Olga Pontes · Hailing Jin *Editors*

Nuclear Functions in Plant Transcription, Signaling and Development

Springer

Nuclear Functions in Plant Transcription, Signaling and Development

Olga Pontes • Hailing Jin
Editors

Nuclear Functions in Plant Transcription, Signaling and Development

Springer

Editors
Olga Pontes
Shanghai Center for Plant Stress Biology (PSC)
Shanghai
China

Hailing Jin
Plant Pathology and Microbiology
University of California
Riverside
California
USA

ISBN 978-1-4939-2385-4 ISBN 978-1-4939-2386-1 (eBook)
DOI 10.1007/978-1-4939-2386-1

Library of Congress Control Number: 2015933016

Springer New York Heidelberg Dordrecht London

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface for Nuclear Functions (eds. Pontes and Jin)

Plants provide us with food and are the source of several other by-products such as compounds used in pharmaceutics or biofuels. Therefore, it comes as no surprise that plants are crucial in solving major challenges now facing humanity, namely, food productivity/security, increasing energy demands, and environmental changes. There has been a dramatic increase for plant-derived food and feed products as the world population grows exponentially. Plants are also playing a role in filling our ever-increasing energy needs and these bio-energy crops are expected to provide a sustainable, CO_2-neutral emission solution in the near future. Yet such crops will need to be compatible with food and feed agriculture production and must preserve Earth's ecosystems. So the question is: how can we face all of these challenges?

To meet our planet's needs, we need to improve and further develop sustainable methods for plant production by incorporating both biotechnology and sustainable agricultural practices. In this context, we must first establish a baseline understanding of different molecular and cellular mechanisms underlying plant development and response to stress, so we can then apply it to practical advances in plant production across the globe. Surprisingly, the biological networks underpinning plant yield are still poorly understood, particularly regarding the master regulator of cellular function: the nucleus.

The nucleus harbors the large majority of the plant's DNA, the linear sequence of which is the blueprint of every living organism. However, this is only the beginning: how this DNA is expressed and regulated depends on a variety of other interacting factors that play crucial roles in shaping its organization and function. With the sequencing of several plant genomes and recent advances in high-throughput technologies, plant nuclear biologists have been able to unveil many of the mechanisms underlying genome regulation.

For instance, epigenetic modifications, such as histone post-translational modifications and DNA methylation, directly impact gene expression and genome defense by regulating the organization and function of the genome. Importantly, while evolutionary processes take place at a timescale that does not allow plants to respond and adapt to climate-induced stress, we are starting to recognize that epigenetic mechanisms can confer phenotypic plasticity. Epigenetics enables a heritable control of

phenotypes that can change rapidly in response to environmental cues—sometimes over the course of just two to three generations. This epigenetic timescale of change has tremendous implications for how environmentally altered phenotypes are acquired and inherited at the organism and eventually at the population levels.

Another exciting recent discovery that came about through plant biology research is the previously unacknowledged role of noncoding small RNAs in gene expression. These small molecules have been increasingly recognized as players in the establishment of epigenetic modifications, as well as in genome defense and integrity. Noncoding small RNAs impact normal growth, development, and stress responses in diverse plant species, including staple crops such as rice and maize. Small noncoding RNAs are already playing key roles in plant biotechnology applications including directing the specific and enhanced expression of selected genes. These molecules are therefore of great interest in the context of bioengineering, and have enormous potential for enhancing crop productivity in a wide range of ecosystems. Yet, there is still a great deal left to learn about how small noncoding RNAs are integrated into plants' feedback loops, which direct epigenetic modifications throughout development and the stress response process.

Finally, genomes are dynamic structures as their functional properties are strongly determined by their spatial organization over time. Similarly, changes in higher order nuclear organization alter the functional properties of genomic regions. Various types of subcellular physical domains have been identified in the nucleus, the known nuclear bodies or subcompartments, and these structures are associated with transcription factors, RNA-processing proteins, and epigenetic regulators. Interestingly, these nuclear domains display different behaviors in response to the environment, yet it is still a matter of debate how nuclear organization functionally relates to plant biological processes.

The mechanisms and processes described above make it clear that a true understanding of genome function requires integrating the genomic sequence with what we are still discovering about how epigenetics, small noncoding RNAs, and dynamic nuclear organization modify genomes. It is the goal of this book to compile a series of landmark discussions on the recent advances in plant nuclear biology research and offer new perspectives into the functional relevance of the arrangement of genomes and nuclear processes that impact plant physiology and development. The following chapters will provide insights as to how genes are switched on/off and are tuned to specific expression levels, which will allow us to better predict plant phenotypes. Overall, a better understanding of the fundamentals of plant gene expression will aid in the more efficient design of numerous biotechnological applications and plant-breeding programs. This new knowledge will thus provide a foundation for solving both agricultural and environmental problems as well as developing practices that enable global sustainability. Lastly, plant biology is also relevant to human biology as several aspects of underlying mechanisms are conserved between both organisms. Understanding this shared biology will shed light on human diseases and could lead to better therapies for cancer and genetic diseases.

O. Pontes
H. Jin

Contents

Contributors

Rita Abranches Plant Cell Biology laboratory, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal

Jörg D. Becker Instituto Gulbenkian de Ciencia, Oeiras, Portugal

Frédéric Berger Temasek Lifesciences Laboratory, 1 Research Link, National University of Singapore, Singapore, Republic of Singapore

Department of Biological Sciences, National University of Singapore, Singapore, Republic of Singapore

Leonor C. Boavida Instituto Gulbenkian de Ciencia, Oeiras, Portugal

Xian-Yang Deng National Institute of Biological Sciences, Beijing, China

Yuda Fang National Key Laboratory of Plant Molecular Genetics, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China

Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China

Christy M. Foust Department of Integrative Biology, University of South Florida, Tampa, FL, USA

Xin-Jian He National Institute of Biological Sciences, Beijing, China

Marcela Hernandez-Coronado Instituto Gulbenkian de Ciencia, Oeiras, Portugal

Yoko Ikeda Institute of Plant Science and Resources, Okayama University, Okayama, Japan

Xuncheng Liu South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, Guangdong, China

Yin Liu National Key Laboratory of Plant Molecular Genetics, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai, China

Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai, China

Ming Luo Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, Guangdong, China

Taisuke Nishimura PRESTO, Japanese Science and Technology Agency, Tokyo, Japan

Bioscience and Biotechnology Center, Nagoya University, Nagoya, Japan

Nagaoka University of Technology, Nagaoka, Japan

Margarida Oliveira Genomics of Plant Stress Unit, Instituto de Biologia Experimental e Tecnológica, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal

Christina L. Richards Department of Integrative Biology, University of South Florida, Tampa, FL, USA

Ana Paula Santos Genomics of Plant Stress Unit, Instituto de Biologia Experimental e Tecnológica, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Oeiras, Portugal

Aaron W. Schrey Department of Biology, Armstrong State University, Savannah, GA, USA

Peter Shaw Department of Cell and Developmental Biology, John Innes Centre, Norwich, UK

Lisa M. Smith Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK

Heike Wollmann Temasek Lifesciences Laboratory, 1 Research Link, National University of Singapore, Singapore, Republic of Singapore

Genome Institute of Singapore, Singapore, Republic of Singapore

Keqiang Wu Institute of Plant Biology, College of Life Science, National Taiwan University, Taipei, Taiwan

Songguang Yang South China Botanical Garden, Chinese Academy of Sciences, Guangzhou, Guangdong, China

Chapter 1
RNA-Directed DNA Methylation and Transcriptional Silencing in *Arabidopsis*

Xian-Yang Deng and Xin-Jian He

Introduction

DNA methylation is an important chromatin marker that is involved in transcriptional regulation in plants, fungi, and mammals [1, 2]. In plants, most DNA methylation occurs at transposable elements and other repetitive DNA sequences and is required for the transcriptional silencing of these regions [3, 4]. DNA methylation occurs in all the three cytosine contexts: the symmetric CG and CHG contexts (in which H = A, T, or C) and the asymmetric CHH context. CG methylation is maintained by methyltransferase 1 (MET1), a homolog of mammalian DNA methyltransferase 1 (DNMT1) [5, 6]. The plant-specific chromomethylase 3 (CMT3) specifically catalyzes CHG methylation [7, 8]. Domains rearranged methyltransferase 2 (DRM2) and its homologs are responsible for establishing CHH methylation and to a lesser extent CG and CHG methylation [9, 10]. CG methylation is present in transposable elements and DNA repeats, as well as in genic regions, but CHG and CHH methylation is almost exclusively present in transposable elements and DNA repeats [3, 4].

Small interfering RNAs (siRNAs) and long noncoding RNAs (ncRNAs) are responsible for establishing DNA methylation and/or repressive histone H3K9 methylation at transposable elements and DNA repeats in plants, fungi, and mammals [1, 11, 12]. In *Arabidopsis,* DNA methylation can be established through a well-described RNA-directed DNA methylation (RdDM) pathway [2]. RdDM plays important roles in development, stress response, and genome evolution [2, 12, 13]. RdDM requires canonical components in the conserved RNA interference (RNAi) machinery; these components are members of the Dicer and Argonaute families and RNA-dependent RNA polymerase 2 (RDR2) [14, 15]. Moreover, plant-specific DNA-dependent RNA polymerases IV and V (Pol IV and Pol V), DNA methyltransferase DRM2, and several other proteins are required for RdDM [13, 14, 16–20]. In the past few years, our knowledge of RdDM has been

X.-J. He (✉) · X.-Y. Deng
National Institute of Biological Sciences, Beijing 102206, China
e-mail: hexinjian@nibs.ac.cn

O. Pontes, H. Jin (eds.), *Nuclear Functions in Plant Transcription, Signaling and Development,* DOI 10.1007/978-1-4939-2386-1_1

greatly improved by genetic, biochemical, and structural studies. In this chapter, we describe the recent insights into the mechanisms underlying RdDM.

Pol IV-Dependent siRNAs

Two atypical, plant-specific polymerases, Pol IV and Pol V, are required for the biogenesis of siRNAs and long ncRNAs, respectively; both Pol IV and Pol V consist of multiple subunits [20, 21]. Nuclear RNA polymerase D1 (NRPD1; formerly named NRPD1a) is the largest subunit of Pol IV, and nuclear RNA polymerase E1 (NRPE1; formerly named NRPD1b) is the largest subunit of Pol V [20]. Nuclear RNA polymerase D/E2 (NRPD/E2) is the second largest subunit of both Pol IV and Pol V. Some subunits are only present in one RNA polymerase, but others are shared by Pol IV, Pol V, and/or Pol II [16, 17, 19, 20, 22].

RdDM is thought to be initiated by 24-nt siRNAs. As indicated in Fig. 1.1, the biogenesis of the 24-nt siRNAs is dependent on Pol IV [14, 16, 18], RDR2, and Dicer-like 3 (DCL3) [14, 16, 18, 23]. Pol IV is responsible for producing single-stranded RNAs (ssRNAs), which are converted into double-stranded RNAs

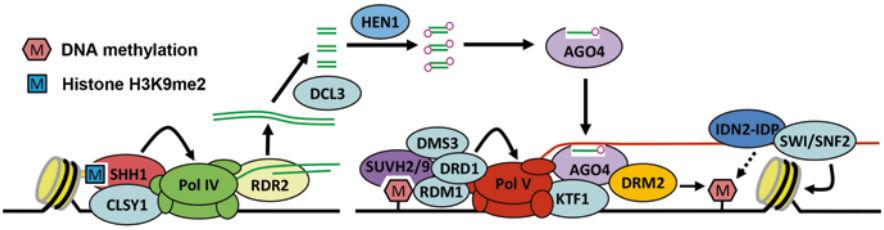

Fig. 1.1 Model for RNA-directed DNA methylation (RdDM) in *Arabidopsis*. In the RdDM pathway, RNA polymerase IV (Pol IV) transcribes single-stranded RNAs that are immediately converted into double-stranded RNAs by RDR2. DCL3 cleaves the double-stranded RNAs into 24-nt siRNAs that are methylated at their 3′-ends by HEN1. A single strand of the siRNA duplex associates with AGO4 and forms the AGO4–siRNA complex. SHH1 recognizes histone H3K9 methylation and then recruits Pol IV to RdDM target loci. CLSY1 interacts with SHH1 and may assist Pol IV recruitment. RNA polymerase V (Pol V) transcribes long noncoding RNAs that act as scaffold RNAs. The recruitment of Pol V to target loci is dependent on the DDR complex, which is composed of DMS3, DRD1, and RDM1. The SRA domain- and SET domain-containing proteins SUVH2 and SUVH9 bind methylated DNA and interact with the DDR complex, thereby facilitating the recruitment of Pol V. Pol V-produced noncoding scaffold RNAs base-pair with siRNAs in the AGO4–siRNA complex, whereas NRPE1 and KTF1 interact with AGO4 through their C-terminal WG/GW domains. The AGO4–siRNA complex guides the de novo DNA methyltransferase DRM2, which catalyzes DNA methylation at the loci. IDN2 interacts with its paralogs IDP1 and IDP2, and forms tetramers that bind Pol V-produced scaffold RNAs. IDN2 associates with the SWI/SNF complex and mediates nucleosome positioning for Pol V-stabilized nucleosomes. DCL3 - Dicer-like 3, RDR2 - RNA-dependent RNA polymerase 2, siRNA - small interfering RNAs, HEN1 - hua enhancer 1, AGO4 - Argonaute protein 4, SHH1 Sawadee homeodomain homolog, NRPE1 - nuclear RNA polymerase E1, IDN2 - involved in de novo 2, SWI/SNF switch/sucrose nonfermentable, CLSY1 - chromatin-remodeling protein CLASSY1, DMS3 defective in meristem silencing 3, DRD1 - defective in RdDM,SRA SET- and RING-associated, SUVH SU(VAR) homologs, KTF1 - KOW-containing transcription factor 1

(dsRNAs) by RDR2. The dsRNAs are cleaved into 24-nt siRNAs mainly by DCL3, which is partially redundant with two other DCL enzymes, DCL2 and DCL4 [14, 23, 24]. The 3′-OH groups of 24-nt siRNAs are methylated by HUA ENHANCER 1 (HEN1), which stabilizes 24-nt siRNAs in vivo [25]. The 24-nt siRNAs are loaded onto ARGONAUTE proteins AGO4, AGO6, or AGO9 [15, 26–28] and are assembled into RNA-induced transcriptional silencing (RITS) complexes that signal de novo DNA methylation and transcriptional silencing at target regions.

Polymerase activity in vitro has been documented for Pol IV [29], and mutations in the conserved catalytic site of NRPD1 abolish the abundance of 24-nt siRNAs; together, these results suggest that Pol IV is an active RNA polymerase in vivo [29]. Unlike Pol II, Pol IV activity in vitro requires an RNA primer and is insensitive to alpha-amanitin [29]. Pol IV and RDR2 associate in vivo, but Pol IV does not require RDR2 for activity, whereas RDR2 is nonfunctional in the absence of associated Pol IV. The coupling of Pol IV and RDR2 results in the channeled synthesis of double-stranded precursors for 24-nt siRNA biogenesis [29].

Pol IV-dependent 24-nt siRNAs are the most abundant class of small RNAs in *Arabidopsis*. These siRNAs are mainly produced at thousands of discrete transposable elements and repetitive DNA elements located at pericentromeric heterochromatin [18, 30, 31]. It is important to determine how Pol IV-dependent siRNAs are produced at specific chromatin regions rather than at others. Sawadee homeodomain homolog 1 (SHH1)/DNA-binding transcription factor 1 (DTF1) was independently identified by a forward genetic screen and a Pol IV affinity purification [32, 33]. SHH1/DTF1 specifically associates with Pol IV but not with Pol V, and the accumulation of most Pol IV-dependent siRNAs is markedly decreased in the *shh1/dtf1* mutants [32, 34, 35], suggesting a role of SHH1 in Pol IV transcription. SHH1 contains a SAWADEE domain that preferentially binds to unmethylated K4 and methylated K9 modifications on the histone H3 [34]. Pol IV ChIP accompanied by DNA deep sequencing indicated that SHH1 is required for the association of Pol IV with chromatin [34]. When critical residues in the SAWADEE domain are mutated, both 24-nt siRNA and DNA methylation levels are decreased [34]. These results suggest that SHH1 is responsible for targeting Pol IV to chromatin by associating with RdDM target loci that have unmethylated K4 and methylated K9 on histone H3. The chromatin-remodeling protein CLASSY1 (CLSY1) was primarily identified as an RdDM component by a forward genetic screen, and Pol IV-dependent siRNA accumulation is drastically decreased in the *clsy1* mutant [36]. CLSY1 is purified by Pol IV affinity purification [32], suggesting that CLSY1 associates with Pol IV. The functional and physical association of CLSY1 with Pol IV suggests that chromatin remodeling is involved in Pol IV transcription.

Pol V-Dependent ncRNAs

Like Pol IV, Pol V is also required for siRNA accumulation. However, the effect of Pol V on siRNA accumulation is limited to a subset of Pol IV-dependent siRNAs and is likely a result of its effect on DNA methylation [30, 34]. Researchers have demonstrated that Pol V-produced ncRNAs function as scaffolds for the recruitment

of the silencing machinery and help siRNAs recognize their target loci; the latter function is possibly facilitated by base-pairing between AGO4-bound siRNAs and nascent Pol V-produced transcripts [37, 38]. Polymerase activity in vitro and in vivo has been shown for Pol V [29, 37]. That Pol V carries out transcription using the bipartite oligonucleotide template but not the tripartite template suggests an inability to disrupt downstream dsDNA during transcription [29].

Pol V-dependent ncRNAs help recruit the RdDM silencing machinery and are required for siRNA-mediated DNA methylation and transcriptional silencing. A genome-wide ChIP-seq analysis indicated that the largest subunit of Pol V, NRPE1, is enriched at promoters of protein-coding genes and at recently evolved transposons [39]. This localization pattern is highly correlated with Pol V-dependent DNA methylation and 24-nt siRNA accumulation [39–41]. The vast majority of Pol V-enriched regions are usually shorter than 250 bp [39]. The association of Pol V with promoters of protein-coding genes indicates that Pol V is likely originated from ancient RNA polymerase II (Pol II) [39], which is consistent with the finding that Pol II shares several conserved subunits with Pol IV and Pol V [20, 21]. A small proportion of ncRNAs are produced by Pol II and these ncRNAs are involved in DNA methylation and transcriptional gene silencing through the RdDM pathway [6, 42], supporting the notion that Pol IV and Pol V are originally evolved from Pol II.

A number of transcription factors are required for Pol II transcription. It is interesting to determine whether Pol V transcription requires transcription factors. RDM4 (RNA-directed DNA methylation 4) / DMS4 (defective in meristem silencing 4), a homolog of the yeast Pol II-dependent transcription factor IWR1, has been identified as a canonical RdDM component by two independent genetic screens [43, 44]. Pol V-produced ncRNAs are decreased in the *rdm4* mutant, suggesting that RDM4 is required for Pol V transcription [43]. Unlike other canonical RdDM mutants, the *rdm4* mutant has pleiotropic developmental defects. RDM4 is a transcription factor that is shared by Pol II, Pol IV, and Pol V [32, 43].

The production of Pol V-dependent ncRNAs is also dependent on DRD1 (defective in RNA-directed DNA methylation 1), DMS3 (defective in meristem silencing 3), and RDM1 (RNA-directed DNA methylation 1) [37, 45–47], which form a DDR (DRD1, DMS3, and RDM1) complex in vivo [11]. The DDR complex is required for the association of Pol V with chromatin, and the association facilitates the transcription of Pol V-dependent noncoding RNAs [39]. The SU(VAR)3–9 homologs SUVH2 and SUVH9 act redundantly in RdDM and transcriptional silencing [48, 49]. By associating with the DDR complex, SUVH2 and SUVH9 are involved in the association of Pol V with chromatin [50, 51]. SUVH2 and SUVH9 contain an SET- and RING-associated (SRA) domain that directly binds to methylated DNA [48]. SUVH2 and SUVH9 act in RdDM by directing the DDR complex and Pol V to RdDM target loci (Fig. 1.1). The binding of SUVH2 and SUVH9 to methylated DNA facilitates the formation of a self-reinforcing loop of DNA methylation and Pol V transcription.

Microrchidia 6 (MORC6)/DMS11 (defective in meristem silencing 11) and MORC1 are the members of the conserved microrchidia adenosine triphosphatase (ATPase) family with a GHKL (gyrase, Hsp90, histidine kinase, MutL) ATPase domain. The *morc1* and *morc6* mutants show decondensation of pericentromeric

heterochromatin and increased interaction between the pericentromeric regions and the rest of the genome [52], whereas the *morc6*/*dms11* mutant shows slight decreases in siRNA accumulation and DNA methylation [53, 54]. A recent study indicate that MORC1 and MORC6 interact with the DDR complex and with SUVH2 and/or SUVH9 [50], which is consistent with the effect of the *morc6*/*dms11* mutation on Pol V-dependent transcripts [54].

Recruitment of RdDM Effector to Chromatin

Pol IV-dependent siRNAs associate with AGO4, thereby facilitating the formation of an RdDM effecter complex that is required for DNA methylation. The assembly of the RdDM effector complex in *Arabidopsis* is similar to the RITS complex in fission yeast [1, 55]. In *Arabidopsis*, 24-nt siRNAs produced by DCL3 are subjected to a sorting process, and the specificity of RNA sorting may be associated with the terminal nucleotide of the siRNA and duplex properties, such as thermodynamic asymmetry or degree of base-pairing [56, 57]. Twenty-four-nucleotide siRNAs are loaded onto AGO4 [58, 59]. The loading process occurs in the cytoplasm with assistance from the ATP-bound HSP90, and ATP hydrolysis induces the dissociation of the siRNA passenger strand and results in a conformational change in AGO4, leading to the importation of the AGO4–siRNA complex into the nucleus [60]. Pol V-dependent noncoding RNAs recruit the AGO4–siRNA complex to chromatin by base-pairing with siRNAs [38]. DRM2 is a key de novo DNA methyltransferase that is responsible for DNA methylation at RdDM target loci, but how DRM2 is recruited to chromatin is poorly understood. A recent study demonstrated that DRM2 exists in a complex with AGO4 and preferentially methylates one DNA strand, which acts as the template for Pol V transcription [61]. The results support a model, in which DRM2 is guided to target loci in a strand-specific manner. The AGO4–siRNA complex and Pol V-dependent ncRNAs may be required for the recruitment of DRM2 to chromatin (Fig. 1.1).

Immunolocalization experiments with isolated nuclei have shown that RDR2, DCL3, and AGO4 localize in the Cajal body [62, 63]. The Cajal body is involved in a variety of functions including pre-mRNA splicing, rRNA processing, and telomere maintenance [64]. It is possible that the biogenesis of siRNAs and long ncRNAs, and the assembly of protein–RNA complexes may involve the function of the Cajal body. Further study is required to understand the detailed role of the Cajal body in RdDM. A portion of the NRPE1 signals colocalize at target loci, together with DRD1, a protein required for Pol V transcription [59, 62], which is consistent with the function association of Pol V and DRD1 at the downstream step of RdDM.

Assembly of the RdDM effector complexes is mediated by multiple protein–protein and protein–RNA interactions that recruit proteins to specific genomic regions. NRPE1 contains WG/GW repeats in its C-terminal domain, which is thought to act as an AGO4 hook motif [65]. KOW-containing transcription factor 1 (KTF1)/SPT5-like protein (SPT5L) also contains conserved WG/GW repeats in its C-terminal domain [66, 67]. Moreover, KTF1/SPT5L is capable of binding to

noncoding scaffold RNAs produced by Pol V [67]. Chromatin binding of AGO4 and KTF1 occurs downstream to Pol V [38, 68]. KTF1 and AGO4 are recruited to chromatin in parallel and partially independently of each other, whereas KTF1 enhances AGO4 chromatin binding at a subset of RdDM sites [68]. The chromatin binding of KTF1/SPT5L and AGO4 may create a platform for the recruitment of DRM2 to chromatin at RdDM targets [68].

IDN2 (involved in de novo 2)/RDM12 (RNA-directed DNA methylation 12), another factor that is thought to act as a downstream RdDM effector, is homologous to suppressor of gene silencing 3 (SGS3), a protein involved in the accumulation of viral siRNAs, ta-siRNAs, and nat-siRNAs for posttranscriptional gene silencing. IDN2 contains an N-terminal C2H2-type zinc finger domain, an XS domain, a coiled-coil domain, and a specific XH domain [69–71]. The XS domain is required for the binding of IDN2 to dsRNAs [70]. A possible RNA substrate for IDN2 is the duplex that is formed between AGO4-bound siRNAs and Pol V-dependent noncoding transcripts. The duplex could be a signal that aids in the recruitment of DRM2 to establish DNA methylation.

IDN2 can form a homodimer and associate with two IDN2 paralogs, IDP1 (IDN2 paralog 1)/IDNL1 (IDN2-LIKE 1) and IDP2/IDNL2 [69, 71]. IDP1 is required for siRNA accumulation, de novo DNA methylation, and transcriptional gene silencing, whereas the roles of IDP2 partially overlap with those of IDP1 [71, 72]. The coiled-coil domain of IDN2 is essential for the homodimerization of IDN2 with itself but is not required for IDN2 association with IDP1 and IDP2, whereas the uncharacterized XH domain of IDN2 is required for association with IDP1 and IDP2 but not for IDN2 homodimerization [71]. Unlike IDN2, IDP1 and IDP2 are incapable of binding to double-stranded RNA, suggesting that IDP1 and IDP2 have distinct roles in the IDN2–IDP1/IDP2 complex [71]. The IDN2–IDP1/IDP2 complex may facilitate the recruitment of the double-stranded RNA-containing RdDM effector complex to specific chromatin regions at a downstream step of the RdDM pathway.

IDN2 was shown to interact with SWI3B, a component of the SWI/SNF (switch/sucrose nonfermentable) ATP-dependent nucleosome-remodeling complex [73]. The SWI/SNF complex affects Pol V-stabilized nucleosome positioning and contributes to transcriptional silencing. The study suggests that IDN2 acts as an adaptor protein connecting Pol V-produced ncRNAs and the SWI/SNF complex, thereby guiding the SWI/SNF complex and mediating nucleosome positioning (Fig. 1.1). Pol V-produced ncRNAs may be not only required for DNA methylation but also required for repressive nucleosome positioning at RdDM target loci.

Conclusion and Perspective

RdDM is involved in the transcriptional silencing of noncoding genomic regions as well as of protein-coding genes flanked by noncoding transposable elements and other repetitive DNA sequences. Disruption of RdDM components not only affects

transposable element silencing and genome stability but also affects protein-coding gene expression and imprinting, stress response, and various developmental stages. An important objective for future research is to determine how RdDM acts in these processes. Although most RdDM components have been cloned and preliminarily characterized, some critical questions remain. How are Pol IV- and Pol V-dependent noncoding RNAs differentiated for distinct functions? Which chromatin superstructure features are preferentially targeted by RdDM? How is DNA methylation established for introduced unmethylated transgenes? How are DNA methylation, histone modification, heterochromatin condensation, and nucleosome positioning concomitantly regulated at RdDM target loci? Recently, several pre-mRNA splicing factors were demonstrated to be involved in RdDM and transcriptional silencing [74–77]. Similarly, RNA splicing and processing is required for RITS in fission yeast [78, 79]. It will be interesting to determine how these splicing factors coordinate with the RdDM machinery and act in transcriptional silencing. Further studies on RdDM will undoubtedly clarify the roles of noncoding genomic regions in plants and will of course benefit the breeding of agricultural crops.

References

1. Castel SE, Martienssen RA. RNA interference in the nucleus: roles for small RNAs in transcription, epigenetics and beyond. Nat Rev Genet. 2013 Feb;14(2):100–12.
2. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet. 2010 March;11(3):204–20.
3. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, et al. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. Nature. 2008 March 13;452(7184):215–9.
4. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, et al. Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. Cell. 2008 May 2;133(3):523–36.
5. Ronemus MJ, Galbiati M, Ticknor C, Chen J, Dellaporta SL. Demethylation-induced developmental pleiotropy in *Arabidopsis*. Science. 1996 Aug 2;273(5275):654–7.
6. Stroud H, Greenberg MV, Feng S, Bernatavichute YV, Jacobsen SE. Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. Cell. 2013 Jan 17;152(1–2):352–64.
7. Du J, Zhong X, Bernatavichute YV, Stroud H, Feng S, Caro E, et al. Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. Cell. 2012 Sept 28;151(1):167–80.
8. Jackson JP, Lindroth AM, Cao X, Jacobsen SE. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature. 2002 April 4;416(6880):556–60.
9. Cao X, Jacobsen SE. Role of the *arabidopsis* DRM methyltransferases in *de novo* DNA methylation and gene silencing. Curr Biol. 2002 July 9;12(13):1138–44.
10. Henderson IR, Deleris A, Wong W, Zhong X, Chin HG, Horwitz GA, et al. The *de novo* cytosine methyltransferase DRM2 requires intact UBA domains and a catalytically mutated paralog DRM3 during RNA-directed DNA methylation in *Arabidopsis thaliana*. PLoS Genet. 2010 Oct;6(10):e1001182.
11. Law JA, Ausin I, Johnson LM, Vashisht AA, Zhu JK, Wohlschlegel JA, et al. A protein complex required for polymerase V transcripts and RNA- directed DNA methylation in *Arabidopsis*. Curr Biol. 2010 May 25;20(10):951–6.

