle	2.62E-05	40
regulation of cell cycle	3.45E-05	88
response to UV	3.47E-05	77
response to oxidative stress	4.06E-05	133
steroid metabolic process	4.07E-05	76
DNA methylation or demethylation	6.85E-05	61
sepal development	6.95E-05	31
flower calyx development	6.95E-05	31
sepal morphogenesis	7.78E-05	30
petal formation	7.78E-05	30
sepal formation	7.78E-05	30
gene silencing	0.00013374	108
negative regulation of cell death	0.00056285	54
negative regulation of programmed cell death	0.00056285	53
floral organ formation	0.00064442	54
response to hydrogen peroxide	0.00070183	59
petal morphogenesis	0.00141633	31
DNA replication initiation	0.00144747	25
chromatin silencing by small RNA	0.00146008	39
regulation of gene expression, epigenetic	0.00206109	97
epidermis development	0.00206109	15
epidermal cell differentiation	0.00206109	15
epithelial cell differentiation	0.00206109	15
plant-type cell wall biogenesis	0.00212401	51
cell fate specification	0.00291407	22
de-etiolation	0.00373398	7
anthocyanin-containing compound metabolic process	0.00428424	29
regulation of shoot system development	0.00446792	101
petal development	0.00457771	36
corolla development	0.00457771	36
response to light intensity	0.00469053	76
response to UV-B	0.00723291	33
leaf vascular tissue pattern formation	0.0077738	12
cell wall biogenesis	0.00930105	88
phenol-containing compound biosynthetic process	0.01003347	60
response to red or far red light	0.01092618	102
plant epidermis morphogenesis	0.01149309	77
pigment biosynthetic process	0.01187813	70

Biological Processes Enriched in <i>rlt1rlt2</i> mutants	p.adjust	Count
anthocyanin-containing compound biosynthetic process	0.01245552	21
gene silencing by RNA	0.01458355	79
DNA endoreduplication	0.02375704	35
pigmentation	0.02485092	33
pigment accumulation	0.02485092	33
pigment accumulation in response to UV light	0.02485092	33
pigment accumulation in tissues in response to UV light	0.02485092	33
pigment accumulation in tissues	0.02485092	33
anthocyanin accumulation in tissues in response to UV light	0.02485092	33
hydrogen peroxide catabolic process	0.02586035	23
cell cycle DNA replication	0.02643975	35
post-embryonic plant organ development	0.02951197	89

Supplemental Table 4 Selected biological pathways identified with GO term enrichment analysis in *rlt1/2* mutants.

Biological Processes Enriched in <i>chr11chr17</i> mutants	p.adjust	Count
photosynthesis	7.08E-63	218
photosynthesis, light reaction	1.27E-54	178
glyceraldehyde-3-phosphate metabolic process	1.41E-52	201
cellular aldehyde metabolic process	1.96E-45	226
isopentenyl diphosphate biosynthetic process, methylerythritol 4-phosphate pathway	5.20E-40	143
isopentenyl diphosphate biosynthetic process	1.13E-39	144
isopentenyl diphosphate metabolic process	1.13E-39	144
plastid organization	3.20E-34	192
sulfur compound biosynthetic process	7.95E-34	221
pigment biosynthetic process	3.64E-31	145
pigment metabolic process	1.19E-27	166
secondary metabolite biosynthetic process	5.41E-15	130
shoot system morphogenesis	6.30E-15	175
response to far red light	2.08E-13	56
leaf morphogenesis	1.97E-10	88
response to UV	1.68E-08	93
leaf development	1.08E-06	136
hydrogen peroxide metabolic process	1.12E-06	111
regulation of hydrogen peroxide metabolic process	1.42E-06	70
phenol-containing compound metabolic process	4.35E-06	83
oxidation-reduction process	4.74E-06	118
pigmentation	5.15E-06	47
pigment accumulation	5.15E-06	47
pigment accumulation in response to UV light	5.15E-06	47
pigment accumulation in tissues in response to UV light	5.15E-06	47
pigment accumulation in tissues	5.15E-06	47
anthocyanin accumulation in tissues in response to UV light	5.15E-06	47
regulation of programmed cell death	3.15E-05	121
programmed cell death	3.35E-05	134
regulation of cellular response to stress	3.42E-05	118
phenol-containing compound biosynthetic process	3.82E-05	75
flavonoid metabolic process	4.23E-05	81
meristem maintenance	6.42E-05	91
cell death	7.53E-05	144
regulation of cell death	7.82E-05	121
regulation of meristem development	0.00020562	73
meristem growth	0.00020562	61

Biological Processes Enriched in <i>chr11chr17</i> mutants	p.adjust	Count
aging	0.00023456	54
plant ovule development	0.00024564	55
plant-type ovary development	0.00029738	55
anthocyanin-containing compound biosynthetic process	0.00047215	26
root hair cell development	0.00085711	65
root hair elongation	0.00089385	61
plant epidermis morphogenesis	0.00098298	89
regulation of flavonoid biosynthetic process	0.00107034	43
response to red or far red light	0.00146121	116
anthocyanin-containing compound metabolic process	0.00176313	32
phenylpropanoid metabolic process	0.0023744	58
pattern specification process	0.00381566	98
stomatal complex patterning	0.00598352	7
DNA recombination	0.00976644	71
homologous chromosome segregation	0.01187796	30
regulation of growth	0.012717	85
meiotic chromosome segregation	0.0140882	47
cellular response to DNA damage stimulus	0.0157173	83
response to UV-B	0.0168333	34
regulation of developmental growth	0.01862713	77
double-strand break repair	0.01979362	34
meiotic DNA double-strand break formation	0.0241029	30
regulation of secondary metabolic process	0.02441137	12
asymmetric cell division	0.02589976	11
carpel development	0.02652594	71
regulation of anthocyanin biosynthetic process	0.02671472	8
DNA repair	0.03161362	70
regulation of secondary metabolite biosynthetic process	0.03998078	7
response to gibberellin	0.03998078	48
phenylpropanoid biosynthetic process	0.04607956	41

Supplemental Table 5 Selected biological pathways identified with GO term enrichment analysis in *chr11/17* mutants.

Gene	Sequence	Reference
Primers to amplify pH4::H4 for cloning		
At3g53730 (promoter) - F	CTGAGAAGAAGAAGTGAACGTC	This paper
At3g53730 (terminator) - R	GGCGATTGCACCTGATTGGTG	This paper
Primers to amplify endogenous H4 for genotyping		
At1g07660-F	TCCCACCACTTTGTACTIONCGAAG	This paper
At1g07660-R	ACCCTAATCCCCCAAATCGG	This paper
At1g07820-F	ATCACACGGATCAACGCAGT	This paper
At1g07820-R	TCCCCAACTTGCATAAACCCCT	This paper
At2g28740-F	ACGGATCGTTATCTTTGACCGT	This paper
At2g28740-R	AATCCATTACAATGCTGCCCT	This paper
At3g45930-F	TCATTGTTACCCCGGATCGT	This paper
At3g45930-R	CACGAAATTGCAACAACAACCT	This paper
At5g59690-F	AGACTCTCCGTCTTTGCCTA	This paper
At5g59690-R	ACAAGAACGCTACCGCAAAC	This paper
At3g46320-F	TCACTTACAGGCATCACGGG	This paper
At3g46320-R	ACAACGCAAACGCGAATACA	This paper
At5g59970-F	AGCGAAGCTCTGAAGAACACTION	This paper
At5g59970-R	TCCAGAAACAGATCTCCAAACA	This paper
At3g53730-F	GTCCAGTGTTCTCCGCAAGA	This paper
At3g53730-R	GTTGGGCCACTIONTACAAAAGGT	This paper
Primers to amplify PRMT7 for cloning		
AT4G16570 - F	GGGGACAAGTTTGTACAAAAAAGCAGGCTTCATG TCGCCTCTGTCTTCTC	This paper
AT4G16570 - R	GGGGACCACTTTGTACAAGAAAGCTGGGTCTCAA GAAATAGTATGAGTGACG	This paper
Sequencing primers		
At1g07660	TCCCACCACTTTGTACTIONCGAAG	This paper
At1g07820	ATCACACGGATCAACGCAGT	This paper
At2g28740	ACGGATCGTTATCTTTGACCGT	This paper
At3g45930	TCATTGTTACCCCGGATCGT	This paper
At5g59690	AGACTCTCCGTCTTTGCCTA	This paper
At3g46320	TCACTTACAGGCATCACGGG	This paper
At5g59970	AGCGAAGCTCTGAAGAACACTION	This paper
At3g53730	TCAGGCATAATGACGCGGAT	This paper

Gene	Sequence	Reference
Primers for gRNA		
At3g53730 - F	ATTGAGGGAGGAGCCAAGAGACAT	This paper
At3g53730 - R	AAACATGTCTCTTGGCTCCTCCCT	This paper
At5g59970, At3g45930, At3g46320, At5g59690 - F	ATTGGTTCTGAGAGACAACATCCA	This paper
At5g59970, At3g45930, At3g46320, At5g59690 - R	AAACTGGATGTTGTCTCTCAGAAC	This paper
At1g07660, At1g07820 - F	ATTGAGAGGCACAGGAAGGTTCTG	This paper
At1g07660, At1g07820 - R	AAACCAGAACCTTCCTGTGCCTCT	This paper
At2g28740 - F	ATTGCGTCGTCTTGCTCGTAGAGG	This paper
At2g28740 - R	AAACCCTCTACGAGCAAGACGACG	This paper
Site-directed mutagenesis primers for silent mutations in H4		
F1	CGTCGTCTCGCGAGGCGAGGAGGCGTGAAGCG	This paper
R1	CGCTTCACGCCTCCTCGCCTCGCGAGACGACG	This paper
F2	GGAAAGGGAGGAGCGAAAAGACATCGGAAAGTAC	This paper
R2	GTACTTTCCGATGTCTTTTCGCTCCTCCCTTTCC	This paper
F3	GGCGTGAAGCGTATTTCTGGTTTGATCTATGAAG	This paper
R3	CTTCATAGATCAAACCAGAAATACGCTTCACGCC	This paper
F4	GAAAAATGTCAGGCCGTGGGAAGGGAGGAAAAGG	This paper
R4	CCTTTTCCTCCCTTCCCACGGCCTGACATTTTTTC	This paper
F5	CCCAAGAGAAAAATGTCAGGACGTGGGAAGGGAGG	This paper
R5	CCTCCCTTCCCACGTCTGACATTTTTTCTCTTGGG	This paper
Site-directed mutagenesis primers for H4 mutant library		
H4S1A - F	CCCAAGAGAAAAATGGCAGGACGTGGGAAGGG	This paper
H4S1A - R	CCCTTCCCACGTCCTGCCATTTTTTCTCTTGGG	This paper
H4R3K - F	GAAAAATGTCAGGAAAAGGGGAAGGGAGG	This paper
H4R3K - R	CCTCCCTTCCCTTTTCCTGACATTTTTTC	This paper
H4R3A - F	GAAAAATGTCAGGAGCTGGGAAGGGAGG	This paper
H4R3A - R	CCTCCCTTCCCAGCTCCTGACATTTTTTC	This paper
H4K5R - F	GTCAGGACGTGGGAGGGGAGGAAAAGG	This paper

Gene	Sequence	Reference
H4K5R - R	CCTTTTCCTCCCCTCCCACGTCCTGAC	This paper
H4K5A - F	GTCAGGACGTGGGGCGGGAGGAAAAGG	This paper
H4K5A - R	CCTTTTCCTCCCGCCCCACGTCCTGAC	This paper
H4K8R - F	CGTGGGAAGGGAGGAAGAGGATTAGGAAAGGG	This paper
H4K8R - R	CCCTTTCCTAATCCTCTTCCCTTCCCACG	This paper
H4K8A - F	CGTGGGAAGGGAGGAGCAGGATTAGGAAAGGG	This paper
H4K8A - R	CCCTTTCCTAATCCTGCTCCTCCCTTCCCACG	This paper
H4K12R - F	GGAGGAAAAGGATTAGGAAGGGGAGGAGCGAAA AG	This paper
H4K12R - R	CTTTTCGCTCCTCCCCTTCTAATCCTTTTCCTCC	This paper
H4K12A - F	GGAGGAAAAGGATTAGGAGCGGGAGGAGCGAAA AG	This paper
H4K12A - R	CTTTTCGCTCCTCCCGCTCCTAATCCTTTTCCTCC	This paper
H4K16R - F	GGAAAGGGAGGAGCGAGAAGACATCGGAAAG	This paper
H4K16R - R	CTTTCCGATGTCTTCTCGCTCCTCCCTTTCC	This paper
H4K16A - F	GGAAAGGGAGGAGCGGCAAGACATCGGAAAG	This paper
H4K16A - R	CTTTCCGATGTCTTGCCGCTCCTCCCTTTCC	This paper
H4R17K - F	GGGAGGAGCGAAAAACATCGGAAAGTACTC	This paper
H4R17K - R	GAGTACTTTCCGATGTTTTTTCGCTCCTCCC	This paper
H4R17A - F	GGGAGGAGCGAAAGCACATCGGAAAGTACTC	This paper
H4R17A - R	GAGTACTTTCCGATGTGCTTTTCGCTCCTCCC	This paper
H4R19K - F	GGAGCGAAAAGACATAAGAAAGTACTCAGAGAC	This paper
H4R19K -R	GTCTCTGAGTACTTTCTTATGTCTTTTCGCTCC	This paper
H4R19A - F	GGAGCGAAAAGACATGCGAAAGTACTCAGAGAC	This paper
H4R19A - R	GTCTCTGAGTACTTTTCGCATGTCTTTTCGCTCC	This paper
H4K20R - F	GCGAAAAGACATCGGAGAGTACTCAGAGACAAC	This paper
H4K20R - R	GTTGTCTCTGAGTACTCTCCGATGTCTTTTCGC	This paper
H4K20A - F	GCGAAAAGACATCGGGCAGTACTCAGAGACAAC	This paper
H4K20A - R	GTTGTCTCTGAGTACTGCCCGATGTCTTTTCGC	This paper
H4R23K - F	CATCGGAAAGTACTCAAAGACAACATCCAAGGG	This paper
H4R23K - R	CCCTTGATGTTGTCTTTGAGTACTTTCCGATG	This paper
H4R23A - F	CATCGGAAAGTACTCGCAGACAACATCCAAGGG	This paper
H4R23A - R	CCCTTGATGTTGTCTGCGAGTACTTTCCGATG	This paper
H4T30V - F	CAACATCCAAGGGATTGTCAAACCTGCGATTCCG	This paper
H4T30V - R	CGAATCGCAGGTTTGACAATCCCTTGGATGTTG	This paper

Gene	Sequence	Reference
H4K31R - F	CATCCAAGGGATTACCAGACCTGCGATTCGTCTG	This paper
H4K31R - R	CGACGAATCGCAGGTCTGGTAATCCCTTGGATG	This paper
H4K31A - F	CATCCAAGGGATTACCGCACCTGCGATTCGTCTG	This paper
H4K31A - R	CGACGAATCGCAGGTGCGGTAATCCCTTGGATG	This paper
H4R35K - F	CCAAACCTGCGATTAAACGTCTCGCGAGGCCG	This paper
H4R35K - R	CGCCTCGCGAGACGTTTAATCGCAGGTTTGG	This paper
H4R35A - F	CCAAACCTGCGATTGCTCGTCTCGCGAGGCCG	This paper
H4R35A - R	CGCCTCGCGAGACGAGCAATCGCAGGTTTGG	This paper
H4R36K - F	CAAACCTGCGATTCGTAAACTCGCGAGGCCGAGG	This paper
H4R36K - R	CCTCGCCTCGCGAGTTTACGAATCGCAGGTTTGG	This paper
H4R36A - F	CAAACCTGCGATTCGTGCTCTCGCGAGGCCGAGG	This paper
H4R36A - R	CCTCGCCTCGCGAGAGCACGAATCGCAGGTTTGG	This paper
H4R39K - F	GATTCGTCTGCTCTCGCGAAGCGAGGAGGCCGTGAA G	This paper
H4R39K - R	CTTCACGCCTCCTCGCTTCGCGAGACGACGAATC	This paper
H4R39A - F	GATTCGTCTGCTCTCGCGGCCGCGAGGAGGCCGTGAA G	This paper
H4R39A - R	CTTCACGCCTCCTCGCGCCGCGAGACGACGAAT C	This paper
H4R40K - F	CGTCGTCTCGCGAGGAAAGGAGGCCGTGAAGCG	This paper
H4R40K - R	CGCTTCACGCCTCCTTTCCTCGCGAGACGACG	This paper
H4R40A - F	CGTCGTCTCGCGAGGGCAGGAGGCCGTGAAGCG	This paper
H4R40A - R	CGCTTCACGCCTCCTGCCCTCGCGAGACGACG	This paper
H4K44R - F	GGCGAGGAGGCCGTGAGGCCGATTTCTGGTTTGG	This paper
H4K44R - R	CAAACCAGAAATACGCCTCACGCCTCCTCGCC	This paper
H4K44A - F	GGCGAGGAGGCCGTGGCGCGTATTTCTGGTTTGG	This paper
H4K44A - R	CAAACCAGAAATACGCGCCACGCCTCCTCGCC	This paper
H4R45K - F	CGAGGAGGCCGTGAAGAAAATTTCTGGTTTGGATC	This paper
H4R45K - R	GATCAAACCAGAAATTTTCTTCACGCCTCCTCG	This paper
H4R45A - F	CGAGGAGGCCGTGAAGGCTATTTCTGGTTTGGATC	This paper
H4R45A - R	GATCAAACCAGAAATAGCCTTCACGCCTCCTCG	This paper
H4S47A - F	GGCGTGAAGCGTATTGCTGGTTTGGATCTATGAAG	This paper
H4S47A - R	CTTCATAGATCAAACCAGCAATACGCTTCACGCC	This paper
H4Y51F - F	GTATTTCTGGTTTGGATCTTTGAAGAGACTCGCGG C	This paper
H4Y51F - R	GCCGCGAGTCTCTTCAAAGATCAAACCAGAAATA C	This paper

Gene	Sequence	Reference
H4T54V - F	GTTTGATCTATGAAGAGGTTTCGCGGCGTTCTCAA G	This paper
H4T54V - R	CTTGAGAACGCCGCGAACCTCTTCATAGATCAAA C	This paper
H4R55K - F	GATCTATGAAGAGACTAAAGGCGTTCTCAAGATC	This paper
H4R55K - R	GATCTTGAGAACGCCTTTAGTCTCTTCATAGATC	This paper
H4R55A - F	GATCTATGAAGAGACTGCCGCGTTCTCAAGATC	This paper
H4R55A - R	GATCTTGAGAACGCCGGCAGTCTCTTCATAGATC	This paper
H4K59A - F	GACTCGCGGCGTTCTCGCGATCTTTCTCGAGAAC	This paper
H4K59A - R	GTTCTCGAGAAAGATCGCGAGAACGCCGCGAGT C	This paper
H4K59R - F	GACTCGCGGCGTTCTCAGGATCTTTCTCGAGAAC	This paper
H4K59R - R	GTTCTCGAGAAAGATCCTGAGAACGCCGCGAGTC	This paper
H4R67K - F	CTCGAGAACGTGATTAAGACGCCGTTACTTAC	This paper
H4R67K - R	GTAAGTAACGGCGTCTTTAATCACGTTCTCGAG	This paper
H4R67A - F	CTCGAGAACGTGATTGCTGACGCCGTTACTTAC	This paper
H4R67A - R	GTAAGTAACGGCGTCAGCAATCACGTTCTCGAG	This paper
H4T71V - F	GATTCGTGACGCCGTTGTTTACACGGAGCACGC	This paper
H4T71V - R	GCGTGCTCCGTGTAAACAACGGCGTCACGAATC	This paper
H4Y72F - F	CGTGACGCCGTTACTTTACGGAGCACGCTCGC	This paper
H4Y72F - R	GCGAGCGTGCTCCGTGAAAGTAACGGCGTCACG	This paper
H4T73V - F	GACGCCGTTACTTACGTGGAGCACGCTCGCCGG	This paper
H4T73V - R	CCGGCGAGCGTGCTCCACGTAAGTAACGGCGTC	This paper
H4R77K - F	CTTACACGGAGCACGCTAAACGGAAAAGTGTAC G	This paper
H4R77K - R	CGTAACAGTTTTCCGTTTAGCGTGCTCCGTGTAA G	This paper
H4R77A - F	CACGGAGCACGCTGCCCGGAAAAGTGTACG	This paper
H4R77A - R	CGTAACAGTTTTCCGGGACGCGTGCTCCGTG	This paper
H4R78K - F	CGGAGCACGCTCGCAAGAAAAGTGTACGGCG	This paper
H4R78K - R	CGCCGTAACAGTTTTCTTGCGAGCGTGCTCCG	This paper
H4R78A - F	CGGAGCACGCTCGCGCGAAAAGTGTACGGCG	This paper
H4R78A - R	CGCCGTAACAGTTTTCGCGCGAGCGTGCTCCG	This paper
H4K79R - F	GAGCACGCTCGCCGGAGAACTGTTACGGCGATG	This paper
H4K79R - R	CATCGCCGTAACAGTTCTCCGGCGAGCGTGCTC	This paper
H4K79A - F	GAGCACGCTCGCCGGGCAACTGTTACGGCGATG	This paper
H4K79A - R	CATCGCCGTAACAGTTGCCCGGCGAGCGTGCTC	This paper