12. Matzke M, Kanno T, Daxinger L, Huettel B, Matzke AJ. RNA-mediated chromatin-based silencing in plants. Curr Opin Cell Biol. 2009 June;21(3):367–76.
13. Haag JR, Pikaard CS. Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. Nat Rev Mol Cell Biol. 2011 Aug;12(8):483–92.
14. Chan SW, Zilberman D, Xie Z, Johansen LK, Carrington JC, Jacobsen SE. RNA silencing genes control *de novo* DNA methylation. Science. 2004 Feb 27;303(5662):1336.
15. Zilberman D, Cao X, Jacobsen SE. ARGONAUTE4 control of locus-specific siRNA accumulation and DNA and histone methylation. Science. 2003 Jan 31;299(5607):716–9.
16. Herr AJ, Jensen MB, Dalmay T, Baulcombe DC. RNA polymerase IV directs silencing of endogenous DNA. Science. 2005 April 1;308(5718):118–20.
17. Kanno T, Huettel B, Mette MF, Aufsatz W, Jaligot E, Daxinger L, et al. Atypical RNA polymerase subunits required for RNA-directed DNA methylation. Nat Genet. 2005 July;37(7):761–5.
18. Onodera Y, Haag JR, Ream T, Costa Nunes P, Pontes O, Pikaard CS. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. Cell. 2005 March 11;120(5):613–22.
19. Pontier D, Yahubyan G, Vega D, Bulski A, Saez-Vasquez J, Hakimi MA, et al. Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in *Arabidopsis*. Genes Dev. 2005 Sept 1;19(17):2030–40.
20. Ream TS, Haag JR, Wierzbicki AT, Nicora CD, Norbeck AD, Zhu JK, et al. Subunit compositions of the RNA-silencing enzymes Pol IV and Pol V reveal their origins as specialized forms of RNA polymerase II. Mol Cell. 2009 Jan 30;33(2):192–203.
21. Huang L, Jones AM, Searle I, Patel K, Vogler H, Hubner NC, et al. An atypical RNA polymerase involved in RNA silencing shares small subunits with RNA polymerase II. Nat Struct Mol Biol. 2009 Jan;16(1):91–3.
22. He XJ, Hsu YF, Pontes O, Zhu J, Lu J, Bressan RA, et al. NRPD4, a protein related to the RPB4 subunit of RNA polymerase II, is a component of RNA polymerases IV and V and is required for RNA-directed DNA methylation. Genes Dev. 2009 Feb 1;23(3):318–30.
23. Xie Z, Johansen LK, Gustafson AM, Kasschau KD, Lellis AD, Zilberman D, et al. Genetic and functional diversification of small RNA pathways in plants. PLoS Biol. 2004 May;2(5):E104.
24. Henderson IR, Zhang X, Lu C, Johnson L, Meyers BC, Green PJ, et al. Dissecting *Arabidopsis thaliana* DICER function in small RNA processing, gene silencing and DNA methylation patterning. Nat Genet. 2006 June;38(6):721–5.
25. Ji L, Chen X. Regulation of small RNA stability: methylation and beyond. Cell Res. 2012 April;22(4):624–36.
26. Eun C, Lorkovic ZJ, Naumann U, Long Q, Havecker ER, Simon SA, et al. AGO6 functions in RNA-mediated transcriptional gene silencing in shoot and root meristems in *Arabidopsis thaliana*. PloS One. 2011;6(10):e25730.
27. Havecker ER, Wallbridge LM, Hardcastle TJ, Bush MS, Kelly KA, Dunn RM, et al. The *Arabidopsis* RNA-directed DNA methylation argonautes functionally diverge based on their expression and interaction with target loci. Plant Cell. 2010 Feb;22(2):321–34.
28. Zheng X, Zhu J, Kapoor A, Zhu JK. Role of *Arabidopsis* AGO6 in siRNA accumulation, DNA methylation and transcriptional gene silencing. EMBO J. 2007 March 21;26(6):1691–701.
29. Haag JR, Ream TS, Marasco M, Nicora CD, Norbeck AD, Pasa-Tolic L, et al. In vitro transcription activities of Pol IV, Pol V, and RDR2 reveal coupling of Pol IV and RDR2 for dsRNA synthesis in plant RNA silencing. Mol Cell. 2012 Dec 14;48(5):811–8.
30. Mosher RA, Schwach F, Studholme D, Baulcombe DC. PolIVb influences RNA-directed DNA methylation independently of its role in siRNA biogenesis. Proc Natl Acad Sci U S A. 2008 Feb 26;105(8):3145–50.
31. Zhang X, Henderson IR, Lu C, Green PJ, Jacobsen SE. Role of RNA polymerase IV in plant small RNA metabolism. Proc Natl Acad Sci U S A. 2007 March 13;104(11):4536–41.
32. Law JA, Vashisht AA, Wohlschlegel JA, Jacobsen SE. SHH1, a homeodomain protein required for DNA methylation, as well as RDR2, RDM4, and chromatin remodeling factors, associate with RNA polymerase IV. PLoS Genet. 2011 July;7(7):e1002195.

33. Liu J, Bai G, Zhang C, Chen W, Zhou J, Zhang S, et al. An atypical component of RNA-directed DNA methylation machinery has both DNA methylation-dependent and -independent roles in locus-specific transcriptional gene silencing. Cell Res. 2011 Dec;21(12):1691–700.
34. Law JA, Du J, Hale CJ, Feng S, Krajewski K, Palanca AM, et al. Polymerase IV occupancy at RNA-directed DNA methylation sites requires SHH1. Nature. 2013 June 20;498(7454):385–9.
35. Zhang H, Ma ZY, Zeng L, Tanaka K, Zhang CJ, Ma J, et al. DTF1 is a core component of RNA-directed DNA methylation and may assist in the recruitment of Pol IV. Proc Natl Acad Sci U S A. 2013 May 14;110(20):8290–5.
36. Smith LM, Pontes O, Searle I, Yelina N, Yousafzai FK, Herr AJ, et al. An SNF2 protein associated with nuclear RNA silencing and the spread of a silencing signal between cells in *Arabidopsis*. Plant Cell. 2007 May;19(5):1507–21.
37. Wierzbicki AT, Haag JR, Pikaard CS. Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. Cell. 2008 Nov 14;135(4):635–48.
38. Wierzbicki AT, Ream TS, Haag JR, Pikaard CS. RNA polymerase V transcription guides ARGONAUTE4 to chromatin. Nat Genet. 2009 May;41(5):630–4.
39. Zhong X, Hale CJ, Law JA, Johnson LM, Feng S, Tu A, et al. DDR complex facilitates global association of RNA polymerase V to promoters and evolutionarily young transposons. Nat Struct Mol Biol. 2012 Sept;19(9):870–5.
40. Lee TF, Gurazada SG, Zhai J, Li S, Simon SA, Matzke MA, et al. RNA polymerase V-dependent small RNAs in *Arabidopsis* originate from small, intergenic loci including most SINE repeats. Epigenetics. 2012 July;7(7):781–95.
41. Wierzbicki AT, Cocklin R, Mayampurath A, Lister R, Rowley MJ, Gregory BD, et al. Spatial and functional relationships among Pol V-associated loci, Pol IV-dependent siRNAs, and cytosine methylation in the *Arabidopsis* epigenome. Genes Dev. 2012 Aug 15;26(16):1825–36.
42. Zheng B, Wang Z, Li S, Yu B, Liu JY, Chen X. Intergenic transcription by RNA polymerase II coordinates Pol IV and Pol V in siRNA-directed transcriptional gene silencing in *Arabidopsis*. Genes Dev. 2009 Dec 15;23(24):2850–60.
43. He XJ, Hsu YF, Zhu S, Liu HL, Pontes O, Zhu J, et al. A conserved transcriptional regulator is required for RNA-directed DNA methylation and plant development. Genes Dev. 2009 Dec 1;23(23):2717–22.
44. Kanno T, Bucher E, Daxinger L, Huettel B, Kreil DP, Breinig F, et al. RNA-directed DNA methylation and plant development require an IWR1-type transcription factor. EMBO Rep. 2010 Jan;11(1):65–71.
45. Gao Z, Liu HL, Daxinger L, Pontes O, He X, Qian W, et al. An RNA polymerase II- and AGO4-associated protein acts in RNA-directed DNA methylation. Nature. 2010 May 6;465(7294):106–9.
46. Kanno T, Bucher E, Daxinger L, Huettel B, Bohmdorfer G, Gregor W, et al. A structural-maintenance-of-chromosomes hinge domain-containing protein is required for RNA-directed DNA methylation. Nat Genet. 2008 May;40(5):670–5.
47. Kanno T, Mette MF, Kreil DP, Aufsatz W, Matzke M, Matzke AJ. Involvement of putative SNF2 chromatin remodeling protein DRD1 in RNA-directed DNA methylation. Curr Biol. 2004 May 4;14(9):801–5.
48. Johnson LM, Law JA, Khattar A, Henderson IR, Jacobsen SE. SRA-domain proteins required for DRM2-mediated *de novo* DNA methylation. PLoS Genet. 2008 Nov;4(11):e1000280.
49. Kuhlmann M, Mette MF. Developmentally non-redundant SET domain proteins SUVH2 and SUVH9 are required for transcriptional gene silencing in *Arabidopsis* thaliana. Plant Mol Biol. 2012 Aug;79(6):623–33.
50. Liu ZW, Shao CR, Zhang CJ, Zhou JX, Zhang SW, Li L, et al. The SET domain proteins SUVH2 and SUVH9 are required for Pol V occupancy at RNA-directed DNA methylation loci. PLoS Genet. 2014 Jan;10(1):e1003948.
51. Johnson LM, Du J, Hale CJ, Bischof S, Feng S, Chodavarapu RK, et al. SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation. Nature. 2014 March 6;507(7490):124–8.

52. Moissiard G, Cokus SJ, Cary J, Feng S, Billi AC, Stroud H, et al. MORC family ATPases required for heterochromatin condensation and gene silencing. Science. 2012 June 15;336(6087):1448–51.
53. Brabbs TR, He Z, Hogg K, Kamenski A, Li Y, Paszkiewicz KH, et al. The stochastic silencing phenotype of *Arabidopsis* morc6 mutants reveals a role in efficient RNA-directed DNA methylation. Plant J. 2013 Sept;75(5):836–46.
54. Lorkovic ZJ, Naumann U, Matzke AJ, Matzke M. Involvement of a GHKL ATPase in RNA-directed DNA methylation in *Arabidopsis thaliana*. Curr Biol. 2012 May 22;22(10):933–8.
55. Moazed D. Small RNAs in transcriptional gene silencing and genome defence. Nature. 2009 Jan 22;457(7228):413–20.
56. Czech B, Hannon GJ. Small RNA sorting: matchmaking for Argonautes. Nat Rev Genet. 2011 Jan;12(1):19–31.
57. Mi S, Cai T, Hu Y, Chen Y, Hodges E, Ni F, et al. Sorting of small RNAs into *Arabidopsis* argonaute complexes is directed by the 5' terminal nucleotide. Cell. 2008 April 4;133(1):116–27.
58. Qi Y, He X, Wang XJ, Kohany O, Jurka J, Hannon GJ. Distinct catalytic and non-catalytic roles of ARGONAUTE4 in RNA-directed DNA methylation. Nature. 2006 Oct 26;443(7114):1008–12.
59. Pontes O, Li CF, Costa Nunes P, Haag J, Ream T, Vitins A, et al. The *Arabidopsis* chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. Cell. 2006 July 14;126(1):79–92.
60. Ye R, Wang W, Iki T, Liu C, Wu Y, Ishikawa M, et al. Cytoplasmic assembly and selective nuclear import of *Arabidopsis* Argonaute4/siRNA complexes. Mol Cell. 2012 June 29;46(6):859–70.
61. Zhong X, Du J, Hale CJ, Gallego-Bartolome J, Feng S, Vashisht AA, et al. Molecular mechanism of action of plant DRM de novo DNA methyltransferases. Cell. 2014 May 22;157(5):1050–60.
62. Li CF, Pontes O, El-Shami M, Henderson IR, Bernatavichute YV, Chan SW, et al. An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. Cell. 2006 July 14;126(1):93–106.
63. Pontes O, Vitins A, Ream TS, Hong E, Pikaard CS, Costa-Nunes P. Intersection of small RNA pathways in *Arabidopsis thaliana* sub-nuclear domains. PloS One. 2013;8(6):e65652.
64. Pikaard CS. Cell biology of the *Arabidopsis* nuclear siRNA pathway for RNA-directed chromatin modification. Cold Spring Harb Symp Quant Biol. 2006;71:473–80.
65. El-Shami M, Pontier D, Lahmy S, Braun L, Picart C, Vega D, et al. Reiterated WG/GW motifs form functionally and evolutionarily conserved ARGONAUTE-binding platforms in RNAi-related components. Genes Dev. 2007 Oct 15;21(20):2539–44.
66. Bies-Etheve N, Pontier D, Lahmy S, Picart C, Vega D, Cooke R, et al. RNA-directed DNA methylation requires an AGO4-interacting member of the SPT5 elongation factor family. EMBO Rep. 2009 June;10(6):649–54.
67. He XJ, Hsu YF, Zhu S, Wierzbicki AT, Pontes O, Pikaard CS, et al. An effector of RNA-directed DNA methylation in *arabidopsis* is an ARGONAUTE 4- and RNA-binding protein. Cell. 2009 May 1;137(3):498–508.
68. Rowley MJ, Avrutsky MI, Sifuentes CJ, Pereira L, Wierzbicki AT. Independent chromatin binding of ARGONAUTE4 and SPT5 L/KTF1 mediates transcriptional gene silencing. PLoS Genet. 2011 June;7(6):e1002120.
69. Ausin I, Greenberg MV, Simanshu DK, Hale CJ, Vashisht AA, Simon SA, et al. INVOLVED IN DE NOVO 2-containing complex involved in RNA-directed DNA methylation in *Arabidopsis*. Proc Natl Acad Sci U S A. 2012 May 29;109(22):8374–81.
70. Ausin I, Mockler TC, Chory J, Jacobsen SE. IDN1 and IDN2 are required for *de novo* DNA methylation in *Arabidopsis thaliana*. Nat Struct Mol Biol. 2009 Dec;16(12):1325–7.
71. Zhang CJ, Ning YQ, Zhang SW, Chen Q, Shao CR, Guo YW, et al. IDN2 and its paralogs form a complex required for RNA-directed DNA methylation. PLoS Genet. 2012;8(5):e1002693.
72. Xie M, Ren G, Costa-Nunes P, Pontes O, Yu B. A subgroup of SGS3-like proteins act redundantly in RNA-directed DNA methylation. Nucleic Acids Res. 2012 May;40(10):4422–31.

73. Zhu Y, Rowley MJ, Bohmdorfer G, Wierzbicki AT. A SWI/SNF chromatin-remodeling complex acts in noncoding RNA-mediated transcriptional silencing. Mol Cell. 2013 Jan 24;49(2):298–309.
74. Ausin I, Greenberg MV, Li CF, Jacobsen SE. The splicing factor SR45 affects the RNA-directed DNA methylation pathway in *Arabidopsis*. Epigenetics. 2012 Jan 1;7(1):29–33.
75. Dou K, Huang CF, Ma ZY, Zhang CJ, Zhou JX, Huang HW, et al. The PRP6-like splicing factor STA1 is involved in RNA-directed DNA methylation by facilitating the production of Pol V-dependent scaffold RNAs. Nucleic Acids Res. 2013 Oct;41(18):8489–502.
76. Huang CF, Miki D, Tang K, Zhou HR, Zheng Z, Chen W, et al. A Pre-mRNA-splicing factor is required for RNA-directed DNA methylation in *Arabidopsis*. PLoS Genet. 2013;9(9):e1003779.
77. Zhang CJ, Zhou JX, Liu J, Ma ZY, Zhang SW, Dou K, et al. The splicing machinery promotes RNA-directed DNA methylation and transcriptional silencing in *Arabidopsis*. EMBO J. 2013 April 17;32(8):1128–40.
78. Bayne EH, Portoso M, Kagansky A, Kos-Braun IC, Urano T, Ekwall K, et al. Splicing factors facilitate RNAi-directed silencing in fission yeast. Science. 2008 Oct 24;322(5901):602–6.
79. Zofall M, Yamanaka S, Reyes-Turcu FE, Zhang K, Rubin C, Grewal SI. RNA elimination machinery targeting meiotic mRNAs promotes facultative heterochromatin formation. Science. 2012 Jan 6;335(6064):96–100.

Chapter 2
The Role of DNA Methylation in Transposable Element Silencing and Genomic Imprinting

Yoko Ikeda and Taisuke Nishimura

Introduction

DNA methylation is the covalent modification of DNA nucleotides that may act to change chromatin structure and gene regulation. DNA methylation is a stable but reversible modification; thus, many organisms utilize this type of genomic modification for transcriptional regulation. In eukaryotes, methylation of the fifth carbon of cytosine residues is precisely maintained during mitosis and, therefore, is proposed to be involved in epigenetic regulation. Extensive studies of cytosine methylation have uncovered mechanisms for the establishment, maintenance and elimination of this modification in eukaryotes [1]. In both mammals and plants, cytosine methylation plays important roles in the silencing of TEs and genomic imprinting. In the past decade, molecular genetics and genomics approaches using mutants of *Arabidopsis* (*Arabidopsis thaliana*) have revealed many new aspects of TE silencing and genomic imprinting regulated by DNA methylation in plants. In this chapter, we describe the current models derived from these studies.

T. Nishimura (✉)
PRESTO, Japanese Science and Technology Agency, Tokyo, Japan
Bioscience and Biotechnology Center, Nagoya University, Nagoya, Japan
e-mail: tnishi@vos.nagaokaut.ac.jp

Nagaoka University of Technology, Nagaoka, Japan

Y. Ikeda
Institute of Plant Science and Resources, Okayama University, Okayama, Japan

O. Pontes, H. Jin (eds.), *Nuclear Functions in Plant Transcription, Signaling and Development,* DOI 10.1007/978-1-4939-2386-1_2

Transposable Element and DNA Methylation

TEs in Plants

Transposable elements (TEs) are mobile DNA fragments first discovered in the maize (*Zea mays*) genome by Barbara McClintock [2] that are now known to be present in nearly all eukaryotes. Some TEs contain genes whose products facilitate autonomous movement in combination with factors encoded by the host genome, whereas nonautonomous TEs lack movement genes and rely on factors encoded by other TEs. According to their mode of mobility, TEs can be divided into two main classes [3]: class I TEs, called retrotransposons, use a "copy-and-paste" mechanism in which the element replicates via reverse transcription of messenger RNAs (mRNAs), and the duplicates integrate into other chromosomal locations and class II TEs, also called DNA transposons, transpose by a "cut-and-paste" mechanism in which the elements are excised from the chromosome and integrate into new locations. A subclass of DNA transposons, the *helitrons*, have a different mechanism to transpose, called the "rolling-circle" mechanism, in which a single strand of the element is nicked and invades another chromosomal location. Plant species have all classes/subclasses of TEs, although the most common ones in plant genomes are class I TEs.

High TE content and genome size are strongly correlated. In the relatively small genome (120 Mb) of *Arabidopsis*, TE fragments constitute approximately 17% of the genome [4]. In contrast, maize, with a 20-fold larger genome, 2.3 Gb, has TE fragments represent 85% of the genome [5]. TE activities can potentially change the expression and function of genes near their insertion site and can also cause chromosome breakage leading to genome rearrangement. Such changes in the genome may generate genetic variations that contribute to the adaptation and evolution of host plants [6]. Furthermore, plant genomes have developed mechanisms to silence TE activities to prevent the disruption of normal regulation and gene function that would render deleterious effects on plant growth. Mutant studies of *Arabidopsis* in which TE activities are completely silent, have demonstrated that several layers of epigenetic regulation play a major role in repressing TE activities [7]. Notably, the role of one of the layers in TE silencing, DNA methylation, has been investigated extensively. Because DNA methylation can be copied into a newly synthesized strand during DNA replication, the silenced status of TEs is inherited stably. In the remainder of this section, we describe how DNA methylation contributes to the regulation of TE activities.

DNA Methylation in TE Silencing

The involvement of DNA methylation in TE silencing was implicated in studies of maize class II TEs, *Mu, Ac* and *Spm* in the 1980s [8–11]. These investigators com-

Fig. 2.1 Variegated phenotype of a *met1*-derived epiRIL where mobilization of *EVD*/*COPIA93* is active. *EVD*/*COPIA93* transposed into *VAR2* gene somatically, resulting in sectored leaf variegation [21]. *met* dna methyltransferase, *epiRIL* epigenetic recombinant inbred lines, *EVD* Évadé, *VAR* YELLOW VARIEGATE. (Photograph was taken and provided by Dr. Olivier Mathieu)

pared the DNA methylation level of silent TEs with that of active TEs and found that DNAs of the former were hypermethylated and those of the latter were hypomethylated. Consistent with the first observations in maize, genome-wide analyses of DNA methylation in the past decade have demonstrated that the CG and non-CG sequence context (CHG and CHH, H is C, A, T) of silent TEs, both of class I and class II, are hypermethylated in *Arabidopsis* [12–15]. Our knowledge on the roles of DNA methylation in TE silencing has been augmented by the use of *Arabidopsis* mutants defective in DNA methylation, in which some silent TEs were transposed within their genome (Fig. 2.1). The first reported mobile TE was a member of the CACTA family of class II TE found in an inbred *ddm1* (decrease in *dna methylation1*), a mutant of a chromatin remodelling adenosine triphosphatase (ATPase) involved in both CG and non-CG methylation of heterochromatic regions [16, 17]. The CACTA family of TEs is not the only family able to transpose since some other class II TEs of the *Mutator* family are also mobile in *ddm1* inbred lines [18–20]. In more recent studies, transposition of class I TEs, the most abundant TE family in the *Arabidopsis* genome, was observed in inbred progeny of *ddm1* and *methyltransferase* 1 (*met1*), a mutant of a methyltransferase essential for the maintenance of CG methylation [20, 21].

Most TEs contain and are probably silenced by both CG and non-CG methylation [12, 14]. Indeed, expression of *ATGP3*, a class I TE in the *gypsy* family, does not occur in single mutants of *met1* or *chromomethylase3* (*cmt3*), a mutant of a methyltransferase for non-CG methylation maintenance, but does occur in the *ddm1* and *met1 cmt3* double mutants [20]. This finding suggests a redundant function of CG and non-CG methylation in the transcriptional silencing of the TE. For

another example, *EVD* (*Évadé*)/*ATCOPIA93*, a class I TE in the *copia* family, can be activated transcriptionally in the *met1* single mutant, but the double mutant of *met1* with *drm2* (*domains rearranged metyltransferase2*), a mutant of another non-CG methyltransferase, shows a synergic increase in *EVD/ATCOPIA93* transcription [21]. Interestingly, the *met1 cmt3* double mutant did not show such a synergic effect in *EVD/ATCOPIA93* transcription. CMT3 and DRM2 function as DNA methyltransferases downstream of histone methylation and siRNA activities, respectively [1]. Thus, *ATGP3* and *EVD/ATCOPIA93* are transcriptionally inactivated by CG and non-CG methylation, although the modes for non-CG methylation are different.

Selective Regulation for Mobilization of TE

In the *met1* mutant, CG methylation of TEs is eliminated resulting in their transcriptional activation, but most TEs are not transposed [12, 14]. The majority of TEs would have some mutations in their protein coding sequences allowing for post-transcriptional regulation in transposition, but some TE sequences seem to encode intact proteins. This indicates that there are other processes repressing transposition in addition to transcriptional silencing. Transposition of *EVD/ATCOPIA93* can be observed in the *met1* mutant, but not in the first generation of homozygous plants [21]. *EVD/ATCOPIA93* transposition is detected beginning with the second generation and increases as the generations progress. Such a progressive mobilization was also observed in *met1*-derived epigenetic recombinant inbred lines (epiRILs). The *met1* epiRILs are genetically identical to wild type but display mosaic DNA methylation patterns in their genomes that were generated by self-inbreeding wild-type *MET1*+/+plants selected in the F2 generation after a cross between wild-type and *met1* plants [22]. *EVD/ATCOPIA93* began to mobilize beginning with the F4 generation in an epiRIL in which their transcription is active [21]. In contrast, no mobilization of *EVD/ATCOPIA93* was observed even beyond the F9 generation in an epiRIL in which DNA methylation of this TE was retained to maintain transcriptional silencing. Mobilization requires transcription, but mobilization activity is not associated with the transcription level as described in the next paragraph. In addition to *EVD/ATCOPIA93*, class II TEs of the CACTA family can be mobilized progressively in some epiRILs, although they cannot be progressively mobilized in the parental *met1* single mutant [22]. Similarly, in inbred *ddm1* mutants, class I TEs in the *gypsy* and *copia* families including *EVD/ATCOPIA93* are transposed, and the transposition of each TE occurs stochastically and independently in the inbred lines [20]. Taken together, these observations suggest that there are additional and specific repressing processes for the transposition of each TE in contrast to their transcriptional silencing maintained ubiquitously by DNA methylation and that inbreeding releases this repression somehow.

Although the repression mechanisms are still unknown, inbreeding can be skipped for the transposition of some combinations of mutants [21]. In double mutants of *met1* and *nuclear RNA polymerase d2a* (*nrpd2a*), a mutation in the common

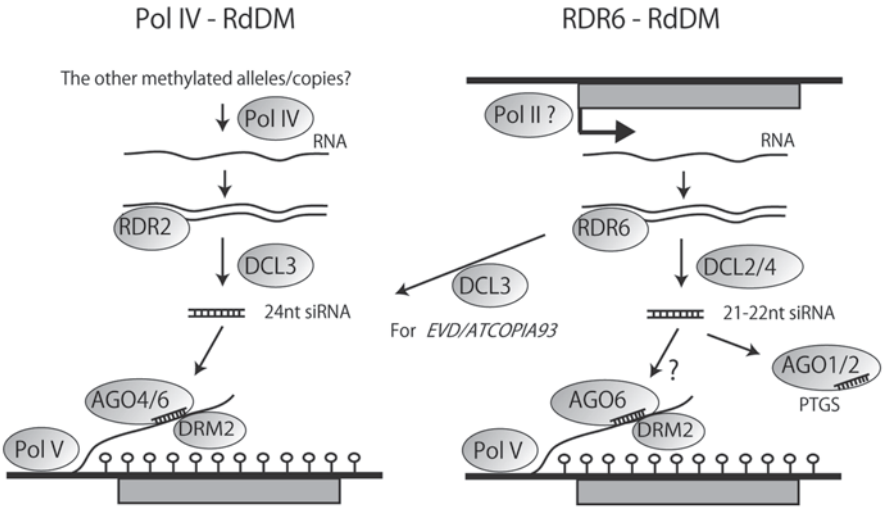

Fig. 2.2 Two pathways for the remethylation of active TEs in demethylated loci when methylation activity is restored. A 24-nt siRNA produced from the PolIV–RDR2–DCL3 pathway and a 21-nt siRNA from the RDR6–DCL2/4 pathway act on DNA methylation. In conventional RNA-directed DNA methylation (PolIV-RdDM; *left*), a 24-nt siRNA is incorporated with AGO4/6 and recruits a DRM2 *de novo* methyltransferase to a target guided by a PolV-transcribed RNA. In RDR6–RdDM, a 21–22-nt siRNA probably induces methylation of the target in cooperation with AGO6 and PolV [27] in a similar way to PolIV–RdDM. The 24-nt siRNA may be from other methylated alleles/loci homologous to the target since PolIV transcribes RNA from methylated loci. In contrast, the 21-nt siRNA would be derived from PolII-dependent RNA from a demethylated target. In *EVD/ATCOPIA93* remethylation, a 24-nt siRNA can be produced from RDR6-dependent dsRNA [28]. *TE* transposable element, *nt* nucleotide, *siRNA* small interfering RNA, *Pol* RNA polymerase, *RDR* RNA-DEPENDENT RNA POLYMERASE, *DCL* DICER-LIKE PROTEIN, *RdDM* RNA-dependent DNA methylation, *AGO* ARGONAUTE, *DRM* DOMAIN REARRANGED METHYLTRANSFERASE, *dsRNA* double-stranded RNA

subunit of plant-specific RNA polymerases, PolIV and PolV, involved in RdDM [1], as the details are described again in the next subsection (Fig. 2.2), transposition of *EVD/ATCOPIA93* was observed as well as a synergic increase in transcription. In contrast, transcription increased but TEs were not transposed in the *met1 drm2* double mutant, although DRM2 acts with PolV and siRNA produced by PolIV in RdDM. This result suggests an unknown RdDM-independent mechanism for repression of transposition. Another mutant combination showing a synergic effect for *EVD/ATCOPIA93* transposition is *met1* and *kryptonite* (*kyp*)/*suvh4*, a histone methyltransferase mutant; although a synergic increase in transcription was not observed, unlike the *met1 nrpd2a* double mutant. KYP/SUVH4 histone methyltransferase acts in the non-CG methylation pathway via CMT3 methylation activity [1]. In the *met1 cmt3* double mutant, however, there was no transposition of *EVD/ATCOPIA93*. These results suggest that the *post-transcriptional* activity of *EVD/ATCOPIA93* transposition is regulated by unknown mechanisms involving PolIV and/or PolV, and KYP/SUVH4 methyltransferase, but probably not via DNA methylation as previously proposed [21].

Resilencing Dynamics for Active TE

Transcription of subsets of TEs activated in mutants deficient in DNA methylation is repressed again when maintenance of DNA methylation is restored. A study using *ddm1*-derived epiRILs and mutants of RdDM components revealed that transcriptional silencing is achieved by re-establishment of DNA methylation by the action of RdDM [23]. In RdDM, 24-nt siRNAs produced by DICER-LIKE PROTEIN3 (DCL3) from double-stand (dsRNA) originated by Pol IV and RNA-dependent RNA POLYMERASE2 (RDR2), recruit Pol V, ARGONAUTE4 (AGO4), and in some instances AGO6, to guide DRM2 in the establishment of de novo DNA methylation (Fig. 2.2) [1]. In addition, components of post-transcriptional gene silencing (PTGS) that act in antiviral and antibacterial defences are suggested to function in DNA methylation to silence other TEs. In contrast to RdDM, a 21–22-nt siRNA produced by RDR6, DCL2 and DCL4 is associated with AGO1 and AGO2 to degrade target RNAs in PTGS [24]. A 21–22-nt siRNA produced by PTGS components, not a 24-nt siRNA in conventional RdDM, is required for transcriptional silencing in the *helitron* and *copia* elements [25, 26]. Genome-wide analyses for siRNAs in the *rdr6* and *ddm1* mutants revealed that RDR6-dependent 21–22-nt siRNAs triggered the resilencing of subsets of TEs activated in the *ddm1* mutant [27]. Thus, resilencing of TEs can be accomplished by two kinds of siRNAs produced from dsRNAs synthesized with PolIV–RDR2 and those with RDR6 (Fig. 2.2). Once DNA methylation is established, TE silencing is maintained by MET1 for CG methylation and the conventional RdDM for the others.

In contrast, other TEs cannot be remethylated even after the restoration of DNA methylation activity, and, in some cases, TEs like *EVD/ATCOPIA93* begin to transpose within the genome [21, 28]. Host plants have a mechanism to repress the activity of mobile TEs preventing an excess invasion of the genome. In *met1*- and *ddm1*-derived epiRILs, *EVD/ATCOPIA93* copy number increases with proceeding generations, but this increase stops at approximately 40 copies per genome [28]. This peak copy number seems to be fixed because several lines of both *met1* and *ddm1*-derived epiRILs attain the same copy number, whereas the number of generations required to reach the peak vary among the lines. During the process, there are several steps for silencing *EVD/ATCOPIA93* [28]. First, PTGS is activated against *EVD/ATCOPIA93* expression in which 21–22-nt siRNAs for the coding region are produced, but the RNA of *EVD/ATCOPIA93* is resistant to PTGS, thereby keeping their increased copy number. When approximately 40 copies are reached, 24-nt siRNA produced by DCL3 from RDR6-dependent dsRNA induces RdDM in the coding region (Fig. 2.2), and 24-nt siRNA and DNA methylation are spread into the promoter region. Eventually, *EVD/ATCOPIA93* is silenced again at around 40 copies [28].

Reinforcement of TE Silencing during Gametogenesis

DNA methylation is important for transposon silencing; however, DNA methylation status is not always stably maintained during the life cycle. In mammals, where

DNA methylation is primarily in the CG context, methylation is erased and re-established in the germ line. Thus, DNA methylation is reprogrammed for each generation [29]. In contrast, the reprogramming of DNA methylation in sperm and egg cells has not been reported in plants, except that only CHH methylation of transposons was partly erased in sperm cells [30, 31]. Genome-wide DNA demethylation was observed in a limited plant germ cells that are not of embryonic origin. One such example is the vegetative nucleus of pollen, which is not fertilized but controls pollen germination and pollen tube elongation [32]. DNA demethylation in the vegetative nucleus is accompanied by siRNA production from transcriptionally active TEs [30, 31, 33]. The siRNA originating from the vegetative nucleus is proposed to be transported into the sperm cells where it is involved in the silencing of TEs by the *de novo* RdDM pathway [30, 31, 33]. DNA demethylation also occurs in the central cell, which is the origin of endosperm, in the female gametophyte [31]. Although the DNA methylation pattern of plant egg cells has not been investigated, siRNAs from demethylated TEs in the central cell are predicted to reinforce DNA methylation of TEs in the egg cell by a mode of action similar to the sperm cell. Indeed, the small RNAs expressed by the central cell can affect transcription of the target in the egg cell [31]. Mechanisms of demethylation in the plant companion cell in the female gamete, the central cell, are common to those for regulating genomic imprinting as described in the next section. Thus, TE silencing in plants can be reinforced at each generation.

Genomic Imprinting and DNA Methylation

Molecular Mechanism of Genomic Imprinting

Genomic imprinting occurs in mammals and plants and results from differential gene expression caused by differences in the epigenetic status of parental genomes, including asymmetric DNA methylation. In plants, genomic imprinting primarily occurs in the endosperm and controls seed development. The endosperm is the embryo-nourishing tissue whose genetic information is not directly inherited by the offspring. Many endosperm-imprinted genes have been identified in plants (Table 2.1). Plant genomic imprinting mechanisms have been mostly elucidated by genetic approaches using *Arabidopsis* mutants. Endosperm imprinting is regulated by several epigenetic pathways involving DNA methylation and histone modifications. *FWA* is known as a maternally imprinted gene regulated by DNA methylation [34]. The 5′ region of *FWA* contains two tandem repeats, and this region is highly methylated in adult tissues. In contrast, the maternal *FWA* allele in the endosperm is hypomethylated and expressed (Fig. 2.3) [35]. This asymmetric DNA methylation status between the maternal and paternal genomes is established in the central cells before fertilization. In the central cell, genome-wide DNA demethylation was proposed to occur by both passive and active mechanisms [36, 37]. In somatic tissues, CG methylation was maintained by the MET1 methyltransferase, whereas

Table 2.1 Examples of the endosperm-imprinted genes

Gene	Species	Protein family	Expression	Regulation	Reference for imprinting regulation
MEA	*Arabidopsis*	PcG protein E(z)	Maternal	DNAme, H3K27me3	[39, 40, 51]
Mez1	Maize	PcG protein E(z)	Maternal	DNAme, H3K27me3	[79, 80]
FIS2	*Arabidopsis*	PcG protein Su(z)12	Maternal	DNAme	[38]
Fie1	Maize	PcG protein Esc	Maternal	DNAme, H3K27me3	[80, 81]
OsFIE1	Rice	PcG protein Esc	Maternal	DNAme	[82]
FWA	*Arabidopsis*	HD-ZIP	Maternal	DNAme	[34]
PHE1	*Arabidopsis*	Type I MADS box	Paternal	DNAme, H3K27me3	[48]
Mee1	Maize	Hypothetical protein	Maternal in the endosperm and in the early embryo	DNAme	[56]
SDC	*Arabidopsis*	F-box	Maternal	RdDM	[62]
MOP9.5	*Arabidopsis*	Phosphatidylinositol kinase	Maternal	RdDM	[62]

during the development of reproductive cells, *MET1* gene expression was abolished in the central cell depending on the complex consisting of RETINOBLASTOMA RELATED 1 (RBR1) and MULTICOPY SUPPRESSOR OF IRA1 (MSI) [38]. This finding suggested that passive demethylation happened in the central cell. Furthermore, DEMETER (DME), a cytosine demethylase, is expressed and contributes to active DNA demethylation in the central cell [39]. In *dme* endosperm, the DNA methylation level is higher than in wild-type endosperm [36, 37, 40]. Further analysis using single-nucleotide polymorphisms (SNPs) between different *Arabidopsis* accessions revealed the precise DNA methylation pattern of maternal and paternal genomes in the wild-type and *dme* endosperms [31]. DME seems to target TEs that are small, AT-rich and enriched in euchromatic regions. *DME* encodes a DNA glycosylase protein that is involved in the base excision repair (BER) pathway. REPRESSOR OF SILENCING1 (ROS1), DEMETER LIKE2 (DML2) and DML3 are also cytosine demethylases that are in the same family as DME but are not expressed during the reproductive phase, and their triple mutant does not show any phenotype affecting genomic imprinting. Therefore, DNA demethylation in the central cell seems to be catalyzed by DME. In *Arabidopsis*, AtLIG1 (*Arabidopsis thaliana* DNA LIGASE1) has been also reported to affect imprinting involved in the BER pathway [41]. Recently, DNA 3′ phosphatase AtZDP (*A.*

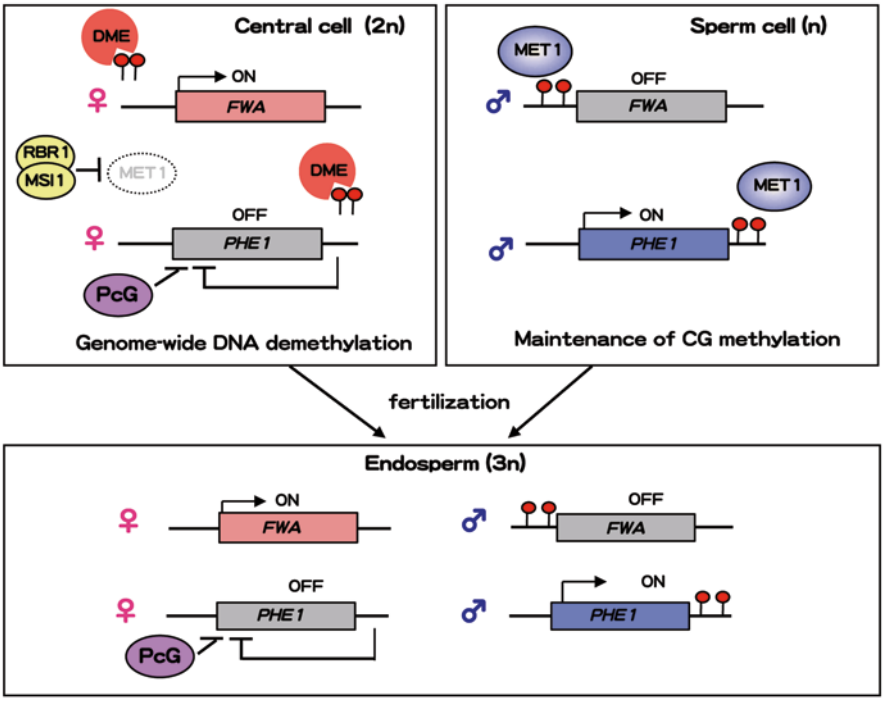

Fig. 2.3 Control of maternally imprinted *FWA* and paternally imprinted *PHE1* in *Arabidopsis*. In the central cell, genome-wide demethylation occurs by *MET1* inactivation and *DME* activation. Loss of DNA methylation in the *FWA* promoter region induces the transcriptional activation of *FWA*, whereas that in the 3′ region of *PHE1* causes PRC2-dependent H3K27me3 resulting in the *PHE1* silencing. In the sperm cell, *MET1* activity maintains CG methylation, resulting in the *FWA* silencing and the *PHE1* activation. The expression status of each allele is maintained in the endosperm after the fertilization. *PHE* PHERES, *DME* DEMETER, *PRC* polycomb repressive complex, *H3K27me3* tri-methylation at the 27th lysine of histone H3, *CG* cytosine-guanine sequence context

thaliana ZINC FINGER DNA 3′ PHOSPHOESTERASE) and DNA repair protein XRCC1 (X-RAY CROSS-COMPLEMENTING GROUP PROTEIN1) were shown to affect DNA demethylation [42, 43]. These proteins bind to ROS1 in vitro and act downstream of ROS1 involved in BER; however, it is not yet known whether they also act with DME.

MEA (*MEDEA*) is the first identified imprinted gene in *Arabidopsis* and is well characterized as a maternally imprinted gene (Table 2.1) [44, 45]. Endosperm of the *mea* mutant shows an over proliferation phenotype. Maternal *MEA* expression is controlled by DME and MET1 antagonistically [46]. However, a 200-bp minimum *MEA* imprinting control region (ICR) was recently identified, and this region was independent of DNA methylation control [47]. This result suggests an unknown regulatory mechanism for the imprinted expression of the minimum *MEA* ICR *cis*-regulatory region.

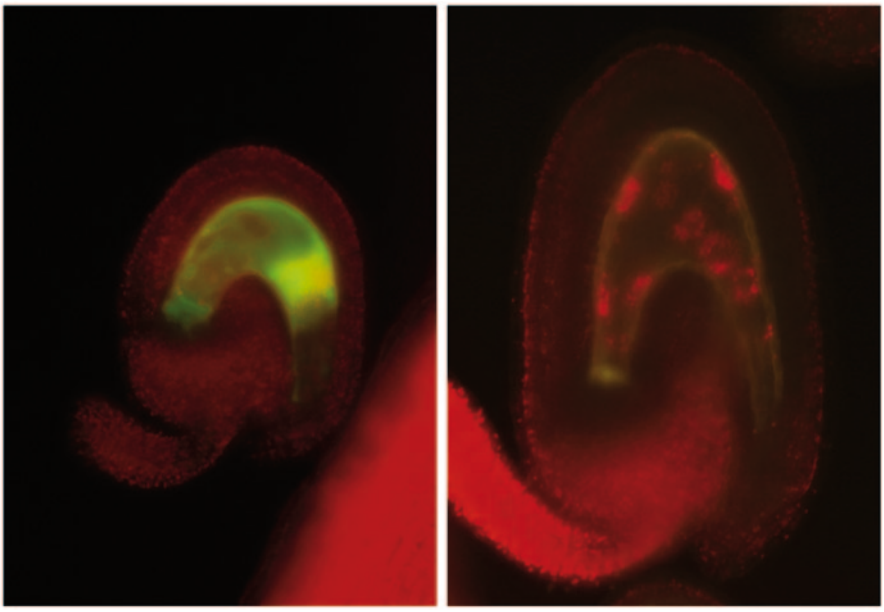

Fig. 2.4 FWA–GFP fluorescence in wild-type (*left*) and *ssrp1–3* (*right*) *Arabidopsis* ovules observed 5 days after emasculation. Fluorescence was detected in the nucleus of the wild-type central cell. FWA–GFP in the *ssrp1–3* ovule was not activated and dividing nuclei were observed without fertilization in the *ssrp1–3* ovule. Accompanying autonomous central cell proliferation, *ssrp1–3* ovules become slightly expanded. *GFP* green fluorescent protein, *ssrp* structure-specific recognition protein

Not only DNA methylation, but also histone methylation (tri-methylation at the 27th lysine of histone H3; H3K27me3) also contributes to genomic imprinting. A paternally expressed imprinted gene, *PHE1* (*PHERES1*), has a DNA-methylated region located in the 3′ region of the gene, and hypomethylation of this region is necessary for silencing via H3K27me3 by polycomb repressive complex 2 (PRC2) [48]. The hypomethylated status of the 3′ region of maternal *PHE1* makes it possible for PRC2 to access the gene and be silenced by H3K27me3 (Fig. 2.3). Genome-wide H3K27me3 profiling revealed many H3K27me3 targets like *PHE1* in the endosperm, and hypomethylation of the target region seems to be a trigger for H3K27me3 by PRC2 [49]. Interestingly, components of PRC2, FERTILIZATION-INDEPENDENT SEED2 (FIS2) and MEA are maternally imprinted, and the paternal *MEA* allele is also silenced by PRC2 in the endosperm [40, 50, 51].

Recently, additional evidence for the controlling mechanism of genomic imprinting was reported; STRUCTURE-SPECIFIC RECOGNITION PROTEIN1 (*SSRP1*) was identified as being required for the activation of maternally imprinted genes and the mutant showed an autonomous endosperm development phenotype (Fig. 2.4) [52]. In the *ssrp1* mutant, the DNA methylation level in the endosperm is higher than in the wild-type plant. *SSRP1* encodes a component of facilitating chromatin-mediated transcription (FACT), an H2A/H2B histone chaperone. FACT

controls chromatin structure during transcription and DNA replication in mammals and yeast. SSRP1 may control imprinted expression with DME by a chromatin-based mechanism. Histone H1 was identified by yeast two-hybrid screening to be a protein that interacts with DME [53]. The triple mutant of *Arabidopsis* histone H1 homologs exhibited deficiencies in DNA demethylation and expression of maternally imprinted genes in the endosperm like *ssrp1* and *dme*. Histone H1 binds to the linker region between nucleosomes and is related to chromatin structure through nucleosome compaction. These reports suggest that DNA demethylation by DME cytosine demethylase are affected by chromatin structure. Histone acetyltransferase of INCREASED DNA METHYLATION1 (IDM1) and the small RNA-binding protein REPRESSOR OF SILENCING3 (ROS3) were reported to affect DNA demethylation upstream of another demethylase, ROS1 [54, 55]. Thus, histone acetylation and small RNA-guiding mechanisms might also be related to DNA demethylation by DME.

Resetting Mechanism in the Embryo by RdDM

Unlike mammals, genomic imprinting in plants is not observed in the fully developed embryo and adult tissues. The maize *MEE1* (*MATERNALLY EXPRESSED IN EMBRYO1*) gene shows maternally imprinted expression in the early embryo and in the endosperm [56]. The maternal *MEE1* allele is hypomethylated in the early embryo and endosperm, but remethylation occurs during embryo development, and expression disappears in the mature embryo. Imprinted genes in rice embryos were reported [57]; however, imprinted expression of these genes was not detected in adult plants. In *Arabidopsis*, genome-wide evaluation of the parental genome's contribution to the early embryo (~32 cells) revealed some parental-origin specifically expressed genes, but these genes are not among the allele-specific expressed genes in the fully developed embryo [58–60]. As mentioned in the former section, Calarco et al. showed a decrease in the CHH methylation level of TEs in sperm cells and suggested that restoration of DNA methylation in the embryo depended on siRNA [30]. Indeed, the expression of DRM2 *de novo* methyltransferase was detected in the egg cell and the early embryo, and recovery of CHH methylation was observed depending on the stage of embryo development [61]. Consistent with maize *MEE1*, the resetting mechanism of DNA methylation exists in the early embryo, thus imprinted expression in the plant's adult phase may not be found.

Moreover, Vu et al. found several maternally imprinted genes controlled by RdDM. In the mutants involved in the RdDM pathway, the *drm1 drm2* double mutant and *nrpd2a*, maternal imprinted genes, *SDC* (*SUPPRESSOR OF DRM1 DRM2 CMT3*) and *MOP9.5* were mis-expressed by both paternal and maternal alleles [62]. Paternal *SDC* and *MOP9.5* may also be silenced by RdDM depending on the siRNA produced in the vegetative cells and endosperm.

DNA demethylation caused siRNA production in the central cell and in the endosperm. Indeed, maternal 24-nt siRNA accumulated in *Arabidopsis* endosperm [63]. Furthermore, in crosses between different ploidy levels of *Arabidopsis*, a maternal

siRNA level was altered and affected expression of specific AGL *(AGAMOUS-LIKE)* genes that are targets of interploidy crosses [64]. A model was proposed in which siRNA produced in the central cell and endosperm are transported to the egg and embryo, respectively. This proposal provides an explanation for why genomic imprinting is observed in the endosperm but not in the embryo.

Divergence of Imprinted Genes and Evolutionary Aspects

Genome-wide information about the transcriptome and methylome of the endosperm of several plant species is updated frequently. As summarized in the reviews [65, 66], many candidates for imprinted genes have been identified in *Arabidopsis*, rice and maize [37, 57, 58, 67–69]. To identify maternally and paternally expressed alleles, different accessions or inbred lines are used to detect SNPs between the alleles. In *Arabidopsis*, many nuclear proteins, such as transcription factors and chromatin-related proteins, and hormone signal transduction proteins were imprinted. Interestingly, components of PRC2 are commonly maternally imprinted in some plant species. These maternally imprinted genes may have positively evolved to control endosperm development.

Very few imprinted genes are common in *Arabidopsis*, rice and maize [57, 67], implying rapid evolution of imprinted genes in the plant genome. Positive selection of the *MEA* gene [70–72] and a comparison of the ratio of non-synonymous to synonymous substitutions between different species show evidence of rapid change and positive selection of imprinted genes [67, 68]. Moreover, Wolff et al. also suggested that gene duplication affects the evolution of imprinted genes [68]. For the type I MADS-box gene family, including the paternally imprinted gene *PHE1*, gene duplication and imprinted expression have a positive relationship [73]. Not only gene duplication but also TE insertion and subsequent silencing were proposed to be the driving force for imprinted expression. This theory can apply for imprinted genes regulated by DNA methylation. A theory for the evolution of imprinted genes is still controversial, but genome-wide information makes it possible to compare many imprinted genes among many species. As a result, a comprehensive view about the evolution of imprinted genes may be forthcoming.

Perspective

Genetic and genome-wide analyses using *Arabidopsis* mutants provide many new insights into the role of DNA methylation in TE silencing and genomic imprinting. However, there remain many unsolved issues in these research fields. DNA methylation acts on silencing TE at the transcriptional level, but other layers mediated by unknown mechanisms act to repress transposition. The mechanism for DNA remethylation of activated TEs has just begun to be understood. Uncovering

activation and repression systems for TEs by the host genome would lead to an understanding of the function of TEs in the genome from an evolutionary point of view. We still cannot elucidate the clear biological and evolutionary significance of genomic imprinting. Dissecting the function of newly identified imprinted genes in seed development would provide some answers for this unresolved issue.

We have discussed about the action of DNA methylation in TE silencing and genomic imprinting. However the functions of DNA methylation in plant development and environmental responses have not been identified so much compared with those in TE silencing and genomic imprinting. In mammals, DNA methylation plays pivotal roles in several developmental steps and responses to the environment [74]. DNA methylation in plants may be dynamically regulated depending on the developmental stages, specific tissues, circumstances, and so on. Plants deficient in DNA methylation in their genomes, like epiRILs, show quantitative phenotypes for development, indicating some unknown functions of DNA methylation during plant development [22, 75]. In addition, several reports describe the role of DNA methylation in several steps of plant development [76, 77]. Quantitative and comprehensive analyses with high-resolution patterns of genome-wide DNA methylation enable DNA methylation dynamics during biotic stress to be visible [78]. Thus, rapid technological progress in genome-wide analyses of DNA methylation will reveal more roles for DNA methylation than we currently know. Molecular action of DNA methylation revealed from studies of TE silencing and genomic imprinting will likely be applied to understanding the function of DNA methylation in plant development and environmental responses.

Acknowledgments We thank Dr. Olivier Mathieu (Centre National de la Recherche Scientifique, France) for his kind permission to use the photograph shown in Fig. 2.1 and Dr. Tatsuo Kanno (The National Institute of Agrobiological Sciences, Japan) for critical reading of the manuscript. Preparation of this manuscript was supported by the Programme to Disseminate Tenure Tracking System, MEXT, Japan (for YI) and Japanese Science and Technology Agency, PRESTO (for TN).

References

1. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet. 2010;11(3):204–20.
2. McClintock B. Chromosome organization and genic expression. Cold Spring Harbor Symp Quant Biol. 1951;16:13–47.
3. Wicker T, Sabot F, Hua-Van A, Bennetzen JL, Capy P, Chalhoub B, et al. A unified classification system for eukaryotic transposable elements. Nat Rev Genet. 2007;8(12):973–82.
4. Buisine N, Quesneville H, Colot V. Improved detection and annotation of transposable elements in sequenced genomes using multiple reference sequence sets. Genomics. 2008;91(5):467–75.
5. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, et al. The B73 maize genome: complexity, diversity, and dynamics. Science. 2009;326(5956):1112–5.
6. Lisch D. How important are transposons for plant evolution? Nat Rev Genet. 2013;14(1):49–61.
7. Rigal M, Mathieu O. A “mille-feuille” of silencing: epigenetic control of transposable elements. Biochim Biophys Acta. 2011;1809(8):452–8.

8. Banks JA, Masson P, Fedoroff N. Molecular mechanisms in the developmental regulation of the maize *Suppressor-mutator* transposable element. Genes Dev. 1988;2(11):1364–80.
9. Chandler VL, Walbot V. DNA Modification of a maize transposable element correlates with loss of activity. Proc Natl Acad Sci U S A. 1986;83(6):1767–71.
10. Chomet PS, Wessler S, Dellaporta SL. Inactivation of the maize transposable element *Activator* (*Ac*) is associated with its DNA modification. EMBO J. 1987;6(2):295–302.
11. Schwartz D, Dennis E. Transposase activity of the *Ac* controlling element in maize is regulated by its degree of methylation. Mol Gen Genet. 1986;205(3):476–82.
12. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, et al. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. Nature. 2008;452(7184):215–9.
13. Lippman Z, Martienssen R. The role of RNA interference in heterochromatic silencing. Nature. 2004;431(7006):364–70.
14. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, et al. Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. Cell. 2008;133(3):523–36.
15. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. Cell. 2006;126(6):1189–201.
16. Miura A, Yonebayashi S, Watanabe K, Toyama T, Shimada H, Kakutani T. Mobilization of transposons by a mutation abolishing full DNA methylation in *Arabidopsis*. Nature. 2001;411(6834):212–4.
17. Zemach A, Kim MY, Hsieh PH, Coleman-Derr D, Eshed-Williams L, Thao K, et al. The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. Cell. 2013;153(1):193–205.
18. Fu Y, Kawabe A, Etcheverry M, Ito T, Toyoda A, Fujiyama A, et al. Mobilization of a plant transposon by expression of the transposon-encoded anti-silencing factor. EMBO J. 2013;32(17):2407–17.
19. Singer T, Yordan C, Martienssen RA. Robertson's *Mutator* transposons in *A. thaliana* are regulated by the chromatin-remodeling gene *Decrease in DNA Methylation* (*DDM1*). Genes Dev. 2001;15(5):591–602.
20. Tsukahara S, Kobayashi A, Kawabe A, Mathieu O, Miura A, Kakutani T. Bursts of retrotransposition reproduced in *Arabidopsis*. Nature. 2009;461(7262):423–6.
21. Mirouze M, Reinders J, Bucher E, Nishimura T, Schneeberger K, Ossowski S, et al. Selective epigenetic control of retrotransposition in *Arabidopsis*. Nature. 2009;461(7262):427–30.
22. Reinders J, Wulff BB, Mirouze M, Mari-Ordonez A, Dapp M, Rozhon W, et al. Compromised stability of DNA methylation and transposon immobilization in mosaic *Arabidopsis* epigenomes. Genes Dev. 2009;23(8):939–50.
23. Teixeira FK, Heredia F, Sarazin A, Roudier F, Boccara M, Ciaudo C, et al. A role for RNAi in the selective correction of DNA methylation defects. Science. 2009;323(5921):1600–4.
24. Pumplin N, Voinnet O. RNA silencing suppression by plant pathogens: defence, counter-defence and counter-counter-defence. Nat Rev Genet. 2013;11(11):745–60.
25. Garcia D, Garcia S, Pontier D, Marchais A, Renou JP, Lagrange T, et al. Ago hook and RNA helicase motifs underpin dual roles for SDE3 in antiviral defense and silencing of nonconserved intergenic regions. Mol Cell. 2012;48(1):109–20.
26. Pontier D, Picart C, Roudier F, Garcia D, Lahmy S, Azevedo J, et al. NERD, a plant-specific GW protein, defines an additional RNAi-dependent chromatin-based pathway in *Arabidopsis*. Mol Cell. 2012;48(1):121–32.
27. Nuthikattu S, McCue AD, Panda K, Fultz D, DeFraia C, Thomas EN, et al. The initiation of epigenetic silencing of active transposable elements is triggered by RDR6 and 21–22 nucleotide small interfering RNAs. Plant Physiol. 2013;162(1):116–31.
28. Mari-Ordonez A, Marchais A, Etcheverry M, Martin A, Colot V, Voinnet O. Reconstructing de novo silencing of an active plant retrotransposon. Nat Genet. 2013;45(9):1029–39.

29. Feng S, Jacobsen SE, Reik W. Epigenetic reprogramming in plant and animal development. Science. 2010;330(6004):622–7.
30. Calarco JP, Borges F, Donoghue MT, Van Ex F, Jullien PE, Lopes T, et al. Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. Cell. 2012;151(1):194–205.
31. Ibarra CA, Feng X, Schoft VK, Hsieh TF, Uzawa R, Rodrigues JA, et al. Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. Science. 2012;337(6100):1360–4.
32. Schoft VK, Chumak N, Choi Y, Hannon M, Garcia-Aguilar M, Machlicova A, et al. Function of the DEMETER DNA glycosylase in the *Arabidopsis thaliana* male gametophyte. Proc Natl Acad Sci U S A. 2011;108(19):8042–7.
33. Slotkin RK, Vaughn M, Borges F, Tanurdzic M, Becker JD, Feijo JA, et al. Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell. 2009;136(3):461–72.
34. Kinoshita T, Miura A, Choi Y, Kinoshita Y, Cao X, Jacobsen SE, et al. One-way control of *FWA* imprinting in *Arabidopsis* endosperm by DNA methylation. Science. 2004;303(5657):521–3.
35. Kinoshita Y, Saze H, Kinoshita T, Miura A, Soppe WJ, Koornneef M, et al. Control of *FWA* gene silencing in *Arabidopsis thaliana* by SINE-related direct repeats. Plant J. 2007;49(1):38–45.
36. Gehring M, Bubb KL, Henikoff S. Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science. 2009;324(5933):1447–51.
37. Hsieh TF, Ibarra CA, Silva P, Zemach A, Eshed-Williams L, Fischer RL, et al. Genome-wide demethylation of *Arabidopsis* endosperm. Science. 2009;324(5933):1451–4.
38. Jullien PE, Mosquna A, Ingouff M, Sakata T, Ohad N, Berger F. Retinoblastoma and its binding partner MSI1 control imprinting in *Arabidopsis*. PLoS Biol. 2008;6(8):e194.
39. Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, et al. DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in *Arabidopsis*. Cell. 2002;110(1):33–42.
40. Gehring M, Huh JH, Hsieh TF, Penterman J, Choi Y, Harada JJ, et al. DEMETER DNA glycosylase establishes *MEDEA* polycomb gene self-imprinting by allele-specific demethylation. Cell. 2006;124(3):495–506.
41. Andreuzza S, Li J, Guitton AE, Faure JE, Casanova S, Park JS, et al. DNA LIGASE I exerts a maternal effect on seed development in *Arabidopsis thaliana*. Development. 2010;137(1):73–81.
42. Martinez-Macias MI, Cordoba-Canero D, Ariza RR, Roldan-Arjona T. The DNA repair protein XRCC1 functions in the plant DNA demethylation pathway by stimulating cytosine methylation (5-meC) excision, gap tailoring, and DNA ligation. J Biol Chem. 2013;288(8):5496–505.
43. Martinez-Macias MI, Qian W, Miki D, Pontes O, Liu Y, Tang K, et al. A DNA 3′ phosphatase functions in active DNA demethylation in *Arabidopsis*. Mol Cell. 2012;45(3):357–70.
44. Kinoshita T, Yadegari R, Harada JJ, Goldberg RB, Fischer RL. Imprinting of the *MEDEA* polycomb gene in the *Arabidopsis* endosperm. Plant Cell. 1999;11(10):1945–52.
45. Vielle-Calzada JP, Thomas J, Spillane C, Coluccio A, Hoeppner MA, Grossniklaus U. Maintenance of genomic imprinting at the *Arabidopsis medea* locus requires zygotic DDM1 activity. Genes Dev. 1999;13(22):2971–82.
46. Xiao W, Gehring M, Choi Y, Margossian L, Pu H, Harada JJ, et al. Imprinting of the *MEA* Polycomb gene is controlled by antagonism between MET1 methyltransferase and DME glycosylase. Dev Cell. 2003;5(6):891–901.
47. Wohrmann HJ, Gagliardini V, Raissig MT, Wehrle W, Arand J, Schmidt A, et al. Identification of a DNA methylation-independent imprinting control region at the *Arabidopsis MEDEA* locus. Genes Dev. 2012;26(16):1837–50.
48. Makarevich G, Villar CB, Erilova A, Kohler C. Mechanism of *PHERES1* imprinting in *Arabidopsis*. J Cell Sci. 2008;121(6):906–12.

49. Weinhofer I, Hehenberger E, Roszak P, Hennig L, Kohler C. H3K27me3 profiling of the endosperm implies exclusion of polycomb group protein targeting by DNA methylation. PLoS Genet. 2010;6(10):e1001152.
50. Baroux C, Gagliardini V, Page DR, Grossniklaus U. Dynamic regulatory interactions of Polycomb group genes: *MEDEA* autoregulation is required for imprinted gene expression in *Arabidopsis*. Genes Dev. 2006;20(9):1081–6.
51. Jullien PE, Katz A, Oliva M, Ohad N, Berger F. Polycomb group complexes self-regulate imprinting of the Polycomb group gene *MEDEA* in *Arabidopsis*. Curr Biol. 2006;16(5):486–92.
52. Ikeda Y, Kinoshita Y, Susaki D, Iwano M, Takayama S, Higashiyama T, et al. HMG domain containing *SSRP1* is required for DNA demethylation and genomic imprinting in *Arabidopsis*. Dev Cell. 2011;21(3):589–96.
53. Rea M, Zheng W, Chen M, Braud C, Bhangu D, Rognan TN, et al. Histone H1 affects gene imprinting and DNA methylation in *Arabidopsis*. Plant J. 2012;71(5):776–86.
54. Qian W, Miki D, Zhang H, Liu Y, Zhang X, Tang K, et al. A histone acetyltransferase regulates active DNA demethylation in *Arabidopsis*. Science. 2012;336(6087):1445–8.
55. Zheng X, Pontes O, Zhu J, Miki D, Zhang F, Li WX, et al. ROS3 is an RNA-binding protein required for DNA demethylation in *Arabidopsis*. Nature. 2008;455(7217):1259–62.
56. Jahnke S, Scholten S. Epigenetic resetting of a gene imprinted in plant embryos. Curr Biol. 2009;19(19):1677–81.
57. Luo M, Taylor JM, Spriggs A, Zhang H, Wu X, Russell S, et al. A genome-wide survey of imprinted genes in rice seeds reveals imprinting primarily occurs in the endosperm. PLoS Genet. 2011;7(6):e1002125.
58. Gehring M, Missirian V, Henikoff S. Genomic analysis of parent-of-origin allelic expression in *Arabidopsis thaliana* seeds. PLoS One. 2011;6(8):e23687.
59. Hsieh TF, Shin J, Uzawa R, Silva P, Cohen S, Bauer MJ, et al. Regulation of imprinted gene expression in *Arabidopsis* endosperm. Proc Natl Acad Sci U S A. 2011;108(5):1755–62.
60. Nodine MD, Bartel DP. Maternal and paternal genomes contribute equally to the transcriptome of early plant embryos. Nature. 2012;482(7383):94–7.
61. Jullien PE, Susaki D, Yelagandula R, Higashiyama T, Berger F. DNA methylation dynamics during sexual reproduction in *Arabidopsis thaliana*. Curr Biol. 2012;22(19):1825–30.
62. Vu TM, Nakamura M, Calarco JP, Susaki D, Lim PQ, Kinoshita T, et al. RNA-directed DNA methylation regulates parental genomic imprinting at several loci in *Arabidopsis*. Development. 2013;140(14):2953–60.
63. Mosher RA, Melnyk CW, Kelly KA, Dunn RM, Studholme DJ, Baulcombe DC. Uniparental expression of PolIV-dependent siRNAs in developing endosperm of *Arabidopsis*. Nature. 2009;460(7252):283–6.
64. Lu J, Zhang C, Baulcombe DC, Chen ZJ. Maternal siRNAs as regulators of parental genome imbalance and gene expression in endosperm of *Arabidopsis* seeds. Proc Natl Acad Sci U S A. 2012;109(14):5529–34.
65. Ikeda Y. Plant imprinted genes identified by genome-wide approaches and their regulatory mechanisms. Plant Cell Physiol. 2012;53(5):809–16.
66. Pignatta D, Gehring M. Imprinting meets genomics: new insights and new challenges. Curr Opin Plant Biol. 2012;15(5):530–5.
67. Waters AJ, Bilinski P, Eichten SR, Vaughn MW, Ross-Ibarra J, Gehring M, et al. Comprehensive analysis of imprinted genes in maize reveals allelic variation for imprinting and limited conservation with other species. Proc Natl Acad Sci U S A. 2013;110(48):19639–44.
68. Wolff P, Weinhofer I, Seguin J, Roszak P, Beisel C, Donoghue MT, et al. High-resolution analysis of parent-of-origin allelic expression in the *Arabidopsis* Endosperm. PLoS Genet. 2011;7(6):e1002126.
69. Zhang M, Zhao H, Xie S, Chen J, Xu Y, Wang K, et al. Extensive, clustered parental imprinting of protein-coding and noncoding RNAs in developing maize endosperm. Proc Natl Acad Sci U S A. 2011;108(50):20042–7.
70. Kawabe A, Fujimoto R, Charlesworth D. High diversity due to balancing selection in the promoter region of the *Medea* gene in *Arabidopsis lyrata*. Curr Biol. 2007;17(21):1885–9.