Gene	Sequence	Reference
H4T80V - F	CACGCTCGCCGGAAAGTTGTTACGGCGATGG	This paper
H4T80V - R	CCATCGCCGTAACAACCTTCCGGCGAGCGTG	This paper
H4T82V - F	CGACGTCCATCGCCACAACAGTTTTCCGGCG	This paper
H4T82V - R	CGCCGGAAAACGTGTGTGGCGATGGACGTCG	This paper
H4Y88F - F	GCGATGGACGTCGTTTTCCGCTCTCAAGAGAC	This paper
H4Y88F - R	GTCTCTTGAGAGCGAAAACGACGTCCATCGC	This paper
H4K91R - F	GTCGTTTACGCTCTCAGGAGACAAGGACGAAC	This paper
H4K91R - R	GTTTCGTCCTTGTCTCCTGAGAGCGTAAACGAC	This paper
H4K91A - F	GTCGTTTACGCTCTCGCGAGACAAGGACGAAC	This paper
H4K91A - R	GTTTCGTCCTTGTCTCGCGAGAGCGTAAACGAC	This paper
H4R92K - F	GTTTACGCTCTCAAGAAACAAGGACGAACTTTG	This paper
H4R92K - R	CAAAGTTCGTCCTTGTCTTCTTGAGAGCGTAAAC	This paper
H4R92A - F	GTTTACGCTCTCAAGGCACAAGGACGAACTTTG	This paper
H4R92A - R	CAAAGTTCGTCCTTGTGCCTTGAGAGCGTAAAC	This paper
H4R95K - F	CTCAAGAGACAAGGAAAAACTTTGTATGGATTC	This paper
H4R95K - R	GAATCCATACAAAGTTTTTCTTGTCTCTTGAG	This paper
H4R95A - F	CTCAAGAGACAAGGAGCAACTTTGTATGGATTC	This paper
H4R95A - R	GAATCCATACAAAGTTGCTCCTTGTCTCTTGAG	This paper
H4T96V - F	GAGACAAGGACGAGTTTTGTATGGATTCGGCGGC	This paper
H4T96V - R	GCCGCCGAATCCATACAAAACCTCGTCCTTGTCTC	This paper
H4Y98F - F	CAAGGACGAACTTTGTTTGGATTCGGCGGCTAA	This paper
H4Y98F - R	TTAGCCGCCGAATCCAAACAAAGTTCGTCCTTG	This paper
qRT-PCR primers		
Actin - F	TCGTGGTGGTGAGTTTGTAC	(Dong et al., 2021)
Actin - R	CAGCATCATCACAAGCATCC	(Dong et al., 2021)
H4 (At3g53730) - F	GAGAACGTGATTCGTGACGC	This paper
H4 (At3g53730) - R	GCCGCCGAATCCATACAAAG	This paper
BRCA1 - F	CATGTGCCTTTTGTGAGTGTTT	(Dong et al., 2021)
BRCA1 - R	TGGAGCCCATTCAGCACAGTTT	(Dong et al., 2021)
TSI - F	ATCCAGTCCGAAGAACGCGAACTA	(Dong et al., 2021)

Gene	Sequence	Reference
TSI - R	TCACTTGTGAGTGTTTCGTGAGGTC	(Dong et al., 2021)
PRMT7 5' – F1	GAACTTCCACTGACGGCCTA	This paper
PRMT7 5' – R1	TCACCTCTCATTGCTCTCACA	This paper
PRMT7 3' – F2	TTGACTTCTCCAAGCCCATC	This paper
PRMT7 3' – R2	TCACCCAATCCATCCACAAT	This paper
FT - F	ATCTCCATTGGTTGGTGACTGATA	(Wu et al., 2008)
FT - R	GCCAAAGGTTGTTCCAGTTGTAG	(Wu et al., 2008)
SOC1 - F	AACAACCTCGAAGCTTCTAAACGTAA	(Richter et al., 2019)
SOC1 - R	CCTCGATTGAGCATGTTCCCT	(Richter et al., 2019)

Supplemental Table 6 Cloning and PCR Primers.

Samples	Experiment	Total Reads	Mapped Reads	Percentage of Mapped Reads	Average Insert Size
Col rep1	RNA-seq	35317497	34487327	97.65	261.2
Col rep2	RNA-seq	37358839	36612747	98.00	255
rH4-1 rep1	RNA-seq	35411859	34686565	97.95	241.3
rH4-1 rep2	RNA-seq	34498024	33936248	98.37	233.7
rH4-2 rep1	RNA-seq	38750965	37663933	97.19	166.5
rH4-2 rep2	RNA-seq	44891415	43556242	97.03	172.0
rH4R17A-1 rep1	RNA-seq	41304934	39291281	95.12	235.4
rH4R17A-1 rep2	RNA-seq	39680698	37804646	95.27	235.4
rH4R17A-2 rep1	RNA-seq	40557837	39704087	97.89	253.7
rH4R17A-2 rep2	RNA-seq	33428391	32774618	98.04	256.1
arid5 rep1	RNA-seq	37304631	35159757	94.25	239.8
arid5 rep2	RNA-seq	54013709	50976592	94.38	235.6
rit1/2 rep1	RNA-seq	31215439	30353801	97.24	262.8
rit1/2 rep2	RNA-seq	19013353	18547256	97.55	264.1
chr11/17 rep1	RNA-seq	32853678	32036677	97.51	247.6
chr11/17 rep2	RNA-seq	34842702	34293590	98.42	246.5
pie1 rep1	RNA-seq	48159095	47040083	97.68	178.5
pie1 rep2	RNA-seq	46390059	45169492	97.37	164.0
Col rep1	MNase-seq	74760430	74321275	99.41	166.6
Col rep2	MNase-seq	60979268	60569848	99.33	170.0
rH4-1 rep1	MNase-seq	74420518	73858813	99.25	168.8
rH4-1 rep2	MNase-seq	65077948	64608215	99.28	167.8
rH4R17A-1 rep1	MNase-seq	64392968	63727785	98.97	168.2
rH4R17A-1 rep2	MNase-seq	60674356	60012743	98.91	167.5
arid5 rep1	MNase-seq	81972020	75271258	91.83	143.3
arid5 rep2	MNase-seq	78396464	72967660	93.08	149.9
rit1/2 rep1	MNase-seq	76009948	71361811	93.88	149.8
rit1/2 rep2	MNase-seq	91163646	82393229	90.38	153.7

Supplemental Table 7 Statistics for mapping and coverage of the NGS data.

References

- Adam, M., Robert, F., Laroche, M., and Gaudreau, L. (2001). H2A.Z is required for global chromatin integrity and for recruitment of RNA polymerase II under specific conditions. *Mol Cell Biol* 21, 6270-6279.
- Adams, C.R., and Kamakaka, R.T. (1999). Chromatin assembly: biochemical identities and genetic redundancy. *Curr Opin Genet Dev* 9, 185-190.
- Aguilera, A. (2002). The connection between transcription and genomic instability. *EMBO J* 21, 195-201.
- Aguilera, A., and Garcia-Muse, T. (2012). R loops: from transcription byproducts to threats to genome stability. *Mol Cell* 46, 115-124.
- Aguilera, A., and Garcia-Muse, T. (2013). Causes of genome instability. *Annu Rev Genet* 47, 1-32.
- Aguilera, A., and Gomez-Gonzalez, B. (2008). Genome instability: a mechanistic view of its causes and consequences. *Nat Rev Genet* 9, 204-217.
- Akawi, N., McRae, J., Ansari, M., Balasubramanian, M., Blyth, M., Brady, A.F., Clayton, S., Cole, T., Deshpande, C., Fitzgerald, T.W., *et al.* (2015). Discovery of four recessive developmental disorders using probabilistic genotype and phenotype matching among 4,125 families. *Nat Genet* 47, 1363-1369.
- Alabert, C., and Groth, A. (2012). Chromatin replication and epigenome maintenance. *Nat Rev Mol Cell Biol* 13, 153-167.
- Allis, C.D., and Jenuwein, T. (2016). The molecular hallmarks of epigenetic control. *Nat Rev Genet* 17, 487-500.
- Allshire, R.C., and Madhani, H.D. (2018). Ten principles of heterochromatin formation and function. *Nat Rev Mol Cell Biol* 19, 229-244.
- Almouzni, G., and Cedar, H. (2016). Maintenance of Epigenetic Information. *Cold Spring Harb Perspect Biol* 8.
- Alonso, J.M., Stepanova, A.N., Leisse, T.J., Kim, C.J., Chen, H., Shinn, P., Stevenson, D.K., Zimmerman, J., Barajas, P., Cheuk, R., *et al.* (2003). Genome-wide insertional mutagenesis of *Arabidopsis thaliana*. *Science* 301, 653-657.
- Alvarez-Venegas, R., Pien, S., Sadder, M., Witmer, X., Grossniklaus, U., and Avramova, Z. (2003). ATX-1, an *Arabidopsis* homolog of trithorax, activates flower homeotic genes. *Curr Biol* 13, 627-637.
- Amasino, R.M. (2005). Vernalization and flowering time. *Curr Opin Biotechnol* 16, 154-158.
- An, H., Roussot, C., Suarez-Lopez, P., Corbesier, L., Vincent, C., Pineiro, M., Hepworth, S., Mouradov, A., Justin, S., Turnbull, C., *et al.* (2004). CONSTANS acts in the phloem to regulate a systemic signal that induces photoperiodic flowering of *Arabidopsis*. *Development* 131, 3615-3626.
- Anderson, J.D., and Widom, J. (2001). Poly(dA-dT) promoter elements increase the equilibrium accessibility of nucleosomal DNA target sites. *Mol Cell Biol* 21, 3830-3839.
- Andres, F., and Coupland, G. (2012). The genetic basis of flowering responses to seasonal cues. *Nat Rev Genet* 13, 627-639.
- Arents, G., Burlingame, R.W., Wang, B.C., Love, W.E., and Moudrianakis, E.N. (1991). The nucleosomal core histone octamer at 3.1 Å resolution: a tripartite protein assembly and a left-handed superhelix. *Proc Natl Acad Sci U S A* 88, 10148-10152.
- Armache, K.J., Garlick, J.D., Canzio, D., Narlikar, G.J., and Kingston, R.E. (2011). Structural basis of silencing: Sir3 BAH domain in complex with a nucleosome at 3.0 Å resolution. *Science* 334, 977-982.
- Aukerman, M.J., and Sakai, H. (2003). Regulation of flowering time and floral organ identity by a MicroRNA and its APETALA2-like target genes. *Plant Cell* 15, 2730-2741.

Aydin, O.Z., Vermeulen, W., and Lans, H. (2014). ISWI chromatin remodeling complexes in the DNA damage response. *Cell Cycle* 13, 3016-3025.

Baekelandt, A., Pauwels, L., Wang, Z., Li, N., De Milde, L., Natran, A., Vermeersch, M., Li, Y., Goossens, A., Inze, D., *et al.* (2018). Arabidopsis Leaf Flatness Is Regulated by PPD2 and NINJA through Repression of CYCLIN D3 Genes. *Plant Physiol* 178, 217-232.

Bai, L., and Morozov, A.V. (2010). Gene regulation by nucleosome positioning. *Trends Genet* 26, 476-483.

Balakrishnan, L., and Milavetz, B. (2010). Decoding the histone H4 lysine 20 methylation mark. *Crit Rev Biochem Mol Biol* 45, 440-452.

Balasubramanian, S., Sureshkumar, S., Lempe, J., and Weigel, D. (2006). Potent induction of Arabidopsis thaliana flowering by elevated growth temperature. *PLoS Genet* 2, e106.

Bannister, A.J., Zegerman, P., Partridge, J.F., Miska, E.A., Thomas, J.O., Allshire, R.C., and Kouzarides, T. (2001). Selective recognition of methylated lysine 9 on histone H3 by the HP1 chromo domain. *Nature* 410, 120-124.

Bar, S., and Benvenisty, N. (2019). Epigenetic aberrations in human pluripotent stem cells. *EMBO J* 38.

Barski, A., Cuddapah, S., Cui, K., Roh, T.Y., Schones, D.E., Wang, Z., Wei, G., Chepelev, I., and Zhao, K. (2007). High-resolution profiling of histone methylations in the human genome. *Cell* 129, 823-837.

Bastow, R., Mylne, J.S., Lister, C., Lippman, Z., Martienssen, R.A., and Dean, C. (2004). Vernalization requires epigenetic silencing of FLC by histone methylation. *Nature* 427, 164-167.

Baubec, T., and Defossez, P.A. (2020). Reading DNA Modifications. *J Mol Biol*.

Benhamed, M., Bertrand, C., Servet, C., and Zhou, D.X. (2006). Arabidopsis GCN5, HD1, and TAF1/HAF2 interact to regulate histone acetylation required for light-responsive gene expression. *Plant Cell* 18, 2893-2903.

Bennett, R.L., Bele, A., Small, E.C., Will, C.M., Nabet, B., Oyer, J.A., Huang, X., Ghosh, R.P., Grzybowski, A.T., Yu, T., *et al.* (2019). A Mutation in Histone H2B Represents a New Class of Oncogenic Driver. *Cancer Discov* 9, 1438-1451.

Berger, S.L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. *Genes Dev* 23, 781-783.

Bernstein, B.E., Mikkelsen, T.S., Xie, X., Kamal, M., Huebert, D.J., Cuff, J., Fry, B., Meissner, A., Wernig, M., Plath, K., *et al.* (2006). A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125, 315-326.

Berry, S., and Dean, C. (2015). Environmental perception and epigenetic memory: mechanistic insight through FLC. *Plant J* 83, 133-148.

Bertrand, C., Bergounioux, C., Domenichini, S., Delarue, M., and Zhou, D.X. (2003). Arabidopsis histone acetyltransferase AtGCN5 regulates the floral meristem activity through the WUSCHEL/AGAMOUS pathway. *J Biol Chem* 278, 28246-28251.

Biasini, M., Bienert, S., Waterhouse, A., Arnold, K., Studer, G., Schmidt, T., Kiefer, F., Gallo Cassarino, T., Bertoni, M., Bordoli, L., *et al.* (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res* 42, W252-258.

Bilichak, A., Yao, Y., Titov, V., Golubov, A., and Kovalchuk, I. (2014). Genome stability in the *uvh6* mutant of Arabidopsis thaliana. *Plant Cell Rep* 33, 979-991.

Bilokapic, S., Strauss, M., and Halic, M. (2018). Structural rearrangements of the histone octamer translocate DNA. *Nat Commun* 9, 1330.

Bird, A. (2007). Perceptions of epigenetics. *Nature* 447, 396-398.

Blanc, R.S., Vogel, G., Chen, T., Crist, C., and Richard, S. (2016). PRMT7 Preserves Satellite Cell Regenerative Capacity. *Cell Rep* 14, 1528-1539.

Blazquez, M.A., Green, R., Nilsson, O., Sussman, M.R., and Weigel, D. (1998). Gibberellins promote flowering of arabidopsis by activating the LEAFY promoter. *Plant Cell* 10, 791-800.

Bolouri Moghaddam, M.R., and Van den Ende, W. (2013). Sugars, the clock and transition to flowering. *Front Plant Sci* 4, 22.

Bonasio, R., Tu, S., and Reinberg, D. (2010). Molecular signals of epigenetic states. *Science* 330, 612-616.

Bond, D.M., Dennis, E.S., Pogson, B.J., and Finnegan, E.J. (2009). Histone acetylation, VERNALIZATION INSENSITIVE 3, FLOWERING LOCUS C, and the vernalization response. *Mol Plant* 2, 724-737.

Bonke, M., Thitamadee, S., Mahonen, A.P., Hauser, M.T., and Helariutta, Y. (2003). APL regulates vascular tissue identity in Arabidopsis. *Nature* 426, 181-186.

Borg, M., Jacob, Y., Susaki, D., LeBlanc, C., Buendia, D., Axelsson, E., Kawashima, T., Voigt, P., Boavida, L., Becker, J., *et al.* (2020). Targeted reprogramming of H3K27me3 resets epigenetic memory in plant paternal chromatin. *Nat Cell Biol.*

Bostick, M., Kim, J.K., Esteve, P.O., Clark, A., Pradhan, S., and Jacobsen, S.E. (2007). UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science* 317, 1760-1764.

Botuyan, M.V., Lee, J., Ward, I.M., Kim, J.E., Thompson, J.R., Chen, J., and Mer, G. (2006). Structural basis for the methylation state-specific recognition of histone H4-K20 by 53BP1 and Crb2 in DNA repair. *Cell* 127, 1361-1373.

Bouche, F., Lobet, G., Tocquin, P., and Perilleux, C. (2016). FLOR-ID: an interactive database of flowering-time gene networks in Arabidopsis thaliana. *Nucleic Acids Res* 44, D1167-1171.

Boudreault, A.A., Cronier, D., Selleck, W., Lacoste, N., Utley, R.T., Allard, S., Savard, J., Lane, W.S., Tan, S., and Cote, J. (2003). Yeast enhancer of polycomb defines global Esa1-dependent acetylation of chromatin. *Genes Dev* 17, 1415-1428.

Bourguet, P., Picard, C.L., Yelagandula, R., Pelissier, T., Lorkovic, Z.J., Feng, S., Pouch-Pelissier, M.N., Schmucker, A., Jacobsen, S.E., Berger, F., *et al.* (2021). The histone variant H2A.W and linker histone H1 co-regulate heterochromatin accessibility and DNA methylation. *Nat Commun* 12, 2683.

Brabencova, S., Ihnatova, I., Potesil, D., Fojtova, M., Fajkus, J., Zdrahal, Z., and Lochmanova, G. (2017). Variations of Histone Modification Patterns: Contributions of Inter-plant Variability and Technical Factors. *Front Plant Sci* 8, 2084.

Brignon, P., and Chaubet, N. (1993). Constitutive and cell-division-inducible protein-DNA interactions in two maize histone gene promoters. *Plant J* 4, 445-457.

Brownell, J.E., Zhou, J., Ranalli, T., Kobayashi, R., Edmondson, D.G., Roth, S.Y., and Allis, C.D. (1996). Tetrahymena histone acetyltransferase A: a homolog to yeast Gcn5p linking histone acetylation to gene activation. *Cell* 84, 843-851.

Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G., *et al.* (2002). T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites. *EMBO Rep* 3, 1152-1157.

Bu, Z., Yu, Y., Li, Z., Liu, Y., Jiang, W., Huang, Y., and Dong, A.W. (2014). Regulation of arabidopsis flowering by the histone mark readers MRG1/2 via interaction with CONSTANS to modulate FT expression. *PLoS Genet* 10, e1004617.

Cao, Y., Dai, Y., Cui, S., and Ma, L. (2008). Histone H2B monoubiquitination in the chromatin of FLOWERING LOCUS C regulates flowering time in Arabidopsis. *Plant Cell* 20, 2586-2602.

Carrozza, M.J., Li, B., Florens, L., Suganuma, T., Swanson, S.K., Lee, K.K., Shia, W.J., Anderson, S., Yates, J., Washburn, M.P., *et al.* (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. *Cell* 123, 581-592.

Cedar, H., and Bergman, Y. (2009). Linking DNA methylation and histone modification: patterns and paradigms. *Nat Rev Genet* 10, 295-304.

Chaboute, M.E., Chaubet, N., Philipps, G., Ehling, M., and Gigot, C. (1987). Genomic organization and nucleotide sequences of two histone H3 and two histone H4 genes of *Arabidopsis thaliana*. *Plant Mol Biol* 8, 179-191.

Chatterjee, N., and Walker, G.C. (2017). Mechanisms of DNA damage, repair, and mutagenesis. *Environ Mol Mutagen* 58, 235-263.

Chaubet, N., Flenet, M., Clement, B., Brignon, P., and Gigot, C. (1996). Identification of cis-elements regulating the expression of an *Arabidopsis* histone H4 gene. *Plant J* 10, 425-435.

Cheedipudi, S., Genolet, O., and Dobрева, G. (2014). Epigenetic inheritance of cell fates during embryonic development. *Front Genet* 5, 19.

Cheng, C.Y., Krishnakumar, V., Chan, A.P., Thibaud-Nissen, F., Schobel, S., and Town, C.D. (2017). Araport11: a complete reannotation of the *Arabidopsis thaliana* reference genome. *Plant J* 89, 789-804.

Chereji, R.V., Bryson, T.D., and Henikoff, S. (2019). Quantitative MNase-seq accurately maps nucleosome occupancy levels. *Genome Biol* 20, 198.