71. Miyake T, Takebayashi N, Wolf DE. Possible diversifying selection in the imprinted gene, *MEDEA*, in *Arabidopsis*. Mol Biol Evol. 2009;26(4):843–57.
72. Spillane C, Schmid KJ, Laoueille-Duprat S, Pien S, Escobar-Restrepo JM, Baroux C, et al. Positive Darwinian selection at the imprinted *MEDEA* locus in plants. Nature. 2007;448(7151):349–52.
73. Yoshida T, Kawabe A. Importance of gene duplication in the evolution of genomic imprinting revealed by molecular evolutionary analysis of the type I MADS-box gene family in *Arabidopsis* species. PLoS One. 2013;8(9):e73588.
74. Smith ZD, Meissner A. DNA methylation: roles in mammalian development. Nat Rev Genet. 2013;14(3):204–20.
75. Johannes F, Porcher E, Teixeira FK, Saliba-Colombani V, Simon M, Agier N, et al. Assessing the impact of transgenerational epigenetic variation on complex traits. PLoS Genet. 2009;5(6):e1000530.
76. Li W, Liu H, Cheng ZJ, Su YH, Han HN, Zhang Y, et al. DNA methylation and histone modifications regulate de novo shoot regeneration in *Arabidopsis* by modulating *WUSCHEL* expression and auxin signaling. PLoS Genet. 2011;7(8):e1002243.
77. Iwasaki M, Takahashi H, Iwakawa H, Nakagawa A, Ishikawa T, Tanaka H, et al. Dual regulation of *ETTIN* (*ARF3*) gene expression by AS1-AS2, which maintains the DNA methylation level, is involved in stabilization of leaf adaxial-abaxial partitioning in *Arabidopsis*. Development. 2013;140(9):1958–69.
78. Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, et al. Widespread dynamic DNA methylation in response to biotic stress. Proc Natl Acad Sci U S A. 2012;109(32):E2183–91.
79. Haun WJ, Laoueille-Duprat S, O'Connell M J, Spillane C, Grossniklaus U, Phillips AR, et al. Genomic imprinting, methylation and molecular evolution of maize *Enhancer of zeste* (*Mez*) homologs. Plant J. 2007;49(2):325–37.
80. Haun WJ, Springer NM. Maternal and paternal alleles exhibit differential histone methylation and acetylation at maize imprinted genes. Plant J. 2008;56(6):903–12.
81. Gutierrez-Marcos JF, Costa LM, Dal Pra M, Scholten S, Kranz E, Perez P, et al. Epigenetic asymmetry of imprinted genes in plant gametes. Nat Genet. 2006;38(8):876–8.
82. Ishikawa R, Ohnishi T, Kinoshita Y, Eiguchi M, Kurata N, Kinoshita T. Rice interspecies hybrids show precocious or delayed developmental transitions in the endosperm without change to the rate of syncytial nuclear division. Plant J. 2011;65(5):798–806.

Chapter 3
Nuclear Bodies and Responses to the Environments

Yin Liu and Yuda Fang

Introduction

The nucleus is a structurally complicated organelle containing numbers of different machines essential for the function and maintenance of genome, which is folded into chromosomes and resided in chromosome territories nonrandomly [1–3]. Microscopy analysis showed that nucleus is quite heterogeneous and harbors various structurally and functionally distinct subnuclear compartments called nuclear bodies (NBs) [4, 5].

NBs are special subcellular domains containing protein–protein or protein–RNA complexes without membranes, where the components freely exchange with the surrounding nucleoplasm to regulate biological processes like the cell cycle, RNA processing, signal transduction, DNA damage repair, cell death, or stress response [6–10]. Many NBs have been found, including nucleoli, Cajal bodies (CBs), and nuclear speckles, which exist both in animals and plants. NBs like paraspeckles, promyelocytic leukemia NB (PML-NB), polycomb group (PcG) body, and 53BP1 NB have been characterized in mammalian cells. In plants, plant-specific NBs such as dicing bodies (D-bodies), photobodies, cyclophilin-containing bodies, and A-kinase-interacting protein 1 (AKIP1)-concentrated bodies are also characterized [9, 11, 12]. Increasing reports have elucidated the assembly and maintenance of NBs [13–17]. Since these NBs quite differ from each other in size, numbers, morphology, and functions [12, 18], which are regulated by different internal or external cues, characterizing new components and explaining the biogenesis mechanism will be of great importance to understand their precise functions in signal transduction.

Y. Fang (✉) · Y. Liu
National Key Laboratory of Plant Molecular Genetics, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai 200032, China
e-mail: yfang@sippe.ac.cn

Shanghai Institutes for Biological Sciences, Chinese Academy of Sciences, Shanghai 200032, China

O. Pontes, H. Jin (eds.), *Nuclear Functions in Plant Transcription, Signaling and Development,* DOI 10.1007/978-1-4939-2386-1_3

Following, we focus on plant NBs, their relations with environmental cues, the cross talk between them, and proposed assembly models to elucidate multiple functions of NBs during the growth and development of plants.

Nuclear Bodies and Response to Environment

Without membrane enclosed, NBs are still structurally intact, suggesting a maintenance mechanism to exert many important biological functions [12, 18, 19]. NBs can act as reaction sites to promote cellular processes by concentrating proteins or RNA required, such as the nucleolus and CBs, where many RNAs and proteins are enriched to form complexes with a higher efficiency [20–22]. NBs can also serve as hubs to recruit many gene loci to regulate their expression by stabilizing their interactions, like PcG bodies in *Drosophila* where PcG proteins are concentrated [23, 24]. Additionally, NBs are storage and modification sites where many components are modified such as phosphorylation in nuclear speckles and sumoylation in PcG bodies [25–28].

Nucleolus

The nucleolus is the most prominent NB containing three distinct regions: small and lightly stained fibrillar center (FC), intensively stained dense fibrillar component (DFC), and the granular component (GC) [6, 12, 29]. Ribosomal DNAs (rDNAs) localize in the FCs; rDNA transcriptions take place at the boundary of FCs and DFC; precursor ribosomal RNAs (pre-rRNAs) are processed in the DFC where fibrillarin and small nucleolar ribonucleoproteins (snoRNPs) are located; GC is the site for accumulation and assembly of pre-ribosomal subunits. So the nucleolus provides sites for rDNA transcription, pre-rRNA processing, and pre-ribosome assembly [20, 29]. Compared with mammalian cells, plant DFC is less densely stained but much more extensive, which covers up to 70% volume of the nucleolus. In addition, the plant nucleolus is often spherical, quite regular in morphology [11].

Altering ribosome biogenesis may induce reorganization of the nucleolus, which is also affected by various stresses because of altered rDNA transcription [6, 30]. In addition to ribosome assembly, the nucleolus has many other functions. Processing of RNAs such as the signal recognition particle RNAs, precursor transfer RNAs (pre-tRNAs), and U6 small nuclear RNA (U6 snRNA) occurs in the nucleolus [31–35]. Besides, it is the site for biogenesis and modification of RNP complexes [20, 29, 36]. The telomerase RNP assembly also happens in the nucleolus [29, 33]. Posttranslational modification of some nuclear proteins is essential to sequester and repress their activities to adjust multiple cellular processes such as cell cycle, stress response, or aging [12, 29, 35]. Many plant and animal viruses target the nucleolus to produce and export viral RNPs [37–39]. For example, in plants, many viruses interact with the nucleolus, which is important for viral RNPs assembly, virus replication, and transportation, to antagonize the defense system [39]. Open

reading frame 3 (ORF3), a protein of *Groundnut rosette virus,* disrupts CBs and relocates to the nucleolus from CBs, recruiting fibrillarin to assemble cytoplasmic infectious viral particles. Nucleolar localization of ORF3 is important for systemic infection [40, 41]. Viral proteins like the coat proteins and their read-through factors of *Potato leafroll virus* (PLRV) also locate to the nucleolus to interact with fibrillin. This interaction is essential for the long-distance transportation of PLRV [42]. NIa, a viral protein of *Potato virus A,* interacts with fibrillarin in the nucleolus and CBs. Depleting fibrillarin can cause a decrease in concentrating the virus, suggesting the involvement of fibrillarin during the virus infection [43]. Additionally, targeting viral proteins to the nucleolus may also disturb the host defenses as their interactions with host proteins like fibrillarin may disrupt the normal function of the nucleolus [35]. Many evidences suggested that the nucleolus plays a role in response to stresses [6]. In mammalian cells, the nucleolus is related to the p53 DNA damage pathway. Under normal condition, p53 is at low level because of interaction, ubiquitination, nuclear exporting, and degradation involved in an E3 ligase mouse double minute 2 homolog (MDM2) [6, 44]. p14ARF, predominantly located in the nucleolus, can interact and repress the MDM2 activity [45]. When oncogenic stress appears, p14ARF is induced, then sequesters and inhibits MDM2, causing an increase of p53 level in the nucleus [45–47]. Nucleolar stress like disrupting ribosome biogenesis may cause failure in degradation of p53 because of direct association between MDM2 and many ribosomal proteins released from the nucleolus, which breaks the repression of MDM2 on p53 [48]. Besides, rDNA may be a sensor for DNA damage as the genome size has much correlation with the number of rDNA repeats [49]. In *Arabidopsis,* depletion of chromatin assembly factor 1 (CAF1) causes hypersensitivity to DNA damage and loss of telomeric and rDNA copies in successive generations [50]. These data suggest that the nucleolus may be a direct stress sensor [9, 35]. Around the nucleolus organizer regions (NORs), there exist satellite DNA and silent rDNA repeats [51]. This heterochromatin region is highly compact and called perinucleolar region, which may involve in maintaining the silence of nonribosomal genomic regions [12].

Cajal Bodies

CBs, also called coiled bodies, are one of the nonnucleolar subcellular bodies, which contain many components involved in the assembly, modification, and trafficking of small nuclear RNPs (snRNPs) and snoRNPs [9, 52, 53]. CBs are also involved in the 3′ end processing of histone pre-messenger RNA (pre-mRNA) and telomerase assembly [9, 54, 55].

During the assembly of spliceosomal snRNPs, spliceosomal snRNAs are transported to the cytoplasm after nuclear transcription. In the cytoplasm, the survival motor neuron (SMN) complexes assemble because of interactions between these snRNAs and the core Sm proteins. The newly formed snRNPs are relocated into the nucleus and enter CBs for maturation before moving to nuclear speckles or transcription sites [9, 56]. CBs also have roles in spliceosome recycle by increasing the

assembly rate through concentrating components required for reactions [53]. Cells containing at least one CB assemble the U4/U6 di-snRNP at a rate of about 11-fold faster than the cells without CBs [22]. So CBs may act as depots which greatly enhance the snRNPs assembly rate [9]. A human disease like spinal muscular atrophy (SMA) is caused by the functional deficiency of the SMN1 [57]. SMNs with low levels may disrupt the CBs and greatly decrease the activity of snRNP assembly, suggesting association between CBs and spliceosomal snRNP assembly [58]. CBs also have functions in regulating telomeres [54, 55]. Telomerase RNA (TR) and telomerase reverse transcriptase (TERT) compose the telomerase RNP enzyme, contributing to the homeostasis of telomere length [59]. WRAP53/TCAB1, a small CB-specific RNP (scaRNP) component, recognizes the human TR (hTR) through the CAB box in hTR and targets it to CBs, where the telomerase RNP assembles and the association between telomerase and telomeres is enhanced [60, 61].

Studies showed that the phosphorylation and dephosphorylation of coilin and SMN have functional impacts on the formation of CBs and components recruiting [62, 63]. The C-terminal of coilin is hyperphosphorylated in cells that lack CBs, decreasing its ability for self-interaction. While in cells with many CBs, hypophosphorylation happens on coilin, promoting its self-interaction ability [63]. The cytoplasmic SMN is hyperphosphorylated and able to target proteins to the SMN complex [57]. Hypophosphorylated coilin recruits the SMN complex containing snRNPs to CBs, then coilin is hyperphosphorylated and disrupts interaction with SMN, which accelerates SMN releasing from CBs and strengthens snRNP–coilin interaction. These snRNPs translocated from SMN complexes to CBs will be modified in CBs [9, 62]. Phosphatase PPM1G is involved in the dephosphorylation of coilin [62].

In plants, live cell imaging analysis of green fluorescent protein (GFP)-fused spliceosomal protein U2B″ (U2B″–GFP) and use of fluorescence recovery after photobleaching (FRAP) reveal the dynamics of CBs [64, 65]. Besides, plant cell type, developmental stage, and cell cycles may also change the numbers of CBs. In G1, there are many CBs in the nucleus, while in late G2, two large CBs are showed [66]. Using U2B″–GFP transgenic *Arabidopsis* for forward genetic screening, mutants with no CB formation (*no CB-1, ncb-1*) or smaller CBs (*ncb-2, ncb-3*) were obtained and the mutated gene was named *Atcoilin* (At1g13030) based on alignment with vertebrate coilin [67]. In *Arabidopsis* and *Drosophila,* mutants lacking CBs because of deficient coilin show no significant development defects, while in zebra fish, the knockdown of coilin leads to loss of CBs and developmental arrest in the embryo, which can be rescued by injecting preassembly snRNPs [67–69]. These data suggest that the efficiency of biological processes may be decreased without CBs as it can concentrate factors essential for snRNPs complex assembly [67–69].

Nuclear Speckles

Most genes in plants and animals require removing introns and fusing exons to generate mature mRNAs [70, 71]. During the splicing process, there are two patterns according to whether the splice sites are the same. One is constitutive splicing

and the other is alternative splicing. Both two splicing ways are critical for proper expression of genes containing introns [28, 71, 72]. Splicing regulators interact with numerous spliceosomal proteins and are recruited to the pre-mRNAs to regulate splicing process [71].

The splicing of pre-mRNAs coupled with transcription happens in the spliceosome, a large complex containing snRNAs and numbers of proteins [73, 74]. In *Arabidopsis,* there exist more than 200 RNA-binding proteins (RBPs). Some belong to the snRNPs, while numerous non-snRNP proteins also present in the spliceosome to regulate splicing [75]. Among these non-snRNPs, serine-/arginine-rich proteins (SR proteins) are widely investigated in animals. In plants, more and more reports appear to characterize the plant SR proteins [72, 76, 77].

SR proteins are a family of splicing factors containing RNA recognition motifs (RRMs) in the N-terminus and arginine–serine (RS)-rich domains in the C-terminus [78, 79]. The RRM motif is involved in recognizing and binding to specific sequences in pre-mRNAs and the RS domain is essential for interacting with and recruiting other proteins [80, 81]. In addition, signals for subcellular localization and shuttling also locate in the RS domain [72, 77, 82, 83]. In either constitutive or alternative splicing processes, SR proteins are important regulators both in animals and plants. In animals, SR proteins have roles in mRNA transportation, localization, and translation. Regulating genome stability and microRNA (miRNA) biogenesis also require the participation of SR proteins [84]. Since some SR proteins are plant specific, it may suggest some different functions of plant SR proteins [79, 85].

At the cellular level, many splicing factors and snRNPs in mammalian cells are targeted to irregular subnuclear domains named nuclear speckles, in addition to nucleoplasmic signals or sometimes CBs distribution [28, 86]. SR proteins relocate to active transcription sites from the nuclear speckles, suggesting a storage function of nuclear speckles in spliceosome assembly [11, 28]. Plant SR proteins have similar nuclear speckle positions. Interestingly, SR proteins belonging to different subfamilies localize to different speckles, indicating that different SR proteins may interact with distinct pre-mRNAs or mRNAs to regulate splicing and transportation [11, 87]. FRAP and fluorescence loss in photobleaching (FLIP) analysis show that the SR proteins rapidly shuttle between speckles and the surrounding environment [71, 88, 89]. Fluorescent protein-based methods are used to elucidate the dynamics of components in nuclear speckles, whose morphologies are usually changed because of internal or external cues [90]. Inhibiting transcription by drug or heat shock or repressing the activities of kinase or phosphatase may cause a decrease in nucleoplasmic distribution of plant SR proteins and an increased localization in large speckles [91, 92]. In *Arabidopsis,* depletion of SAD1, a protein similar to Sm-like snRNPs required for mRNA splicing and exporting, caused hypersensitivity to abscisic acid (ABA) and drought stress during seed germination [93]. Mutation of transportin-SR (TRN-SR), a member functioning as the nuclear import receptor for SR proteins, impaired the nuclear localization of some SR proteins, which leads to compromised resistance because of altered splicing patterns of some resistance genes, providing an evidence linking splicing and plant immunity [94]. Mutation of SR45, a plant-specific SR-related protein, displayed many developmental defects and altered response to ABA and glucose [95, 96]. Some splicing regulators like

SR proteins also undergo extensive alternative splicing. Abiotic and biotic stress may affect the spliced products and induce some other variants to adjust the stress [97]. Cold and heat stress may change the splicing patterns of some *Arabidopsis* SR genes, causing alterations on the pre-mRNAs splicing [98]. In animals, the serine residues in the RS domain of SR proteins are usually phosphorylated, which has much correlation with the subcellular localization of SR proteins and their interactions with other components in splicing [80]. In *Arabidopsis, fus3-complementing gene 2* (*AFC2*), a Clk/Sty protein kinase, has been characterized to phosphorylate and interact with many SR proteins. The interactions are modulated by phosphorylation [99]. Another kinase PK12, belonging to the tobacco LAMMER type and induced by ethylene, is showed to phosphorylate and interact with SR34, implicating a connection between plant hormones and splicing regulation through PK12 kinase [100]. High-throughput identification for potential substrates of mitogen-activated protein kinases (MAPKs) in *Arabidopsis* showed that stress-activated MPK6 and MPK3 can phosphorylate some SR proteins, providing linking between stress and splicing regulation [101]. Together, these data suggest that SR proteins are important splicing regulators in response to different environmental cues.

Dicing Bodies

miRNAs are one class of small noncoding regulatory RNAs in eukaryotes containing approximately 21–22 nucleotides, which are first transcribed into primary transcripts called primary miRNAs (pri-miRNAs) by RNA polymerase II [102–104]. The processing of pri-miRNAs into miRNA duplex is different in animals and plants. In animals, this process happens at two different places by distinct endonucleolytic enzymes. Converting pri-mRNAs into precursor miRNAs (pre-miRNAs) takes place in the nucleus by a complex containing the ribonuclease (RNase) III family protein Drosha and its partner Pasha, a double-strand RNA-binding domain (dsRBD) protein [105–107]. Then exportin 5 recognizes and transports these pre-miRNAs to the cytoplasm, where the miRNA/miRNA* duplex are produced from these pre-miRNAs by the RNase III enzyme Dicer and its partner TAR RBP 2 (TRBP2) [108–111]. While in plants, these two steps take place only in the nucleus by a single plant RNase III enzyme DICER-LIKE1 (DCL1) and its partner HYPONASTIC LEAVES1 (HYL1) [112]. Emerging reports have discovered many other proteins involved, such as the zinc finger protein SERRATE (SE), the nuclear cap-binding protein (CBP) complex, the transcription factor NOT2 and CDC5, the phosphatase C-TERMINAL DOMAIN PHOSPHATASE-LIKE 1 (CPL1), and the RBPs DAWDLE, TOUGH, and MOS2 [113–123].

As one class of small RNAs, which are ubiquitous modulators involved in plant growth and development, miRNAs are reported as essential players responding to various plant stress [124–127]. Many stress-response genes are miRNAs targets, and the miRNAs levels are also regulated by stress. This kind of small RNAs-guided regulatory network makes plants better to adjust to different abiotic and

biotic stress [128]. CSD1 and CSD2, two Cu–Zn superoxide dismutases involved in scavenging superoxide radicals to defense oxidative stress, are reported as targets of miR398. Under oxidative stress condition, the *CSD1* and *CSD2* mRNAs levels are accumulated because of downregulation of miR398 transcript, which suppresses *CSD1* and *CSD2* genes under normal conditions [129]. UBC24 or PHO2, an ubiquitin-conjugating enzyme involved in protein turnover in inorganic phosphate (Pi) homeostasis, was characterized as a target of miR399. The miR399 transcription is downregulated by sufficient Pi and UBC24 is accumulated, which decreased the expressions of phosphate transporters and components regulating root growth through proteasome pathway to prevent excessive absorption of Pi [130, 131]. During sulfur homeostasis pathway, miR395 is reported to be involved by targeting adenosine triphosphate (ATP) sulfurylases like APS1, APS3, and APS4 [132]. Many miRNAs response to aluminum (Al) stress have been identified, such as miR319, miR390, and miR393 [133]. Small RNA sequencing analysis revealed that the regulation of some miRNAs are stress specific such as miR319c, which is induced by cold but not salt, ABA, or dehydration [134]. In addition, some targets of miRNAs encode E3 ligases and are regulated by different stress, which provides cues to elucidate the mechanisms for plants to respond to stress through proteolysis-guided pathway [128]. In auxin signaling pathway, some auxin response factors (ARFs) are showed as targets of miRNAs. ARF6 and ARF8 are the targets of miR167; ARF10, ARF16, and ARF17 are regulated by miR160 [135, 136]. Additionally, miR160 and miR167 are reported as important players in drought and ABA response, indicating miRNAs and their targets are essential players during the cross talk among various environmental cues [137, 138]. Mutants of key regulators in miRNA biogenesis pathway like *HYL1, DCL1, SE, HEN1,* and *HASTY* displayed ABA-hypersensitive phenotypes [139]. Depletion of *HYL1* made *Arabidopsis* more sensitive to ABA in germination; the *dcl1* and *hen1* mutants also exhibited increased sensitivity to ABA; loss of functions of *SE* and *HASTY* showed enhanced ABA sensitivity [140, 141]. Increasing evidences have revealed important contributions of components in small RNA pathways during the defense of plant immunity [142–148]. Mutations of *DCL1, HYL1,* or *HEN1* cause plants impressionable to bacterial pathogen because of deficient miRNAs [149–151]. In addition, *DCL1* and *HYL1* are reported to participate in producing some endogenous small interfering RNAs (siRNAs) which regulate effector-triggered immunity (ETI) mediated by R genes [150, 151].

Live cell imaging analysis of the localization of DCL1 and HYL1 showed that DCL1 is concentrated in round bodies and HYL1 displays a similar pattern. The DCL1-/HYL1-containing bodies are distinct from CBs as no overlay occurred between the marker protein Atcoilin of CBs and DCL1/HYL1 [152]. These DCL1-/HYL1-concentrated bodies are called dicing bodies or D-bodies [104, 153]. In vivo tracking of a pri-miRNA indicated that the pri-miRNAs can be recruited to D-bodies [153]. In addition, the two C-terminal dsRBDs of DCL1 have essential roles in pri-miRNA binding and protein–protein interactions, respectively. Interestingly, the two C-terminal dsRBDs of DCL1 can substitute the functions of HYL1 in D-bodies [154–156], supporting a model in which *Arabidopsis* pri-miRNAs are recruited to D-bodies through functionally divergent dsRBDs of microprocessor for accurate

processing. As numerous reports showed the close relationships between miRNAs and environmental stress, it is of interest to study the potential roles of D-bodies in plant stress responses.

Photobodies

Light is an important environmental cue for plant growth and development [157–159]. In *Arabidopsis,* there are mainly three classes of photoreceptors to perceive different light conditions, including the red- and far-red-absorbing phytochromes (phyA–phyE), the UV-A/blue light-sensing receptors like cryptochromes (cry1 and cry2), phototropins (phot1 and phot2), and the ZEITLUPE (ZTL)/ FLAVIN-BINDING, KELCH, F-BOX1(FKF1)/LOV KELCH PROTEIN 2 (LKP2) family, and the newly characterized UV-B receptor UV RESISTANCE LOCUS 8 (UVR8), which regulate almost the whole life of plants [160–168].

After light activation, the phytochromes relocate from cytoplasm to the nucleus to form discrete subcellular domains called photobodies [162, 169–175]. Cryptochromes and UVR8 are also relocated to these subnuclear foci [176–180]. The size, number, and morphology of photobodies are dynamic and regulated by different light fluence rates [171, 181, 182]. The morphologies of phytochrome B (phyB) NBs are changed under different light conditions; there exist four localization patterns: diffused nuclear location, numerous small bodies, small and large bodies coexisting, and only a few large bodies [181]. Large phyB photobodies come up with high fluence rate of red light, at which the phyB response is the strongest. Moreover, the inhibition of hypocotyl elongation of phyB–GFP transgenic plant correlates with the morphology of phyB photobodies. Along with the increasing of fluence rate, the phyB photobodies grow larger and the hypocotyls are much shorter [171, 181]. FRAP analysis revealed that the phyB signal in photobodies exchanged with the nucleoplasmic phyB, which took a few minutes for partial recovery [183]. When plants are grown from dark to light condition, phytochromes relocate to photobodies very quickly [172, 174]. Some can be observed after light exposure for a few minutes and then disappear, such as early photobodies like phyA photobodies [184]. While some reappear and remain in the light after exposure for 2 h. These bodies are late photobodies like phyB photobodies [173, 184].

As light is essential for plant growth and development, many components in photobodies have been showed to involve in numerous other pathways regulated by different environmental cues like plant hormones, temperature, nitric oxide, and many stresses [185–196]. Phytochrome-interacting factors (PIFs), one small subset of basic helix-loop-helix (bHLH) transcription factors, act as a cellular hub to intersect different signaling pathways [197–199]. GAI and RGA, two DELLA proteins, interact with PIF3 and PIF4 to inhibit their DNA-binding activities [200, 201]. High temperature can induce a rapid increase of PIF4 transcript to sense the temperature shift [202]. ETHYLENE-INSENSITIVE 3 (EIN3) and EIN3-like 1 (EIN1), two downstream activators in the ethylene signaling pathway, activate the expression of PIF3 by directly binding to its promoter [203]. LONG HYPOCOTYL 5 (HY5),

a major positive regulator in photomorphogenesis, is characterized to be a master in connecting multiple plant hormones [189, 194, 204–211]. CONSTITUTIVE PHOTOMORPHOGENIC 1 (COP1), an E3 ubiquitin ligase, is a central switch in photomorphogenesis [212, 213]. By analyzing the substrates of COP1, proteins in other pathways are found such as GATA2 in brassinosteroid pathway and the R protein HRT in plant defense, suggesting a broad cross talk between light and other signaling pathways [214, 215]. Though it is still not clear about the functions of photobodies, some models have been proposed. One is the storage model, in which photoactivated phytochromes are sequestrated to balance their levels in the nucleoplasm [183, 216, 217]. Another is the degradation model, where turnover of proteins takes place, because many components colocalize with the E3 ligase COP1 [179, 180, 218–224]. The third is the transcription model. Some transcription factors localize in photobodies, so they may bring target genes to photobodies for regulation [225]. In addition, photobodies may be sites for phytochrome signaling transduction as the morphology is affected by light intensities [181, 226].

Cyclophilin-Containing Bodies

Cyclophilins are an amusing class of proteins having prolyl-peptidyl isomerase activity essential for protein folding [227, 228]. In *Arabidopsis*, some cyclophilins are characterized to have an arginine/serine (SR) and serine/proline (SP) dipeptides-rich domain at the C-terminal essential for protein interactions [11, 71, 229]. Yeast two-hybrid assay revealed an interaction between cyclophilins and *Arabidopsis* SR proteins. AtCypRS64 and AtCypRS92 interact with SR proteins such as SCL28, SCL30, SRp30, and SRp34. Additionally, they also interact with U1-70K and U11-35K, two proteins specific to U1 snRNP and U11 snRNP, indicating an involvement of cyclophilins in pre-mRNA processing [230]. AtCyp59, another cyclophilin with a peptidyl-prolyl cis/trans isomerase (PPIase) domain, is found to interact with SR33 and the C-terminal domain of RNA polymerase II. The RNA binding of AtCyp59 inhibits its PPIase activity, suggesting an essential role of AtCyp59 in connecting splicing and transcription [231, 232]. Subcellular localization showed that AtCypRS64 located in many small NBs distinct from CBs. When co-expressed with the interactive SR proteins, AtCypRS64 relocated to the nuclear speckles, suggesting that the cyclophilin-concentrated bodies may be storage sites essential for their targeting to nuclear speckles to participate in splicing with SR proteins or other specific snRNPs [11, 230].

AKIP1-Containing Bodies

Phytohormone ABA plays important roles in regulating gene expression and ion channel activity [233, 234]. Mutants of some RBPs display altered sensitivity to ABA, which indicated a correlation between ABA pathway and RNA metabolism

[93, 140, 235, 236]. Time course analysis showed that AKIP1, a heterogeneous nuclear RNP (hnRNP)-like RBP, was recruited to numerous discrete NBs by ABA treatment as ABA can activate a Ca^{+}-independent kinase AAPK, which is essential for the phosphorylation of AKIP1 [237, 238]. After phosphorylation, AKIP1 gained ability to bind dehydrin mRNAs, which encode a class of proteins to respond to stress [237]. Treating with transcriptional inhibitors, ABA-induced relocation of AKIP1 to NBs was decreased. These data imply that the phosphorylation of AKIP1 induced by ABA and transcriptional activation of ABA-response genes are required for the formation of AKIP1-concentrated bodies [237]. The precise function of AKIP1-containing bodies is not clear, but they may involve in mRNA metabolism [11].

Cross Talk Between NBs

Since NBs are nonmembranous complexes with various components exchanging among the surrounding environment, it is quite common for NBs to talk with each other [11, 12, 18]. Analyzing the proteomic data of *Arabidopsis* nucleoli, many non-ribosomal and nonnucleolar proteins were found, including translation factors and splicing regulators, suggesting multiple functions of the nucleolus [239]. eIF4A-III, a component of the core exon junction complex, was induced to relocate to the nucleolus from the nucleoplasm and finally enter nuclear speckles by hypoxia stress [240]. RSZ22, one member of plant SR proteins, was showed to have a similar redistribution. When ATP is depleted, RSZ22 is located to the nucleolus [90]. As amount of snRNPs were characterized in the nucleolar proteome, the nucleolus may have roles in producing some specific snRNPs in plants [241]. Additionally, the nucleolus can also provide a site responsible for production of some siRNAs which involve in silencing transcription, such as miRNAs [242–244]. Though the maturation of most miRNAs usually takes place in D-bodies, where microprocessors concentrate, the nucleolus may have functions in miRNAs maturation and possible cross talk with D-bodies [9, 35, 104]. In plants, some mRNAs and small regulatory RNAs are present in the nucleolus, indicating potential roles of the nucleolus in regulating gene expressions [245]. Subcellular localization of CB components showed that CBs were close to the nucleolus spatially and changed dynamically in the nucleus, usually targeting to the nucleolar periphery or staying in the nucleolus [64, 246]. Many components, especially some small nucleolar RNAs, RNA processing proteins like dyskerin and fibrillarin, are shared by CBs and the nucleolus, suggesting the correlation between these two structures [35].