Chereji, R.V., and Clark, D.J. (2018). Major Determinants of Nucleosome Positioning. *Biophys J* 114, 2279-2289.

Choi, K., Kim, S., Kim, S.Y., Kim, M., Hyun, Y., Lee, H., Choe, S., Kim, S.G., Michaels, S., and Lee, I. (2005). SUPPRESSOR OF FRIGIDA3 encodes a nuclear ACTIN-RELATED PROTEIN6 required for floral repression in *Arabidopsis*. *Plant Cell* 17, 2647-2660.

Choi, K., Park, C., Lee, J., Oh, M., Noh, B., and Lee, I. (2007). *Arabidopsis* homologs of components of the SWR1 complex regulate flowering and plant development. *Development* 134, 1931-1941.

Chong, S.Y., Cutler, S., Lin, J.J., Tsai, C.H., Tsai, H.K., Biggins, S., Tsukiyama, T., Lo, Y.C., and Kao, C.F. (2020). H3K4 methylation at active genes mitigates transcription-replication conflicts during replication stress. *Nat Commun* 11, 809.

Clapier, C.R., and Cairns, B.R. (2009). The biology of chromatin remodeling complexes. *Annu Rev Biochem* 78, 273-304.

Clapier, C.R., and Cairns, B.R. (2012). Regulation of ISWI involves inhibitory modules antagonized by nucleosomal epitopes. *Nature* 492, 280-284.

Clapier, C.R., Iwasa, J., Cairns, B.R., and Peterson, C.L. (2017). Mechanisms of action and regulation of ATP-dependent chromatin-remodelling complexes. *Nat Rev Mol Cell Biol* 18, 407-422.

Clapier, C.R., Langst, G., Corona, D.F., Becker, P.B., and Nightingale, K.P. (2001). Critical role for the histone H4 N terminus in nucleosome remodeling by ISWI. *Mol Cell Biol* 21, 875-883.

Clapier, C.R., Nightingale, K.P., and Becker, P.B. (2002). A critical epitope for substrate recognition by the nucleosome remodeling ATPase ISWI. *Nucleic Acids Res* 30, 649-655.

Clark-Adams, C.D., Norris, D., Osley, M.A., Fassler, J.S., and Winston, F. (1988). Changes in histone gene dosage alter transcription in yeast. *Genes Dev* 2, 150-159.

Clouaire, T., Webb, S., Skene, P., Illingworth, R., Kerr, A., Andrews, R., Lee, J.H., Skalnik, D., and Bird, A. (2012). Cfp1 integrates both CpG content and gene activity for accurate H3K4me3 deposition in embryonic stem cells. *Genes Dev* 26, 1714-1728.

Clough, S.J., and Bent, A.F. (1998). Floral dip: a simplified method for *Agrobacterium*-mediated transformation of *Arabidopsis thaliana*. *Plant J* 16, 735-743.

Congdon, L.M., Houston, S.I., Veerappan, C.S., Spektor, T.M., and Rice, J.C. (2010). PR-Set7-mediated monomethylation of histone H4 lysine 20 at specific genomic regions induces transcriptional repression. *J Cell Biochem* 110, 609-619.

Corona, D.F., Langst, G., Clapier, C.R., Bonte, E.J., Ferrari, S., Tamkun, J.W., and Becker, P.B. (1999). ISWI is an ATP-dependent nucleosome remodeling factor. *Mol Cell* 3, 239-245.

Corona, D.F., Siriaco, G., Armstrong, J.A., Snarskaya, N., McClymont, S.A., Scott, M.P., and Tamkun, J.W. (2007). ISWI regulates higher-order chromatin structure and histone H1 assembly in vivo. *PLoS Biol* 5, e232.

Corona, D.F., and Tamkun, J.W. (2004). Multiple roles for ISWI in transcription, chromosome organization and DNA replication. *Biochim Biophys Acta* 1677, 113-119.

Crespo, M., Luense, L.J., Arlotto, M., Hu, J., Dorsey, J., Garcia-Oliver, E., Shah, P.P., Pflieger, D., Berger, S.L., and Govin, J. (2020). Systematic genetic and proteomic screens during gametogenesis identify H2BK34 methylation as an evolutionary conserved meiotic mark. *Epigenetics Chromatin* 13, 35.

Crevillen, P., Gomez-Zambrano, A., Lopez, J.A., Vazquez, J., Pineiro, M., and Jarillo, J.A. (2019). *Arabidopsis* YAF9 histone readers modulate flowering time through NuA4-complex-dependent H4 and H2A.Z histone acetylation at FLC chromatin. *New Phytol* 222, 1893-1908.

Crevillen, P., Yang, H., Cui, X., Greeff, C., Trick, M., Qiu, Q., Cao, X., and Dean, C. (2014). Epigenetic reprogramming that prevents transgenerational inheritance of the vernalized state. *Nature* 515, 587-590.

Cui, X., Lu, F., Qiu, Q., Zhou, B., Gu, L., Zhang, S., Kang, Y., Cui, X., Ma, X., Yao, Q., *et al.* (2016). REF6 recognizes a specific DNA sequence to demethylate H3K27me3 and regulate organ boundary formation in *Arabidopsis*. *Nat Genet* 48, 694-699.

Curtis, M.D., and Grossniklaus, U. (2003). A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant Physiol* 133, 462-469.

Dai, J., Hyland, E.M., Yuan, D.S., Huang, H., Bader, J.S., and Boeke, J.D. (2008). Probing nucleosome function: a highly versatile library of synthetic histone H3 and H4 mutants. *Cell* 134, 1066-1078.

Dann, G.P., Liszczak, G.P., Bagert, J.D., Muller, M.M., Nguyen, U.T.T., Wojcik, F., Brown, Z.Z., Bos, J., Panchenko, T., Pihl, R., *et al.* (2017). ISWI chromatin remodellers sense nucleosome modifications to determine substrate preference. *Nature* 548, 607-611.

Davey, C.A., Sargent, D.F., Luger, K., Maeder, A.W., and Richmond, T.J. (2002). Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 Å resolution. *J Mol Biol* 319, 1097-1113.

Davila Lopez, M., and Samuelsson, T. (2008). Early evolution of histone mRNA 3' end processing. *RNA* 14, 1-10.

De Buck, S., Podevin, N., Nolf, J., Jacobs, A., and Depicker, A. (2009). The T-DNA integration pattern in *Arabidopsis* transformants is highly determined by the transformed target cell. *Plant J* 60, 134-145.

De Koning, L., Corpet, A., Haber, J.E., and Almouzni, G. (2007). Histone chaperones: an escort network regulating histone traffic. *Nat Struct Mol Biol* 14, 997-1007.

de la Paz Sanchez, M., and Gutierrez, C. (2009). *Arabidopsis* ORC1 is a PHD-containing H3K4me3 effector that regulates transcription. *Proc Natl Acad Sci U S A* 106, 2065-2070.

De Lucia, F., Crevillen, P., Jones, A.M., Greb, T., and Dean, C. (2008). A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of FLC during vernalization. *Proc Natl Acad Sci U S A* *105*, 16831-16836.

De Veylder, L., Beeckman, T., Beemster, G.T., Krols, L., Terras, F., Landrieu, I., van der Schueren, E., Maes, S., Naudts, M., and Inze, D. (2001). Functional analysis of cyclin-dependent kinase inhibitors of Arabidopsis. *Plant Cell* *13*, 1653-1668.

Deal, R.B., and Henikoff, S. (2010). A simple method for gene expression and chromatin profiling of individual cell types within a tissue. *Dev Cell* *18*, 1030-1040.

Deal, R.B., Kandasamy, M.K., McKinney, E.C., and Meagher, R.B. (2005). The nuclear actin-related protein ARP6 is a pleiotropic developmental regulator required for the maintenance of FLOWERING LOCUS C expression and repression of flowering in Arabidopsis. *Plant Cell* *17*, 2633-2646.

Deal, R.B., Topp, C.N., McKinney, E.C., and Meagher, R.B. (2007). Repression of flowering in Arabidopsis requires activation of FLOWERING LOCUS C expression by the histone variant H2A.Z. *Plant Cell* *19*, 74-83.

Deem, A.K., Li, X., and Tyler, J.K. (2012). Epigenetic regulation of genomic integrity. *Chromosoma* *121*, 131-151.

Deng, W., Liu, C., Pei, Y., Deng, X., Niu, L., and Cao, X. (2007). Involvement of the histone acetyltransferase AtHAC1 in the regulation of flowering time via repression of FLOWERING LOCUS C in Arabidopsis. *Plant Physiol* *143*, 1660-1668.

Deng, X., Gu, L., Liu, C., Lu, T., Lu, F., Lu, Z., Cui, P., Pei, Y., Wang, B., Hu, S., *et al.* (2010). Arginine methylation mediated by the Arabidopsis homolog of PRMT5 is essential for proper pre-mRNA splicing. *Proc Natl Acad Sci U S A* *107*, 19114-19119.

Dewitte, W., Scofield, S., Alcasabas, A.A., Maughan, S.C., Menges, M., Braun, N., Collins, C., Nieuwland, J., Prinsen, E., Sundaresan, V., *et al.* (2007). Arabidopsis CYCD3 D-type cyclins link cell proliferation and endocycles and are rate-limiting for cytokinin responses. *Proc Natl Acad Sci U S A* *104*, 14537-14542.

Deyter, G.M., Hildebrand, E.M., Barber, A.D., and Biggins, S. (2017). Histone H4 Facilitates the Proteolysis of the Budding Yeast CENP-ACse4 Centromeric Histone Variant. *Genetics* *205*, 113-124.

Dias, B.G., and Ressler, K.J. (2014). Parental olfactory experience influences behavior and neural structure in subsequent generations. *Nat Neurosci* *17*, 89-96.

Dion, M.F., Altschuler, S.J., Wu, L.F., and Rando, O.J. (2005). Genomic characterization reveals a simple histone H4 acetylation code. *Proc Natl Acad Sci U S A* *102*, 5501-5506.

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

Dong, J., Gao, Z., Liu, S., Li, G., Yang, Z., Huang, H., and Xu, L. (2013). SLIDE, the protein interacting domain of Imitation Switch remodelers, binds DDT-domain proteins of different subfamilies in chromatin remodeling complexes. *J Integr Plant Biol* *55*, 928-937.

Dong, J., LeBlanc, C., Poulet, A., Mermaz, B., Villarino, G., Webb, K.M., Joly, V., Mendez, J., Voigt, P., and Jacob, Y. (2020). H3.1K27me1 maintains transcriptional silencing and genome stability by preventing GCN5-mediated histone acetylation. *Plant Cell*.

Dong, J., LeBlanc, C., Poulet, A., Mermaz, B., Villarino, G., Webb, K.M., Joly, V., Mendez, J., Voigt, P., and Jacob, Y. (2021). H3.1K27me1 maintains transcriptional silencing and genome stability by preventing GCN5-mediated histone acetylation. *Plant Cell* *33*, 961-979.

Drane, P., Brault, M.E., Cui, G., Meghani, K., Chaubey, S., Detappe, A., Parnandi, N., He, Y., Zheng, X.F., Botuyan, M.V., *et al.* (2017). TIRR regulates 53BP1 by masking its histone methyl-lysine binding function. *Nature* *543*, 211-216.

Draviam, V.M., Xie, S., and Sorger, P.K. (2004). Chromosome segregation and genomic stability. *Curr Opin Genet Dev* 14, 120-125.

Du, H.N., and Briggs, S.D. (2010). A nucleosome surface formed by histone H4, H2A, and H3 residues is needed for proper histone H3 Lys36 methylation, histone acetylation, and repression of cryptic transcription. *J Biol Chem* 285, 11704-11713.

Du, L.L., Nakamura, T.M., and Russell, P. (2006). Histone modification-dependent and -independent pathways for recruitment of checkpoint protein Crb2 to double-strand breaks. *Genes Dev* 20, 1583-1596.

Duran-Figueroa, N., and Vielle-Calzada, J.P. (2010). ARGONAUTE9-dependent silencing of transposable elements in pericentromeric regions of Arabidopsis. *Plant Signal Behav* 5, 1476-1479.

Durkin, S.G., and Glover, T.W. (2007). Chromosome fragile sites. *Annu Rev Genet* 41, 169-192.

Duronio, R.J., and Marzluff, W.F. (2017). Coordinating cell cycle-regulated histone gene expression through assembly and function of the Histone Locus Body. *RNA Biol* 14, 726-738.

Earley, K.W., Shook, M.S., Brower-Toland, B., Hicks, L., and Pikaard, C.S. (2007). In vitro specificities of Arabidopsis co-activator histone acetyltransferases: implications for histone hyperacetylation in gene activation. *Plant J* 52, 615-626.

Ebel, C., Mariconti, L., and Gruissem, W. (2004). Plant retinoblastoma homologues control nuclear proliferation in the female gametophyte. *Nature* 429, 776-780.

Ebert, A., Schotta, G., Lein, S., Kubicek, S., Krauss, V., Jenuwein, T., and Reuter, G. (2004). Su(var) genes regulate the balance between euchromatin and heterochromatin in Drosophila. *Genes Dev* 18, 2973-2983.

Edwards, C.R., Dang, W., and Berger, S.L. (2011). Histone H4 lysine 20 of *Saccharomyces cerevisiae* is monomethylated and functions in subtelomeric silencing. *Biochemistry* 50, 10473-10483.

El-Brolosy, M.A., Kontarakis, Z., Rossi, A., Kuenne, C., Gunther, S., Fukuda, N., Kikhi, K., Boezio, G.L.M., Takacs, C.M., Lai, S.L., *et al.* (2019). Genetic compensation triggered by mutant mRNA degradation. *Nature* 568, 193-197.

Erdel, F., and Rippe, K. (2011). Chromatin remodelling in mammalian cells by ISWI-type complexes--where, when and why? *FEBS J* 278, 3608-3618.

Eriksson, P.R., Ganguli, D., Nagarajavel, V., and Clark, D.J. (2012). Regulation of histone gene expression in budding yeast. *Genetics* 191, 7-20.

Eriksson, S., Bohlenius, H., Moritz, T., and Nilsson, O. (2006). GA4 is the active gibberellin in the regulation of LEAFY transcription and Arabidopsis floral initiation. *Plant Cell* 18, 2172-2181.

Espinosa-Cores, L., Bouza-Morcillo, L., Barrero-Gil, J., Jimenez-Suarez, V., Lazaro, A., Piqueras, R., Jarillo, J.A., and Pineiro, M. (2020). Insights Into the Function of the NuA4 Complex in Plants. *Front Plant Sci* 11, 125.

Fan, Y., Nikitina, T., Zhao, J., Fleury, T.J., Bhattacharyya, R., Bouhassira, E.E., Stein, A., Woodcock, C.L., and Skoultchi, A.I. (2005). Histone H1 depletion in mammals alters global chromatin structure but causes specific changes in gene regulation. *Cell* 123, 1199-1212.

Faucher, D., and Wellinger, R.J. (2010). Methylated H3K4, a transcription-associated histone modification, is involved in the DNA damage response pathway. *PLoS Genet* 6.

Fausser, F., Roth, N., Pacher, M., Ilg, G., Sanchez-Fernandez, R., Biesgen, C., and Puchta, H. (2012). In planta gene targeting. *Proc Natl Acad Sci U S A* 109, 7535-7540.

Fazio, T.G., Gelbart, M.E., and Tsukiyama, T. (2005). Two distinct mechanisms of chromatin interaction by the Isw2 chromatin remodeling complex in vivo. *Mol Cell Biol* 25, 9165-9174.

Fei, J., Torigoe, S.E., Brown, C.R., Khuong, M.T., Kassavetis, G.A., Boeger, H., and Kadonaga, J.T. (2015). The prenucleosome, a stable conformational isomer of the nucleosome. *Genes Dev* 29, 2563-2575.

Feng, J.X., and Riddle, N.C. (2020). Epigenetics and genome stability. *Mamm Genome*.

Feng, S., Cokus, S.J., Schubert, V., Zhai, J., Pellegrini, M., and Jacobsen, S.E. (2014a). Genome-wide Hi-C analyses in wild-type and mutants reveal high-resolution chromatin interactions in Arabidopsis. *Mol Cell* 55, 694-707.

Feng, W., Hale, C.J., Over, R.S., Cokus, S.J., Jacobsen, S.E., and Michaels, S.D. (2017). Large-scale heterochromatin remodeling linked to overreplication-associated DNA damage. *Proc Natl Acad Sci U S A* 114, 406-411.

Feng, Y., Hadjikyriacou, A., and Clarke, S.G. (2014b). Substrate specificity of human protein arginine methyltransferase 7 (PRMT7): the importance of acidic residues in the double E loop. *J Biol Chem* 289, 32604-32616.

Feng, Y., Maity, R., Whitelegge, J.P., Hadjikyriacou, A., Li, Z., Zurita-Lopez, C., Al-Hadid, Q., Clark, A.T., Bedford, M.T., Masson, J.Y., *et al.* (2013). Mammalian protein arginine methyltransferase 7 (PRMT7) specifically targets RXR sites in lysine- and arginine-rich regions. *J Biol Chem* 288, 37010-37025.

Feng, Z., Mao, Y., Xu, N., Zhang, B., Wei, P., Yang, D.L., Wang, Z., Zhang, Z., Zheng, R., Yang, L., *et al.* (2014c). Multigeneration analysis reveals the inheritance, specificity, and patterns of CRISPR/Cas-induced gene modifications in Arabidopsis. *Proc Natl Acad Sci U S A* 111, 4632-4637.

Fenley, A.T., Anandakrishnan, R., Kidane, Y.H., and Onufriev, A.V. (2018). Modulation of nucleosomal DNA accessibility via charge-altering post-translational modifications in histone core. *Epigenetics Chromatin* 11, 11.

Fingerman, I.M., Li, H.C., and Briggs, S.D. (2007). A charge-based interaction between histone H4 and Dot1 is required for H3K79 methylation and telomere silencing: identification of a new trans-histone pathway. *Genes Dev* 21, 2018-2029.

Finnegan, E.J., Genger, R.K., Peacock, W.J., and Dennis, E.S. (1998). DNA Methylation in Plants. *Annu Rev Plant Physiol Plant Mol Biol* 49, 223-247.

Fischer, A., Hofmann, I., Naumann, K., and Reuter, G. (2006). Heterochromatin proteins and the control of heterochromatic gene silencing in Arabidopsis. *J Plant Physiol* 163, 358-368.

Foltz, D.R., Jansen, L.E., Bailey, A.O., Yates, J.R., 3rd, Bassett, E.A., Wood, S., Black, B.E., and Cleveland, D.W. (2009). Centromere-specific assembly of CENP-a nucleosomes is mediated by HJURP. *Cell* 137, 472-484.

Fornara, F., Panigrahi, K.C., Gissot, L., Sauerbrunn, N., Ruhl, M., Jarillo, J.A., and Coupland, G. (2009). Arabidopsis DOF transcription factors act redundantly to reduce CONSTANS expression and are essential for a photoperiodic flowering response. *Dev Cell* 17, 75-86.

Freidkin, I., and Katcoff, D.J. (2001). Specific distribution of the *Saccharomyces cerevisiae* linker histone homolog HHO1p in the chromatin. *Nucleic Acids Res* 29, 4043-4051.

Fromm, M., Taylor, L.P., and Walbot, V. (1985). Expression of genes transferred into monocot and dicot plant cells by electroporation. *Proc Natl Acad Sci U S A* 82, 5824-5828.

Fu, Y., Foden, J.A., Khayter, C., Maeder, M.L., Reyon, D., Joung, J.K., and Sander, J.D. (2013). High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells. *Nat Biotechnol* 31, 822-826.

Fu, Y., Zhu, Z., Meng, G., Zhang, R., and Zhang, Y. (2021). A CRISPR-Cas9 based shuffle system for endogenous histone H3 and H4 combinatorial mutagenesis. *Sci Rep* 11, 3298.

Gaillard, H., and Aguilera, A. (2016). Transcription as a Threat to Genome Integrity. *Annu Rev Biochem* 85, 291-317.

Gaillard, P.H.L., Martini, E.M., Kaufman, P.D., Stillman, B., Moustacchi, E., and Almouzni, G. (1996). Chromatin assembly coupled to DNA repair: a new role for chromatin assembly factor I. *Cell* 86, 887-896.

Galbraith, D.W., Harkins, K.R., and Knapp, S. (1991). Systemic Endopolyploidy in *Arabidopsis thaliana*. *Plant Physiol* 96, 985-989.

Gamarra, N., Johnson, S.L., Trnka, M.J., Burlingame, A.L., and Narlikar, G.J. (2018). The nucleosomal acidic patch relieves auto-inhibition by the ISWI remodeler SNF2h. *Elife* 7.

Gelvin, S.B. (2017). Integration of *Agrobacterium* T-DNA into the Plant Genome. *Annu Rev Genet* 51, 195-217.

Gendall, A.R., Levy, Y.Y., Wilson, A., and Dean, C. (2001). The VERNALIZATION 2 gene mediates the epigenetic regulation of vernalization in *Arabidopsis*. *Cell* 107, 525-535.

Geraldo, N., Baurle, I., Kidou, S., Hu, X., and Dean, C. (2009). FRIGIDA delays flowering in *Arabidopsis* via a cotranscriptional mechanism involving direct interaction with the nuclear cap-binding complex. *Plant Physiol* 150, 1611-1618.

Gil, K.E., Park, M.J., Lee, H.J., Park, Y.J., Han, S.H., Kwon, Y.J., Seo, P.J., Jung, J.H., and Park, C.M. (2017). Alternative splicing provides a proactive mechanism for the diurnal CONSTANS dynamics in *Arabidopsis* photoperiodic flowering. *Plant J* 89, 128-140.

Gkikopoulos, T., Schofield, P., Singh, V., Pinskaya, M., Mellor, J., Smolle, M., Workman, J.L., Barton, G.J., and Owen-Hughes, T. (2011). A role for Snf2-related nucleosome-spacing enzymes in genome-wide nucleosome organization. *Science* 333, 1758-1760.

Goldberg, A.D., Banaszynski, L.A., Noh, K.M., Lewis, P.W., Elsaesser, S.J., Stadler, S., Dewell, S., Law, M., Guo, X., Li, X., *et al.* (2010). Distinct factors control histone variant H3.3 localization at specific genomic regions. *Cell* 140, 678-691.

Gomez-Zambrano, A., Merini, W., and Calonje, M. (2019). The repressive role of *Arabidopsis* H2A.Z in transcriptional regulation depends on AtBMI1 activity. *Nat Commun* 10, 2828.

Gonsalvez, G.B., Tian, L., Ospina, J.K., Boisvert, F.M., Lamond, A.I., and Matera, A.G. (2007). Two distinct arginine methyltransferases are required for biogenesis of Sm-class ribonucleoproteins. *J Cell Biol* 178, 733-740.

Goodrich, J., Puangsomlee, P., Martin, M., Long, D., Meyerowitz, E.M., and Coupland, G. (1997). A Polycomb-group gene regulates homeotic gene expression in *Arabidopsis*. *Nature* 386, 44-51.

Govin, J., Dorsey, J., Gaucher, J., Rousseaux, S., Khochbin, S., and Berger, S.L. (2010a). Systematic screen reveals new functional dynamics of histones H3 and H4 during gametogenesis. *Genes Dev* 24, 1772-1786.

Govin, J., Schug, J., Krishnamoorthy, T., Dorsey, J., Khochbin, S., and Berger, S.L. (2010b). Genome-wide mapping of histone H4 serine-1 phosphorylation during sporulation in *Saccharomyces cerevisiae*. *Nucleic Acids Res* 38, 4599-4606.

Greenberg, M.V.C., and Bourc'his, D. (2019). The diverse roles of DNA methylation in mammalian development and disease. *Nat Rev Mol Cell Biol* 20, 590-607.

Grewal, S.I., and Jia, S. (2007). Heterochromatin revisited. *Nat Rev Genet* 8, 35-46.

Groth, A., Corpet, A., Cook, A.J., Roche, D., Bartek, J., Lukas, J., and Almouzni, G. (2007). Regulation of replication fork progression through histone supply and demand. *Science* 318, 1928-1931.

Gu, X., Jiang, D., Wang, Y., Bachmair, A., and He, Y. (2009). Repression of the floral transition via histone H2B monoubiquitination. *Plant J* 57, 522-533.

Gunesdogan, U., Jackle, H., and Herzig, A. (2010). A genetic system to assess in vivo the functions of histones and histone modifications in higher eukaryotes. *EMBO Rep* 11, 772-776.

Gunesdogan, U., Jackle, H., and Herzig, A. (2014). Histone supply regulates S phase timing and cell cycle progression. *Elife* 3, e02443.

Gunjan, A., and Verreault, A. (2003). A Rad53 kinase-dependent surveillance mechanism that regulates histone protein levels in *S. cerevisiae*. *Cell* 115, 537-549.

Gusti, A., Baumberger, N., Nowack, M., Pusch, S., Eisler, H., Potuschak, T., De Veylder, L., Schnittger, A., and Genschik, P. (2009). The Arabidopsis thaliana F-box protein FBL17 is essential for progression through the second mitosis during pollen development. *PLoS One* 4, e4780.

Hale, C.J., Potok, M.E., Lopez, J., Do, T., Liu, A., Gallego-Bartolome, J., Michaels, S.D., and Jacobsen, S.E. (2016). Identification of Multiple Proteins Coupling Transcriptional Gene Silencing to Genome Stability in Arabidopsis thaliana. *PLoS Genet* 12, e1006092.

Hamiche, A., Kang, J.G., Dennis, C., Xiao, H., and Wu, C. (2001). Histone tails modulate nucleosome mobility and regulate ATP-dependent nucleosome sliding by NURF. *Proc Natl Acad Sci U S A* 98, 14316-14321.

Hansen, K.H., Bracken, A.P., Pasini, D., Dietrich, N., Gehani, S.S., Monrad, A., Rappsilber, J., Lerdrup, M., and Helin, K. (2008). A model for transmission of the H3K27me3 epigenetic mark. *Nat Cell Biol* 10, 1291-1300.

Hartley, P.D., and Madhani, H.D. (2009). Mechanisms that specify promoter nucleosome location and identity. *Cell* 137, 445-458.

Hartmann, U., Hohmann, S., Nettesheim, K., Wisman, E., Saedler, H., and Huijser, P. (2000). Molecular cloning of SVP: a negative regulator of the floral transition in Arabidopsis. *Plant J* 21, 351-360.

Hashimoto, H., Takami, Y., Sonoda, E., Iwasaki, T., Iwano, H., Tachibana, M., Takeda, S., Nakayama, T., Kimura, H., and Shinkai, Y. (2010). Histone H1 null vertebrate cells exhibit altered nucleosome architecture. *Nucleic Acids Res* 38, 3533-3545.

Hauk, G., McKnight, J.N., Nodelman, I.M., and Bowman, G.D. (2010). The chromodomains of the Chd1 chromatin remodeler regulate DNA access to the ATPase motor. *Mol Cell* 39, 711-723.

Hauser, M.T., Aufsatz, W., Jonak, C., and Luschnig, C. (2011). Transgenerational epigenetic inheritance in plants. *Biochim Biophys Acta* 1809, 459-468.

He, F., and Jacobson, A. (2015). Nonsense-Mediated mRNA Decay: Degradation of Defective Transcripts Is Only Part of the Story. *Annu Rev Genet* 49, 339-366.

He, Y. (2009). Control of the transition to flowering by chromatin modifications. *Mol Plant* 2, 554-564.

He, Y., and Amasino, R.M. (2005). Role of chromatin modification in flowering-time control. *Trends Plant Sci* 10, 30-35.

He, Y., Doyle, M.R., and Amasino, R.M. (2004). PAF1-complex-mediated histone methylation of FLOWERING LOCUS C chromatin is required for the vernalization-responsive, winter-annual habit in Arabidopsis. *Genes Dev* 18, 2774-2784.

Hecht, A., Strahl-Bolsinger, S., and Grunstein, M. (1996). Spreading of transcriptional repressor SIR3 from telomeric heterochromatin. *Nature* 383, 92-96.

Heintzman, N.D., Stuart, R.K., Hon, G., Fu, Y., Ching, C.W., Hawkins, R.D., Barrera, L.O., Van Calcar, S., Qu, C., Ching, K.A., et al. (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. *Nat Genet* 39, 311-318.

Helliwell, C.A., Wood, C.C., Robertson, M., James Peacock, W., and Dennis, E.S. (2006). The Arabidopsis FLC protein interacts directly in vivo with SOC1 and FT chromatin and is part of a high-molecular-weight protein complex. *Plant J* 46, 183-192.

Henderson, I.R., and Jacobsen, S.E. (2007). Epigenetic inheritance in plants. *Nature* 447, 418-424.

Hennig, B.P., Bendrin, K., Zhou, Y., and Fischer, T. (2012). Chd1 chromatin remodelers maintain nucleosome organization and repress cryptic transcription. *EMBO Rep* 13, 997-1003.

Hepworth, S.R., Valverde, F., Ravenscroft, D., Mouradov, A., and Coupland, G. (2002). Antagonistic regulation of flowering-time gene *SOC1* by *CONSTANS* and *FLC* via separate promoter motifs. *EMBO J* 21, 4327-4337.

Hermann, A., Goyal, R., and Jeltsch, A. (2004). The Dnmt1 DNA-(cytosine-C5)-methyltransferase methylates DNA processively with high preference for hemimethylated target sites. *J Biol Chem* 279, 48350-48359.

Herz, H.M., Morgan, M., Gao, X., Jackson, J., Rickels, R., Swanson, S.K., Florens, L., Washburn, M.P., Eissenberg, J.C., and Shilatifard, A. (2014). Histone H3 lysine-to-methionine mutants as a paradigm to study chromatin signaling. *Science* 345, 1065-1070.

Heslop-Harrison, J.S., Brandes, A., and Schwarzacher, T. (2003). Tandemly repeated DNA sequences and centromeric chromosomal regions of *Arabidopsis* species. *Chromosome Res* 11, 241-253.

Higgs, M.R., Sato, K., Reynolds, J.J., Begum, S., Bayley, R., Goula, A., Vernet, A., Paquin, K.L., Skalnik, D.G., Kobayashi, W., *et al.* (2018). Histone Methylation by SETD1A Protects Nascent DNA through the Nucleosome Chaperone Activity of FANCD2. *Mol Cell* 71, 25-41 e26.

Hirochika, H., Okamoto, H., and Kakutani, T. (2000). Silencing of retrotransposons in *Arabidopsis* and reactivation by the *ddm1* mutation. *Plant Cell* 12, 357-369.

Hisamatsu, T., and King, R.W. (2008). The nature of floral signals in *Arabidopsis*. II. Roles for *FLOWERING LOCUS T* (*FT*) and gibberellin. *J Exp Bot* 59, 3821-3829.

Hisanaga, T., Ferjani, A., Horiguchi, G., Ishikawa, N., Fujikura, U., Kubo, M., Demura, T., Fukuda, H., Ishida, T., Sugimoto, K., *et al.* (2013). The ATM-dependent DNA damage response acts as an upstream trigger for compensation in the *fas1* mutation during *Arabidopsis* leaf development. *Plant Physiol* 162, 831-841.

Hodl, M., and Basler, K. (2009). Transcription in the absence of histone H3.3. *Curr Biol* 19, 1221-1226.

Hodl, M., and Basler, K. (2012). Transcription in the absence of histone H3.2 and H3K4 methylation. *Curr Biol* 22, 2253-2257.

Hoek, M., and Stillman, B. (2003). Chromatin assembly factor 1 is essential and couples chromatin assembly to DNA replication in vivo. *Proc Natl Acad Sci U S A* 100, 12183-12188.

Holmes, W.F., Braastad, C.D., Mitra, P., Hampe, C., Doenecke, D., Albig, W., Stein, J.L., van Wijnen, A.J., and Stein, G.S. (2005). Coordinate control and selective expression of the full complement of replication-dependent histone H4 genes in normal and cancer cells. *J Biol Chem* 280, 37400-37407.

Howe, F.S., Fischl, H., Murray, S.C., and Mellor, J. (2017). Is H3K4me3 instructive for transcription activation? *Bioessays* 39, 1-12.

Hu, J., Donahue, G., Dorsey, J., Govin, J., Yuan, Z., Garcia, B.A., Shah, P.P., and Berger, S.L. (2015). H4K44 Acetylation Facilitates Chromatin Accessibility during Meiosis. *Cell Rep* 13, 1772-1780.

Hu, J., Gu, L., Ye, Y., Zheng, M., Xu, Z., Lin, J., Du, Y., Tian, M., Luo, L., Wang, B., *et al.* (2018). Dynamic placement of the linker histone H1 associated with nucleosome arrangement and gene transcription in early *Drosophila* embryonic development. *Cell Death Dis* 9, 765.

Huang, T.K., and Puchta, H. (2019). CRISPR/Cas-mediated gene targeting in plants: finally a turn for the better for homologous recombination. *Plant Cell Rep* 38, 443-453.

Huen, M.S., Sy, S.M., van Deursen, J.M., and Chen, J. (2008). Direct interaction between SET8 and proliferating cell nuclear antigen couples H4-K20 methylation with DNA replication. *J Biol Chem* 283, 11073-11077.

Hughes, A.L., Jin, Y., Rando, O.J., and Struhl, K. (2012). A functional evolutionary approach to identify determinants of nucleosome positioning: a unifying model for establishing the genome-wide pattern. *Mol Cell* 48, 5-15.

Huther, P., Schandry, N., Jandrasits, K., Bezrukov, I., and Becker, C. (2020). ARADEEPOPSIS, an Automated Workflow for Top-View Plant Phenomics using Semantic Segmentation of Leaf States. *Plant Cell* 32, 3674-3688.

Hwang, W.L., Deindl, S., Harada, B.T., and Zhuang, X. (2014). Histone H4 tail mediates allosteric regulation of nucleosome remodelling by linker DNA. *Nature* 512, 213-217.

Hyland, E.M., Cosgrove, M.S., Molina, H., Wang, D., Pandey, A., Cottee, R.J., and Boeke, J.D. (2005). Insights into the role of histone H3 and histone H4 core modifiable residues in *Saccharomyces cerevisiae*. *Mol Cell Biol* 25, 10060-10070.

Imaizumi, T., Schultz, T.F., Harmon, F.G., Ho, L.A., and Kay, S.A. (2005). FKF1 F-box protein mediates cyclic degradation of a repressor of CONSTANS in *Arabidopsis*. *Science* 309, 293-297.

Imaizumi, T., Tran, H.G., Swartz, T.E., Briggs, W.R., and Kay, S.A. (2003). FKF1 is essential for photoperiodic-specific light signalling in *Arabidopsis*. *Nature* 426, 302-306.

Ioshikhes, I., Bolshoy, A., Derenshteyn, K., Borodovsky, M., and Trifonov, E.N. (1996). Nucleosome DNA sequence pattern revealed by multiple alignment of experimentally mapped sequences. *J Mol Biol* 262, 129-139.

Ioshikhes, I.P., Albert, I., Zanton, S.J., and Pugh, B.F. (2006). Nucleosome positions predicted through comparative genomics. *Nat Genet* 38, 1210-1215.

Ishiyama, S., Nishiyama, A., Saeki, Y., Moritsugu, K., Morimoto, D., Yamaguchi, L., Arai, N., Matsumura, R., Kawakami, T., Mishima, Y., *et al.* (2017). Structure of the Dnmt1 Reader Module Complexed with a Unique Two-Mono-Ubiquitin Mark on Histone H3 Reveals the Basis for DNA Methylation Maintenance. *Mol Cell* 68, 350-360 e357.

Iwasaki, W., Miya, Y., Horikoshi, N., Osakabe, A., Taguchi, H., Tachiwana, H., Shibata, T., Kagawa, W., and Kurumizaka, H. (2013). Contribution of histone N-terminal tails to the structure and stability of nucleosomes. *FEBS Open Bio* 3, 363-369.

Jablonka, E., Lamb, M.J., and Avital, E. (1998). 'Lamarckian' mechanisms in darwinian evolution. *Trends Ecol Evol* 13, 206-210.

Jacob, Y., Bergamin, E., Donoghue, M.T., Mongeon, V., LeBlanc, C., Voigt, P., Underwood, C.J., Brunzelle, J.S., Michaels, S.D., Reinberg, D., *et al.* (2014). Selective methylation of histone H3 variant H3.1 regulates heterochromatin replication. *Science* 343, 1249-1253.

Jacob, Y., Feng, S., LeBlanc, C.A., Bernatavichute, Y.V., Stroud, H., Cokus, S., Johnson, L.M., Pellegrini, M., Jacobsen, S.E., and Michaels, S.D. (2009). ATXR5 and ATXR6 are H3K27 monomethyltransferases required for chromatin structure and gene silencing. *Nat Struct Mol Biol* 16, 763-768.

Jacob, Y., Stroud, H., Leblanc, C., Feng, S., Zhuo, L., Caro, E., Hassel, C., Gutierrez, C., Michaels, S.D., and Jacobsen, S.E. (2010). Regulation of heterochromatic DNA replication by histone H3 lysine 27 methyltransferases. *Nature* 466, 987-991.

Jaeger, K.E., and Wigge, P.A. (2007). FT protein acts as a long-range signal in *Arabidopsis*. *Curr Biol* 17, 1050-1054.

Jain, K., and Clarke, S.G. (2019). PRMT7 as a unique member of the protein arginine methyltransferase family: A review. *Arch Biochem Biophys* 665, 36-45.

Jain, K., Jin, C.Y., and Clarke, S.G. (2017). Epigenetic control via allosteric regulation of mammalian protein arginine methyltransferases. *Proc Natl Acad Sci U S A* *114*, 10101-10106.

Jansen, L.E., Black, B.E., Foltz, D.R., and Cleveland, D.W. (2007). Propagation of centromeric chromatin requires exit from mitosis. *J Cell Biol* *176*, 795-805.

Jasencakova, Z., Scharf, A.N., Ask, K., Corpet, A., Imhof, A., Almouzni, G., and Groth, A. (2010). Replication stress interferes with histone recycling and predeposition marking of new histones. *Mol Cell* *37*, 736-743.

Jeong, H.J., Lee, H.J., Vuong, T.A., Choi, K.S., Choi, D., Koo, S.H., Cho, S.C., Cho, H., and Kang, J.S. (2016). Prmt7 Deficiency Causes Reduced Skeletal Muscle Oxidative Metabolism and Age-Related Obesity. *Diabetes* *65*, 1868-1882.

Jia, H., and Wang, N. (2014). Targeted genome editing of sweet orange using Cas9/sgRNA. *PLoS One* *9*, e93806.

Jiang, C., and Pugh, B.F. (2009). Nucleosome positioning and gene regulation: advances through genomics. *Nat Rev Genet* *10*, 161-172.

Jiang, D., Wang, Y., Wang, Y., and He, Y. (2008). Repression of FLOWERING LOCUS C and FLOWERING LOCUS T by the Arabidopsis Polycomb repressive complex 2 components. *PLoS One* *3*, e3404.

Jiang, S., Liu, Y., Wang, A., Qin, Y., Luo, M., Wu, Q., Boeke, J.D., and Dai, J. (2017). Construction of Comprehensive Dosage-Matching Core Histone Mutant Libraries for *Saccharomyces cerevisiae*. *Genetics* *207*, 1263-1273.

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J.A., and Charpentier, E. (2012). A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* *337*, 816-821.

Jing, Y., Sun, H., Yuan, W., Wang, Y., Li, Q., Liu, Y., Li, Y., and Qian, W. (2016). SUVH2 and SUVH9 Couple Two Essential Steps for Transcriptional Gene Silencing in Arabidopsis. *Mol Plant* *9*, 1156-1167.

Johanson, U., West, J., Lister, C., Michaels, S., Amasino, R., and Dean, C. (2000). Molecular analysis of FRIGIDA, a major determinant of natural variation in Arabidopsis flowering time. *Science* *290*, 344-347.

Johnson, L., Mollah, S., Garcia, B.A., Muratore, T.L., Shabanowitz, J., Hunt, D.F., and Jacobsen, S.E. (2004). Mass spectrometry analysis of Arabidopsis histone H3 reveals distinct combinations of post-translational modifications. *Nucleic Acids Res* *32*, 6511-6518.

Jones, P.A. (2012). Functions of DNA methylation: islands, start sites, gene bodies and beyond. *Nat Rev Genet* *13*, 484-492.

Jorgensen, S., Schotta, G., and Sorensen, C.S. (2013). Histone H4 lysine 20 methylation: key player in epigenetic regulation of genomic integrity. *Nucleic Acids Res* *41*, 2797-2806.

Joshi, A.A., and Struhl, K. (2005). Eaf3 chromodomain interaction with methylated H3-K36 links histone deacetylation to Pol II elongation. *Mol Cell* *20*, 971-978.

Jung, C., Pillen, K., Staiger, D., Coupland, G., and von Korff, M. (2016). Editorial: Recent Advances in Flowering Time Control. *Front Plant Sci* *7*, 2011.

Jung, G.A., Shin, B.S., Jang, Y.S., Sohn, J.B., Woo, S.R., Kim, J.E., Choi, G., Lee, K.M., Min, B.H., Lee, K.H., *et al.* (2011). Methylation of eukaryotic elongation factor 2 induced by basic fibroblast growth factor via mitogen-activated protein kinase. *Exp Mol Med* *43*, 550-560.

Jung, J.H., Seo, Y.H., Seo, P.J., Reyes, J.L., Yun, J., Chua, N.H., and Park, C.M. (2007). The GIGANTEA-regulated microRNA172 mediates photoperiodic flowering independent of CONSTANS in Arabidopsis. *Plant Cell* *19*, 2736-2748.