As CBs play important roles in the assembly and maturation of some spliceosomal snRNPs and snoRNPs, nuclear speckles and the nucleolus may have talks with CBs during recycle of these components among CBs and the two compartments [53]. Cellularly, the processing of some heterochromatic siRNAs (hc-siRNAs) also happened in a special subnuclear domain reminiscent of CBs in addition to the nucleolar position [247, 248]. Immunofluorescence assay on the localizations

of proteins involved in siRNA or miRNA production showed that some proteins concentrated and colocalized in a subcellular domain around the nucleolus, which have some overlap with the marker protein of CBs. When CBs are disrupted, the proper position of these proteins may be altered, indicating the structural integrity of CBs is required [249]. Taking together, CBs may provide sites where some Dicer-produced RNA precursors are processed and sorted to different pathways [249]. So CBs may have some relationship with D-bodies to some extent. Besides, some splicing factors like STA1 and SR45 are reported to take part in the RNA-direct DNA methylation (RdDM) pathway, in which many components have colocalization with CBs under immunolocalization assay, implicating the connections between nuclear speckles and CBs in RdDM pathway [250–252]. In miRNA processing, SE interacts with HYL1 and DCL1 to promote the cleavage and accuracy of pri-miRNAs [113–115, 122]. HYL1 and DCL1 concentrate in D-bodies while SE has a small overlap with D-bodies [153]. Colocalization between SE and SR33 is observed in *Arabidopsis* [153]. These data suggest that SE may be one of the linkers to connect the splicing and processing of miRNA [253]. In *Arabidopsis,* AtCBP20 and AtCBP80, two components of the nuclear CBPs, interact with SE to participate in alternative splicing even though the CBPs and SE are characterized to be involved in the biogenesis of miRNAs [115]. Moreover, splicing regulators like SF2/ASF and STA1 are showed to involve in miRNA processing [251, 254]. So to find out the direct relationship between nuclear speckles and D-bodies is of great interest. Genetic screening showed RRC1, a homolog of the human potential splicing factor SR140, is essential to regulate the alternative splicing of some other SR protein genes and phyB pathway-related genes through its RS domain and this regulation is phyB dependent, which provides cues to explore the cross talk between nuclear speckles and photobodies [255, 256].

The Biogenesis of NBs

Without membrane enveloped, NBs still have abilities to maintain integrate structures, suggesting a distinct mechanism for their formation and maintenance compared with those cytoplasmic membrane-around organelles [9, 12, 18]. Elucidating how these NBs are assembled and maintained in details is of great interest to understand their physiological functions during signaling transduction and their correlations with the environment. In mammalian cells, a bacterial Lac operator/repressor (LacO/LacI) tethering system is quite popular to be used to study the assembly of NBs [14, 257, 258]. In this system, LacO array is first integrated into the genome, then LacI fusion protein tagged by a fluorescent protein is artificially recruited and tethered to the LacO position because of specific interactions between LacI/LacO. As many reports revealed a significant function of RNA in the biogenesis of NBs, there are some modifications about this system [257, 258]. LacI is fused with the MS2 coat protein, which derives from the bacteriophage and is able to recognize and specifically bind to a RNA stem loop. The RNA specific to a NB is usually fused with the MS2 stem loop. This fusion is tethered to LacO repeats through

LacI-LacO interaction and specific recognization of MS2 coat protein on the MS2 stem loop. Then fluorescence in situ hybridization (FISH) is used to visualize this kind of RNA fusion [259–261]. Another system for RNA in vivo visualization, named λN22 RNA stem-loop binding system, is also built in plant cells [262].

Three models for the assembly of NBs have been proposed based on the interaction network among components in these structures [12, 18]. The first is a stochastic assembly model. Each component interacts stochastically and is able to build the NBs randomly, equal in the assembly order [12]. Tethering analysis revealed that CBs might be assembled in a stochastic pattern. Immobilizing different components of CBs to the LacO array was sufficient to recruit most other components to the tethered site to assemble a de novo subnuclear structure, which has similar size, morphology, composition, and dynamics compared with the endogenous CBs. As individual components can initiate newly formed CBs, though with different efficiencies, CBs are integrated stochastically [16]. The second is an ordered assembly model proposing that each component participates in the formation of a NB in a sequential manner. Tethering system is not suitable for this kind of model [12, 18]. The third is a seeding model that a component or subset may be a seed initiating the biogenesis of NBs. A protein or RNA related to NBs can act as a seed in the assembly process. Histone locus bodies (HLBs) are the sites of components involved in histone pre-mRNA 3′ end processing. Artificially immobilizing H2b-MS2 transcripts to the LacO position, newly formed HLBs can be detected there as NPAT and FLASH, two essential markers of HLBs involved in production of histone genes, were successfully recruited to the tethered site, indicating a seeding role of histone pre-mRNA in initiation of HLBs assembly [14]. Similarly, tethering the noncoding human satellite (sat) III-MS2 transcripts to the LacO array results in de novo formation of nuclear stress bodies (nSBs), where components like heat-shock transcription factors HSF1, SAF-B, and splicing factor SF2/ASF can be detected, implying sat III may act as scaffold to seed the nSBs formation [14]. The immobilization of MS2-fused NEAT1, a noncoding RNA (ncRNA) present in paraspeckles, on the LacO array leads to new formation of paraspeckles in which many marker proteins such as PSP1, PSF, and NONO concentrated, suggesting that the nascent NEAT1 transcript plays a seed role to assemble paraspeckles [264]. In addition, immobilizing β-globin–MS2 to the LacO repeats promotes a fusion towards the existing nuclear speckles or production of newly formed speckles. As well, tethering SC35 can also trigger concentration of speckles, indicating that the spliced RNA polymerase II transcript has scaffold function for recruiting splicing factors to contribute speckles formation [14]. Except for the seeding function of RNAs in the biogenesis of NBs, proteins can also initiate the formation process. Inactivation of tumor-suppressor protein PML, a marker protein of the PML-NBs, causes failure of PML-NBs assembly, while the upregulation of PML by viral infection or senescence triggers an increase on the size and number of PML-NBs, implying PML acts as a seed to initiate PML-NBs formation [265, 266]. Recently, a two-step hybrid assembly model for NB biogenesis was put forward. A seeding component like specific protein or RNA initiates the formation process, which is not random but induced by a cellular activity like transcription. Then other components are recruited to complete the assembly of NBs randomly or in a self-organization way [267].

As LacI/LacO tethering system is not used in plants easily, methods applied to elucidate the assembly mechanisms of plant NBs are quite limited. Based on a genetic forward screening, increasing components essential for plants NBs have been characterized. The *no CB 1* (*ncb-1*) mutant from screening of U2B″–GFP transgenic plants revealed depletion of Atcoilin, a homolog of the vertebrate coilin gene, leading to loss of CBs. While overexpressing Atcoilin in the *ncb-1* background, it successfully restored CBs, where U2B″ was concentrated and the size was larger than that in wild type, suggesting that Atcoilin is required for the formation of CBs and its expression has much correlation with the morphology of CBs [67]. Genetic screening of phyB mislocalization mutants in phyB–GFP transgenic lines, one mutant with smaller or nonphotobodies was obtained, named *hmr* mutant, which provides genetic evidence showing the function of photobodies in regulating protein turnover [268, 269]. *Dcl1-9,* a mutant with a T-DNA insertion on the second dsRBD of DCL1, displayed many severe defects in the miRNA biogenesis pathway [152, 270]. Live cell imaging showed that *dcl1-9* failed to target to D-bodies but was retained diffusely in the nucleoplasm, suggesting the second dsRBD of DCL1 is required for its location to D-bodies [104, 153, 154]. The N-terminal dsRBDs of HYL1 are also reported to have roles in pre-miRNAs processing and entering to D-bodies [156]. These data indicate that the dsRBDs of microprocessors may have major function in the formation of D-bodies. As for the assembly of nuclear speckles, it is well characterized in mammalian cells, the biogenesis of plant nuclear speckles is still not clear [14, 71]. For photobody biogenesis, a nucleolus-tethering system (NoTS) has been proposed and revealed that components of photobodies tethered to nucleolus have the capacity to form body-like structures at the periphery of nucleolus, which contained other components of photobodies, suggesting a self-organization model for the assembly of photobodies [263].

Conclusion and Future Perspectives

Increasing evidences provide broad insights on the nuclear integrity, nuclear compartments biogenesis, and functions, which contribute great significance to understand the nuclear activities both from cellular and molecular levels. NBs are one class of prominent subnuclear structures in the nucleus, assembling in a stochastic fashion or initiated by a seeding component like a protein or RNA related to the bodies. A majority of regulatory complexes locate in NBs to efficiently control the gene expressions. Elucidating the mechanisms of interaction network, dynamic exchange of components in NBs and the maintenance and biogenesis of these structures will be essential to understand their precise functions during growth and development and the potential correlations with the surrounding environment.

More and more plant NBs have been discovered, and they may be reaction, storage, modification, or degradation sites for the components to cooperate with each other to efficiently regulate internal or external signaling pathways, leading to a better growth and development for plants. Though accumulating components in plant NBs are identified, the assembly mechanisms are poorly understood because of

limitation of the methods. Strategies specific to plants are required to be proposed to elucidate the biogenesis mechanisms and their precise functions. Genetic approaches, cellular biology, biochemistry, as well as proteomics need to be combined together to solve this problem in the future. As components in NBs exchanged dynamically and rapidly with the surrounding environment, it is sure that these structures exist much cross talk during signaling transduction. Future work on exploring the nodes coupling different pathways through these bodies will be of great interest. In addition, the morphology of NBs is usually regulated by different internal or external cues, and to find out how the cells use these signals to build up a regulatory network in NBs to adjust to the change of environmental cues is also essential for us to explain the functions of NBs in the future.

Acknowledgments We thank members of Fang's laboratory for insightful discussions. This work was supported by grants to Y. F. from National Natural Science Foundation of China (91319304 and 31171168), National Basic Research Program of China (973 Program, 2012CB910503).

References

1. Cremer T, Cremer M. Chromosome territories. Cold Spring Harb Perspect Biol. 2010 Mar;2(3):a003889.
2. Cremer T, Cremer M, Dietzel S, Muller S, Solovei I, Fakan S. Chromosome territories–a functional nuclear landscape. Curr Opin Cell Biol. 2006 Jun;18(3):307–16.
3. Tiang CL, He Y, Pawlowski WP. Chromosome organization and dynamics during interphase, mitosis, and meiosis in plants. Plant Physiol. 2012 Jan;158(1):26–34.
4. Handwerger KE, Gall JG. Subnuclear organelles: new insights into form and function. Trends Cell Biol. 2006 Jan;16(1):19–26.
5. Meldi L, Brickner JH. Compartmentalization of the nucleus. Trends Cell Biol. 2011 Dec;21(12):701–8.
6. Boulon S, Westman BJ, Hutten S, Boisvert FM, Lamond AI. The nucleolus under stress. Mol Cell. 2010 Oct 22;40(2):216–27.
7. Harrigan JA, Belotserkovskaya R, Coates J, Dimitrova DS, Polo SE, Bradshaw CR, et al. Replication stress induces 53BP1-containing OPT domains in G1 cells. J Cell Biol. 2011 Apr 4;193(1):97–108.
8. Kepkay R, Attwood KM, Ziv Y, Shiloh Y, Dellaire G. KAP1 depletion increases PML nuclear body number in concert with ultrastructural changes in chromatin. Cell Cycle. 2011 Jan 15;10(2):308–22.
9. Dundr M. Nuclear bodies: multifunctional companions of the genome. Curr Opin Cell Biol. 2012 Jun;24(3):415–22.
10. Bernardi R, Pandolfi PP. Role of PML and the PML-nuclear body in the control of programmed cell death. Oncogene. 2003 Dec 8;22(56):9048–57.
11. Shaw PJ, Brown JW. Plant nuclear bodies. Curr Opin Plant Biol. 2004 Dec;7(6):614–20.
12. Mao YS, Zhang B, Spector DL. Biogenesis and function of nuclear bodies. Trends Genet. 2011 Aug;27(8):295–306.
13. Naganuma T, Hirose T. Paraspeckle formation during the biogenesis of long non-coding RNAs. RNA Biol. 2013 Mar 1;10(3):456–61.
14. Shevtsov SP, Dundr M. Nucleation of nuclear bodies by RNA. Nat Cell Biol. 2011 Feb;13(2):167–73.
15. Sasaki YT, Ideue T, Sano M, Mituyama T, Hirose T. MENepsilon/beta noncoding RNAs are essential for structural integrity of nuclear paraspeckles. Proc Natl Acad Sci USA. 2009 Feb 24;106(8):2525–30.

16. Kaiser TE, Intine RV, Dundr M. De novo formation of a subnuclear body. Science. 2008 Dec 12;322(5908):1713–7.
17. Shen TH, Lin HK, Scaglioni PP, Yung TM, Pandolfi PP. The mechanisms of PML-nuclear body formation. Mol Cell. 2006 Nov 3;24(3):331–9.
18. Dundr M, Misteli T. Biogenesis of nuclear bodies. Cold Spring Harb Perspect Biol. 2010 Dec;2(12):a000711.
19. Matera AG, Izaguire-Sierra M, Praveen K, Rajendra TK. Nuclear bodies: random aggregates of sticky proteins or crucibles of macromolecular assembly? Dev Cell. 2009 Nov;17(5):639–47.
20. Boisvert FM, van Koningsbruggen S, Navascues J, Lamond AI. The multifunctional nucleolus. Nat Rev Mol Cell Biol. 2007 Jul;8(7):574–85.
21. Novotny I, Blazikova M, Stanek D, Herman P, Malinsky J. In vivo kinetics of U4/U6.U5 tri-snRNP formation in Cajal bodies. Mol Biol Cell. 2011 Feb 15;22(4):513–23.
22. Klingauf M, Stanek D, Neugebauer KM. Enhancement of U4/U6 small nuclear ribonucleoprotein particle association in Cajal bodies predicted by mathematical modeling. Mol Biol Cell. 2006 Dec;17(12):4972–81.
23. Grimaud C, Bantignies F, Pal-Bhadra M, Ghana P, Bhadra U, Cavalli G. RNAi components are required for nuclear clustering of polycomb group response elements. Cell. 2006 Mar 10;124(5):957–71.
24. Lanzuolo C, Roure V, Dekker J, Bantignies F, Orlando V. Polycomb response elements mediate the formation of chromosome higher-order structures in the bithorax complex. Nat Cell Biol. 2007 Oct;9(10):1167–74.
25. MacPherson MJ, Beatty LG, Zhou W, Du M, Sadowski PD. The CTCF insulator protein is posttranslationally modified by SUMO. Mol Cell Biol. 2009 Feb;29(3):714–25.
26. Kagey MH, Melhuish TA, Wotton D. The polycomb protein Pc2 is a SUMO E3. Cell. 2003 Apr 4;113(1):127–37.
27. Kang X, Qi Y, Zuo Y, Wang Q, Zou Y, Schwartz RJ, et al. SUMO-specific protease 2 is essential for suppression of polycomb group protein-mediated gene silencing during embryonic development. Mol Cell. 2010 Apr 23;38(2):191–201.
28. Spector DL, Lamond AI. Nuclear speckles. Cold Spring Harb Perspect Biol. 2011 Feb;3(2):a000646.
29. Pederson T. The nucleolus. Cold Spring Harb Perspect Biol. 2011 Mar;3(3):a000638.
30. Dousset T, Wang C, Verheggen C, Chen D, Hernandez-Verdun D, Huang S. Initiation of nucleolar assembly is independent of RNA polymerase I transcription. Mol Biol Cell. 2000 Aug;11(8):2705–17.
31. Leung E, Brown JD. Biogenesis of the signal recognition particle. Biochem Soc Trans. 2010 Aug;38(4):1093–8.
32. Politz JC, Yarovoi S, Kilroy SM, Gowda K, Zwieb C, Pederson T. Signal recognition particle components in the nucleolus. Proc Natl Acad Sci U S A. 2000 Jan 4;97(1):55–60.
33. Pederson T. The plurifunctional nucleolus. Nucleic Acids Res. 1998 Sep 1;26(17):3871–6.
34. Shaw PJ, Highett MI, Beven AF, Jordan EG. The nucleolar architecture of polymerase I transcription and processing. EMBO J. 1995 Jun 15;14(12):2896–906.
35. Shaw P, Brown J. Nucleoli: composition, function, and dynamics. Plant Physiol. 2012 Jan;158(1):44–51.
36. Gerbi SA, Borovjagin AV, Lange TS. The nucleolus: a site of ribonucleoprotein maturation. Curr Opin Cell Biol. 2003 Jun;15(3):318–25.
37. Hiscox JA. RNA viruses: hijacking the dynamic nucleolus. Nat Rev Microbiol. 2007 Feb;5(2):119–27.
38. Hiscox JA, Whitehouse A, Matthews DA. Nucleolar proteomics and viral infection. Proteomics. 2010 Nov;10(22):4077–86.
39. Taliansky ME, Brown JW, Rajamaki ML, Valkonen JP, Kalinina NO. Involvement of the plant nucleolus in virus and viroid infections: parallels with animal pathosystems. Adv Virus Res. 2010;77:119–58.
40. Kim SH, Macfarlane S, Kalinina NO, Rakitina DV, Ryabov EV, Gillespie T, et al. Interaction of a plant virus-encoded protein with the major nucleolar protein fibrillarin is required for systemic virus infection. Proc Natl Acad Sci USA. 2007 Jun 26;104(26):11115–20.

41. Canetta E, Kim SH, Kalinina NO, Shaw J, Adya AK, Gillespie T, et al. A plant virus movement protein forms ringlike complexes with the major nucleolar protein, fibrillarin, in vitro. J Mol Biol. 2008 Feb 29;376(4):932–7.
42. Haupt S, Stroganova T, Ryabov E, Kim SH, Fraser G, Duncan G, et al. Nucleolar localization of potato leafroll virus capsid proteins. J Gen Virol. 2005 Oct; 86(Pt 10):2891–6.
43. Rajamaki ML, Valkonen JP. Control of nuclear and nucleolar localization of nuclear inclusion protein a of picorna-like potato virus A in Nicotiana species. Plant Cell. 2009 Aug;21(8):2485–502.
44. Miliani de Marval PL, Zhang Y. The RP-Mdm2-p53 pathway and tumorigenesis. Oncotarget. 2011 Mar;2(3):234–8.
45. Savchenko A, Yurchenko M, Snopok B, Kashuba E. Study on the spatial architecture of p53, MDM2, and p14ARF containing complexes. Mol Biotechnol. 2009 Mar;41(3):270–7.
46. Ivanchuk SM, Mondal S, Rutka JT. p14ARF interacts with DAXX: effects on HDM2 and p53. Cell Cycle. 2008 Jun 15;7(12):1836–50.
47. van Leeuwen IM, Higgins M, Campbell J, McCarthy AR, Sachweh MC, Navarro AM, et al. Modulation of p53 C-terminal acetylation by mdm2, p14ARF, and cytoplasmic SirT2. Mol Cancer Ther. 2013 Apr;12(4):471–80.
48. Chakraborty A, Uechi T, Kenmochi N. Guarding the 'translation apparatus': defective ribosome biogenesis and the p53 signaling pathway. Wiley Interdiscip Rev RNA. 2011 Jul–Aug;2(4):507–22.
49. Prokopowich CD, Gregory TR, Crease TJ. The correlation between rDNA copy number and genome size in eukaryotes. Genome. 2003 Feb;46(1):48–50.
50. Mozgova I, Mokros P, Fajkus J. Dysfunction of chromatin assembly factor 1 induces shortening of telomeres and loss of 45S rDNA in Arabidopsis thaliana. Plant Cell. 2010 Aug;22(8):2768–80.
51. Zhao R, Bodnar MS, Spector DL. Nuclear neighborhoods and gene expression. Curr Opin Genet Dev. 2009 Apr;19(2):172–9.
52. Morimoto M, Boerkoel CF. The role of nuclear bodies in gene expression and disease. Biology (Basel). 2013 Jul 9;2(3):976–1033.
53. Nizami Z, Deryusheva S, Gall JG. The cajal body and histone locus body. Cold Spring Harb Perspect Biol. 2010 Jul;2(7):a000653.
54. Zhu Y, Tomlinson RL, Lukowiak AA, Terns RM, Terns MP. Telomerase RNA accumulates in cajal bodies in human cancer cells. Mol Biol Cell. 2004 Jan;15(1):81–90.
55. Jady BE, Bertrand E, Kiss T. Human telomerase RNA and box H/ACA scaRNAs share a common cajal body-specific localization signal. J Cell Biol. 2004 Mar 1;164(5):647–52.
56. Matera AG, Terns RM, Terns MP. Non-coding RNAs: lessons from the small nuclear and small nucleolar RNAs. Nat Rev Mol Cell Biol. 2007 Mar;8(3):209–20.
57. Cauchi RJ. SMN and Gemins: 'we are family'… or are we?: insights into the partnership between gemins and the spinal muscular atrophy disease protein SMN. Bioessays. 2010 Dec;32(12):1077–89.
58. Buhler D, Raker V, Luhrmann R, Fischer U. Essential role for the tudor domain of SMN in spliceosomal U snRNP assembly: implications for spinal muscular atrophy. Hum Mol Genet. 1999 Dec;8(13):2351–7.
59. Blackburn EH, Collins K. Telomerase: an RNP enzyme synthesizes DNA. Cold Spring Harb Perspect Biol. 2011 May;3(5):a003558.
60. Venteicher AS, Abreu EB, Meng Z, McCann KE, Terns RM, Veenstra TD, et al. A human telomerase holoenzyme protein required for cajal body localization and telomere synthesis. Science. 2009 Jan 30;323(5914):644–8.
61. Mahmoudi S, Henriksson S, Weibrecht I, Smith S, Soderberg O, Stromblad S, et al. WRAP53 is essential for cajal body formation and for targeting the survival of motor neuron complex to cajal bodies. PLoS Biol. 2010;8(11):e1000521.
62. Toyota CG, Davis MD, Cosman AM, Hebert MD. Coilin phosphorylation mediates interaction with SMN and SmB'. Chromosoma. 2010 Apr;119(2):205–15.

63. Hearst SM, Gilder AS, Negi SS, Davis MD, George EM, Whittom AA, et al. Cajal-body formation correlates with differential coilin phosphorylation in primary and transformed cell lines. J Cell Sci. 2009 Jun 1; 122(Pt 11):1872–81.
64. Boudonck K, Dolan L, Shaw PJ. The movement of coiled bodies visualized in living plant cells by the green fluorescent protein. Mol Biol Cell. 1999 Jul;10(7):2297–307.
65. Platani M, Goldberg I, Swedlow JR, Lamond AI. In vivo analysis of cajal body movement, separation, and joining in live human cells. J Cell Biol. 2000 Dec 25;151(7):1561–74.
66. Boudonck K, Dolan L, Shaw PJ. Coiled body numbers in the Arabidopsis root epidermis are regulated by cell type, developmental stage and cell cycle parameters. J Cell Sci. 1998 Dec 18; 111(Pt 24):3687–94.
67. Collier S, Pendle A, Boudonck K, van Rij T, Dolan L, Shaw P. A distant coilin homologue is required for the formation of cajal bodies in Arabidopsis. Mol Biol Cell. 2006 Jul;17(7):2942–51.
68. Strzelecka M, Trowitzsch S, Weber G, Luhrmann R, Oates AC, Neugebauer KM. Coilin-dependent snRNP assembly is essential for zebrafish embryogenesis. Nat Struct Mol Biol. 2010 Apr;17(4):403–9.
69. Liu JL, Wu Z, Nizami Z, Deryusheva S, Rajendra TK, Beumer KJ, et al. Coilin is essential for cajal body organization in *Drosophila melanogaster*. Mol Biol Cell. 2009 Mar;20(6):1661–70.
70. Labadorf A, Link A, Rogers MF, Thomas J, Reddy AS, Ben-Hur A. Genome-wide analysis of alternative splicing in *Chlamydomonas reinhardtii*. BMC Genomics. 2010;11:114.
71. Reddy AS, Day IS, Gohring J, Barta A. Localization and dynamics of nuclear speckles in plants. Plant Physiol. 2012 Jan;158(1):67–77.
72. Reddy AS, Shad Ali G. Plant serine/arginine-rich proteins: roles in precursor messenger RNA splicing, plant development, and stress responses. Wiley Interdiscip Rev RNA. 2011 Nov–Dec;2(6):875–89.
73. Wahl MC, Will CL, Luhrmann R. The spliceosome: design principles of a dynamic RNP machine. Cell. 2009 Feb 20;136(4):701–18.
74. Valadkhan S, Jaladat Y. The spliceosomal proteome: at the heart of the largest cellular ribonucleoprotein machine. Proteomics. 2010 Nov;10(22):4128–41.
75. Lorkovic ZJ. Role of plant RNA-binding proteins in development, stress response and genome organization. Trends Plant Sci. 2009 Apr;14(4):229–36.
76. Richardson DN, Rogers MF, Labadorf A, Ben-Hur A, Guo H, Paterson AH, et al. Comparative analysis of serine/arginine-rich proteins across 27 eukaryotes: insights into sub-family classification and extent of alternative splicing. PLoS One. 2011;6(9):e24542.
77. Reddy AS. Plant serine/arginine-rich proteins and their role in pre-mRNA splicing. Trends Plant Sci. 2004 Nov;9(11):541–7.
78. Manley JL, Krainer AR. A rational nomenclature for serine/arginine-rich protein splicing factors (SR proteins). Genes Dev. 2010 Jun 1;24(11):1073–4.
79. Barta A, Kalyna M, Reddy AS. Implementing a rational and consistent nomenclature for serine/arginine-rich protein splicing factors (SR proteins) in plants. Plant Cell. 2010 Sep;22(9):2926–9.
80. Long JC, Caceres JF. The SR protein family of splicing factors: master regulators of gene expression. Biochem J. 2009 Jan 1;417(1):15–27.
81. Shen H, Green MR. RS domains contact splicing signals and promote splicing by a common mechanism in yeast through humans. Genes Dev. 2006 Jul 1;20(13):1755–65.
82. Dauwalder B, Mattox W. Analysis of the functional specificity of RS domains in vivo. EMBO J. 1998 Oct 15;17(20):6049–60.
83. Birney E, Kumar S, Krainer AR. Analysis of the RNA-recognition motif and RS and RGG domains: conservation in metazoan pre-mRNA splicing factors. Nucleic Acids Res. 1993 Dec 25;21(25):5803–16.
84. Twyffels L, Gueydan C, Kruys V. Shuttling SR proteins: more than splicing factors. FEBS J. 2011 Sep;278(18):3246–55.
85. Barta A, Kalyna M, Lorkovic ZJ. Plant SR proteins and their functions. Curr Top Microbiol Immunol. 2008;326:83–102.

86. Lamond AI, Spector DL. Nuclear speckles: a model for nuclear organelles. Nat Rev Mol Cell Biol. 2003 Aug;4(8):605–12.
87. Lorkovic ZJ, Hilscher J, Barta A. Co-localisation studies of Arabidopsis SR splicing factors reveal different types of speckles in plant cell nuclei. Exp Cell Res. 2008 Oct 15;314(17):3175–86.
88. Fang Y, Hearn S, Spector DL. Tissue-specific expression and dynamic organization of SR splicing factors in Arabidopsis. Mol Biol Cell. 2004 Jun;15(6):2664–73.
89. Rausin G, Tillemans V, Stankovic N, Hanikenne M, Motte P. Dynamic nucleocytoplasmic shuttling of an Arabidopsis SR splicing factor: role of the RNA-binding domains. Plant Physiol. 2010 May;153(1):273–84.
90. Tillemans V, Leponce I, Rausin G, Dispa L, Motte P. Insights into nuclear organization in plants as revealed by the dynamic distribution of Arabidopsis SR splicing factors. Plant Cell. 2006 Nov;18(11):3218–34.
91. Ali GS, Golovkin M, Reddy AS. Nuclear localization and in vivo dynamics of a plant-specific serine/arginine-rich protein. Plant J. 2003 Dec;36(6):883–93.
92. Docquier S, Tillemans V, Deltour R, Motte P. Nuclear bodies and compartmentalization of pre-mRNA splicing factors in higher plants. Chromosoma. 2004 Feb;112(5):255–66.
93. Xiong L, Gong Z, Rock CD, Subramanian S, Guo Y, Xu W, et al. Modulation of abscisic acid signal transduction and biosynthesis by an Sm-like protein in Arabidopsis. Dev Cell. 2001 Dec;1(6):771–81.
94. Xu S, Zhang Z, Jing B, Gannon P, Ding J, Xu F, et al. Transportin-SR is required for proper splicing of resistance genes and plant immunity. PLoS Genet. 2011 Jun;7(6):e1002159.
95. Ali GS, Palusa SG, Golovkin M, Prasad J, Manley JL, Reddy AS. Regulation of plant developmental processes by a novel splicing factor. PLoS One. 2007;2(5):e471.
96. Carvalho RF, Carvalho SD, Duque P. The plant-specific SR45 protein negatively regulates glucose and ABA signaling during early seedling development in Arabidopsis. Plant Physiol. 2010 Oct;154(2):772–83.
97. Zhang XN, Mount SM. Two alternatively spliced isoforms of the Arabidopsis SR45 protein have distinct roles during normal plant development. Plant Physiol. 2009 Jul;150(3):1450–8.
98. Palusa SG, Ali GS, Reddy AS. Alternative splicing of pre-mRNAs of Arabidopsis serine/arginine-rich proteins: regulation by hormones and stresses. Plant J. 2007 Mar;49(6):1091–107.
99. Golovkin M, Reddy AS. An SC35-like protein and a novel serine/arginine-rich protein interact with Arabidopsis U1–70K protein. J Biol Chem. 1999 Dec 17;274(51):36428–38.
100. Savaldi-Goldstein S, Sessa G, Fluhr R. The ethylene-inducible PK12 kinase mediates the phosphorylation of SR splicing factors. Plant J. 2000 Jan;21(1):91–6.
101. Feilner T, Hultschig C, Lee J, Meyer S, Immink RG, Koenig A, et al. High throughput identification of potential Arabidopsis mitogen-activated protein kinases substrates. Mol Cell Proteomics. 2005 Oct;4(10):1558–68.
102. Chen X. Plant microRNAs at a glance. Semin Cell Dev Biol. 2010 Oct;21(8):781.
103. Tang G. Plant microRNAs: an insight into their gene structures and evolution. Semin Cell Dev Biol. 2010 Oct;21(8):782–9.
104. Liu Q, Shi L, Fang Y. Dicing bodies. Plant Physiol. 2012 Jan;158(1):61–6.
105. Denli AM, Tops BB, Plasterk RH, Ketting RF, Hannon GJ. Processing of primary microRNAs by the microprocessor complex. Nature. 2004 Nov 11;432(7014):231–5.
106. Han J, Lee Y, Yeom KH, Nam JW, Heo I, Rhee JK, et al. Molecular basis for the recognition of primary microRNAs by the Drosha-DGCR8 complex. Cell. 2006 Jun 2;125(5):887–901.
107. Zeng Y, Yi R, Cullen BR. Recognition and cleavage of primary microRNA precursors by the nuclear processing enzyme Drosha. EMBO J. 2005 Jan 12;24(1):138–48.
108. Bohnsack MT, Czaplinski K, Gorlich D. Exportin 5 is a RanGTP-dependent dsRNA-binding protein that mediates nuclear export of pre-miRNAs. RNA. 2004 Feb;10(2):185–91.
109. Yi R, Qin Y, Macara IG, Cullen BR. Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. Genes Dev. 2003 Dec 15;17(24):3011–6.

110. Benoit MP, Imbert L, Palencia A, Perard J, Ebel C, Boisbouvier J, et al. The RNA-binding region of human TRBP interacts with microRNA precursors through two independent domains. Nucleic Acids Res. 2013 Apr;41(7):4241–52.
111. Chendrimada TP, Gregory RI, Kumaraswamy E, Norman J, Cooch N, Nishikura K, et al. TRBP recruits the Dicer complex to Ago2 for microRNA processing and gene silencing. Nature. 2005 Aug 4;436(7051):740–4.
112. Kurihara Y, Takashi Y, Watanabe Y. The interaction between DCL1 and HYL1 is important for efficient and precise processing of pri-miRNA in plant microRNA biogenesis. RNA. 2006 Feb;12(2):206–12.
113. Dong Z, Han MH, Fedoroff N. The RNA-binding proteins HYL1 and SE promote accurate in vitro processing of pri-miRNA by DCL1. Proc Natl Acad Sci USA. 2008 Jul 22;105(29):9970–5.
114. Machida S, Chen HY, Adam Yuan Y. Molecular insights into miRNA processing by Arabidopsis thaliana SERRATE. Nucleic Acids Res. 2011 Sep 1;39(17):7828–36.
115. Laubinger S, Sachsenberg T, Zeller G, Busch W, Lohmann JU, Ratsch G, et al. Dual roles of the nuclear cap-binding complex and SERRATE in pre-mRNA splicing and microRNA processing in Arabidopsis thaliana. Proc Natl Acad Sci USA. 2008 Jun 24;105(25):8795–800.
116. Yang L, Liu Z, Lu F, Dong A, Huang H. SERRATE is a novel nuclear regulator in primary microRNA processing in Arabidopsis. Plant J. 2006 Sep;47(6):841–50.
117. Wang L, Song X, Gu L, Li X, Cao S, Chu C, et al. NOT2 proteins promote polymerase II-dependent transcription and interact with multiple MicroRNA biogenesis factors in Arabidopsis. Plant Cell. 2013 Feb;25(2):715–27.
118. Manavella PA, Hagmann J, Ott F, Laubinger S, Franz M, Macek B, et al. Fast-forward genetics identifies plant CPL phosphatases as regulators of miRNA processing factor HYL1. Cell. 2012 Nov 9;151(4):859–70.
119. Yu B, Bi L, Zheng B, Ji L, Chevalier D, Agarwal M, et al. The FHA domain proteins DAWDLE in Arabidopsis and SNIP1 in humans act in small RNA biogenesis. Proc Natl Acad Sci USA. 2008 Jul 22;105(29):10073–8.
120. Ren G, Xie M, Dou Y, Zhang S, Zhang C, Yu B. Regulation of miRNA abundance by RNA binding protein TOUGH in Arabidopsis. Proc Natl Acad Sci USA. 2012 Jul 31;109(31):12817–21.
121. Wu X, Shi Y, Li J, Xu L, Fang Y, Li X, et al. A role for the RNA-binding protein MOS2 in microRNA maturation in Arabidopsis. Cell Res. 2013 May;23(5):645–57.
122. Iwata Y, Takahashi M, Fedoroff NV, Hamdan SM. Dissecting the interactions of SERRATE with RNA and DICER-LIKE 1 in Arabidopsis microRNA precursor processing. Nucleic Acids Res. 2013 Oct;41(19):9129–40.
123. Zhang S, Xie M, Ren G, Yu B. CDC5, a DNA binding protein, positively regulates post-transcriptional processing and/or transcription of primary microRNA transcripts. Proc Natl Acad Sci USA. 2013 Oct 22;110(43):17588–93.
124. Khraiwesh B, Zhu JK, Zhu J. Role of miRNAs and siRNAs in biotic and abiotic stress responses of plants. Biochim Biophys Acta. 2012 Feb;1819(2):137–48.
125. Sunkar R. MicroRNAs with macro-effects on plant stress responses. Semin Cell Dev Biol. 2010 Oct;21(8):805–11.
126. Sunkar R, Li YF, Jagadeeswaran G. Functions of microRNAs in plant stress responses. Trends Plant Sci. 2012 Apr;17(4):196–203.
127. Shukla LI, Chinnusamy V, Sunkar R. The role of microRNAs and other endogenous small RNAs in plant stress responses. Biochim Biophys Acta. 2008 Nov;1779(11):743–8.
128. Sunkar R, Chinnusamy V, Zhu J, Zhu JK. Small RNAs as big players in plant abiotic stress responses and nutrient deprivation. Trends Plant Sci. 2007 Jul;12(7):301–9.
129. Sunkar R, Kapoor A, Zhu JK. Posttranscriptional induction of two Cu/Zn superoxide dismutase genes in Arabidopsis is mediated by downregulation of miR398 and important for oxidative stress tolerance. Plant Cell. 2006 Aug;18(8):2051–65.
130. Chiou TJ, Aung K, Lin SI, Wu CC, Chiang SF, Su CL. Regulation of phosphate homeostasis by MicroRNA in Arabidopsis. Plant Cell. 2006 Feb;18(2):412–21.

131. Bari R, Datt Pant B, Stitt M, Scheible WR. PHO2, microRNA399, and PHR1 define a phosphate-signaling pathway in plants. Plant Physiol. 2006 Jul;141(3):988–99.
132. Lappartient AG, Vidmar JJ, Leustek T, Glass AD, Touraine B. Inter-organ signaling in plants: regulation of ATP sulfurylase and sulfate transporter genes expression in roots mediated by phloem-translocated compound. Plant J. 1999 Apr;18(1):89–95.
133. He H, He L, Gu M. Role of microRNAs in aluminum stress in plants. Plant Cell Rep. 2014:33(6):831–6.
134. Sunkar R, Zhu JK. Novel and stress-regulated microRNAs and other small RNAs from Arabidopsis. Plant Cell. 2004 Aug;16(8):2001–19.
135. Mallory AC, Bartel DP, Bartel B. MicroRNA-directed regulation of Arabidopsis AUXIN RESPONSE FACTOR17 is essential for proper development and modulates expression of early auxin response genes. Plant Cell. 2005 May;17(5):1360–75.
136. Teotia PS, Mukherjee SK, Mishra NS. Fine tuning of auxin signaling by miRNAs. Physiol Mol Biol Plants. 2008 Apr;14(1–2):81–90.
137. Liu HH, Tian X, Li YJ, Wu CA, Zheng CC. Microarray-based analysis of stress-regulated microRNAs in Arabidopsis thaliana. RNA. 2008 May;14(5):836–43.
138. Liu PP, Montgomery TA, Fahlgren N, Kasschau KD, Nonogaki H, Carrington JC. Repression of AUXIN RESPONSE FACTOR10 by microRNA160 is critical for seed germination and post-germination stages. Plant J. 2007 Oct;52(1):133–46.
139. Ding Y, Tao Y, Zhu C. Emerging roles of microRNAs in the mediation of drought stress response in plants. J Exp Bot. 2013 Aug;64(11):3077–86.
140. Lu C, Fedoroff N. A mutation in the Arabidopsis HYL1 gene encoding a dsRNA binding protein affects responses to abscisic acid, auxin, and cytokinin. Plant Cell. 2000 Dec;12(12):2351–66.
141. Zhang JF, Yuan LJ, Shao Y, Du W, Yan DW, Lu YT. The disturbance of small RNA pathways enhanced abscisic acid response and multiple stress responses in Arabidopsis. Plant Cell Environ. 2008 Apr;31(4):562–74.
142. Pelaez P, Sanchez F. Small RNAs in plant defense responses during viral and bacterial interactions: similarities and differences. Front Plant Sci. 2013;4:343.
143. Seo JK, Wu J, Lii Y, Li Y, Jin H. Contribution of small RNA pathway components in plant immunity. Mol Plant Microbe Interact. 2013 Jun;26(6):617–25.
144. Lopez A, Ramirez V, Garcia-Andrade J, Flors V, Vera P. The RNA silencing enzyme RNA polymerase v is required for plant immunity. PLoS Genet. 2011 Dec;7(12):e1002434.
145. Zhang X, Zhao H, Gao S, Wang WC, Katiyar-Agarwal S, Huang HD, et al. Arabidopsis Argonaute 2 regulates innate immunity via miRNA393(*)-mediated silencing of a Golgi-localized SNARE gene, MEMB12. Mol Cell. 2011 May 6;42(3):356–66.
146. Garcia-Ruiz H, Takeda A, Chapman EJ, Sullivan CM, Fahlgren N, Brempelis KJ, et al. Arabidopsis RNA-dependent RNA polymerases and dicer-like proteins in antiviral defense and small interfering RNA biogenesis during Turnip mosaic virus infection. Plant Cell. 2010 Feb;22(2):481–96.
147. Agorio A, Vera P. ARGONAUTE4 is required for resistance to *Pseudomonas syringae* in Arabidopsis. Plant Cell. 2007 Nov;19(11):3778–90.
148. Deleris A, Gallego-Bartolome J, Bao J, Kasschau KD, Carrington JC, Voinnet O. Hierarchical action and inhibition of plant Dicer-like proteins in antiviral defense. Science. 2006 Jul 7;313(5783):68–71.
149. Navarro L, Jay F, Nomura K, He SY, Voinnet O. Suppression of the microRNA pathway by bacterial effector proteins. Science. 2008 Aug 15;321(5891):964–7.
150. Katiyar-Agarwal S, Gao S, Vivian-Smith A, Jin H. A novel class of bacteria-induced small RNAs in Arabidopsis. Genes Dev. 2007 Dec 1;21(23):3123–34.
151. Katiyar-Agarwal S, Morgan R, Dahlbeck D, Borsani O, Villegas A Jr, Zhu JK, et al. A pathogen-inducible endogenous siRNA in plant immunity. Proc Natl Acad Sci USA. 2006 Nov 21;103(47):18002–7.
152. Song L, Han MH, Lesicka J, Fedoroff N. Arabidopsis primary microRNA processing proteins HYL1 and DCL1 define a nuclear body distinct from the Cajal body. Proc Natl Acad Sci USA. 2007 Mar 27;104(13):5437–42.

153. Fang Y, Spector DL. Identification of nuclear dicing bodies containing proteins for microRNA biogenesis in living Arabidopsis plants. Curr Biol. 2007 May 1;17(9):818–23.
154. Liu Q, Yan Q, Liu Y, Hong F, Sun Z, Shi L, et al. Complementation of HYPONASTIC LEAVES1 by double-strand RNA-binding domains of DICER-LIKE1 in nuclear dicing bodies. Plant Physiol. 2013 Sep;163(1):108–17.
155. Rasia RM, Mateos J, Bologna NG, Burdisso P, Imbert L, Palatnik JF, et al. Structure and RNA interactions of the plant MicroRNA processing-associated protein HYL1. Biochemistry. 2010 Sep 28;49(38):8237–9.
156. Wu F, Yu L, Cao W, Mao Y, Liu Z, He Y. The N-terminal double-stranded RNA binding domains of Arabidopsis HYPONASTIC LEAVES1 are sufficient for pre-microRNA processing. Plant Cell. 2007 Mar;19(3):914–25.
157. Arsovski AA, Galstyan A, Guseman JM, Nemhauser JL. Photomorphogenesis. Arabidopsis Book. 2012;10:e0147.
158. Kami C, Lorrain S, Hornitschek P, Fankhauser C. Light-regulated plant growth and development. Curr Top Dev Biol. 2010;91:29–66.
159. Jiao Y, Lau OS, Deng XW. Light-regulated transcriptional networks in higher plants. Nat Rev Genet. 2007 Mar;8(3):217–30.
160. Rizzini L, Favory JJ, Cloix C, Faggionato D, O'Hara A, Kaiserli E, et al. Perception of UV-B by the Arabidopsis UVR8 protein. Science. 2011 Apr 1;332(6025):103–6.
161. Kaiserli E, Jenkins GI. UV-B promotes rapid nuclear translocation of the Arabidopsis UV-B specific signaling component UVR8 and activates its function in the nucleus. Plant Cell. 2007 Aug;19(8):2662–73.
162. Quail PH. Phytochromes. Curr Biol. 2010 Jun 22;20(12):R504–7.
163. Yu X, Liu H, Klejnot J, Lin C. The cryptochrome blue light receptors. Arabidopsis Book. 2010 Sep 23;8:e0135.
164. Briggs WR, Christie JM. Phototropins 1 and 2: versatile plant blue-light receptors. Trends Plant Sci. 2002 May;7(5):204–10.
165. Mizoguchi T, Coupland G. ZEITLUPE and FKF1: novel connections between flowering time and circadian clock control. Trends Plant Sci. 2000 Oct;5(10):409–11.
166. Nelson DC, Lasswell J, Rogg LE, Cohen MA, Bartel B. FKF1, a clock-controlled gene that regulates the transition to flowering in Arabidopsis. Cell. 2000 Apr 28;101(3):331–40.
167. Somers DE, Schultz TF, Milnamow M, Kay SA. ZEITLUPE encodes a novel clock-associated PAS protein from Arabidopsis. Cell. 2000 Apr 28;101(3):319–29.
168. Schultz TF, Kiyosue T, Yanovsky M, Wada M, Kay SA. A role for LKP2 in the circadian clock of Arabidopsis. Plant Cell. 2001 Dec;13(12):2659–70.
169. Kevei E, Schafer E, Nagy F. Light-regulated nucleo-cytoplasmic partitioning of phytochromes. J Exp Bot. 2007;58(12):3113–24.
170. Nagatani A. Light-regulated nuclear localization of phytochromes. Curr Opin Plant Biol. 2004 Dec;7(6):708–11.
171. Van Buskirk EK, Decker PV, Chen M. Photobodies in light signaling. Plant Physiol. 2012 Jan;158(1):52–60.
172. Kircher S, Kozma-Bognar L, Kim L, Adam E, Harter K, Schafer E, et al. Light quality-dependent nuclear import of the plant photoreceptors phytochrome A and B. Plant Cell. 1999 Aug;11(8):1445–56.
173. Yamaguchi R, Nakamura M, Mochizuki N, Kay SA, Nagatani A. Light-dependent translocation of a phytochrome B-GFP fusion protein to the nucleus in transgenic Arabidopsis. J Cell Biol. 1999 May 3;145(3):437–45.
174. Kircher S, Gil P, Kozma-Bognar L, Fejes E, Speth V, Husselstein-Muller T, et al. Nucleocytoplasmic partitioning of the plant photoreceptors phytochrome A, B, C, D, and E is regulated differentially by light and exhibits a diurnal rhythm. Plant Cell. 2002 Jul;14(7):1541–55.
175. Chen M. Phytochrome nuclear body: an emerging model to study interphase nuclear dynamics and signaling. Curr Opin Plant Biol. 2008 Oct;11(5):503–8.
176. Favory JJ, Stec A, Gruber H, Rizzini L, Oravecz A, Funk M, et al. Interaction of COP1 and UVR8 regulates UV-B-induced photomorphogenesis and stress acclimation in Arabidopsis. EMBO J. 2009 Mar 4;28(5):591–601.

177. Yu X, Sayegh R, Maymon M, Warpeha K, Klejnot J, Yang H, et al. Formation of nuclear bodies of Arabidopsis CRY2 in response to blue light is associated with its blue light-dependent degradation. Plant Cell. 2009 Jan;21(1):118–30.
178. Gu NN, Zhang YC, Yang HQ. Substitution of a conserved glycine in the PHR domain of Arabidopsis cryptochrome 1 confers a constitutive light response. Mol Plant. 2012 Jan;5(1):85–97.
179. Lian HL, He SB, Zhang YC, Zhu DM, Zhang JY, Jia KP, et al. Blue-light-dependent interaction of cryptochrome 1 with SPA1 defines a dynamic signaling mechanism. Genes Dev. 2011 May 15;25(10):1023–8.
180. Liu B, Zuo Z, Liu H, Liu X, Lin C. Arabidopsis cryptochrome 1 interacts with SPA1 to suppress COP1 activity in response to blue light. Genes Dev. 2011 May 15;25(10):1029–34.
181. Chen M, Schwab R, Chory J. Characterization of the requirements for localization of phytochrome B to nuclear bodies. Proc Natl Acad Sci USA. 2003 Nov 25;100(24):14493–8.
182. Chen M, Chory J. Phytochrome signaling mechanisms and the control of plant development. Trends Cell Biol. 2011 Nov;21(11):664–71.
183. Rausenberger J, Hussong A, Kircher S, Kirchenbauer D, Timmer J, Nagy F, et al. An integrative model for phytochrome B mediated photomorphogenesis: from protein dynamics to physiology. PLoS One. 2010;5(5):e10721.
184. Bauer D, Viczian A, Kircher S, Nobis T, Nitschke R, Kunkel T, et al. Constitutive photomorphogenesis 1 and multiple photoreceptors control degradation of phytochrome interacting factor 3, a transcription factor required for light signaling in Arabidopsis. Plant Cell. 2004 Jun;16(6):1433–45.
185. Feng CM, Qiu Y, Van Buskirk EK, Yang EJ, Chen M. Light-regulated gene repositioning in Arabidopsis. Nat Commun. 2014 Jan 6;5:3027.
186. Zhang X, Min JH, Huang P, Chung JS, Lee KH, Kim CS. AtSKIP functions as a mediator between cytokinin and light signaling pathway in Arabidopsis thaliana. Plant Cell Rep. 2013 Nov 22:33(3):401–9.
187. Lariguet P, Ranocha P, De Meyer M, Barbier O, Penel C, Dunand C. Identification of a hydrogen peroxide signalling pathway in the control of light-dependent germination in Arabidopsis. Planta. 2013 Aug;238(2):381–95.
188. Svyatyna K, Riemann M. Light-dependent regulation of the jasmonate pathway. Protoplasma. 2012 Jun; 249(Suppl 2):S137–45.
189. Weller JL, Hecht V, Vander Schoor JK, Davidson SE, Ross JJ. Light regulation of gibberellin biosynthesis in pea is mediated through the COP1/HY5 pathway. Plant Cell. 2009 Mar;21(3):800–13.
190. Hua J. Modulation of plant immunity by light, circadian rhythm, and temperature. Curr Opin Plant Biol. 2013 Aug;16(4):406–13.
191. Kaur N, Li J, Hu J. Peroxisomes and photomorphogenesis. Subcell Biochem. 2013;69:195–211.
192. Hu J, Aguirre M, Peto C, Alonso J, Ecker J, Chory J. A role for peroxisomes in photomorphogenesis and development of Arabidopsis. Science. 2002 Jul 19;297(5580):405–9.
193. Carvalho RF, Campos ML, Azevedo RA. The role of phytochrome in stress tolerance. J Integr Plant Biol. 2011 Dec;53(12):920–9.
194. Lau OS, Deng XW. Plant hormone signaling lightens up: integrators of light and hormones. Curr Opin Plant Biol. 2010 Oct;13(5):571–7.
195. Franklin KA. Light and temperature signal crosstalk in plant development. Curr Opin Plant Biol. 2009 Feb;12(1):63–8.
196. Bai S, Yao T, Li M, Guo X, Zhang Y, Zhu S, et al. PIF3 is involved in the primary root growth inhibition of Arabidopsis induced by nitric oxide in the light. Mol Plant. 2014;7(4):616–25.
197. Jeong J, Choi G. Phytochrome-interacting factors have both shared and distinct biological roles. Mol Cells. 2013 May;35(5):371–80.
198. Leivar P, Quail PH. PIFs: pivotal components in a cellular signaling hub. Trends Plant Sci. 2011 Jan;16(1):19–28.
199. Quail PH. Phytochrome-interacting factors. Semin Cell Dev Biol. 2000 Dec;11(6):457–66.

200. Feng S, Martinez C, Gusmaroli G, Wang Y, Zhou J, Wang F, et al. Coordinated regulation of Arabidopsis thaliana development by light and gibberellins. Nature. 2008 Jan 24;451(7177):475–9.
201. de Lucas M, Daviere JM, Rodriguez-Falcon M, Pontin M, Iglesias-Pedraz JM, Lorrain S, et al. A molecular framework for light and gibberellin control of cell elongation. Nature. 2008 Jan 24;451(7177):480–4.
202. Koini MA, Alvey L, Allen T, Tilley CA, Harberd NP, Whitelam GC, et al. High temperature-mediated adaptations in plant architecture require the bHLH transcription factor PIF4. Curr Biol. 2009 Mar 10;19(5):408–13.
203. Zhong S, Shi H, Xue C, Wang L, Xi Y, Li J, et al. A molecular framework of light-controlled phytohormone action in Arabidopsis. Curr Biol. 2012 Aug 21;22(16):1530–5.
204. Yu Y, Wang J, Zhang Z, Quan R, Zhang H, Deng XW, et al. Ethylene promotes hypocotyl growth and HY5 degradation by enhancing the movement of COP1 to the nucleus in the light. PLoS Genet. 2013 Dec;9(12):e1004025.
205. Ram H, Chattopadhyay S. Molecular interaction of bZIP domains of GBF1, HY5 and HYH in Arabidopsis seedling development. Plant Signal Behav. 2012 Nov 3;8(1):e22703.
206. Shi QM, Yang X, Song L, Xue HW. Arabidopsis MSBP1 is activated by HY5 and HYH and is involved in photomorphogenesis and brassinosteroid sensitivity regulation. Mol Plant. 2011 Nov;4(6):1092–104.
207. Zhang Y, Liu Z, Liu R, Hao H, Bi Y. Gibberellins negatively regulate low temperature-induced anthocyanin accumulation in a HY5/HYH-dependent manner. Plant Signal Behav. 2011 May;6(5):632–4.
208. Chen H, Xiong L. Genetic interaction of two abscisic acid signaling regulators, HY5 and FIERY1, in mediating lateral root formation. Plant Signal Behav. 2011 Jan;6(1):123–5.
209. Chen H, Xiong L. Role of HY5 in abscisic acid response in seeds and seedlings. Plant Signal Behav. 2008 Nov;3(11):986–8.
210. Vandenbussche F, Habricot Y, Condiff AS, Maldiney R, Van der Straeten D, Ahmad M. HY5 is a point of convergence between cryptochrome and cytokinin signalling pathways in *Arabidopsis thaliana*. Plant J. 2007 Feb;49(3):428–41.
211. Sibout R, Sukumar P, Hettiarachchi C, Holm M, Muday GK, Hardtke CS. Opposite root growth phenotypes of hy5 versus hy5 hyh mutants correlate with increased constitutive auxin signaling. PLoS Genet. 2006 Nov 24;2(11):e202.
212. Yi C, Deng XW. COP1– from plant photomorphogenesis to mammalian tumorigenesis. Trends Cell Biol. 2005 Nov;15(11):618–25.
213. Lau OS, Deng XW. The photomorphogenic repressors COP1 and DET1: 20 years later. Trends Plant Sci. 2012 Oct;17(10):584–93.
214. Luo XM, Lin WH, Zhu S, Zhu JY, Sun Y, Fan XY, et al. Integration of light- and brassinosteroid-signaling pathways by a GATA transcription factor in Arabidopsis. Dev Cell. 2010 Dec 14;19(6):872–83.
215. Jeong RD, Chandra-Shekara AC, Barman SR, Navarre D, Klessig DF, Kachroo A, et al. Cryptochrome 2 and phototropin 2 regulate resistance protein-mediated viral defense by negatively regulating an E3 ubiquitin ligase. Proc Natl Acad Sci U S A. 2010 Jul 27;107(30):13538–43.
216. Matsushita T, Mochizuki N, Nagatani A. Dimers of the N-terminal domain of phytochrome B are functional in the nucleus. Nature. 2003 Jul 31;424(6948):571–4.
217. Palagyi A, Terecskei K, Adam E, Kevei E, Kircher S, Merai Z, et al. Functional analysis of amino-terminal domains of the photoreceptor phytochrome B. Plant Physiol. 2010 Aug;153(4):1834–45.
218. Zuo Z, Liu H, Liu B, Liu X, Lin C. Blue light-dependent interaction of CRY2 with SPA1 regulates COP1 activity and floral initiation in Arabidopsis. Curr Biol. 2011 May 24;21(10):841–7.
219. Chang CS, Maloof JN, Wu SH. COP1-mediated degradation of BBX22/LZF1 optimizes seedling development in Arabidopsis. Plant Physiol. 2011 May;156(1):228–39.

220. Jang IC, Henriques R, Seo HS, Nagatani A, Chua NH. Arabidopsis PHYTOCHROME INTERACTING FACTOR proteins promote phytochrome B polyubiquitination by COP1 E3 ligase in the nucleus. Plant Cell. 2010 Jul;22(7):2370–83.
221. Yu JW, Rubio V, Lee NY, Bai S, Lee SY, Kim SS, et al. COP1 and ELF3 control circadian function and photoperiodic flowering by regulating GI stability. Mol Cell. 2008 Dec 5;32(5):617–30.
222. Liu LJ, Zhang YC, Li QH, Sang Y, Mao J, Lian HL, et al. COP1-mediated ubiquitination of CONSTANS is implicated in cryptochrome regulation of flowering in Arabidopsis. Plant Cell. 2008 Feb;20(2):292–306.
223. Jang IC, Yang JY, Seo HS, Chua NH. HFR1 is targeted by COP1 E3 ligase for post-translational proteolysis during phytochrome A signaling. Genes Dev. 2005 Mar 1;19(5):593–602.
224. Saijo Y, Sullivan JA, Wang H, Yang J, Shen Y, Rubio V, et al. The COP1-SPA1 interaction defines a critical step in phytochrome A-mediated regulation of HY5 activity. Genes Dev. 2003 Nov 1;17(21):2642–7.
225. Yang SW, Jang IC, Henriques R, Chua NH. FAR-RED ELONGATED HYPOCOTYL1 and FHY1-LIKE associate with the Arabidopsis transcription factors LAF1 and HFR1 to transmit phytochrome A signals for inhibition of hypocotyl elongation. Plant Cell. 2009 May;21(5):1341–59.
226. Su YS, Lagarias JC. Light-independent phytochrome signaling mediated by dominant GAF domain tyrosine mutants of Arabidopsis phytochromes in transgenic plants. Plant Cell. 2007 Jul;19(7):2124–39.
227. Kumari S, Roy S, Singh P, Singla-Pareek SL, Pareek A. Cyclophilins: proteins in search of function. Plant Signal Behav. 2012 Nov 3;8(1):e22734.
228. Wang P, Heitman J. The cyclophilins. Genome Biol. 2005;6(7):226.
229. Schiene C, Fischer G. Enzymes that catalyse the restructuring of proteins. Curr Opin Struct Biol. 2000 Feb;10(1):40–5.
230. Lorkovic ZJ, Lopato S, Pexa M, Lehner R, Barta A. Interactions of Arabidopsis RS domain containing cyclophilins with SR proteins and U1 and U11 small nuclear ribonucleoprotein-specific proteins suggest their involvement in pre-mRNA Splicing. J Biol Chem. 2004 Aug 6;279(32):33890–8.
231. Bannikova O, Zywicki M, Marquez Y, Skrahina T, Kalyna M, Barta A. Identification of RNA targets for the nuclear multidomain cyclophilin atCyp59 and their effect on PPIase activity. Nucleic Acids Res. 2013 Feb 1;41(3):1783–96.
232. Gullerova M, Barta A, Lorkovic ZJ. AtCyp59 is a multidomain cyclophilin from *Arabidopsis thaliana* that interacts with SR proteins and the C-terminal domain of the RNA polymerase II. RNA. 2006 Apr;12(4):631–43.
233. Finkelstein RR, Gampala SS, Rock CD. Abscisic acid signaling in seeds and seedlings. Plant Cell. 2002;14 (Suppl):S15–45.
234. Lee SC, Lim CW, Lan W, He K, Luan S. ABA signaling in guard cells entails a dynamic protein-protein interaction relay from the PYL-RCAR family receptors to ion channels. Mol Plant. 2013 Mar;6(2):528–38.
235. Lesicka-Gorecka J, Szarzynska B, Sawczak M, Bagdiul I, Gorski P, Jarmolowski A, et al. Abscisic acid does not influence the subcellular distribution of the HYL1 protein from *Arabidopsis thaliana*. Acta Biochim Pol. 2008;55(3):517–24.
236. Hugouvieux V, Kwak JM, Schroeder JI. An mRNA cap binding protein, ABH1, modulates early abscisic acid signal transduction in Arabidopsis. Cell. 2001 Aug 24;106(4):477–87.
237. Ng CK, Kinoshita T, Pandey S, Shimazaki K, Assmann SM. Abscisic acid induces rapid subnuclear reorganization in guard cells. Plant Physiol. 2004 Apr;134(4):1327–31.
238. Li J, Wang XQ, Watson MB, Assmann SM. Regulation of abscisic acid-induced stomatal closure and anion channels by guard cell AAPK kinase. Science. 2000 Jan 14;287(5451):300–3.
239. Pendle AF, Clark GP, Boon R, Lewandowska D, Lam YW, Andersen J, et al. Proteomic analysis of the Arabidopsis nucleolus suggests novel nucleolar functions. Mol Biol Cell. 2005 Jan;16(1):260–9.