Kaimori, J.Y., Maehara, K., Hayashi-Takanaka, Y., Harada, A., Fukuda, M., Yamamoto, S., Ichimaru, N., Umehara, T., Yokoyama, S., Matsuda, R., *et al.* (2016). Histone H4 lysine 20 acetylation is associated with gene repression in human cells. *Sci Rep* 6, 24318.

Kalashnikova, A.A., Porter-Goff, M.E., Muthurajan, U.M., Luger, K., and Hansen, J.C. (2013). The role of the nucleosome acidic patch in modulating higher order chromatin structure. *J R Soc Interface* 10, 20121022.

Kamakaka, R.T., and Biggins, S. (2005). Histone variants: deviants? *Genes Dev* 19, 295-310.

Kan, P.Y., Caterino, T.L., and Hayes, J.J. (2009). The H4 tail domain participates in intra- and internucleosome interactions with protein and DNA during folding and oligomerization of nucleosome arrays. *Mol Cell Biol* 29, 538-546.

Kapahnke, M., Banning, A., and Tikkanen, R. (2016). Random Splicing of Several Exons Caused by a Single Base Change in the Target Exon of CRISPR/Cas9 Mediated Gene Knockout. *Cells* 5.

Kaplan, N., Moore, I.K., Fondufe-Mittendorf, Y., Gossett, A.J., Tillo, D., Field, Y., LeProust, E.M., Hughes, T.R., Lieb, J.D., Widom, J., *et al.* (2009). The DNA-encoded nucleosome organization of a eukaryotic genome. *Nature* 458, 362-366.

Karachentsev, D., Sarma, K., Reinberg, D., and Steward, R. (2005). PR-Set7-dependent methylation of histone H4 Lys 20 functions in repression of gene expression and is essential for mitosis. *Genes Dev* 19, 431-435.

Kawashima, T., and Berger, F. (2014). Epigenetic reprogramming in plant sexual reproduction. *Nat Rev Genet* 15, 613-624.

Kaygun, H., and Marzluff, W.F. (2005a). Regulated degradation of replication-dependent histone mRNAs requires both ATR and Upf1. *Nat Struct Mol Biol* 12, 794-800.

Kaygun, H., and Marzluff, W.F. (2005b). Translation termination is involved in histone mRNA degradation when DNA replication is inhibited. *Mol Cell Biol* 25, 6879-6888.

Kayne, P.S., Kim, U.J., Han, M., Mullen, J.R., Yoshizaki, F., and Grunstein, M. (1988). Extremely conserved histone H4 N terminus is dispensable for growth but essential for repressing the silent mating loci in yeast. *Cell* 55, 27-39.

Kebede, A.F., Schneider, R., and Daujat, S. (2015). Novel types and sites of histone modifications emerge as players in the transcriptional regulation contest. *FEBS J* 282, 1658-1674.

Keogh, M.C., Kurdistani, S.K., Morris, S.A., Ahn, S.H., Podolny, V., Collins, S.R., Schuldiner, M., Chin, K., Punna, T., Thompson, N.J., *et al.* (2005). Cotranscriptional set2 methylation of histone H3 lysine 36 recruits a repressive Rpd3 complex. *Cell* 123, 593-605.

Khorasanizadeh, S. (2004). The nucleosome: from genomic organization to genomic regulation. *Cell* 116, 259-272.

Kim, G.T., Tsukaya, H., and Uchimiya, H. (1998). The CURLY LEAF gene controls both division and elongation of cells during the expansion of the leaf blade in *Arabidopsis thaliana*. *Planta* 206, 175-183.

Kim, S.I., Veena, and Gelvin, S.B. (2007). Genome-wide analysis of *Agrobacterium* T-DNA integration sites in the *Arabidopsis* genome generated under non-selective conditions. *Plant J* 51, 779-791.

Kim, S.Y., He, Y., Jacob, Y., Noh, Y.S., Michaels, S., and Amasino, R. (2005). Establishment of the vernalization-responsive, winter-annual habit in *Arabidopsis* requires a putative histone H3 methyl transferase. *Plant Cell* 17, 3301-3310.

Kimura, A., Umehara, T., and Horikoshi, M. (2002). Chromosomal gradient of histone acetylation established by Sas2p and Sir2p functions as a shield against gene silencing. *Nat Genet* 32, 370-377.

Kimura, H., and Cook, P.R. (2001). Kinetics of core histones in living human cells: little exchange of H3 and H4 and some rapid exchange of H2B. *J Cell Biol* 153, 1341-1353.

Kitevski-LeBlanc, J.L., Yuwen, T., Dyer, P.N., Rudolph, J., Luger, K., and Kay, L.E. (2018). Investigating the Dynamics of Destabilized Nucleosomes Using Methyl-TROSY NMR. *J Am Chem Soc* 140, 4774-4777.

Knizewski, L., Ginalski, K., and Jerzmanowski, A. (2008). Snf2 proteins in plants: gene silencing and beyond. *Trends Plant Sci* 13, 557-565.

Koncz, C., Martini, N., Mayerhofer, R., Koncz-Kalman, Z., Korber, H., Redei, G.P., and Schell, J. (1989). High-frequency T-DNA-mediated gene tagging in plants. *Proc Natl Acad Sci U S A* 86, 8467-8471.

Konev, A.Y., Tribus, M., Park, S.Y., Podhraski, V., Lim, C.Y., Emelyanov, A.V., Vershilova, E., Pirrotta, V., Kadonaga, J.T., Lusser, A., *et al.* (2007). CHD1 motor protein is required for deposition of histone variant H3.3 into chromatin in vivo. *Science* 317, 1087-1090.

Koornneef, M., Alonso-Blanco, C., Blankestijn-de Vries, H., Hanhart, C.J., and Peeters, A.J. (1998). Genetic interactions among late-flowering mutants of Arabidopsis. *Genetics* 148, 885-892.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell* 128, 693-705.

Krishnamoorthy, T., Chen, X., Govin, J., Cheung, W.L., Dorsey, J., Schindler, K., Winter, E., Allis, C.D., Guacci, V., Khochbin, S., *et al.* (2006). Phosphorylation of histone H4 Ser1 regulates sporulation in yeast and is conserved in fly and mouse spermatogenesis. *Genes Dev* 20, 2580-2592.

Kruger, W., Peterson, C.L., Sil, A., Coburn, C., Arents, G., Moudrianakis, E.N., and Herskowitz, I. (1995). Amino acid substitutions in the structured domains of histones H3 and H4 partially relieve the requirement of the yeast SWI/SNF complex for transcription. *Genes Dev* 9, 2770-2779.

Kubo, H., Peeters, A.J., Aarts, M.G., Pereira, A., and Koornneef, M. (1999). ANTHOCYANINLESS2, a homeobox gene affecting anthocyanin distribution and root development in Arabidopsis. *Plant Cell* 11, 1217-1226.

Kumar, S.V., and Wigge, P.A. (2010). H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis. *Cell* 140, 136-147.

Kuo, A.J., Song, J., Cheung, P., Ishibe-Murakami, S., Yamazoe, S., Chen, J.K., Patel, D.J., and Gozani, O. (2012). The BAH domain of ORC1 links H4K20me2 to DNA replication licensing and Meier-Gorlin syndrome. *Nature* 484, 115-119.

Lachner, M., O'Carroll, D., Rea, S., Mechtler, K., and Jenuwein, T. (2001). Methylation of histone H3 lysine 9 creates a binding site for HP1 proteins. *Nature* 410, 116-120.

Lafarge, S., and Montane, M.H. (2003). Characterization of Arabidopsis thaliana ortholog of the human breast cancer susceptibility gene 1: AtBRCA1, strongly induced by gamma rays. *Nucleic Acids Res* 31, 1148-1155.

Lalonde, S., Stone, O.A., Lessard, S., Lavertu, A., Desjardins, J., Beaudoin, M., Rivas, M., Stainier, D.Y.R., and Lettre, G. (2017). Frameshift indels introduced by genome editing can lead to in-frame exon skipping. *PLoS One* 12, e0178700.

Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9, 357-359.

Langst, G., and Becker, P.B. (2001). Nucleosome mobilization and positioning by ISWI-containing chromatin-remodeling factors. *J Cell Sci* 114, 2561-2568.

Lanzotti, D.J., Kaygun, H., Yang, X., Duronio, R.J., and Marzluff, W.F. (2002). Developmental control of histone mRNA and dSLBP synthesis during Drosophila embryogenesis and the role of dSLBP in histone mRNA 3' end processing in vivo. *Mol Cell Biol* 22, 2267-2282.

Larkin, M.A., Blackshields, G., Brown, N.P., Chenna, R., McGettigan, P.A., McWilliam, H., Valentin, F., Wallace, I.M., Wilm, A., Lopez, R., *et al.* (2007). Clustal W and Clustal X version 2.0. *Bioinformatics* 23, 2947-2948.

Lauberth, S.M., Nakayama, T., Wu, X., Ferris, A.L., Tang, Z., Hughes, S.H., and Roeder, R.G. (2013). H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* 152, 1021-1036.

Laubinger, S., Marchal, V., Le Gourrierec, J., Wenkel, S., Adrian, J., Jang, S., Kulajta, C., Braun, H., Coupland, G., and Hoecker, U. (2006). Arabidopsis SPA proteins regulate photoperiodic flowering and interact with the floral inducer CONSTANS to regulate its stability. *Development* 133, 3213-3222.

Lawrence, M., Daujat, S., and Schneider, R. (2016). Lateral Thinking: How Histone Modifications Regulate Gene Expression. *Trends Genet* 32, 42-56.

LeBlanc, C., Zhang, F., Mendez, J., Lozano, Y., Chatpar, K., Irish, V., and Jacob, Y. (2017). Increased efficiency of targeted mutagenesis by CRISPR/Cas9 in plants using heat stress. *Plant J*.

Lee, J., Oh, M., Park, H., and Lee, I. (2008). SOC1 translocated to the nucleus by interaction with AGL24 directly regulates leafy. *Plant J* 55, 832-843.

Lee, J.H., Yoo, S.J., Park, S.H., Hwang, I., Lee, J.S., and Ahn, J.H. (2007). Role of SVP in the control of flowering time by ambient temperature in Arabidopsis. *Genes Dev* 21, 397-402.

Lei, B., and Berger, F. (2020). H2A Variants in Arabidopsis: Versatile Regulators of Genome Activity. *Plant Commun* 1, 100015.

Levy, Y.Y., Mesnage, S., Mylne, J.S., Gendall, A.R., and Dean, C. (2002). Multiple roles of Arabidopsis VRN1 in vernalization and flowering time control. *Science* 297, 243-246.

Lewis, P.W., Muller, M.M., Koletsky, M.S., Cordero, F., Lin, S., Banaszynski, L.A., Garcia, B.A., Muir, T.W., Becher, O.J., and Allis, C.D. (2013). Inhibition of PRC2 activity by a gain-of-function H3 mutation found in pediatric glioblastoma. *Science* 340, 857-861.

Li, D., Liu, C., Shen, L., Wu, Y., Chen, H., Robertson, M., Helliwell, C.A., Ito, T., Meyerowitz, E., and Yu, H. (2008). A repressor complex governs the integration of flowering signals in Arabidopsis. *Dev Cell* 15, 110-120.

Li, D., Wan, C., Bai, B., Cao, H., Liu, C., and Zhang, Q. (2019). Identification of histone acetylation markers in human fetal brains and increased H4K5ac expression in neural tube defects. *Mol Genet Genomic Med* 7, e1002.

Li, G., Liu, S., Wang, J., He, J., Huang, H., Zhang, Y., and Xu, L. (2014). ISWI proteins participate in the genome-wide nucleosome distribution in Arabidopsis. *Plant J* 78, 706-714.

Li, G., Zhang, J., Li, J., Yang, Z., Huang, H., and Xu, L. (2012). Imitation Switch chromatin remodeling factors and their interacting RINGLET proteins act together in controlling the plant vegetative phase in Arabidopsis. *Plant J* 72, 261-270.

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., Marth, G., Abecasis, G., Durbin, R., and Genome Project Data Processing, S. (2009). The Sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078-2079.

Li, H., Ilin, S., Wang, W., Duncan, E.M., Wysocka, J., Allis, C.D., and Patel, D.J. (2006). Molecular basis for site-specific read-out of histone H3K4me3 by the BPTF PHD finger of NURF. *Nature* 442, 91-95.

Li, J., Ou-Lee, T.M., Raba, R., Amundson, R.G., and Last, R.L. (1993). Arabidopsis Flavonoid Mutants Are Hypersensitive to UV-B Irradiation. *Plant Cell* 5, 171-179.

Li, J.F., Norville, J.E., Aach, J., McCormack, M., Zhang, D., Bush, J., Church, G.M., and Sheen, J. (2013). Multiplex and homologous recombination-mediated genome editing in Arabidopsis and *Nicotiana benthamiana* using guide RNA and Cas9. *Nat Biotechnol* 31, 688-691.

Li, Y., Sun, L., Zhang, Y., Wang, D., Wang, F., Liang, J., Gui, B., and Shang, Y. (2011a). The histone modifications governing TFF1 transcription mediated by estrogen receptor. *J Biol Chem* 286, 13925-13936.

Li, Z., Nie, F., Wang, S., and Li, L. (2011b). Histone H4 Lys 20 monomethylation by histone methylase SET8 mediates Wnt target gene activation. *Proc Natl Acad Sci U S A* 108, 3116-3123.

Liang, D., Burkhart, S.L., Singh, R.K., Kabbaj, M.H., and Gunjan, A. (2012). Histone dosage regulates DNA damage sensitivity in a checkpoint-independent manner by the homologous recombination pathway. *Nucleic Acids Res* 40, 9604-9620.

Liao, Y., Smyth, G.K., and Shi, W. (2014). featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. *Bioinformatics* 30, 923-930.

Lifton, R.P., Goldberg, M.L., Karp, R.W., and Hogness, D.S. (1978). The organization of the histone genes in *Drosophila melanogaster*: functional and evolutionary implications. *Cold Spring Harb Symp Quant Biol* 42 Pt 2, 1047-1051.

Lim, M.H., Kim, J., Kim, Y.S., Chung, K.S., Seo, Y.H., Lee, I., Kim, J., Hong, C.B., Kim, H.J., and Park, C.M. (2004). A new *Arabidopsis* gene, FLK, encodes an RNA binding protein with K homology motifs and regulates flowering time via FLOWERING LOCUS C. *Plant Cell* 16, 731-740.

Liu, C., Chen, H., Er, H.L., Soo, H.M., Kumar, P.P., Han, J.H., Liou, Y.C., and Yu, H. (2008a). Direct interaction of AGL24 and SOC1 integrates flowering signals in *Arabidopsis*. *Development* 135, 1481-1491.

Liu, C.L., Kaplan, T., Kim, M., Buratowski, S., Schreiber, S.L., Friedman, N., and Rando, O.J. (2005). Single-nucleosome mapping of histone modifications in *S. cerevisiae*. *PLoS Biol* 3, e328.

Liu, L.J., Zhang, Y.C., Li, Q.H., Sang, Y., Mao, J., Lian, H.L., Wang, L., and Yang, H.Q. (2008b). COP1-mediated ubiquitination of CONSTANS is implicated in cryptochrome regulation of flowering in *Arabidopsis*. *Plant Cell* 20, 292-306.

Liu, Z., Hong, S.W., Escobar, M., Vierling, E., Mitchell, D.L., Mount, D.W., and Hall, J.D. (2003). *Arabidopsis* UVH6, a homolog of human XPD and yeast RAD3 DNA repair genes, functions in DNA repair and is essential for plant growth. *Plant Physiol* 132, 1405-1414.

Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2^{-ΔΔC(T)} Method. *Methods* 25, 402-408.

Love, M.I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol* 15, 550.

Lu, D., Wang, T., Persson, S., Mueller-Roeber, B., and Schippers, J.H. (2014). Transcriptional control of ROS homeostasis by KUODA1 regulates cell expansion during leaf development. *Nat Commun* 5, 3767.

Lu, X., Simon, M.D., Chodaparambil, J.V., Hansen, J.C., Shokat, K.M., and Luger, K. (2008). The effect of H3K79 dimethylation and H4K20 trimethylation on nucleosome and chromatin structure. *Nat Struct Mol Biol* 15, 1122-1124.

Ludwigsen, J., Pfennig, S., Singh, A.K., Schindler, C., Harrer, N., Forne, I., Zacharias, M., and Mueller-Planitz, F. (2017). Concerted regulation of ISWI by an autoinhibitory domain and the H4 N-terminal tail. *Elife* 6.

Luense, L.J., Wang, X., Schon, S.B., Weller, A.H., Lin Shiao, E., Bryant, J.M., Bartolomei, M.S., Coutifaris, C., Garcia, B.A., and Berger, S.L. (2016). Comprehensive analysis of histone post-translational modifications in mouse and human male germ cells. *Epigenetics Chromatin* 9, 24.

Luger, K., Mader, A.W., Richmond, R.K., Sargent, D.F., and Richmond, T.J. (1997). Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature* 389, 251-260.

Luo, Y.X., Hou, X.M., Zhang, C.J., Tan, L.M., Shao, C.R., Lin, R.N., Su, Y.N., Cai, X.W., Li, L., Chen, S., *et al.* (2020). A plant-specific SWR1 chromatin-remodeling complex couples histone H2A.Z deposition with nucleosome sliding. *EMBO J* 39, e102008.

Lusser, A., Urwin, D.L., and Kadonaga, J.T. (2005). Distinct activities of CHD1 and ACF in ATP-dependent chromatin assembly. *Nat Struct Mol Biol* 12, 160-166.

Machida, S., Takizawa, Y., Ishimaru, M., Sugita, Y., Sekine, S., Nakayama, J.I., Wolf, M., and Kurumizaka, H. (2018). Structural Basis of Heterochromatin Formation by Human HP1. *Mol Cell* 69, 385-397 e388.

Maeshima, K., Tamura, S., Hansen, J.C., and Itoh, Y. (2020). Fluid-like chromatin: Toward understanding the real chromatin organization present in the cell. *Curr Opin Cell Biol* 64, 77-89.

Magdalou, I., Lopez, B.S., Pasero, P., and Lambert, S.A. (2014). The causes of replication stress and their consequences on genome stability and cell fate. *Semin Cell Dev Biol* 30, 154-164.

Maquat, L.E., and Li, X. (2001). Mammalian heat shock p70 and histone H4 transcripts, which derive from naturally intronless genes, are immune to nonsense-mediated decay. *RNA* 7, 445-456.

March-Diaz, R., Garcia-Dominguez, M., Lozano-Juste, J., Leon, J., Florencio, F.J., and Reyes, J.C. (2008). Histone H2A.Z and homologues of components of the SWR1 complex are required to control immunity in Arabidopsis. *Plant J* 53, 475-487.

Marzluff, W.F., Wagner, E.J., and Duronio, R.J. (2008). Metabolism and regulation of canonical histone mRNAs: life without a poly(A) tail. *Nat Rev Genet* 9, 843-854.

Masel, J., and Siegal, M.L. (2009). Robustness: mechanisms and consequences. *Trends Genet* 25, 395-403.

Mathieu, J., Yant, L.J., Murdter, F., Kuttner, F., and Schmid, M. (2009). Repression of flowering by the miR172 target SMZ. *PLoS Biol* 7, e1000148.

Matsoukas, I.G., Massiah, A.J., and Thomas, B. (2012). Florigenic and antiflorigenic signaling in plants. *Plant Cell Physiol* 53, 1827-1842.

Matsubara, K., Sano, N., Umehara, T., and Horikoshi, M. (2007). Global analysis of functional surfaces of core histones with comprehensive point mutants. *Genes Cells* 12, 13-33.

Mavrich, T.N., Ioshikhes, I.P., Venters, B.J., Jiang, C., Tomsho, L.P., Qi, J., Schuster, S.C., Albert, I., and Pugh, B.F. (2008a). A barrier nucleosome model for statistical positioning of nucleosomes throughout the yeast genome. *Genome Res* 18, 1073-1083.

Mavrich, T.N., Jiang, C., Ioshikhes, I.P., Li, X., Venters, B.J., Zanton, S.J., Tomsho, L.P., Qi, J., Glaser, R.L., Schuster, S.C., *et al.* (2008b). Nucleosome organization in the Drosophila genome. *Nature* 453, 358-362.

McKay, D.J., Klusza, S., Penke, T.J., Meers, M.P., Curry, K.P., McDaniel, S.L., Malek, P.Y., Cooper, S.W., Tatomer, D.C., Lieb, J.D., *et al.* (2015a). Interrogating the function of metazoan histones using engineered gene clusters. *Dev Cell* 32, 373-386.

McKay, Daniel J., Klusza, S., Penke, Taylor J.R., Meers, Michael P., Curry, Kaitlin P., McDaniel, Stephen L., Malek, Pamela Y., Cooper, Stephen W., Tatomer, Deirdre C., Lieb, Jason D., *et al.* (2015b). Interrogating the Function of Metazoan Histones using Engineered Gene Clusters. *Developmental Cell* 32, 373-386.