240. Koroleva OA, Calder G, Pendle AF, Kim SH, Lewandowska D, Simpson CG, et al. Dynamic behavior of Arabidopsis eIF4A-III, putative core protein of exon junction complex: fast relocation to nucleolus and splicing speckles under hypoxia. Plant Cell. 2009 May;21(5):1592–606.
241. Lorkovic ZJ, Barta A. Role of Cajal bodies and nucleolus in the maturation of the U1 snRNP in Arabidopsis. PLoS One. 2008;3(12):e3989.
242. Politz JC, Hogan EM, Pederson T. MicroRNAs with a nucleolar location. RNA. 2009 Sep;15(9):1705–15.
243. Taft RJ, Simons C, Nahkuri S, Oey H, Korbie DJ, Mercer TR, et al. Nuclear-localized tiny RNAs are associated with transcription initiation and splice sites in metazoans. Nat Struct Mol Biol. 2010 Aug;17(8):1030–4.
244. Scott MS, Avolio F, Ono M, Lamond AI, Barton GJ. Human miRNA precursors with box H/ACA snoRNA features. PLoS Comput Biol. 2009 Sep;5(9):e1000507.
245. Kim SH, Koroleva OA, Lewandowska D, Pendle AF, Clark GP, Simpson CG, et al. Aberrant mRNA transcripts and the nonsense-mediated decay proteins UPF2 and UPF3 are enriched in the Arabidopsis nucleolus. Plant Cell. 2009 Jul;21(7):2045–57.
246. Platani M, Goldberg I, Lamond AI, Swedlow JR. Cajal body dynamics and association with chromatin are ATP-dependent. Nat Cell Biol. 2002 Jul;4(7):502–8.
247. Onodera Y, Haag JR, Ream T, Costa Nunes P, Pontes O, Pikaard CS. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. Cell. 2005 Mar 11;120(5):613–22.
248. Herr AJ, Jensen MB, Dalmay T, Baulcombe DC. RNA polymerase IV directs silencing of endogenous DNA. Science. 2005 Apr 1;308(5718):118–20.
249. Pontes O, Vitins A, Ream TS, Hong E, Pikaard CS, Costa-Nunes P. Intersection of small RNA pathways in *Arabidopsis thaliana* sub-nuclear domains. PLoS One. 2013;8(6):e65652.
250. Dou K, Huang CF, Ma ZY, Zhang CJ, Zhou JX, Huang HW, et al. The PRP6-like splicing factor STA1 is involved in RNA-directed DNA methylation by facilitating the production of Pol V-dependent scaffold RNAs. Nucleic Acids Res. 2013 Oct;41(18):8489–502.
251. Ben Chaabane S, Liu R, Chinnusamy V, Kwon Y, Park JH, Kim SY, et al. STA1, an Arabidopsis pre-mRNA processing factor 6 homolog, is a new player involved in miRNA biogenesis. Nucleic Acids Res. 2013 Feb 1;41(3):1984–97.
252. Ausin I, Greenberg MV, Li CF, Jacobsen SE. The splicing factor SR45 affects the RNA-directed DNA methylation pathway in Arabidopsis. Epigenetics. 2012 Jan 1;7(1):29–33.
253. Raczynska KD, Stepien A, Kierzkowski D, Kalak M, Bajczyk M, McNicol J, et al. The SERRATE protein is involved in alternative splicing in *Arabidopsis thaliana*. Nucleic Acids Res. 2013 Oct 16;42(2):1224–44.
254. Wu H, Sun S, Tu K, Gao Y, Xie B, Krainer AR, et al. A splicing-independent function of SF2/ASF in microRNA processing. Mol Cell. 2010 Apr 9;38(1):67–77.
255. Shikata H, Nakashima M, Matsuoka K, Matsushita T. Deletion of the RS domain of RRC1 impairs phytochrome B signaling in Arabidopsis. Plant Signal Behav. 2012 Aug;7(8):933–6.
256. Shikata H, Shibata M, Ushijima T, Nakashima M, Kong SG, Matsuoka K, et al. The RS domain of Arabidopsis splicing factor RRC1 is required for phytochrome B signal transduction. Plant J. 2012 Jun;70(5):727–38.
257. Dundr M. Nucleation of nuclear bodies. Methods Mol Biol. 2013;1042:351–64.
258. Carmo-Fonseca M, Rino J. RNA seeds nuclear bodies. Nat Cell Biol. 2011 Feb;13(2):110–2.
259. Belaya K, St Johnston D. Using the mRNA-MS2/MS2CP-FP system to study mRNA transport during Drosophila oogenesis. Methods Mol Biol. 2011;714:265–83.
260. Keryer-Bibens C, Barreau C, Osborne HB. Tethering of proteins to RNAs by bacteriophage proteins. Biol Cell. 2008 Feb;100(2):125–38.
261. Querido E, Chartrand P. Using fluorescent proteins to study mRNA trafficking in living cells. Methods Cell Biol. 2008;85:273–92.

262. Schonberger J, Hammes UZ, Dresselhaus T. In vivo visualization of RNA in plants cells using the lambdaN(2)(2) system and a GATEWAY-compatible vector series for candidate RNAs. Plant J. 2012 Jul;71(1):173–81.
263. Liu Y, Liu Q, Yan Q, Shi L, Fang Y. Nucleolus-tethering system (NoTS) reveals that assembly of photobodies follows a self-organization model. Mol Biol Cell. 2014 Apr;25(8):1366–73.
264. Mao YS, Sunwoo H, Zhang B, Spector DL. Direct visualization of the co-transcriptional assembly of a nuclear body by noncoding RNAs. Nat Cell Biol. 2011 Jan;13(1):95–101.
265. Bernardi R, Pandolfi PP. Structure, dynamics and functions of promyelocytic leukaemia nuclear bodies. Nat Rev Mol Cell Biol. 2007 Dec;8(12):1006–16.
266. Lallemand-Breitenbach V, de The H. PML nuclear bodies. Cold Spring Harb Perspect Biol. 2010 May;2(5):a000661.
267. Dundr M. Seed and grow: a two-step model for nuclear body biogenesis. J Cell Biol. 2011 May 16;193(4):605–6.
268. Galvao RM, Li M, Kothadia SM, Haskel JD, Decker PV, Van Buskirk EK, et al. Photoactivated phytochromes interact with HEMERA and promote its accumulation to establish photomorphogenesis in Arabidopsis. Genes Dev. 2012 Aug 15;26(16):1851–63.
269. Chen M, Galvao RM, Li M, Burger B, Bugea J, Bolado J, et al. Arabidopsis HEMERA/pTAC12 initiates photomorphogenesis by phytochromes. Cell. 2010 Jun 25;141(7):1230–40.
270. Jacobsen SE, Running MP, Meyerowitz EM. Disruption of an RNA helicase/RNAse III gene in Arabidopsis causes unregulated cell division in floral meristems. Development. 1999 Dec;126(23):5231–43.

Chapter 4
Plasticity of Chromatin Organization in the Plant Interphase Nucleus

Three-Dimensional Organization of the Interphase Nucleus

Ana Paula Santos, Rita Abranches, Margarida Oliveira and Peter Shaw

Abbreviations

3D	Three dimensional
3C	Chromosome conformation capture
5-AC	5-azacytidine
BAC	Bacterial artificial chromosome
CTs	Chromosome territories
BiFC	Bimolecular fluorescence complementation
BrUTP	Bromouridine triphosphate
FISH	Fluorescence in situ hybridization
GFP	Green fluorescent protein
GISH	Genomic in situ hybridization
HPT	Hygromycin phosphotransferase
ISH	In situ hybridization
lacO	Lactose operator
NGS	Next generation sequencing
NOR	Nucleolus organizing region
SCD	"Spherical 1-Mb chromatin domain" model
tetO	Tetracycline operator
TSA	Trichostatin A

A. P. Santos (✉) · M. Oliveira
Genomics of Plant Stress Unit, Instituto de Biologia Experimental e Tecnológica, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Av. da República, 2780–157 Oeiras, Portugal
e-mail: apsantos@itqb.unl.pt

R. Abranches
Plant Cell Biology laboratory, Instituto de Tecnologia Química e Biológica, Universidade Nova de Lisboa, Av. da República, 2780–157 Oeiras, Portugal

P. Shaw
Department of Cell and Developmental Biology, John Innes Centre, NR4 7UH Norwich, UK

O. Pontes, H. Jin (eds.), *Nuclear Functions in Plant Transcription, Signaling and Development,* DOI 10.1007/978-1-4939-2386-1_4

Historical Landmarks Along the History of Three-Dimensional (3D) Organization of the Interphase Nucleus

The eukaryotic cell nucleus is one of the best known but probably the least understood of cellular organelles. As highlighted by Raices and D'Angelo [1], the term nucleus (from Latin nucleus or nuculeus, meaning kernel) was first used by Robert Brown to describe an opaque spot with a round or oval shape that was present in all plant cells. The term chromosome comes from the Greek words *khroma*, meaning colour, and *soma* meaning body and was first coined by Waldeyer-Hartz [2]. The first observations of chromosomes were made in the second half of the nineteenth century, mainly thanks to the use of conventional DNA-staining techniques using various dyes. Until the late 1960s, these observations were restricted to plant chromosomes during mitotic and meiotic divisions [3]. Animal cytogenetics only became established later on after solving specific experimental problems related to the maintenance of cell cultures and the use of cytoskeleton drugs to obtain metaphase plates. The interphase nucleus remained a mystery for a long time, although the use of conventional microscopy and specific dyes started to enable the identification of distinct chromatin domains based on the labelling intensity; the euchromatin fraction and the nucleoli were weakly labelled, while the heterochromatin regions had characteristically high staining intensity [4].

The Origin of the Concept of Interphase Chromosome Territories

The structure and functional organization of the nucleus has been the subject of much debate, much of it centred on questions such as: To what extent is the nucleus organized? Do chromosomes occupy specific nuclear territories and are they mutually exclusive or interwoven? Do chromosomes have a random or non-random distribution in the nucleus? What are the rules underlying chromatin organization and at what levels can the organization be plastic and changeable?

The work of Carl Rabl [5] contributed decisively to the understanding of nuclear organization. Based on observations of mitotic cells from *Salamandra maculata*, Rabl [5] was able to predict the existence of chromosome territories (CTs) in the interphase nucleus even without being able to actually see them. Rabl's crucial prediction was that the arrangement of anaphase chromosomes would persist during interphase. At that time, the terms centromeres and telomeres had not yet emerged, but Rabl [5] proposed that the chromosomal regions connected to the spindle would be clustered at one side of the nucleus while the terminal regions of chromosomes would be positioned at the opposite poles of the interphase nucleus. This configuration of chromosomes during interphase was later termed the "Rabl configuration". Boveri [6] refined the concept of CTs by postulating that each chromosome would occupy specific domains within the interphase nucleus, maintaining its identity and integrity without intermingling with other chromosomes. Nevertheless, there was a period between the 1950s and the 1970s in which the existence of CTs was ques-

tioned, mainly due to the failure of electron microscopy to detect them [7]. As a consequence, during the 1970s and 1980s, the interphase nucleus was mainly seen as a poorly organized organelle containing intermingling chromatin fibres, often described by the analogy of "spaghetti in a bowl". Only later was the hypothesis of Boveri [6] experimentally validated by Cremer et al. [8]. Their experiment consisted of irradiating defined regions of the interphase nucleus of mammalian cells with a laser–UV microbeam; it was then found that only a few chromosomes of the subsequent mitosis sustained highly localized damage due to the irradiation [8]. The interpretation of these results favoured a model in which chromosomes occupy specific non-overlapping domains or territories rather than a random model. This work was absolutely essential to the foundation of the new research area of interphase cytogenetics in which a major goal was the *in situ* visualization of CTs. This was achieved by the development of the *in situ* hybridization (ISH) technique with specific DNA probes together with important improvements on microscope resolution. Initially, ISH was based on using radiolabeled probes [9] and only later, in the 1980s, did fluorescence *in situ* hybridization (FISH) emerge as a technique for DNA or RNA detection using fluorescent tags. Furthermore, fluorescence microscopy was developed to detect and visualize *in situ* specific DNA or RNA sequences in squashed nuclei or in tissue sections. FISH was first applied in mammalian interphase nuclei to visualize individual interphase chromosomes, giving rise to the term "chromosome painting" [10, 11]. In summary, the observation that chromosomes, in both animal and plant cell interphase nuclei, occupy distinct territories was made possible by the emergence of ISH techniques. These observations represented the unequivocal confirmation of Rabl's and Boveri's hypotheses that chromosomes occupy distinct CTs, rather than being diffusely dispersed in the interphase nucleus as reviewed by Cremer et al. [12, 13], Cremer and Cremer [14–16] and Meaburn and Misteli [17]. It should be emphasised that in plants the observation of chromosome territorial organization was not straightforward mainly due to the inefficient blocking of dispersed repeats and insufficient signal intensity of short unique sequences [18]. The history of nuclear organization in plants is addressed later in this chapter.

Models for the Organization of Chromosome Territories

Shortly after the discovery that chromosomes occupy distinct territories in the interphase nucleus, it was realized that interphase chromosome organization was more complex than revealed by *in situ* chromosome paints, and that it was difficult to understand mechanistically the chromosome architecture during the interphase. The question of whether different CTs are completely separated or instead could intermingle with each other was addressed by Visser and Aten [19], who analysed the boundary areas of individual interphase chromosomes with a sensitive method based on *in vivo* replication labelling by DNA incorporation of two different thymidine analogues during S-phase of cultured Chinese hamster fibroblasts. Differ-

ently DNA labelled cells were fused and the analysis of resulting interphase nuclei by immunocytochemistry and confocal imaging revealed a very irregular shape of chromosomes with fibre-like structures penetrating into other chromosomes [19]. The visualization of CTs and sub-chromosomal domains by electron microscopy showed the chromosome domains as local regions of mostly condensed chromatin with chromatin of adjacent chromosomes contacting in limited regions, implying chromosome–chromosome interactions [20].

Some models were proposed to explain the chromosome organization in the interphase nucleus [12, 14, 15]. The ICD "interchromosome domain" proposed by Cremer et al. [12] presupposes that CTs are separated by a three-dimensional interchromosomal compartment rich in transcriptional machinery. The IC "interchromatin domain" is in agreement with the previous ICD model but suggests the existence of loops at the periphery of CTs for enabling contact with the transcriptional machinery. Initially, it was thought that active genes would concentrate at the surface of the CTs to facilitate accessibility to the transcriptional machinery, while the inactive genes would localize in the interior of CTs. However, in wheat interphase meristematic nuclei, the simultaneous visualization of CTs and transcription sites by detection of nascent RNA by bromouridine triphosphate (BrUTP) incorporation showed the presence of active transcription sites even inside CTs, rather than exclusively at their surface [21]. CTs are viewed as having a high permeability to proteins, with an interchromatin compartment meandering into and folding through the territories [22]. Additionally, it has been shown that chromatin fibres from the CTs periphery are intermingled in interphase nuclei [23]. Such an organization would allow some gene loci to loop out resulting in intermingling of different CTs. Accordingly, genes can be anywhere within a chromosome territory independently of their transcriptional state, and genes located in the interior of CTs may also become accessible to the transcription machinery through DNA-free channels [15].

The Concept of Chromatin Loops

Chromatin loops may allow the closer contact of distal genomic regions with implications on regulation of gene transcription. This view is based on the "transcription factories" model in which transcription takes place at specific nuclear sites, termed transcription factories, to which different genes are recruited to be transcribed [24–26]. This would explain why in BrUTP transcription labelling experiments fewer transcription foci than transcribed genes are seen [27, 28]. However, the mechanisms for bringing specific loci together remain an open question. The relocation of genes in the interphase nucleus has been extensively associated with specific transcription states [29–31]. For example, in *Drosophila*, the insertion of a heterochromatin block at the brown locus caused the recruitment of the gene (and its unaltered allele) to the centromeric heterochromatin in the interphase nucleus [32]. In mammals, silenced genes can be recruited close to ikaros proteins and centromeric heterochromatin while active genes are not [33, 34]. Francastel et al. [35]

showed that a functional transcriptional enhancer is able to suppress transgene silencing by preventing its localization at centromeric heterochromatin and instead the transgene was recruited into an active nuclear compartment. Furthermore, mutations in the enhancer led to increased transgene silencing which was associated with transgene co-localization with the centromeric heterochromatin [35]. Possibly, transcriptional activators or enhancer-associated factors (e.g. histone acetyltransferases) may disrupt interactions between transgene and heterochromatin proteins leaving the transgene free to move away from the centromeric compartment into an active nuclear compartment [35]. Changes in chromatin structure at *HoxB* gene cluster also lead to re-localisation of chromatin loops away from their chromosome territory. In this case, extensive chromatin decondensation and nuclear reorganization of the *HoxB* cluster were correlated with transcription induction associated with cell differentiation [36, 37]. The looping-out of the activated *HoxB* gene locus shows increased interchromosomal interactions, as compared with the inactive locus, which interacts preferentially with other loci on the same chromosome [38]. 3C techniques can provide new insights into physical chromatin interactions *in vivo* [39–41]. For example, Osborne and colleagues [42] used a combination of 3D FISH, immunofluorescence and chromosome conformation capture (3C) to assess the spatial organization for several genes in a mouse chromosome and were able to show that upon transcription physically distant genes co-localize to the same transcription factory, whereas identical, temporarily non-transcribed alleles do not. These authors determined the percentage of co-localization of the RNA-FISH and the corresponding DNA signals, as well as the co-localization of widely separated genes when these genes were being transcribed, and concluded that co-localization of genes was transcription dependent. Also, using a modification of 3C method and FISH, Ling et al. [43] detected the co-localization of distinct DNA segments located on distinct chromosomes. For a long time, it was assumed that chromatin loops would be always linked to chromatin decondensation. However, this correlation may not be so straightforward, since even when chromatin is more condensed, genes can be activated; the opposite may also occur. Indeed, the analysis of gene density distribution and activity in distinct chromatin fractions showed a correlation between chromatin structure and gene density, but was independent of the status of gene activity [44]. Open chromatin fibres correlate with highest gene density, but not with expression levels of those genes, whereas compact chromatin fibres generally have a low gene density, but can also contain active genes [44]. Moreover, in the *Hoxd* locus, decondensed alleles can be found within the CTs and, conversely, looped-out gene loci can still be condensed [45].

The dynamics of chromatin loops has also been addressed in plants but is hampered by the difficulty of visualization by FISH of small, single-copy DNA sequences in plants, mainly due to the large proportion of non-target repeat sequences. Thus, transgenic plants with multiple insertion copies are good tools to study chromatin dynamics within interphase nuclei [46]. Abranches et al. [47] used wheat transgenic lines containing multiple transgene integration sites to determine the precise location of the transgenes within the interphase nucleus since it is during this stage that trans(gene) expression mainly occurs. First, the transgenes in different

transgenic lines were mapped on metaphase chromosomes to their various single and multiple chromosomal locations, in one line even on opposite arms of the chromosome. However, during interphase, these copies were brought together and often were seen as a single labelled focus [47]. Additional studies on these wheat transgenic lines showed that the transgenes were all transcriptionally silenced to some extent [48]. The silencing could be reduced and the transcription level increased by treatment with 5-azacytidine (5-AC), which reduces cytosine methylation, or with trichostatin-A (TSA), which inhibits histone deacetylases and thus increases histone acetylation levels, associated with transcriptional activation. These induced changes in epigenetic modifications were associated with separation of the transgene copies, which were then visualized as a string of dots [48]. This result raises the question of what causes the co-localization of transgenes; is it because of their recruitment to a common transcription factory as the transcription factory model predicts, or, on the contrary, is it a consequence of transcriptional silencing? Messenger RNA can be observed by FISH using antisense RNA as a probe, providing information on gene expression. Wegel and colleagues [49] used wheat transgenic lines, carrying multiple copies of the high molecular weight glutenin genes under the control of their native promoters, to examine in more detail the changes occurring on chromatin organization in relation to transcriptional activation in the endosperm. This was assessed by double ISH with DNA and RNA probes followed by confocal imaging which allowed the simultaneous visualization of genes and corresponding transcripts. In tissues where the transcription is silenced, such as root tissue or at earlier stages in endosperm development, in a line carrying about 20 copies of the transgene in a single tandem array, the transgenes were condensed into a single large focus. Upon transcription initiation, which only occurs in the endosperm from eight days after anthesis until the end of grain filling, it was possible to detect large-scale chromatin reorganization and decondensation into many foci [49]. In another study using mRNA and DNA ISH techniques, this time for endogenous genes, showed that in bristle oat (*Avena strigosa*) the activation of the avenacin gene cluster is accompanied by chromatin decondensation. Specifically, the *Sad1* and *Sad2* genes moved apart from approximately 0.45 to 0.90 μm on transcriptional activation. Labelling each end of the *Sad1* gene with a different colour showed that the *Sad1* gene itself decondensed on activation [50].

Organization of Interphase Nucleus in Plants

In this section, our current knowledge of nuclear organization in plants is described with particular emphasis on the cereals rice and wheat, since their large nuclei are easy to visualize by microscopy and moreover wheat tolerates well the introgression of alien chromosomes. These features have provided good tools to study the organization of interphase chromatin while deepening our knowledge on economically important crops. In fact, wheat is the most widely grown crop in the world and provides 20 % of the daily protein and food calories for 4.5 billion people. It is also

the second most important food crop in the developing world after rice [51]. Rice is particularly important in terms of nutrition and calorie intake, especially in Asia, where it is the primary food source for over 70% of the population [52]. Rice is considered a model for the cereal crop plants due to specific genome features such as a moderately small genome size (≈430 Mbp), and a high-quality reference genome sequence. Importantly, T-DNA insertion mutant collections in genes involved in chromatin remodelling and ordered BAC libraries are available [53, 54]. In addition, rice has substantial synteny with other cereal crops such as maize, sorghum and wheat [55], allowing the results of studies on this model to be more readily adapted to the remaining cereal crops. A better knowledge of the 3D organization of crop genomes in the interphase nucleus may provide some clues about functional mechanisms underlying the regulation of gene expression.

In plants, the painting of CTs by chromosome ISH turned out to be particularly difficult mainly due to the high level of dispersed repetitive sequences, which hampers the generation of chromosome probes with sufficient specificity [18, 56]. These types of chromosome paints may never be feasible for large complex plant genomes such as wheat, although they have been achieved for *Arabidopsis* (see below). Interspecific and intergeneric hybrids, e.g. wheat/rye translocation or addition lines, in combination with genomic *in situ* hybridization (GISH), have been an important tool for studying the structure of chromosomes in plant interphase nuclei. These studies unequivocally demonstrated for the first time the existence of CTs during interphase in plant cell nuclei [57, 58]. In GISH, a total genomic probe is used to label the alien chromosomes or chromosome arms in lines hybrids of species such as wheat containing added or substituted chromosomes or chromosome arms from other closely related species [58–60, 21]. Although GISH is an ingenious methodology to visualize interphase chromosomes, it should be remembered that it does not generate a "true chromosome painting" since what is labelled refers to introgressed chromatin, which in principle may behave differently from the native chromosomes. The majority of studies of interphase nuclei have been performed on squashed preparations in which the three-dimensional organization is lost. However, more faithful protocols have been developed to apply FISH in three-dimensionally (3D) preserved plant tissue sections derived from intact plant structures such as roots or florets [61–63]. These methods are based on the combination of vibratome sectioning with confocal microscopy and allow the visualization of intact and well-preserved 3D nuclei. GISH coupled with confocal microscopy in wheat/rye addition and/or translocation lines showed that interphase chromosomes appeared as elongated domains stretched across the diameter of the nucleus with both arms of each chromosome physically close together (Fig. 4.1). Additionally, in wheat, the centromeres and telomeres were shown to be located at opposite poles of the nucleus, displaying a typical Rabl configuration (Fig. 4.2c).

In rice, a "true" chromosome painting during interphase has not yet fully succeeded and thus the detailed interphase chromosome territorial organization is still unknown. However, chromosome-specific BAC-FISH, centromeric, telomeric and ribosomal probes have been used to investigate chromosomal arrangement in rice interphase nuclei [64–66]. In most cells of diploid rice, the telomeres and centro-

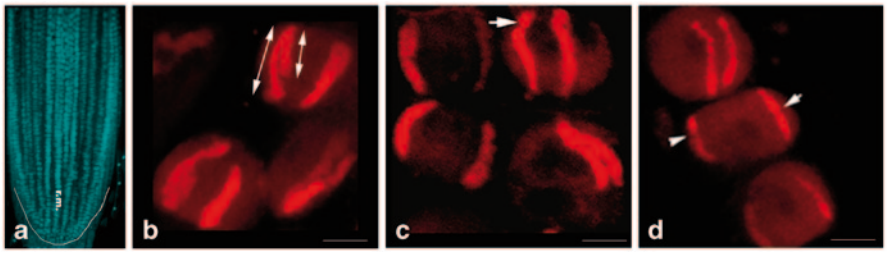

Fig. 4.1 Wheat root longitudinal tissue section (approximately 30 μm thick) labeled with DAPI is shown in (**a**). GISH applied on root tissue sections of wheat disomic addition lines (CS + 5R5R, CS + 1R1R) and a wheat translocation line ($1A^l/1R^S$) enabled the visualization of rye chromosomes in interphase nuclei (**b**–**d**). Single pairs of 5R chromosomes are shown in (**b**), 1R chromosome pairs in (**c**) and short arms of 1R chromosomes ($1R^S$) in (**d**). Interphase chromosome arms are generally seen close to each other and not distinguishable but in some cases the long (e.g. $5R^l$) and the short arm (e.g. $5R^S$) can be assigned (**b**, *arrows*). In the 1R chromosome, the nucleolar organiser region (*NOR*) is assigned (**c, d**, *arrows*). Projections of individual confocal sections taken with an interval of 1 μm are shown. Bar = 10 μm. *GISH* genomic *in situ* hybridization

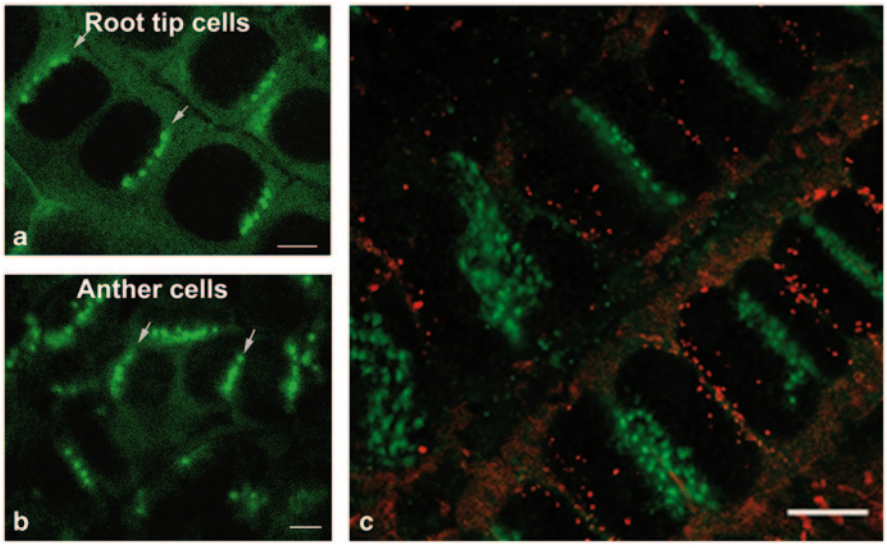

Fig. 4.2 Fluorescence *in situ* hybridization in wheat interphase nuclei with probes to centromeres (*green*) and telomeres (*red*). Centromeres are labelled in interphase nuclei of wheat root sections (**a**) and of anther sections (40 μm thick) at an early premeiotic stage (**b**). In both cell types, the centromeres are clustered and regularly arranged at the nuclear membrane (**a, b**). The Rabl configuration, with the centromeres clustered on one side of the nuclear periphery and the telomeres somewhat more dispersed on the other side of the nuclear periphery, is shown in interphase nuclei of a wheat root section (**c**). A common (alternating) polarity is often maintained through the lines of cells as seen in (**c**). Projections of confocal optical sections are shown (focal distance between original sections = 1 μm). Bar = 10 μm

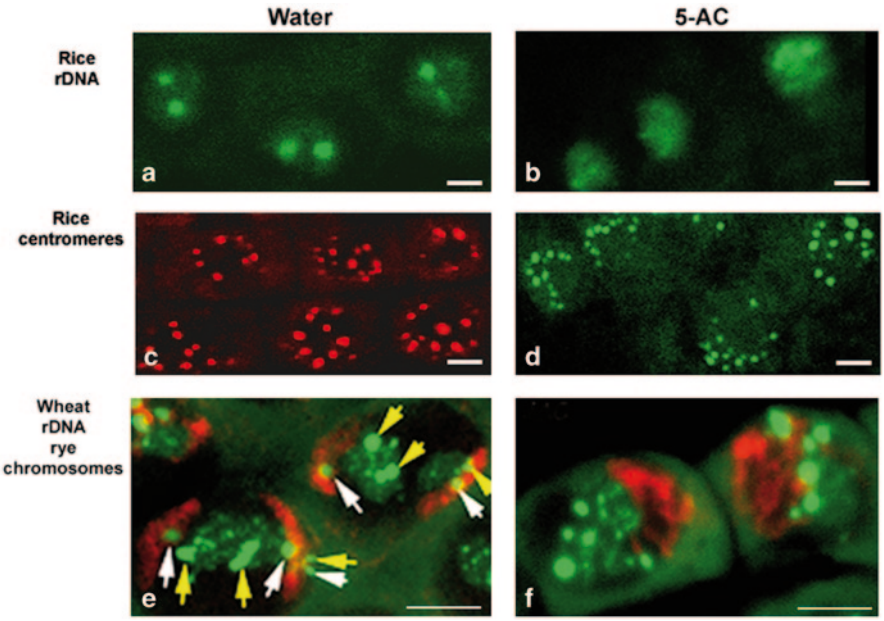

Fig. 4.3 5-AC effects on chromatin organization in interphase nuclei of rice and wheat root sections. In rice, the organization of rDNA chromatin (**a, b**) and of centromeres (**c, d**) after seedlings germination in water (**a, c**) or in 5-AC drug (**b, d**) is shown. In rice interphase nuclei, the rDNA chromatin is organized as two hetrochromatic knobs (**a**) which clearly decondense upon induced DNA hypomethylation (**b**). Centromeres in rice (**c**) are circularly distributed at the nucleus periphery illustrating the absence of Rabl configuration of chromosomes in these root meristematic cells; however, the 5-AC drug was able to modulate the interphase position of rice centromeres namely by inducing its polarization (**d**). In a wheat addition line, the 1R rye chromosome pair is shown in interphase nuclei after seedlings germination in water being seen as domains extending across the nucleus (**e**, *red*). After 5-AC germination, the interphase rye chromosomes show an irregular appearance and the two arms of a single chromosome do not remain alongside one another and follow a meandering path across the nucleus (**f**, *red*). In (**e**), the rRNA genes from rye origin and wheat origin are indicated with *white* and *yellow arrows*, respectively. The 1R rye chromosome pair can be in close association to the nucleolus and two rDNA chromatin knobs of uncertain origin can be seen (**e**, two sided *white* and *yellow arrows*). The rDNA chromatin of rye origin rye appears as a condensed region on each rye homologue (**e**, *white arrows*) but after 5-AC, the rDNA chromatin organization appears more diffuse with less evident knobs of heterochromatin presumably due to knob decondensation (**f**). Root meristematic cells were imaged and consecutive individual confocal sections were recorded with a section spacing of 1 μm being shown in the 3D projections. Bar = 10 μm

meres are dispersed around the nuclear periphery (Fig. 4.3c) rather than showing a Rabl configuration [64, 65]. Nevertheless, in some cells, such as endoreplicated xylem–vessel precursor cells and premeiotic anther cells, a Rabl configuration of chromosomes is displayed with the centromeres and telomeres at opposite poles of the nucleus [64, 65]. These observations emphasize the plastic nature of the CTs organization in relation to cell development and differentiation.

So far, *Arabidopsis thaliana* and its close relative *A. lyrata* are the only plants where "true" chromosome painting, using chromosome-specific probe sets, has been performed [67–69]. The fact that the *Arabidopsis* model plant has a small genome size with a low level of repeated sequences, about 10–12 %, was crucial to successfully visualize interphase chromosomes. The strategy consisted of using a series of contiguous BACs (bacterial artificial chromosomes) as FISH DNA probes and showed that in this species the Rabl configuration is not present and telomeres are clustered around nucleolus while the chromocentres, which consist of single or associated groups of centromeres are located at the nuclear periphery [69, 70]. The localization of centromeres at the nuclear periphery was also confirmed in different cell types by *in situ* visualization, 3D restoration, and quantitative analysis [68]. A model for *Arabidopsis* chromosomes has been proposed in which chromatin loops (0.2–2 Mb in length) emanate from the chromocenters, generating a rosette-like structure of *Arabidopsis* CTs [69, 71]. Exhaustive labelling of multiple CTs with specific paints in *Arabidopsis* has failed to show any preferred order of the chromosomes. The only exceptions are the chromosomes carrying the nucleolar organizing regions (NORs), which contain tandemly arranged copies of rRNA genes [68, 72], and which are frequently seen close to each other, presumably because they are both attached to the nucleolus via the decondensed NORs.