McKinley, K.L., and Cheeseman, I.M. (2016). The molecular basis for centromere identity and function. *Nat Rev Mol Cell Biol* 17, 16-29.

McKnight, J.N., Jenkins, K.R., Nodelman, I.M., Escobar, T., and Bowman, G.D. (2011). Extranucleosomal DNA binding directs nucleosome sliding by Chd1. *Mol Cell Biol* 31, 4746-4759.

Meeks-Wagner, D., and Hartwell, L.H. (1986). Normal stoichiometry of histone dimer sets is necessary for high fidelity of mitotic chromosome transmission. *Cell* 44, 43-52.

Megee, P.C., Morgan, B.A., Mittman, B.A., and Smith, M.M. (1990). Genetic analysis of histone H4: essential role of lysines subject to reversible acetylation. *Science* 247, 841-845.

Mejlvang, J., Feng, Y., Alabert, C., Neelsen, K.J., Jasencakova, Z., Zhao, X., Lees, M., Sandelin, A., Pasero, P., Lopes, M., *et al.* (2014). New histone supply regulates replication fork speed and PCNA unloading. *J Cell Biol* 204, 29-43.

Melaragno, J.E., Mehrotra, B., and Coleman, A.W. (1993). Relationship between Endopolyploidy and Cell Size in Epidermal Tissue of Arabidopsis. *Plant Cell* 5, 1661-1668.

Meselson, M., and Stahl, F.W. (1958). The Replication of DNA in Escherichia Coli. *Proc Natl Acad Sci U S A* 44, 671-682.

Michaels, S.D., and Amasino, R.M. (1999). FLOWERING LOCUS C encodes a novel MADS domain protein that acts as a repressor of flowering. *Plant Cell* 11, 949-956.

Millan-Zambrano, G., Santos-Rosa, H., Puddu, F., Robson, S.C., Jackson, S.P., and Kouzarides, T. (2018). Phosphorylation of Histone H4T80 Triggers DNA Damage Checkpoint Recovery. *Mol Cell* 72, 625-635 e624.

Millar, C.B., Xu, F., Zhang, K., and Grunstein, M. (2006). Acetylation of H2AZ Lys 14 is associated with genome-wide gene activity in yeast. *Genes Dev* 20, 711-722.

Minami, M., Huh, G.H., Yang, P., and Iwabuchi, M. (1993). Coordinate gene expression of five subclass histones and the putative transcription factors, HBP-1a and HBP-1b, of histone genes in wheat. *Plant Mol Biol* 23, 429-434.

Minami, M., Meshi, T., and Iwabuchi, M. (2000). S phase-specific DNA-binding proteins interacting with the Hex and Oct motifs in type I element of the wheat histone H3 promoter. *Gene* 241, 333-339.

Mitra, D., Parnell, E.J., Landon, J.W., Yu, Y., and Stillman, D.J. (2006). SWI/SNF binding to the HO promoter requires histone acetylation and stimulates TATA-binding protein recruitment. *Mol Cell Biol* 26, 4095-4110.

Mizuguchi, G., Shen, X., Landry, J., Wu, W.H., Sen, S., and Wu, C. (2004). ATP-driven exchange of histone H2AZ variant catalyzed by SWR1 chromatin remodeling complex. *Science* 303, 343-348.

Mockler, T.C., Yu, X., Shalitin, D., Parikh, D., Michael, T.P., Liou, J., Huang, J., Smith, Z., Alonso, J.M., Ecker, J.R., *et al.* (2004). Regulation of flowering time in Arabidopsis by K homology domain proteins. *Proc Natl Acad Sci U S A* 101, 12759-12764.

Moes, D., Gatti, S., Hoffmann, C., Dieterle, M., Moreau, F., Neumann, K., Schumacher, M., Diederich, M., Grill, E., Shen, W.H., *et al.* (2013). A LIM domain protein from tobacco involved in actin-bundling and histone gene transcription. *Mol Plant* 6, 483-502.

Moon, J., Suh, S.S., Lee, H., Choi, K.R., Hong, C.B., Paek, N.C., Kim, S.G., and Lee, I. (2003). The SOC1 MADS-box gene integrates vernalization and gibberellin signals for flowering in Arabidopsis. *Plant J* 35, 613-623.

Moraes, I., Yuan, Z.F., Liu, S., Souza, G.M., Garcia, B.A., and Casas-Mollano, J.A. (2015). Analysis of Histones H3 and H4 Reveals Novel and Conserved Post-Translational Modifications in Sugarcane. *PLoS One* 10, e0134586.

Morgan, H.D., Sutherland, H.G., Martin, D.I., and Whitelaw, E. (1999). Epigenetic inheritance at the agouti locus in the mouse. *Nat Genet* 23, 314-318.

Mueller-Planitz, F., Klinker, H., Ludwigsen, J., and Becker, P.B. (2013). The ATPase domain of ISWI is an autonomous nucleosome remodeling machine. *Nat Struct Mol Biol* 20, 82-89.

Muller, J., and Verrijzer, P. (2009). Biochemical mechanisms of gene regulation by polycomb group protein complexes. *Curr Opin Genet Dev* 19, 150-158.

Murawska, M., and Brehm, A. (2011). CHD chromatin remodelers and the transcription cycle. *Transcription* 2, 244-253.

Murzina, N.V., Pei, X.Y., Zhang, W., Sparkes, M., Vicente-Garcia, J., Pratap, J.V., McLaughlin, S.H., Ben-Shahar, T.R., Verreault, A., Luisi, B.F., *et al.* (2008). Structural basis for the recognition of histone H4 by the histone-chaperone RbAp46. *Structure* *16*, 1077-1085.

Nabatiyan, A., and Krude, T. (2004). Silencing of chromatin assembly factor 1 in human cells leads to cell death and loss of chromatin assembly during DNA synthesis. *Mol Cell Biol* *24*, 2853-2862.

Nacev, B.A., Feng, L., Bagert, J.D., Lemiesz, A.E., Gao, J., Soshnev, A.A., Kundra, R., Schultz, N., Muir, T.W., and Allis, C.D. (2019). The expanding landscape of 'oncohistone' mutations in human cancers. *Nature* *567*, 473-478.

Nagashima, R., Hibino, K., Ashwin, S.S., Babokhov, M., Fujishiro, S., Imai, R., Nozaki, T., Tamura, S., Tani, T., Kimura, H., *et al.* (2019). Single nucleosome imaging reveals loose genome chromatin networks via active RNA polymerase II. *J Cell Biol* *218*, 1511-1530.

Nakabayashi, K., Okamoto, M., Koshiba, T., Kamiya, Y., and Nambara, E. (2005). Genome-wide profiling of stored mRNA in *Arabidopsis thaliana* seed germination: epigenetic and genetic regulation of transcription in seed. *Plant J* *41*, 697-709.

Nakamura, K., Saredi, G., Becker, J.R., Foster, B.M., Nguyen, N.V., Beyer, T.E., Cesa, L.C., Faull, P.A., Lukauskas, S., Frimurer, T., *et al.* (2019). H4K20me0 recognition by BRCA1-BARD1 directs homologous recombination to sister chromatids. *Nat Cell Biol* *21*, 311-318.

Nakanishi, S., Sanderson, B.W., Delventhal, K.M., Bradford, W.D., Staehling-Hampton, K., and Shilatifard, A. (2008). A comprehensive library of histone mutants identifies nucleosomal residues required for H3K4 methylation. *Nat Struct Mol Biol* *15*, 881-888.

Nardi, I.K., Zasadzinska, E., Stellfox, M.E., Knippler, C.M., and Foltz, D.R. (2016). Licensing of Centromeric Chromatin Assembly through the Mis18alpha-Mis18beta Heterotetramer. *Mol Cell* *61*, 774-787.

Naumann, K., Fischer, A., Hofmann, I., Krauss, V., Phalke, S., Irmeler, K., Hause, G., Aurich, A.C., Dorn, R., Jenuwein, T., *et al.* (2005). Pivotal role of AtSUVH2 in heterochromatic histone methylation and gene silencing in *Arabidopsis*. *EMBO J* *24*, 1418-1429.

Nekrasov, V., Staskawicz, B., Weigel, D., Jones, J.D., and Kamoun, S. (2013). Targeted mutagenesis in the model plant *Nicotiana benthamiana* using Cas9 RNA-guided endonuclease. *Nat Biotechnol* *31*, 691-693.

Nelson, D.M., Ye, X., Hall, C., Santos, H., Ma, T., Kao, G.D., Yen, T.J., Harper, J.W., and Adams, P.D. (2002). Coupling of DNA synthesis and histone synthesis in S phase independent of cyclin/cdk2 activity. *Mol Cell Biol* *22*, 7459-7472.

Nestor, C.E., Ottaviano, R., Reinhardt, D., Cruickshanks, H.A., Mjoseng, H.K., McPherson, R.C., Lentini, A., Thomson, J.P., Dunican, D.S., Pennings, S., *et al.* (2015). Rapid reprogramming of epigenetic and transcriptional profiles in mammalian culture systems. *Genome Biol* *16*, 11.

Neumann, H., Hancock, S.M., Buning, R., Routh, A., Chapman, L., Somers, J., Owen-Hughes, T., van Noort, J., Rhodes, D., and Chin, J.W. (2009). A method for genetically installing site-specific acetylation in recombinant histones defines the effects of H3 K56 acetylation. *Mol Cell* *36*, 153-163.

Ng, T.M., Lenstra, T.L., Duggan, N., Jiang, S., Ceto, S., Holstege, F.C., Dai, J., Boeke, J.D., and Biggins, S. (2013). Kinetochores function and chromosome segregation rely on critical residues in histones H3 and H4 in budding yeast. *Genetics* *195*, 795-807.

Nguyen, N.H., Jeong, C.Y., Kang, G.H., Yoo, S.D., Hong, S.W., and Lee, H. (2015). MYBD employed by HY5 increases anthocyanin accumulation via repression of MYBL2 in *Arabidopsis*. *Plant J* *84*, 1192-1205.

Nichol, J.N., Dupere-Richer, D., Ezponda, T., Licht, J.D., and Miller, W.H., Jr. (2016). H3K27 Methylation: A Focal Point of Epigenetic Deregulation in Cancer. *Adv Cancer Res* 131, 59-95.

Ning, Y.Q., Chen, Q., Lin, R.N., Li, Y.Q., Li, L., Chen, S., and He, X.J. (2019). The HDA19 histone deacetylase complex is involved in the regulation of flowering time in a photoperiod-dependent manner. *Plant J* 98, 448-464.

Nisa, M.U., Huang, Y., Benhamed, M., and Raynaud, C. (2019). The Plant DNA Damage Response: Signaling Pathways Leading to Growth Inhibition and Putative Role in Response to Stress Conditions. *Front Plant Sci* 10, 653.

Nishioka, K., Rice, J.C., Sarma, K., Erdjument-Bromage, H., Werner, J., Wang, Y., Chuikov, S., Valenzuela, P., Tempst, P., Steward, R., *et al.* (2002). PR-Set7 is a nucleosome-specific methyltransferase that modifies lysine 20 of histone H4 and is associated with silent chromatin. *Mol Cell* 9, 1201-1213.

Niu, L., Lu, F., Pei, Y., Liu, C., and Cao, X. (2007). Regulation of flowering time by the protein arginine methyltransferase AtPRMT10. *EMBO Rep* 8, 1190-1195.

Noir, S., Marrocco, K., Masoud, K., Thomann, A., Gusti, A., Bitrian, M., Schnittger, A., and Genschik, P. (2015). The Control of Arabidopsis thaliana Growth by Cell Proliferation and Endoreplication Requires the F-Box Protein FBL17. *Plant Cell* 27, 1461-1476.

Norris, A., Bianchet, M.A., and Boeke, J.D. (2008). Compensatory interactions between Sir3p and the nucleosomal LRS surface imply their direct interaction. *PLoS Genet* 4, e1000301.

North, J.A., Shimko, J.C., Javaid, S., Mooney, A.M., Shoffner, M.A., Rose, S.D., Bundschuh, R., Fishel, R., Ottesen, J.J., and Poirier, M.G. (2012). Regulation of the nucleosome unwrapping rate controls DNA accessibility. *Nucleic Acids Res* 40, 10215-10227.

Nozaki, T., Imai, R., Tanbo, M., Nagashima, R., Tamura, S., Tani, T., Joti, Y., Tomita, M., Hibino, K., Kanemaki, M.T., *et al.* (2017). Dynamic Organization of Chromatin Domains Revealed by Super-Resolution Live-Cell Imaging. *Mol Cell* 67, 282-293 e287.

Oda, H., Okamoto, I., Murphy, N., Chu, J., Price, S.M., Shen, M.M., Torres-Padilla, M.E., Heard, E., and Reinberg, D. (2009). Monomethylation of histone H4-lysine 20 is involved in chromosome structure and stability and is essential for mouse development. *Mol Cell Biol* 29, 2278-2295.

Okada, T., Endo, M., Singh, M.B., and Bhalla, P.L. (2005). Analysis of the histone H3 gene family in Arabidopsis and identification of the male-gamete-specific variant AtMGH3. *Plant J* 44, 557-568.

Oliver, C., Santos, J.L., and Pradillo, M. (2014). On the role of some ARGONAUTE proteins in meiosis and DNA repair in Arabidopsis thaliana. *Front Plant Sci* 5, 177.

Oltmanns, H., Frame, B., Lee, L.Y., Johnson, S., Li, B., Wang, K., and Gelvin, S.B. (2010). Generation of backbone-free, low transgene copy plants by launching T-DNA from the Agrobacterium chromosome. *Plant Physiol* 152, 1158-1166.

Oppikofer, M., Kueng, S., Martino, F., Soeroes, S., Hancock, S.M., Chin, J.W., Fischle, W., and Gasser, S.M. (2011). A dual role of H4K16 acetylation in the establishment of yeast silent chromatin. *EMBO J* 30, 2610-2621.

Osley, M.A. (1991). The regulation of histone synthesis in the cell cycle. *Annu Rev Biochem* 60, 827-861.

Osley, M.A., and Hereford, L.M. (1981). Yeast histone genes show dosage compensation. *Cell* 24, 377-384.

Ou, H.D., Phan, S., Deerinck, T.J., Thor, A., Ellisman, M.H., and O'Shea, C.C. (2017). ChromEMT: Visualizing 3D chromatin structure and compaction in interphase and mitotic cells. *Science* 357.

Pajoro, A., Muino, J.M., Angenent, G.C., and Kaufmann, K. (2018). Profiling Nucleosome Occupancy by MNase-seq: Experimental Protocol and Computational Analysis. *Methods Mol Biol* 1675, 167-181.

Pajoro, A., Severing, E., Angenent, G.C., and Immink, R.G.H. (2017). Histone H3 lysine 36 methylation affects temperature-induced alternative splicing and flowering in plants. *Genome Biol* 18, 102.

Panier, S., and Boulton, S.J. (2014). Double-strand break repair: 53BP1 comes into focus. *Nat Rev Mol Cell Biol* 15, 7-18.

Park, J.H., Cosgrove, M.S., Youngman, E., Wolberger, C., and Boeke, J.D. (2002). A core nucleosome surface crucial for transcriptional silencing. *Nat Genet* 32, 273-279.

Parker, R., and Song, H. (2004). The enzymes and control of eukaryotic mRNA turnover. *Nat Struct Mol Biol* 11, 121-127.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., and Grisel, O. (2011). Scikit-learn: machine learning in python. *JMLR* 12, 2825-2830.

Pei, Y., Niu, L., Lu, F., Liu, C., Zhai, J., Kong, X., and Cao, X. (2007). Mutations in the Type II protein arginine methyltransferase AtPRMT5 result in pleiotropic developmental defects in Arabidopsis. *Plant Physiol* 144, 1913-1923.

Pengelly, A.R., Copur, O., Jackle, H., Herzig, A., and Muller, J. (2013). A histone mutant reproduces the phenotype caused by loss of histone-modifying factor Polycomb. *Science* 339, 698-699.

Pesavento, J.J., Yang, H., Kelleher, N.L., and Mizzen, C.A. (2008). Certain and progressive methylation of histone H4 at lysine 20 during the cell cycle. *Mol Cell Biol* 28, 468-486.

Peterson, B.A., Haak, D.C., Nishimura, M.T., Teixeira, P.J., James, S.R., Dangl, J.L., and Nimchuk, Z.L. (2016). Genome-Wide Assessment of Efficiency and Specificity in CRISPR/Cas9 Mediated Multiple Site Targeting in Arabidopsis. *PLoS One* 11, e0162169.

Petryk, N., Dalby, M., Wenger, A., Stromme, C.B., Strandsby, A., Andersson, R., and Groth, A. (2018). MCM2 promotes symmetric inheritance of modified histones during DNA replication. *Science* 361, 1389-1392.

Plazas-Mayorca, M.D., Zee, B.M., Young, N.L., Fingerman, I.M., LeRoy, G., Briggs, S.D., and Garcia, B.A. (2009). One-pot shotgun quantitative mass spectrometry characterization of histones. *J Proteome Res* 8, 5367-5374.

Pointner, J., Persson, J., Prasad, P., Norman-Axelsson, U., Stralfors, A., Khorosjutina, O., Krietenstein, N., Svensson, J.P., Ekwall, K., and Korber, P. (2012). CHD1 remodelers regulate nucleosome spacing in vitro and align nucleosomal arrays over gene coding regions in *S. pombe*. *EMBO J* 31, 4388-4403.

Pokholok, D.K., Harbison, C.T., Levine, S., Cole, M., Hannett, N.M., Lee, T.I., Bell, G.W., Walker, K., Rolfe, P.A., Herbolsheimer, E., *et al.* (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. *Cell* 122, 517-527.

Pontvianne, F., Blevins, T., and Pikaard, C.S. (2010). Arabidopsis Histone Lysine Methyltransferases. *Adv Bot Res* 53, 1-22.

Prado, F., and Aguilera, A. (2005). Partial depletion of histone H4 increases homologous recombination-mediated genetic instability. *Mol Cell Biol* 25, 1526-1536.

Probst, A.V., Dunleavy, E., and Almouzni, G. (2009). Epigenetic inheritance during the cell cycle. *Nat Rev Mol Cell Biol* 10, 192-206.

Puchta, H. (2005). The repair of double-strand breaks in plants: mechanisms and consequences for genome evolution. *J Exp Bot* 56, 1-14.

Putterill, J., Robson, F., Lee, K., Simon, R., and Coupland, G. (1995). The CONSTANS gene of Arabidopsis promotes flowering and encodes a protein showing similarities to zinc finger transcription factors. *Cell* 80, 847-857.

Quivy, J.P., Roche, D., Kirschner, D., Tagami, H., Nakatani, Y., and Almouzni, G. (2004). A CAF-1 dependent pool of HP1 during heterochromatin duplication. *EMBO J* 23, 3516-3526.

Racki, L.R., Naber, N., Pate, E., Leonard, J.D., Cooke, R., and Narlikar, G.J. (2014). The histone H4 tail regulates the conformation of the ATP-binding pocket in the SNF2h chromatin remodeling enzyme. *J Mol Biol* 426, 2034-2044.

Radwan, A., Younis, A., Luykx, P., and Khuri, S. (2008). Prediction and analysis of nucleosome exclusion regions in the human genome. *BMC Genomics* 9, 186.

Rajavel, K.S., and Neufeld, E.F. (2001). Nonsense-mediated decay of human HEXA mRNA. *Mol Cell Biol* 21, 5512-5519.

Rakyan, V.K., Chong, S., Champ, M.E., Cuthbert, P.C., Morgan, H.D., Luu, K.V., and Whitelaw, E. (2003). Transgenerational inheritance of epigenetic states at the murine Axin(Fu) allele occurs after maternal and paternal transmission. *Proc Natl Acad Sci U S A* 100, 2538-2543.

Ramirez, F., Ryan, D.P., Gruning, B., Bhardwaj, V., Kilpert, F., Richter, A.S., Heyne, S., Dundar, F., and Manke, T. (2016). deepTools2: a next generation web server for deep-sequencing data analysis. *Nucleic Acids Res* 44, W160-165.

Ramirez-Parra, E., and Gutierrez, C. (2007). E2F regulates FASCIATA1, a chromatin assembly gene whose loss switches on the endocycle and activates gene expression by changing the epigenetic status. *Plant Physiol* 144, 105-120.

Ray-Gallet, D., Quivy, J.P., Scamps, C., Martini, E.M., Lipinski, M., and Almouzni, G. (2002). HIRA is critical for a nucleosome assembly pathway independent of DNA synthesis. *Mol Cell* 9, 1091-1100.

Raynaud, C., Sozzani, R., Glab, N., Domenichini, S., Perennes, C., Cella, R., Kondorosi, E., and Bergounioux, C. (2006). Two cell-cycle regulated SET-domain proteins interact with proliferating cell nuclear antigen (PCNA) in Arabidopsis. *Plant J* 47, 395-407.

Redei, G.P. (1962). Supervital Mutants of Arabidopsis. *Genetics* 47, 443-460.

Reichheld, J.P., Gigot, C., and Chaubet-Gigot, N. (1998). Multilevel regulation of histone gene expression during the cell cycle in tobacco cells. *Nucleic Acids Res* 26, 3255-3262.

Ricci, M.A., Manzo, C., Garcia-Parajo, M.F., Lakadamyali, M., and Cosma, M.P. (2015). Chromatin fibers are formed by heterogeneous groups of nucleosomes in vivo. *Cell* 160, 1145-1158.

Rieder, L.E., Koreski, K.P., Boltz, K.A., Kuzu, G., Urban, J.A., Bowman, S.K., Zeidman, A., Jordan, W.T., 3rd, Tolstorukov, M.Y., Marzluff, W.F., *et al.* (2017). Histone locus regulation by the Drosophila dosage compensation adaptor protein CLAMP. *Genes Dev* 31, 1494-1508.

Rizzardi, L.F., Dorn, E.S., Strahl, B.D., and Cook, J.G. (2012). DNA replication origin function is promoted by H3K4 di-methylation in *Saccharomyces cerevisiae*. *Genetics* 192, 371-384.

Robinson, D.O., Coate, J.E., Singh, A., Hong, L., Bush, M., Doyle, J.J., and Roeder, A.H.K. (2018). Ploidy and Size at Multiple Scales in the Arabidopsis Sepal. *Plant Cell* 30, 2308-2329.

Robinson, J.T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E.S., Getz, G., and Mesirov, J.P. (2011). Integrative genomics viewer. *Nat Biotechnol* 29, 24-26.

Robinson, M.D., McCarthy, D.J., and Smyth, G.K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139-140.

Robson, P., Whitelam, G.C., and Smith, H. (1993). Selected Components of the Shade-Avoidance Syndrome Are Displayed in a Normal Manner in Mutants of Arabidopsis thaliana and Brassica rapa Deficient in Phytochrome B. *Plant Physiol* 102, 1179-1184.

Rothbart, S.B., Krajewski, K., Nady, N., Tempel, W., Xue, S., Badeaux, A.I., Barsyte-Lovejoy, D., Martinez, J.Y., Bedford, M.T., Fuchs, S.M., *et al.* (2012). Association of UHRF1 with methylated H3K9 directs the maintenance of DNA methylation. *Nat Struct Mol Biol* 19, 1155-1160.

Ruthenburg, A.J., Li, H., Milne, T.A., Dewell, S., McGinty, R.K., Yuen, M., Ueberheide, B., Dou, Y., Muir, T.W., Patel, D.J., *et al.* (2011). Recognition of a mononucleosomal histone modification pattern by BPTF via multivalent interactions. *Cell* 145, 692-706.

Rutowicz, K., Lirski, M., Mermaz, B., Teano, G., Schubert, J., Mestiri, I., Kroten, M.A., Fabrice, T.N., Fritz, S., Grob, S., *et al.* (2019). Linker histones are fine-scale chromatin architects modulating developmental decisions in Arabidopsis. *Genome Biol* 20, 157.

Sala, A., Toto, M., Pinello, L., Gabriele, A., Di Benedetto, V., Ingrassia, A.M., Lo Bosco, G., Di Gesu, V., Giancarlo, R., and Corona, D.F. (2011). Genome-wide characterization of chromatin binding and nucleosome spacing activity of the nucleosome remodelling ATPase ISWI. *EMBO J* 30, 1766-1777.

Saleh, A., Alvarez-Venegas, R., Yilmaz, M., Le, O., Hou, G., Sadler, M., Al-Abdallat, A., Xia, Y., Lu, G., Ladunga, I., *et al.* (2008). The highly similar Arabidopsis homologs of trithorax ATX1 and ATX2 encode proteins with divergent biochemical functions. *Plant Cell* 20, 568-579.

Salzler, H.R., Tatomer, D.C., Malek, P.Y., McDaniel, S.L., Orlando, A.N., Marzluff, W.F., and Duronio, R.J. (2013). A sequence in the Drosophila H3-H4 Promoter triggers histone locus body assembly and biosynthesis of replication-coupled histone mRNAs. *Dev Cell* 24, 623-634.

Samach, A., and Wigge, P.A. (2005). Ambient temperature perception in plants. *Curr Opin Plant Biol* 8, 483-486.

Sanders, S.L., Portoso, M., Mata, J., Bahler, J., Allshire, R.C., and Kouzarides, T. (2004). Methylation of histone H4 lysine 20 controls recruitment of Crb2 to sites of DNA damage. *Cell* 119, 603-614.

Saredi, G., Huang, H., Hammond, C.M., Alabert, C., Bekker-Jensen, S., Forne, I., Reveron-Gomez, N., Foster, B.M., Mlejnkova, L., Bartke, T., *et al.* (2016). H4K20me0 marks post-replicative chromatin and recruits the TONSL-MMS22L DNA repair complex. *Nature* 534, 714-718.

Sarkies, P., Reams, C., Simpson, L.J., and Sale, J.E. (2010). Epigenetic instability due to defective replication of structured DNA. *Mol Cell* 40, 703-713.

Sawa, M., Nusinow, D.A., Kay, S.A., and Imaizumi, T. (2007). FKF1 and GIGANTEA complex formation is required for day-length measurement in Arabidopsis. *Science* 318, 261-265.

Schmid, M., Davison, T.S., Henz, S.R., Pape, U.J., Demar, M., Vingron, M., Scholkopf, B., Weigel, D., and Lohmann, J.U. (2005). A gene expression map of Arabidopsis thaliana development. *Nat Genet* 37, 501-506.

Schmitz, R.J., Sung, S., and Amasino, R.M. (2008). Histone arginine methylation is required for vernalization-induced epigenetic silencing of FLC in winter-annual Arabidopsis thaliana. *Proc Natl Acad Sci U S A* 105, 411-416.

Schmucker, A., Lei, B., Lorkovic, Z.J., Capella, M., Braun, S., Bourguet, P., Mathieu, O., Mechtler, K., and Berger, F. (2021). Crosstalk between H2A variant-specific modifications impacts vital cell functions. *PLoS Genet* 17, e1009601.

Schones, D.E., Cui, K., Cuddapah, S., Roh, T.Y., Barski, A., Wang, Z., Wei, G., and Zhao, K. (2008). Dynamic regulation of nucleosome positioning in the human genome. *Cell* 132, 887-898.

Schotta, G., Lachner, M., Sarma, K., Ebert, A., Sengupta, R., Reuter, G., Reinberg, D., and Jenuwein, T. (2004). A silencing pathway to induce H3-K9 and H4-K20 trimethylation at constitutive heterochromatin. *Genes Dev* 18, 1251-1262.

Schubert, D., Primavesi, L., Bishopp, A., Roberts, G., Doonan, J., Jenuwein, T., and Goodrich, J. (2006). Silencing by plant Polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. *EMBO J* 25, 4638-4649.

Sequeira-Mendes, J., Araguez, I., Peiro, R., Mendez-Giraldez, R., Zhang, X., Jacobsen, S.E., Bastolla, U., and Gutierrez, C. (2014). The Functional Topography of the Arabidopsis Genome Is Organized in a Reduced Number of Linear Motifs of Chromatin States. *Plant Cell* 26, 2351-2366.

Seroby, V., Kontarakis, Z., El-Brolosy, M.A., Welker, J.M., Tolstenkov, O., Saadeldein, A.M., Retzer, N., Gottschalk, A., Wehman, A.M., and Stainier, D.Y. (2020). Transcriptional adaptation in *Caenorhabditis elegans*. *Elife* 9.

Sharif, J., Muto, M., Takebayashi, S., Suetake, I., Iwamatsu, A., Endo, T.A., Shinga, J., Mizutani-Koseki, Y., Toyoda, T., Okamura, K., *et al.* (2007). The SRA protein Np95 mediates epigenetic inheritance by recruiting Dnmt1 to methylated DNA. *Nature* 450, 908-912.

Shaul, O. (2015). Unique Aspects of Plant Nonsense-Mediated mRNA Decay. *Trends Plant Sci* 20, 767-779.

Shilatifard, A. (2012). The COMPASS family of histone H3K4 methylases: mechanisms of regulation in development and disease pathogenesis. *Annu Rev Biochem* 81, 65-95.

Shim, J.S., Kubota, A., and Imaizumi, T. (2017). Circadian Clock and Photoperiodic Flowering in Arabidopsis: CONSTANS Is a Hub for Signal Integration. *Plant Physiol* 173, 5-15.

Shim, Y.S., Choi, Y., Kang, K., Cho, K., Oh, S., Lee, J., Grewal, S.I., and Lee, D. (2012). Hrp3 controls nucleosome positioning to suppress non-coding transcription in eu- and heterochromatin. *EMBO J* 31, 4375-4387.

Shimko, J.C., North, J.A., Bruns, A.N., Poirier, M.G., and Ottesen, J.J. (2011). Preparation of fully synthetic histone H3 reveals that acetyl-lysine 56 facilitates protein binding within nucleosomes. *J Mol Biol* 408, 187-204.

Shoab, M., Walter, D., Gillespie, P.J., Izard, F., Fahrenkrog, B., Lleres, D., Lerdrup, M., Johansen, J.V., Hansen, K., Julien, E., *et al.* (2018). Histone H4K20 methylation mediated chromatin compaction threshold ensures genome integrity by limiting DNA replication licensing. *Nat Commun* 9, 3704.

Shogren-Knaak, M., Ishii, H., Sun, J.M., Pazin, M.J., Davie, J.R., and Peterson, C.L. (2006). Histone H4-K16 acetylation controls chromatin structure and protein interactions. *Science* 311, 844-847.

Shyu, A.B., Wilkinson, M.F., and van Hoof, A. (2008). Messenger RNA regulation: to translate or to degrade. *EMBO J* 27, 471-481.

Simon, L., Voisin, M., Tatout, C., and Probst, A.V. (2015). Structure and Function of Centromeric and Pericentromeric Heterochromatin in Arabidopsis thaliana. *Front Plant Sci* 6, 1049.

Simpson, G.G., and Dean, C. (2002). Arabidopsis, the Rosetta stone of flowering time? *Science* 296, 285-289.

Singer, T., Yordan, C., and Martienssen, R.A. (2001). Robertson's Mutator transposons in *A. thaliana* are regulated by the chromatin-remodeling gene Decrease in DNA Methylation (DDM1). *Genes Dev* 15, 591-602.

Singh, A.K., Schauer, T., Pfaller, L., Straub, T., and Mueller-Planitz, F. (2021). The biogenesis and function of nucleosome arrays. *Nat Commun* 12, 7011.

Singh, R.K., Kabbaj, M.H., Paik, J., and Gunjan, A. (2009). Histone levels are regulated by phosphorylation and ubiquitylation-dependent proteolysis. *Nat Cell Biol* 11, 925-933.

Singh, R.K., Liang, D., Gajjalaiahvari, U.R., Kabbaj, M.H., Paik, J., and Gunjan, A. (2010). Excess histone levels mediate cytotoxicity via multiple mechanisms. *Cell Cycle* 9, 4236-4244.

Siriaco, G., Deuring, R., Mawla, G.D., and Tamkun, J.W. (2015). A novel approach for studying histone H1 function in vivo. *Genetics* 200, 29-33.

Sittman, D.B., Graves, R.A., and Marzluff, W.F. (1983). Histone mRNA concentrations are regulated at the level of transcription and mRNA degradation. *Proc Natl Acad Sci U S A* 80, 1849-1853.

Siwaszek, A., Ukleja, M., and Dziembowski, A. (2014). Proteins involved in the degradation of cytoplasmic mRNA in the major eukaryotic model systems. *RNA Biol* 11, 1122-1136.

Smith, S., and Stillman, B. (1989). Purification and characterization of CAF-I, a human cell factor required for chromatin assembly during DNA replication in vitro. *Cell* 58, 15-25.

Soares, L.M., He, P.C., Chun, Y., Suh, H., Kim, T., and Buratowski, S. (2017). Determinants of Histone H3K4 Methylation Patterns. *Mol Cell* 68, 773-785 e776.

Song, J.J., Garlick, J.D., and Kingston, R.E. (2008). Structural basis of histone H4 recognition by p55. *Genes Dev* 22, 1313-1318.

Song, Y.H., Shim, J.S., Kinmonth-Schultz, H.A., and Imaizumi, T. (2015). Photoperiodic flowering: time measurement mechanisms in leaves. *Annu Rev Plant Biol* 66, 441-464.

Soppe, W.J., Jasencakova, Z., Houben, A., Kakutani, T., Meister, A., Huang, M.S., Jacobsen, S.E., Schubert, I., and Fransz, P.F. (2002). DNA methylation controls histone H3 lysine 9 methylation and heterochromatin assembly in Arabidopsis. *EMBO J* 21, 6549-6559.

Srikanth, A., and Schmid, M. (2011). Regulation of flowering time: all roads lead to Rome. *Cell Mol Life Sci* 68, 2013-2037.

Steimer, A., Amedeo, P., Afsar, K., Fransz, P., Mittelsten Scheid, O., and Paszkowski, J. (2000). Endogenous targets of transcriptional gene silencing in Arabidopsis. *Plant Cell* 12, 1165-1178.

Sterken, R., Kiekens, R., Boruc, J., Zhang, F., Vercauteren, A., Vercauteren, I., De Smet, L., Dhondt, S., Inze, D., De Veylder, L., *et al.* (2012). Combined linkage and association mapping reveals CYCD5;1 as a quantitative trait gene for endoreduplication in Arabidopsis. *Proc Natl Acad Sci U S A* 109, 4678-4683.

Sternberg, P.W., Stern, M.J., Clark, I., and Herskowitz, I. (1987). Activation of the yeast HO gene by release from multiple negative controls. *Cell* 48, 567-577.

Stewart-Morgan, K.R., Petryk, N., and Groth, A. (2020). Chromatin replication and epigenetic cell memory. *Nat Cell Biol* 22, 361-371.

Stroud, H., Hale, C.J., Feng, S., Caro, E., Jacob, Y., Michaels, S.D., and Jacobsen, S.E. (2012). DNA methyltransferases are required to induce heterochromatic re-replication in Arabidopsis. *PLoS Genet* 8, e1002808.

Struhl, K., and Segal, E. (2013). Determinants of nucleosome positioning. *Nat Struct Mol Biol* 20, 267-273.

Suganuma, T., and Workman, J.L. (2011). Signals and combinatorial functions of histone modifications. *Annu Rev Biochem* 80, 473-499.

Suka, N., Luo, K., and Grunstein, M. (2002). Sir2p and Sas2p opposingly regulate acetylation of yeast histone H4 lysine16 and spreading of heterochromatin. *Nat Genet* 32, 378-383.

Sullivan, E., Santiago, C., Parker, E.D., Dominski, Z., Yang, X., Lanzotti, D.J., Ingledue, T.C., Marzluff, W.F., and Duronio, R.J. (2001). Drosophila stem loop binding protein coordinates accumulation of mature histone mRNA with cell cycle progression. *Genes Dev* 15, 173-187.

Sung, S., and Amasino, R.M. (2004). Vernalization in Arabidopsis thaliana is mediated by the PHD finger protein VIN3. *Nature* 427, 159-164.

Sutherland, G.R. (1977). Fragile sites on human chromosomes: demonstration of their dependence on the type of tissue culture medium. *Science* 197, 265-266.

Suto, R.K., Clarkson, M.J., Tremethick, D.J., and Luger, K. (2000). Crystal structure of a nucleosome core particle containing the variant histone H2A.Z. *Nat Struct Biol* 7, 1121-1124.

Szabados, L., Kovacs, I., Oberschall, A., Abraham, E., Kerekes, I., Zsigmond, L., Nagy, R., Alvarado, M., Krasovskaja, I., Gal, M., *et al.* (2002). Distribution of 1000 sequenced T-DNA tags in the Arabidopsis genome. *Plant J* 32, 233-242.

Szadeczky-Kardoss, I., Gal, L., Auber, A., Taller, J., and Silhavy, D. (2018). The No-go decay system degrades plant mRNAs that contain a long A-stretch in the coding region. *Plant Sci* 275, 19-27.

Sztaf, T.E., and Stainier, D.Y.R. (2020). Transcriptional adaptation: a mechanism underlying genetic robustness. *Development* 147.

Tabata, T., Nakayama, T., Mikami, K., and Iwabuchi, M. (1991). HBP-1a and HBP-1b: leucine zipper-type transcription factors of wheat. *EMBO J* 10, 1459-1467.

Tagami, H., Ray-Gallet, D., Almouzni, G., and Nakatani, Y. (2004). Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. *Cell* 116, 51-61.

Takeo, K. (2016). Stress-induced flowering: the third category of flowering response. *J Exp Bot* 67, 4925-4934.

Talbert, P.B., and Henikoff, S. (2017). Histone variants on the move: substrates for chromatin dynamics. *Nat Rev Mol Cell Biol* 18, 115-126.

Tan, L.M., Liu, R., Gu, B.W., Zhang, C.J., Luo, J., Guo, J., Wang, Y., Chen, L., Du, X., Li, S., *et al.* (2020). Dual Recognition of H3K4me3 and DNA by the ISWI Component ARID5 Regulates the Floral Transition in Arabidopsis. *Plant Cell* 32, 2178-2195.

Tang, X., Liu, G., Zhou, J., Ren, Q., You, Q., Tian, L., Xin, X., Zhong, Z., Liu, B., Zheng, X., *et al.* (2018). A large-scale whole-genome sequencing analysis reveals highly specific genome editing by both Cas9 and Cpf1 (Cas12a) nucleases in rice. *Genome Biol* 19, 84.

Taoka, K., Kaya, H., Nakayama, T., Araki, T., Meshi, T., and Iwabuchi, M. (1999). Identification of three kinds of mutually related composite elements conferring S phase-specific transcriptional activation. *Plant J* 18, 611-623.

Taoka, K., Ohtsubo, N., Fujimoto, Y., Mikami, K., Meshi, T., and Iwabuchi, M. (1998). The modular structure and function of the wheat H1 promoter with S phase-specific activity. *Plant Cell Physiol* 39, 294-306.

Tardat, M., Brustel, J., Kirsh, O., Lefebvre, C., Callanan, M., Sardet, C., and Julien, E. (2010). The histone H4 Lys 20 methyltransferase PR-Set7 regulates replication origins in mammalian cells. *Nat Cell Biol* 12, 1086-1093.

Taunton, J., Hassig, C.A., and Schreiber, S.L. (1996). A mammalian histone deacetylase related to the yeast transcriptional regulator Rpd3p. *Science* 272, 408-411.

Team, R.C. (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Tenea, G.N., Spantzel, J., Lee, L.Y., Zhu, Y., Lin, K., Johnson, S.J., and Gelvin, S.B. (2009). Overexpression of several Arabidopsis histone genes increases agrobacterium-mediated transformation and transgene expression in plants. *Plant Cell* 21, 3350-3367.

Tessadori, F., Giltay, J.C., Hurst, J.A., Massink, M.P., Duran, K., Vos, H.R., van Es, R.M., Deciphering Developmental Disorders, S., Scott, R.H., van Gassen, K.L.I., *et al.* (2017). Germline mutations affecting the histone H4 core cause a developmental syndrome by altering DNA damage response and cell cycle control. *Nat Genet* 49, 1642-1646.

Thastrom, A., Lowary, P.T., Widlund, H.R., Cao, H., Kubista, M., and Widom, J. (1999). Sequence motifs and free energies of selected natural and non-natural nucleosome positioning DNA sequences. *J Mol Biol* 288, 213-229.

toolkit., B.I.P. (2019). doi:<http://broadinstitute.github.io/picard/>.

Toto, M., D'Angelo, G., and Corona, D.F. (2014). Regulation of ISWI chromatin remodelling activity. *Chromosoma* 123, 91-102.

Trojer, P., Li, G., Sims, R.J., 3rd, Vaquero, A., Kalakonda, N., Boccuni, P., Lee, D., Erdjument-Bromage, H., Tempst, P., Nimer, S.D., *et al.* (2007). L3MBTL1, a histone-methylation-dependent chromatin lock. *Cell* 129, 915-928.

Tronick, E., and Hunter, R.G. (2016). Waddington, Dynamic Systems, and Epigenetics. *Front Behav Neurosci* 10, 107.

Turck, F., Roudier, F., Farrona, S., Martin-Magniette, M.L., Guillaume, E., Buisine, N., Gagnot, S., Martienssen, R.A., Coupland, G., and Colot, V. (2007). Arabidopsis TFL2/LHP1 specifically associates with genes marked by trimethylation of histone H3 lysine 27. *PLoS Genet* 3, e86.

Tweedie-Cullen, R.Y., Brunner, A.M., Grossmann, J., Mohanna, S., Sichau, D., Nanni, P., Panse, C., and Mansuy, I.M. (2012). Identification of combinatorial patterns of post-translational modifications on individual histones in the mouse brain. *PLoS One* 7, e36980.

Tyagi, M., Imam, N., Verma, K., and Patel, A.K. (2016). Chromatin remodelers: We are the drivers!! *Nucleus* 7, 388-404.

Tyler, J.K., Adams, C.R., Chen, S.R., Kobayashi, R., Kamakaka, R.T., and Kadonaga, J.T. (1999). The RCAF complex mediates chromatin assembly during DNA replication and repair. *Nature* 402, 555-560.

van Nuland, R., and Gozani, O. (2016). Histone H4 Lysine 20 (H4K20) Methylation, Expanding the Signaling Potential of the Proteome One Methyl Moiety at a Time. *Mol Cell Proteomics* 15, 755-764.

Vandepoele, K., Raes, J., De Veylder, L., Rouze, P., Rombauts, S., and Inze, D. (2002). Genome-wide analysis of core cell cycle genes in Arabidopsis. *Plant Cell* 14, 903-916.

Vavra, K.J., Allis, C.D., and Gorovsky, M.A. (1982). Regulation of histone acetylation in Tetrahymena macro- and micronuclei. *J Biol Chem* 257, 2591-2598.

Vermeulen, M., Mulder, K.W., Denissov, S., Pijnappel, W.W., van Schaik, F.M., Varier, R.A., Baltissen, M.P., Stunnenberg, H.G., Mann, M., and Timmers, H.T. (2007). Selective anchoring of TFIIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* 131, 58-69.

Vlachonasios, K.E., Thomashow, M.F., and Triezenberg, S.J. (2003). Disruption mutations of ADA2b and GCN5 transcriptional adaptor genes dramatically affect Arabidopsis growth, development, and gene expression. *Plant Cell* 15, 626-638.

Vlieghe, K., Boudolf, V., Beemster, G.T., Maes, S., Magyar, Z., Atanassova, A., de Almeida Engler, J., De Groot, R., Inze, D., and De Veylder, L. (2005). The DP-E2F-like gene DEL1 controls the endocycle in Arabidopsis thaliana. *Curr Biol* 15, 59-63.

Voigt, P., Tee, W.W., and Reinberg, D. (2013). A double take on bivalent promoters. *Genes Dev* 27, 1318-1338.

Voong, L.N., Xi, L., Sebeson, A.C., Xiong, B., Wang, J.P., and Wang, X. (2016). Insights into Nucleosome Organization in Mouse Embryonic Stem Cells through Chemical Mapping. *Cell* 167, 1555-1570 e1515.

Waddington, C.H. (1942). The epigenotype. *Endeavor* 1, 18-20.

Waese, J., Fan, J., Pasha, A., Yu, H., Fucile, G., Shi, R., Cumming, M., Kelley, L.A., Sternberg, M.J., Krishnakumar, V., *et al.* (2017). ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology. *Plant Cell* 29, 1806-1821.

Wang, J., Jia, S.T., and Jia, S. (2016). New Insights into the Regulation of Heterochromatin. *Trends Genet* 32, 284-294.

Wang, J., Liu, X., Dou, Z., Chen, L., Jiang, H., Fu, C., Fu, G., Liu, D., Zhang, J., Zhu, T., *et al.* (2014). Mitotic regulator Mis18beta interacts with and specifies the centromeric assembly of molecular chaperone holliday junction recognition protein (HJURP). *J Biol Chem* 289, 8326-8336.

Wang, J.W. (2014). Regulation of flowering time by the miR156-mediated age pathway. *J Exp Bot* 65, 4723-4730.

Wang, J.W., Czech, B., and Weigel, D. (2009a). miR156-regulated SPL transcription factors define an endogenous flowering pathway in *Arabidopsis thaliana*. *Cell* 138, 738-749.

Wang, X., Wu, R., Lin, X., Bai, Y., Song, C., Yu, X., Xu, C., Zhao, N., Dong, Y., and Liu, B. (2013). Tissue culture-induced genetic and epigenetic alterations in rice pure-lines, F1 hybrids and polyploids. *BMC Plant Biol* 13, 77.

Wang, X., Zhang, Y., Ma, Q., Zhang, Z., Xue, Y., Bao, S., and Chong, K. (2007). SKB1-mediated symmetric dimethylation of histone H4R3 controls flowering time in *Arabidopsis*. *EMBO J* 26, 1934-1941.

Wang, Y., Reddy, B., Thompson, J., Wang, H., Noma, K., Yates, J.R., 3rd, and Jia, S. (2009b). Regulation of Set9-mediated H4K20 methylation by a PWWP domain protein. *Mol Cell* 33, 428-437.

Weigel, D., Alvarez, J., Smyth, D.R., Yanofsky, M.F., and Meyerowitz, E.M. (1992). LEAFY controls floral meristem identity in *Arabidopsis*. *Cell* 69, 843-859.

Weissbein, U., Plotnik, O., Vershkov, D., and Benvenisty, N. (2017). Culture-induced recurrent epigenetic aberrations in human pluripotent stem cells. *PLoS Genet* 13, e1006979.

West, J.A., Cook, A., Alver, B.H., Stadtfeld, M., Deaton, A.M., Hochedlinger, K., Park, P.J., Tolstorukov, M.Y., and Kingston, R.E. (2014). Nucleosomal occupancy changes locally over key regulatory regions during cell differentiation and reprogramming. *Nat Commun* 5, 4719.

White, C.L., Suto, R.K., and Luger, K. (2001). Structure of the yeast nucleosome core particle reveals fundamental changes in internucleosome interactions. *EMBO J* 20, 5207-5218.

Whitehouse, I., Rando, O.J., Delrow, J., and Tsukiyama, T. (2007). Chromatin remodelling at promoters suppresses antisense transcription. *Nature* 450, 1031-1035.

Whitehouse, I., and Tsukiyama, T. (2006). Antagonistic forces that position nucleosomes in vivo. *Nat Struct Mol Biol* 13, 633-640.

Wierzbicki, A.T., and Jerzmanowski, A. (2005). Suppression of histone H1 genes in *Arabidopsis* results in heritable developmental defects and stochastic changes in DNA methylation. *Genetics* 169, 997-1008.

Woo, H.R., Dittmer, T.A., and Richards, E.J. (2008). Three SRA-domain methylcytosine-binding proteins cooperate to maintain global CpG methylation and epigenetic silencing in *Arabidopsis*. *PLoS Genet* 4, e1000156.

Woo, H.R., Pontes, O., Pikaard, C.S., and Richards, E.J. (2007). VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization. *Genes Dev* 21, 267-277.

Wood, C.C., Robertson, M., Tanner, G., Peacock, W.J., Dennis, E.S., and Helliwell, C.A. (2006). The *Arabidopsis thaliana* vernalization response requires a polycomb-like protein complex that also includes VERNALIZATION INSENSITIVE 3. *Proc Natl Acad Sci U S A* 103, 14631-14636.

Woodcock, C.L., Skoultchi, A.I., and Fan, Y. (2006). Role of linker histone in chromatin structure and function: H1 stoichiometry and nucleosome repeat length. *Chromosome Res* 14, 17-25.

Workman, J.L., and Kingston, R.E. (1998). Alteration of nucleosome structure as a mechanism of transcriptional regulation. *Annu Rev Biochem* 67, 545-579.

Wu, G., Park, M.Y., Conway, S.R., Wang, J.W., Weigel, D., and Poethig, R.S. (2009). The sequential action of miR156 and miR172 regulates developmental timing in *Arabidopsis*. *Cell* 138, 750-759.

Wu, X., Kriz, A.J., and Sharp, P.A. (2014). Target specificity of the CRISPR-Cas9 system. *Quant Biol* 2, 59-70.

Wyrick, J.J., Holstege, F.C., Jennings, E.G., Causton, H.C., Shore, D., Grunstein, M., Lander, E.S., and Young, R.A. (1999). Chromosomal landscape of nucleosome-dependent gene expression and silencing in yeast. *Nature* 402, 418-421.

Wysocka, J., Swigut, T., Xiao, H., Milne, T.A., Kwon, S.Y., Landry, J., Kauer, M., Tackett, A.J., Chait, B.T., Badenhorst, P., *et al.* (2006). A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. *Nature* 442, 86-90.

Xiao, H., Sandaltzopoulos, R., Wang, H.M., Hamiche, A., Ranallo, R., Lee, K.M., Fu, D., and Wu, C. (2001). Dual functions of largest NURF subunit NURF301 in nucleosome sliding and transcription factor interactions. *Mol Cell* 8, 531-543.

Xiao, J., Zhang, H., Xing, L., Xu, S., Liu, H., Chong, K., and Xu, Y. (2013). Requirement of histone acetyltransferases HAM1 and HAM2 for epigenetic modification of FLC in regulating flowering in *Arabidopsis*. *J Plant Physiol* 170, 444-451.

Xiong, J., Zhang, Z., and Zhu, B. (2016). Polycomb "polypacks" the chromatin. *Proc Natl Acad Sci U S A* 113, 14878-14880.

Xu, E.Y., Bi, X., Holland, M.J., Gottschling, D.E., and Broach, J.R. (2005). Mutations in the nucleosome core enhance transcriptional silencing. *Mol Cell Biol* 25, 1846-1859.

Xu, L., Menard, R., Berr, A., Fuchs, J., Cognat, V., Meyer, D., and Shen, W.H. (2009). The E2 ubiquitin-conjugating enzymes, AtUBC1 and AtUBC2, play redundant roles and are involved in activation of FLC expression and repression of flowering in *Arabidopsis thaliana*. *Plant J* 57, 279-288.

Xu, L., Zhao, Z., Dong, A., Soubigou-Taconnat, L., Renou, J.P., Steinmetz, A., and Shen, W.H. (2008). Di- and tri- but not monomethylation on histone H3 lysine 36 marks active transcription of genes involved in flowering time regulation and other processes in *Arabidopsis thaliana*. *Mol Cell Biol* 28, 1348-1360.

Yadon, A.N., and Tsukiyama, T. (2011). SnapShot: Chromatin remodeling: ISWI. *Cell* 144, 453-453 e451.

Yaish, M.W., Colasanti, J., and Rothstein, S.J. (2011). The role of epigenetic processes in controlling flowering time in plants exposed to stress. *J Exp Bot* 62, 3727-3735.

Yamaguchi, A., Wu, M.F., Yang, L., Wu, G., Poethig, R.S., and Wagner, D. (2009). The microRNA-regulated SBP-Box transcription factor SPL3 is a direct upstream activator of LEAFY, FRUITFULL, and APETALA1. *Dev Cell* 17, 268-278.

Yan, L., Wang, L., Tian, Y., Xia, X., and Chen, Z. (2016). Structure and regulation of the chromatin remodeller ISWI. *Nature* 540, 466-469.

Yan, L., Wei, S., Wu, Y., Hu, R., Li, H., Yang, W., and Xie, Q. (2015). High-Efficiency Genome Editing in *Arabidopsis* Using YAO Promoter-Driven CRISPR/Cas9 System. *Mol Plant* 8, 1820-1823.

Yan, L., Wu, H., Li, X., Gao, N., and Chen, Z. (2019). Structures of the ISWI-nucleosome complex reveal a conserved mechanism of chromatin remodeling. *Nat Struct Mol Biol* 26, 258-266.

Yan, Q., Dutt, S., Xu, R., Graves, K., Juszczynski, P., Manis, J.P., and Shipp, M.A. (2009). BBAP monoubiquitylates histone H4 at lysine 91 and selectively modulates the DNA damage response. *Mol Cell* 36, 110-120.

Yang, X., Yu, W., Shi, L., Sun, L., Liang, J., Yi, X., Li, Q., Zhang, Y., Yang, F., Han, X., *et al.* (2011). HAT4, a Golgi apparatus-anchored B-type histone acetyltransferase, acetylates free histone H4 and facilitates chromatin assembly. *Mol Cell* 44, 39-50.

Ye, J., Ai, X., Eugeni, E.E., Zhang, L., Carpenter, L.R., Jelinek, M.A., Freitas, M.A., and Parthun, M.R. (2005). Histone H4 lysine 91 acetylation a core domain modification associated with chromatin assembly. *Mol Cell* 18, 123-130.

Ye, X., Franco, A.A., Santos, H., Nelson, D.M., Kaufman, P.D., and Adams, P.D. (2003). Defective S phase chromatin assembly causes DNA damage, activation of the S phase checkpoint, and S phase arrest. *Mol Cell* 11, 341-351.

Ying, Z., Mei, M., Zhang, P., Liu, C., He, H., Gao, F., and Bao, S. (2015). Histone Arginine Methylation by PRMT7 Controls Germinal Center Formation via Regulating Bcl6 Transcription. *J Immunol* 195, 1538-1547.

Yoo, S.K., Chung, K.S., Kim, J., Lee, J.H., Hong, S.M., Yoo, S.J., Yoo, S.Y., Lee, J.S., and Ahn, J.H. (2005). CONSTANS activates SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1 through FLOWERING LOCUS T to promote flowering in Arabidopsis. *Plant Physiol* 139, 770-778.

Yoshiyama, K.O., Sakaguchi, K., and Kimura, S. (2013). DNA damage response in plants: conserved and variable response compared to animals. *Biology (Basel)* 2, 1338-1356.

Yu, C., Gan, H., Serra-Cardona, A., Zhang, L., Gan, S., Sharma, S., Johansson, E., Chabes, A., Xu, R.M., and Zhang, Z. (2018). A mechanism for preventing asymmetric histone segregation onto replicating DNA strands. *Science* 361, 1386-1389.

Yu, C.W., Liu, X., Luo, M., Chen, C., Lin, X., Tian, G., Lu, Q., Cui, Y., and Wu, K. (2011a). HISTONE DEACETYLASE6 interacts with FLOWERING LOCUS D and regulates flowering in Arabidopsis. *Plant Physiol* 156, 173-184.

Yu, G., Wang, L.G., Han, Y., and He, Q.Y. (2012). clusterProfiler: an R package for comparing biological themes among gene clusters. *OMICS* 16, 284-287.

Yu, Q., Olsen, L., Zhang, X., Boeke, J.D., and Bi, X. (2011b). Differential contributions of histone H3 and H4 residues to heterochromatin structure. *Genetics* 188, 291-308.

Yu, Y., Srinivasan, M., Nakanishi, S., Leatherwood, J., Shilatifard, A., and Sternglanz, R. (2011c). A conserved patch near the C terminus of histone H4 is required for genome stability in budding yeast. *Mol Cell Biol* 31, 2311-2325.

Yuan, G., and Zhu, B. (2013). Histone variants and epigenetic inheritance. *Biochim Biophys Acta* 1819, 222-229.

Yuan, G.C., Liu, Y.J., Dion, M.F., Slack, M.D., Wu, L.F., Altschuler, S.J., and Rando, O.J. (2005). Genome-scale identification of nucleosome positions in *S. cerevisiae*. *Science* 309, 626-630.

Yudkovsky, N., Logie, C., Hahn, S., and Peterson, C.L. (1999). Recruitment of the SWI/SNF chromatin remodeling complex by transcriptional activators. *Genes Dev* 13, 2369-2374.

Yung, P.Y., Stuetzer, A., Fischle, W., Martinez, A.M., and Cavalli, G. (2015). Histone H3 Serine 28 Is Essential for Efficient Polycomb-Mediated Gene Repression in *Drosophila*. *Cell Rep* 11, 1437-1445.

Zanier, K., Luyten, I., Crombie, C., Muller, B., Schumperli, D., Linge, J.P., Nilges, M., and Sattler, M. (2002). Structure of the histone mRNA hairpin required for cell cycle regulation of histone gene expression. *RNA* 8, 29-46.

Zaratiegui, M., Castel, S.E., Irvine, D.V., Kloc, A., Ren, J., Li, F., de Castro, E., Marin, L., Chang, A.Y., Goto, D., *et al.* (2011). RNAi promotes heterochromatic silencing through replication-coupled release of RNA Pol II. *Nature* 479, 135-138.

Zaret, K.S., and Carroll, J.S. (2011). Pioneer transcription factors: establishing competence for gene expression. *Genes Dev* 25, 2227-2241.

Zasadzinska, E., Huang, J., Bailey, A.O., Guo, L.Y., Lee, N.S., Srivastava, S., Wong, K.A., French, B.T., Black, B.E., and Foltz, D.R. (2018). Inheritance of CENP-A Nucleosomes during DNA Replication Requires HJURP. *Dev Cell* 47, 348-362 e347.

Zegerman, P., Canas, B., Pappin, D., and Kouzarides, T. (2002). Histone H3 lysine 4 methylation disrupts binding of nucleosome remodeling and deacetylase (NuRD) repressor complex. *J Biol Chem* 277, 11621-11624.

Zemach, A., McDaniel, I.E., Silva, P., and Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. *Science* 328, 916-919.

Zhang, H., and Freudenreich, C.H. (2007). An AT-rich sequence in human common fragile site FRA16D causes fork stalling and chromosome breakage in *S. cerevisiae*. *Mol Cell* 27, 367-379.

Zhang, H., Lang, Z., and Zhu, J.K. (2018). Dynamics and function of DNA methylation in plants. *Nat Rev Mol Cell Biol* 19, 489-506.

Zhang, H., Zhang, J., Wei, P., Zhang, B., Gou, F., Feng, Z., Mao, Y., Yang, L., Zhang, H., Xu, N., *et al.* (2014). The CRISPR/Cas9 system produces specific and homozygous targeted gene editing in rice in one generation. *Plant Biotechnol J* 12, 797-807.

Zhang, K., Mosch, K., Fischle, W., and Grewal, S.I. (2008). Roles of the Clr4 methyltransferase complex in nucleation, spreading and maintenance of heterochromatin. *Nat Struct Mol Biol* 15, 381-388.

Zhang, K., Sridhar, V.V., Zhu, J., Kapoor, A., and Zhu, J.K. (2007a). Distinctive core histone post-translational modification patterns in *Arabidopsis thaliana*. *PLoS One* 2, e1210.

Zhang, L., Eugeni, E.E., Parthun, M.R., and Freitas, M.A. (2003). Identification of novel histone post-translational modifications by peptide mass fingerprinting. *Chromosoma* 112, 77-86.

Zhang, R., Eler, J., and Langowski, J. (2017). Histone Acetylation Regulates Chromatin Accessibility: Role of H4K16 in Inter-nucleosome Interaction. *Biophys J* 112, 450-459.

Zhang, T., Zhang, W., and Jiang, J. (2015). Genome-Wide Nucleosome Occupancy and Positioning and Their Impact on Gene Expression and Evolution in Plants. *Plant Physiol* 168, 1406-1416.

Zhang, W., Zhang, X., Xue, Z., Li, Y., Ma, Q., Ren, X., Zhang, J., Yang, S., Yang, L., Wu, M., *et al.* (2019). Probing the Function of Metazoan Histones with a Systematic Library of H3 and H4 Mutants. *Dev Cell* 48, 406-419 e405.

Zhang, X., Clarenz, O., Cokus, S., Bernatavichute, Y.V., Pellegrini, M., Goodrich, J., and Jacobsen, S.E. (2007b). Whole-genome analysis of histone H3 lysine 27 trimethylation in *Arabidopsis*. *PLoS Biol* 5, e129.

Zhang, Y., Moqtaderi, Z., Rattner, B.P., Euskirchen, G., Snyder, M., Kadonaga, J.T., Liu, X.S., and Struhl, K. (2009). Intrinsic histone-DNA interactions are not the major determinant of nucleosome positions in vivo. *Nat Struct Mol Biol* 16, 847-852.

Zhang, Z., Wippo, C.J., Wal, M., Ward, E., Korber, P., and Pugh, B.F. (2011). A packing mechanism for nucleosome organization reconstituted across a eukaryotic genome. *Science* 332, 977-980.

Zhao, X., McKillop-Smith, S., and Muller, B. (2004). The human histone gene expression regulator HBP/SLBP is required for histone and DNA synthesis, cell cycle progression and cell proliferation in mitotic cells. *J Cell Sci* 117, 6043-6051.

Zheng, C., and Hayes, J.J. (2003). Structures and interactions of the core histone tail domains. *Biopolymers* 68, 539-546.

Zheng, S., Hu, H., Ren, H., Yang, Z., Qiu, Q., Qi, W., Liu, X., Chen, X., Cui, X., Li, S., *et al.* (2019). The Arabidopsis H3K27me3 demethylase JUMONJI 13 is a temperature and photoperiod dependent flowering repressor. *Nat Commun* 10, 1303.

Zhou, J.X., Liu, Z.W., Li, Y.Q., Li, L., Wang, B., Chen, S., and He, X.J. (2018). Arabidopsis PWWP domain proteins mediate H3K27 trimethylation on FLC and regulate flowering time. *J Integr Plant Biol* 60, 362-368.

Zhou, K., Gaullier, G., and Luger, K. (2019). Nucleosome structure and dynamics are coming of age. *Nat Struct Mol Biol* 26, 3-13.