The Puzzle of the Rabl Configuration

The diversity of chromosome arrangements in plant interphase nuclei illustrates the difficulty in finding a universal rule or a mechanistic explanation for the specific arrangement of chromosomes in the interphase nucleus. The integration of multiple factors such as genome and chromosome sizes, chromosome interactions with the nuclear envelope, flexibility and organization of the CTs, distribution of heterochromatin and epigenetic factors may all have a role in determining whether the anaphase chromosome arrangement is preserved in interphase nuclei as the Rabl arrangement.

The Effect of Genome and Chromosome Size

How chromosomes are arranged within plant interphase nuclei has been studied in a variety of cell types and plant species. Genome and nuclear sizes are correlated for plants as well as for other organisms; plants with large genomes have larger nuclei while plants with small genomes have smaller nuclei. The size of the nucleus has also been suggested to be a factor in the chromosome arrangement although it has been difficult to ascertain what is the cause or the consequence. The Rabl arrangement of chromosomes in interphase nuclei is found in some plant species with large genomes, e.g. in wheat and barley [73]. However, maize and sorghum are also considered large genomes but there is no evidence for the presence of Rabl

configuration in these plants [73]. On the other hand, small genomes can show a wide range of chromosome arrangements in interphase nuclei. For example, in rice, Rabl and non-Rabl configurations can be present depending on the cell type [64, 65]. In *Arabidopsis*, the interphase chromosomes do not display a Rabl configuration but exhibit a strikingly different type of chromatin arrangement as described above [69]. In other organisms also considered as having a small genome size, like the yeasts, *Schizosaccharomyces pombe* and *Saccharomyces cerevisaea* Rabl-type chromosome arrangement is observed [74–76]. In contrast, mammal cells have not so far shown any evidence for a Rabl arrangement of chromosomes [77–81]. Based on the examples above, it is certain that neither the genome organization nor the relative genome size or various chromosome lengths, per se, can determine the type of chromosome arrangement adopted at interphase. The Rabl configuration may function as a way to ensure a certain level of order and integrity along chromosomal organization which may be particularly relevant for large genomes. For example, Manders et al. [82] by imaging heterochromatin foci of muntjac cell lines (*Muntiacus muntjak*) did not detect significant reorganization of CTs upon the transition from interphase (G2) to metaphase chromosomes. Also, in wheat, chromosomes are rich in heterochromatic regions dispersed along chromosomes, and at specific chromosomal regions chromosome decondensation can occur; in this way, a large reorganization of chromosomes during progression from anaphase to interphase may not be required. Contrastingly, in *Arabidopsis*, the heterochromatin is predominantly located around the centromeres, and the remaining parts of the chromosomes may have greater freedom to adopt a different conformation on entry into interphase. A plausible point of view is to consider the preservation of the anaphase arrangement to give the Rabl configuration as the default. The question then is why some species or cell types lose the Rabl configuration during interphase. The stabilization of chromosome configuration in interphase may rely on specific anchorage points of telomeres and centromeres to some components of the nuclear envelope [83]. However, the nature of these putative interactions or their functional significance is not known. The nuclear lamins have been proposed to be involved in maintaining the Rabl configuration because of their ability to bind DNA, chromosomes, and histones in vitro [84]. In mice and *Drosophila melanogaster*, the loss of a nuclear lamin causes chromatin to detach from the nuclear envelope [84]. The linkage of telomeric regions to nuclear membrane may be necessary to maintain the Rabl configuration. It has been shown that in these chromosome regions, the nucleosomes are more closely spaced than in the rest of the genome suggesting an increased number of histone–lamin interactions which can strengthen the attachment of telomeres [85, 86]. However, it is still an open question whether similar structures to nuclear lamins exist in plant cells. As the interphase nucleus is a dynamic structure that physically interacts with the cytoskeleton, it may be expected that cytoskeleton components play a role in the 3D arrangement of interphase chromosomes [87, 88]. Living cells stabilize their internal cytoskeleton and control their shape and molecular mechanisms using an architectural system known as tensegrity [89]. The cytoskeleton uses this tensegrity architecture, behaving like a mechanical network that maintains cell shape in association with nuclear structure and function [90].

The Effect of Cell Type

The arrangement of chromosomes in the interphase nucleus is in some cases related to changes in ploidy level. For example, in rice root tissue, the typical Rabl configuration, with polarization of centromeres and telomeres, is present in the endoreduplicated xylem vessel cells but not in the surrounding root meristematic cells, showing the plastic nature of interphase chromosome arrangement [64, 65]. In the endoreduplicated nuclei of rice xylem vessel cells, the final mitotic cell division is followed by a further round of replication, before the chromosomes adopt an interphase configuration. This process possibly generates novel centromere interactions favouring the anaphase chromosome configuration. The relation between interphase chromatin organization, cell type specificity and endoreduplication was also investigated in *Arabidopsis*, where the organization of endoreduplicated sister centromeres is clearly different between the larger leaves and root epidermal cells [68, 72]. The fact that leaves and roots have distinct biological functions may have a role in the organization of interphase chromatin.

Response to Transcriptional Requirements, Epigenetic Modulators, and Environmental Stresses

Epigenetic factors are likely to produce alterations in chromosome structure. GISH labelling of CTs after treatment with 5-azacytidine (5-AC, inducing DNA hypomethylation) and trichostatin A (TSA, increasing histone acetylation) has indeed shown changes in interphase CT structure in wheat–rye addition or translocation lines. Following drug treatment, the rye chromosomes or chromosome arms were seen as a series of strongly labelled heterochromatic regions separated by gaps, rather than the smooth, continuous structure normally seen (Fig. 4.3f). The two arms of a single chromosome could be seen to take a meandering path across the nucleus, and generally did not remain alongside one another [48]. This result is consistent with a chromosome organization containing large stretches of heterochromatin interspersed with gene-rich islands that may decondense [91]. Nevertheless, the Rabl configuration was not disrupted by the treatments with drugs, suggesting strong interactions or specific associations between the centromeric and telomeric regions and the nuclear envelope, which are not directly affected by DNA methylation or histone acetylation levels [48]. In contrast, in rice root-tip meristematic cells, DNA hypomethylation induced by 5-AC induced a Rabl configuration in root-tip meristematic nuclei (in which this arrangement is not normally present; Fig. 4.3d). This illustrates that chromosome structure can be altered by epigenetic factors.

Challenging environments may affect the large-scale organization of plant genomes including effects on repetitive DNA, chromatin structure and nuclear organization [92–95]. McClintock (1984) predicted that stress may cause large-scale genomic changes, including transposon activation and other structural modifications

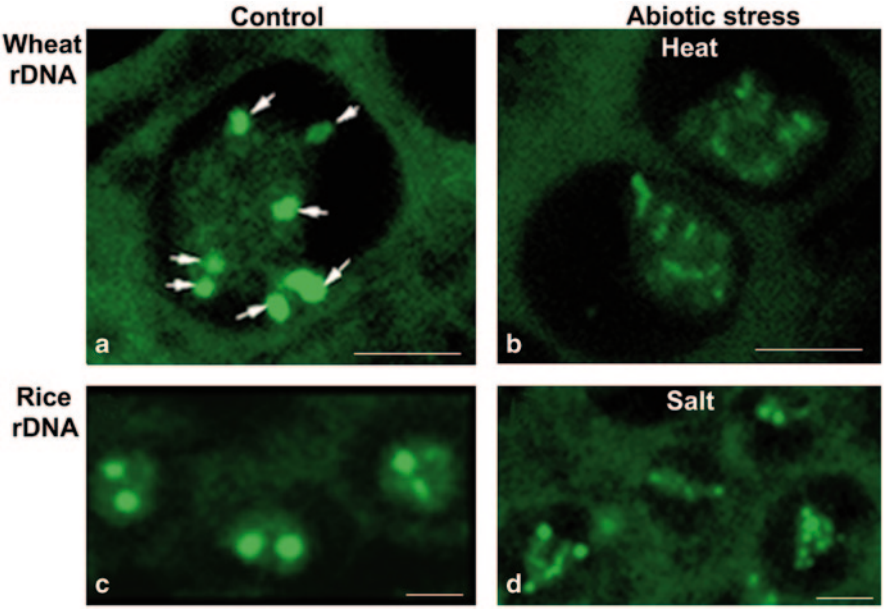

Fig. 4.4 Abiotic stress effects on rDNA chromatin organization in interphase nuclei of wheat and rice after germination in optimal conditions (**a, c**) or upon stress imposition such as heat (42°C, 3 h) or salinity (150 mM, 3 days) treatments during germination (**b, d**, respectively). The heterochromatic knobs presumably inactive are clearly visible in wheat and rice interphase nuclei (**a, c**, respectively) but can decondense after imposition of heat or salt stress conditions during germination (**b, d**, respectively). Confocal image stacks were recorded with a section spacing of 1 μm and the 3D projections are shown. Bar = 10 μm

of the chromosomes [96]. Stress- induced genomic responses have indeed been shown, including transposon activation, transposition and restructuring or rearrangement of chromosomes [97–99]. A hypothetic model to explain genome regulation under normal and stress conditions has been proposed [98]. In this model, under normal conditions, heterochromatin maintenance mechanisms repress transcription of repetitive DNA. Stress can cause the relaxation of epigenetic imprints. RNAi and other heterochromatin maintenance pathways fail, resulting in the activation of transposons. If the cell survives the shock, its genome undergoes epigenetic remodelling and often genetic rearrangements resulting in altered epigenetic marks and in novel gene expression patterns [98]. Cytogenetic analyses using FISH in rice and wheat interphase nuclei of root tissues showed a decondensation of the highly condensed heterochromatic domains of the 45S rDNA loci after short-term salt and heat stresses (Fig. 4.4). The spatial positioning of 5S rDNA genes in the nucleus was also altered under stress conditions; the distance between the two alleles significantly increased when seedlings were exposed to salinity conditions, heat shock or 5-AC [66]. A number of other investigators also reported chromatin decondensation in response to stress conditions or developmental signals [100, 101]. Pecinka et al. [102] further demonstrated heterochromatin decondensation and reduction in

nucleosome occupancy upon activation of several repetitive elements of *Arabidopsis thaliana*. Chromatin plasticity was also related to developmental transitions as in case of *Arabidopsis* mesophill cells of mature leaves where the heterochromatin fraction was found higher than in young leaves [103]. Furthermore, the chromocenters become smaller in leaves prior to the transition to reproductive development recovering their initial size after the elongation of the floral stem [104]. There is still no proved function for chromatin condensation state but it likely plays a role in plant adaptation to challenging stimuli. For example, *Arabidopsis* genotypes acclimated to different latitudes exhibit genetically programmed levels of chromatin compaction, depending on the light intensity of their original habitats [100].

Nuclear Organization in Space and Time

All the studies discussed in the previous sections used cytological techniques, mainly ISH. Although the basic principles of FISH have been established for many years, technical improvements allowing high sensitivity, simultaneous detection of multiple targets and automated data collection and analysis were crucial to the advancement of this field. In the past few years, a set of techniques known as chromosome conformation capture (3C) methods, have been developed [39]. These techniques attempt to detect long-range interactions of sequences within and between chromosomes by determining the frequency with which any pair of loci in the genome is in close enough physical proximity (in the range of 10–100 nm) to become cross-linked. The original 3C methodology was based on chemical (formaldehyde) chromatin cross-linking followed by DNA fragmentation by sonication or restriction digestion, and ligation of the fragments in diluted conditions to favour intramolecular ligation of DNA fragments [105]. In this way, DNA fragments interacting more frequently *in vivo* will be more likely to be ligated together by the cross-linking, thus leading to higher levels of ligation products in the cell population. The techniques differ in the ligation methodology and product analysis. In the original 3C method, the interaction of a known target sequence was tested with a set of other known target sequences. Among the spatial genome mapping methods which derive from the 3C technology are circularized chromosome conformation capture (4C), in which all the sequences ligated to the target sequence (the "bait") are determined, carbon-copy chromosome conformation capture (5C), which allows many bait sequences to be analysed in parallel and Hi-C in which an unbiased high throughput sequencing strategy is used to obtain a statistical cross section of all the ligations produced. These methodologies have allowed studies of chromosome interactions at the genome-wide scale. In plants, the adoption of these technical advances was slower and so far only a few studies have been performed employing 3C technology. In plants, the use of these techniques was pioneered by 3C analysis of chromatin looping and physical chromatin interactions in maize [106]. These authors examined the involvement of chromatin looping in the transcriptional regulation of two epialleles of a maize gene that were tissue-specifically regulated. 3C assays were also used to

investigate the role of DNA looping on the expression of the flowering time regulator locus FLC (FLC), and showed that the loop formation was not dependent on the expression of FLC [107]. However, the loop formation was efficiently disrupted within the first two weeks of cold exposure during vernalization and that was not recovered when plants were transferred back to warm conditions [107].

In Vivo Imaging of Chromatin

Significant breakthroughs in the analysis of chromatin dynamics have been achieved by taking advantage of bacterial repressor/operator interactions together with GFP technology. Two systems are currently in use, one based on the lac operator/repressor [108–110], and the other on the tet operator/repressor [111, 112]. Tandem arrays of the cognate DNA sequences are inserted into the target genome, and then a GFP fusion to the binding protein is expressed, which binds specifically to the target sequence arrays, marking them with a fluorescent label. This technology allows the visualization of specific chromosome regions or even single chromosome loci in live cells. In the initial report of this system, the Belmont group used 256 tandem copies of the lac operator array in mammalian cells. The method was soon used in prokaryotic cells, in yeast, in *Drosophila* and in *Caenorhabditis elegans*.

A few studies have used this method in plants [113, 114]. Kato and Lam [114] used it to show that chromatin in the endoreduplicated epidermal pavement cells in *Arabidopsis* had a greater range of movement than the diploid guard cells. Rosin et al. [115] and Matzke et al. [116, 117] have produced *Arabidopsis* lines with single sites labelled at various chromosomal loci. The system has been used to study the positioning of the *flowering locus C* (FLC) gene in *Arabidopsis*, which is a target of transcriptional repression by a polycomb protein during vernalization [118]. In this study, the alleles of FLC were shown to associate physically in the nucleus during the silencing process. Mutants in *trans*-acting protein factors required for the silencing did not show this clustering behaviour, suggesting that it was functionally related to the silencing mechanism. One drawback of these chromosome tags is that they represent artificial chromosome loci, which may behave differently from native genomic sequences, and eventually generate unusual nuclear interphase arrangements or artifactual interactions [119]. Therefore, these studies must be carefully controlled and confirmed by the use of mutants and if possible by *in situ* labelling. A further problem is that the expression level of the GFP fusion protein needs to be carefully chosen or controlled; the fusion protein is present as a background in the whole nucleus, and is merely concentrated at the sites of the cognate DNA sequences. This means that if the expression level is too high, the background, unbound fusion protein can easily obscure the specific labelled foci.

Similar technology has been developed to visualize RNA directly in living cells. The ability to visualize RNA movement in living cells gives insights as to how and where specific sequences are expressed and the steps by which transcripts are processed and exported from the nucleus and even across cells. In the first such method, the bacteriophage MS2 coat protein which binds to a short specific RNA stem

loop sequence was used, in an entirely analogous way to the DNA tagging methods, incorporating tandem arrays of the MS2 cognate sequence into target RNAs and expressing an MS2-GFP fusion [120]. A major problem with this approach is to find a position within the RNA where enough tandemly repeated sequences can be inserted for visualization without affecting the RNA biological activity. In plants, two methods have been used: the MS2 system and one based on λN_{22}/Box-B [121]. These authors showed that the λN_{22} system has some advantages over the MS2 system, in particular that the MS2 seems to show a degree of multimerization of the RNA binding protein. In any case, it is useful to have two different tagging systems, which can be coupled to differently coloured fluorescent proteins for multiple labelling experiments and for this Schönberger et al. [121] have generated a useful series of gateway-compatible vectors for plant expression.

Another fluorescence tagging strategy for RNAs is the Pumilio system, which has been used successfully in plants [122]. In this system, pumilio homology domain (PUMHD) polypeptides were engineered to recognize RNA sequences in the target RNA and fused to either the C or the N terminal portion of a split mCitrine fluorescent protein. Binding of the target sequences to their cognate pumilio domains then brought together the two halves of mCitrine, resulting in bimolecular fluorescence complementation (BiFC). The advantage of this system is that the background level should be very low, as no fluorescence is produced until the RNA is bound to the proteins, and no modification of the target RNA is necessary, as pumilio domains can be modified to recognize existing sequences in the target RNA. The disadvantages are that considerable knowledge of target RNA and trial and error of different target sequences may be necessary to obtain effective binding. Also, the sensitivity is likely to be low, since in general only one fluorescent molecule is present per RNA. This will probably restrict this approach to abundant RNAs such as viral RNAs. Recently, a method was described for RNA visualization using RNA aptamers that bind to fluorophores resembling the GFP fluorophore and produce RNA–fluorophore complexes that are comparable in brightness to fluorescent proteins [123]. This method however has not yet been applied to plants and cannot be engineered for *in vivo* expression, so it will probably be difficult to overcome the cell wall barrier that makes access to exogenous molecules difficult.

Future work on live imaging will focus on measuring histone mobility and will use single cell analysis. The fusion of histones to GFP provides a way to visualize chromatin movements *in vivo*. Techniques such as fluorescence recovery after photobleaching (FRAP) allow the analysis of the kinetics of molecules in living cells. The use of multi-photon approaches, in which a high-intensity pulsed laser source of longer wavelength photons causes the excitation of the fluorophore by pairs or triplets of photons, has expanded the application of fluorescence imaging. Near infrared excitation light can penetrate biological specimens more deeply and with less toxicity than visible light [124]. More recently, there is also interest in obtaining data from single cells, thus measuring the distribution of behaviours, and not just the average behaviour. Single-cell imaging provides insight in situations where there is static or dynamic heterogeneity, as in a complex system such as the cell nucleus. For instance, structured illumination microscopy (SIM) has been used to compare the nuclear functional organization in a variety of cell types and species

[125]. Through correlative microscopy for sequential studies of individual cells, it is possible to combine live cell microscopy, SIM and transmission electron microscopy (TEM) [126]. Moreover, three-dimensional structured illumination microscopy (3D-SIM) has opened up new possibilities to study nuclear architecture at the ultrastructural level down to the ~100-nm range while super-resolution techniques can go even further and reach a level of resolution of 10–50 nm or better. Future work in plants will certainly make use of these technologies in order to bring new insights into the functional organization of the plant cell nucleus.

Future Perspectives

The way in which chromosome and chromatin arrangement in the interphase nuclei may modulate gene expression is still far from being understood. Studies *in vivo* and at the level of single living cells are needed to understand the detailed mechanisms underlying nuclear activity. New super-resolution light microscopy techniques are allowing imaging beyond the classical optical resolution limit and should help to fit the pieces of the puzzle together. A final applied aim of understanding nuclear organization in plants is to allow manipulation of gene expression in a targeted manner in order to generate better performance of crops in response to unpredictable environments.

Acknowledgments APS is supported by Fundação para a Ciência e a Tecnologia (FCT), Portugal, through grants BPD/74197/2010 and BIA-BCM/111645/2009. RA is supported by FCT, Portugal, through grants PEst-OE/EQB/LA0004/2013 and PTDC/BIA-PLA/2411/2012. PS was supported by BB/J004588/1 from BBSRC and the John Innes Foundation.

References

1. Raices M, D'Angelo M. Nuclear pore complex composition: a new regulator of tissue specific and developmental functions. Nat Rev Mol Cell Biol. 2012;13(11):687–99.
2. Waldeyer W. Über Karyokinese und ihre Beziehungen zu den Befruchtungsvorgängen. Arch Mikr Anat. 1888;32:1–122.
3. Schubert I. Between genes and genomes-future challenges for cytogenetics. Front Genet. 2011;2(30):1–2.
4. Heitz E. Das Heterochromatin der Moose. I. Jahrb wiss Bot. 1928;69:762–818.
5. Rabl C. Über Zellteilung. Morphol Jahrb. 1885;10:214–330.
6. Boveri T. Die Blastomerenkerne von Ascaris megalocephala und die Theorie der Chromosomenindividualität. Arch. Zellforch. 1909;3:181–268.
7. Wischnitzer S. The submicroscopic morphology of the interphase nucleus. Int Rev Cytol. 1973;34:1–48.
8. Cremer T, Cremer C, Schneider T, Baumann H, Hens L, Kirsch-Volders M. Analysis of chromosome positions in the interphase nucleus of Chinese hamster cells by laser-UV-microirradiation experiments. Hum Genet. 1982;62(3):201–9.
9. Gall JG, Pardue ML. Formation and detection of RNA–DNA hybrid molecules in cytological preparations. Proc Natl Acad Sci U S A. 1969;63(2):378–83.

10. Lichter P, Cremer T, Borden J, Manuelidis L, Ward DC. Delineation of individual human chromosomes in metaphase and interphase cells by in situ suppression hybridization using recombinant DNA libraries. Hum Genet. 1988;80(3):224–34.
11. Pinkel D, Gray JW, Trask B, Vanda Engh G, Fuscoe J, Van Dekken H. Cytogenetic analysis by in situ hybridization with fluorescently labelled nucleic acid probes. Cold Spring Harb Symp Quant Biol. 1986;51(1):151–7.
12. Cremer T, Kurz A, Zirbel R, Dietzel S, Rinke B, Schröck E, Speicher MR, Mathieu U, Jauch A, Emmerich P, Scherthan H, Ried T, Cremer C, Lichter P. Role of chromosome territories in the functional compartmentalization of the cell nucleus. Cold Spring Harb Symp Quant Biol. 1993;58:777–92.
13. Cremer T, Cremer C. Rise, fall and resurrection of chromosome territories: a historical perspective. Part I. The rise of chromosome territories. Eur J Histochem. 2006;50(3):161–76.
14. Cremer T, Cremer C. Chromosome territories, nuclear architecture and gene regulation in mammalian cells. Nat Rev Genet. 2001;2:292–301.
15. Cremer T, Cremer M, Dietzel S, Muller S, Solovei I, Fakan S. Chromosome territories—a functional nuclear landscape. Curr Opin Cell Biol. 2006;18(3):307–16.
16. Cremer T, Cremer M. Chromosome territories. Cold Spring Harb Perspect Biol. 2010;2:a003889.
17. Meaburn KJ, Misteli T. Cell biology: chromosome territories. Nature. 2007;445(7126):379–781.
18. Fuchs J, Houben A, Brandes A, Schubert I. Chromosome "painting" in plants—a feasible technique? Chromosoma. 1996;104(5):315–20.
19. Visser AE, Aten JA. Chromosomes as well as chromosomal subdomains constitute distinct units in interphase nuclei. J Cell Sci. 1999;112(19):3353–60.
20. Visser AE, Jaunin F, Fakan S, Aten JA. High resolution analysis of interphase chromosome domains. J Cell Sci. 2000;113(14):2585–93.
21. Abranches R, Santos AP, Wegel E, Williams S, Castilho A, Christou P, Shaw P, Stoger E. Widely separated multiple transgene integration sites in wheat chromosomes are brought together at interphase. Plant J. 2000;24(6):713–23.
22. Verschure PJ, van der Kraan I, Manders EM, Hoogstraten D, Houtsmuller AB, van Driel R. Condensed chromatin domains in the mammalian nucleus are accessible to large macromolecules. EMBO Rep. 2003;4(9):861–6.
23. Branco MR, Pombo A. Intermingling of chromosome territories in interphase suggests a role in translocation and transcription dependent associations. PLoS Biol. 2006;4:38.
24. Cook PR. A chromomeric model for nuclear and chromosome structure. J Cell Sci. 1995;108(9):2927–35.
25. Cook PR. The organization of replication and transcription. Science. 1999;284(5421):1790–5.
26. Cook PR. A model for all genomes: the role of transcription factories. J Mol Biol. 2010;395(1):1–10.
27. Jackson DA, Hassan AB, Errington RJ, Cook PR. Visualization of focal sites of transcription within human nuclei. EMBO J. 1993;12(3):1059–65.
28. Iborra FJ, Pombo A, Jackson DA, Cook PR. Active RNA polymerases are localized within discrete transcription "factories" in human nuclei. J Cell Sci. 1996;109(6):1427–36.
29. Lanctôt C, Cheutin T, Cremer M, Cavalli G, Cremer T. Dynamic genome architecture in the nuclear space: regulation of gene expression in three dimensions. Nat Rev Genet. 2007;8(2):104–15.
30. Misteli T. Spatial positioning: a new dimension in genome function. Cell. 2004;119(2):153–6.
31. Takizawa T, Meaburn KJ, Misteli T. The meaning of gene positioning. Cell. 2008;135(1):9–13.
32. Dernburg AF, Broman KW, Fung JC, Marshall WF, Philips J, Agard DA, Sedat JW. Perturbation of nuclear architecture by long-distance chromosome interactions. Cell. 1996;85(5):745–59.
33. Brown KE, Guest SS, Smale ST, Hahm K, Merkenschlager M, Fisher AG. Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. Cell. 1997;91(6):845–54.

34. Brown KE, Baxter J, Graf D, Merkenschlager M, Fisher AG. Dynamic repositioning of genes in the nucleus of lymphocytes preparing for cell division. Mol Cell. 1999;3(2):207–17.
35. Francastel C, Walters MC, Groudine M, Martin DIK. A functional enhancer suppresses silencing of a transgene and prevents its localization close to centromeric heterochromatin. Cell. 1999;99(3):259–69.
36. Chambeyron S, Bickmore WA. Chromatin decondensation and nuclear reorganization of the HoxB locus upon induction of transcription. Gene Dev. 2004;18(10):1119–30.
37. Chambeyron S, Da Silva NR, Lawson KA, Bickmore WA. Nuclear re-organisation of the Hoxb complex during mouse embryonic development. Development. 2005;132(9):2215–23.
38. Wurtele H, Chartrand P. Genome-wide scanning of HoxB1-associated loci in mouse ES cells using an open-ended chromosome conformation capture methodology. Chromosome Res. 2006;14(5):477–95.
39. Wit E, Laat W. A decade of 3C technologies: insights into nuclear organization. Genes Dev. 2012;26(1):11–24.
40. O'Sullivan JM, Hendy MD, Pichugina T, Wake GC, Langowski J. The statistical-mechanics of chromosome conformation capture. Nucleus. 2013;4(5):390–8.
41. Dekker J, Marti-Renom MA, Mirny LA. Exploring the three-dimensional organization of genomes: interpreting chromatin interaction data. Nat Rev Genet. 2013;14(6):390–403.
42. Osborne CS, Chakalova L, Brown KE, Carter D, Horton A, Debrand E, Goyenechea B, Mitchell J.A, Lopes S, Reik W. Active genes dynamically colocalize to shared sites of ongoing transcription. Nat Genet. 2004;36(10):1065–71.
43. Ling JQ, Li T, Hu JF, Vu TH, Chen HL, Qiu XW, Cherry AM, Hoffman AR. CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb17Nf1. Science. 2006;312(5771):269–72.
44. Gilbert N, Boyle S, Fiegler H, Woodfine K, Carter NP, Bickmore WA. Chromatin architecture of the human genome: gene-rich domains are enriched in open chromatin fibers. Cell. 2004;118(5):555–66.
45. Morey C, Da Silva NR, Perry P, Bickmore WA. Nuclear reorganisation and chromatin decondensation are conserved, but distinct, mechanisms linked to Hox gene activation. Development. 2007;134(5):909–19.
46. Santos AP, Wegel E, Shaw P, Stoger E, Thompson B, Allen G, Abranches R. In situ methods to localize transgenes and transcripts in interphase nuclei: a tool for transgenic plant research. Plant Methods. 2006;2:18.
47. Abranches R, Beven AF, Aragón-Alcaide L, Shaw PJ. Transcription sites are not correlated with chromosome territories in wheat nuclei. J Cell Biol. 1998;143(1):5–12.
48. Santos AP, Abranches R, Stoger E, Beven A, Viegas W. Shaw PJ. The architecture of interphase chromosomes and gene positioning are altered by changes in DNA methylation and histone acetylation. J Cell Sci. 2002;115(23):4597–605.
49. Wegel E, Vallejos RH, Christou P, Stöger E, Shaw P. Large-scale chromatin decondensation induced in a developmentally activated transgene locus. J Cell Sci. 2005;118(11):1021–31.
50. Wegel E, Koumproglou R, Shaw P, Osbourn A. Cell type-specific chromatin decondensation of a metabolic gene cluster in oats. Plant Cell. 2009;21(12):3926–36.
51. The wheat initiative; 2012. An international vision for wheat improvement. Available from http://www.wheatinitiative.org/. Accessed 16 June 2014.
52. Khush GS. Harnessing science and technology for sustainable rice-based production systems. Proceedings of FAO rice conference "Rice is life". Int Rice Comm Newsl. 2004;53:17–23.
53. Devos KM, Gale MD. Genome relationships: the grass model in current research. Plant Cell. 2000;12(5):637–46.
54. Zhang Q, Wing R. Genome studies and molecular genetics: understanding the functional genome based on the rice model. Curr Opin Plant Biol. 2013;16(2):129–32.
55. Gale MD, Devos KM. Comparative genetics in the grasses. Proc Natl Acad Sci U S A. 1998;95:1971–4.
56. Schubert I, Fransz PF, Fuchs J, de Jong JH. Chromosome painting in plants. Methods Cell Sci. 2001;23:57–69.

57. Schwarzacher T, Leitch AR, Bennett MD, Heslop-Harrison JS. In situ localization of parental genomes in a wide hybrid. Ann Bot. 1989;64(3):315–24.
58. Schwarzacher T, Anamthawat-Jónsson K, Harrison GE, Islam AKMR, Jia JZ, King IP, Leitch AR, Miller TE, Reader SM, Rogers WJ, Shi M, Heslop-Harrison JS. Genomic in situ hybridization to identify alien chromosomes and chromosome segments in wheat. Theor App Genet. 1992;84:778–86.
59. Islam-Faridi MN, Mujeeb-Kazi A. Visualization of *Secale cereale* DNA in wheat germplasm by fluorescent in situ hybridization. Theor App Genet. 1995;90(5):595–600.
60. Morais-Cecilio L, Delgado M, Jones RN, Viegas W. Painting rye B chromosomes in wheat: interphase chromatin organization, nuclear disposition and associations in plants with two, three or four Bs. Chromosome Res. 1996;4(3):1–6.
61. Rawlins DJ, Shaw PJ. 3-dimensional organization of ribosomal DNA in interphase nuclei of *Pisum sativum* by in situ hybridization and optical tomography. Chromosoma. 1990;99(2):143–51.
62. Aragón-Alcaide L, Beven A, Moore G, Shaw P. The use of vibratome sections of cereal spikelets to study anther development and meiosis. Plant J. 1998;14(4):503–8.
63. Prieto P, Moore G, Shaw P. Fluorescence in situ hybridization on vibratome sections of plant tissues. Nat Protoc. 2007;2(7):1831–8.
64. Santos AP, Shaw P. Interphase chromosomes and the Rabl configuration: does genome size matter? J Microsc. 2004;214(2):201–6.
65. Prieto P, Santos AP, Moore G, Shaw P. Chromosomes associate premeiotically and in xylem vessel cells via their telomeres and centromeres in diploid rice (*Oryza sativa*). Chromosoma. 2004;112(6):300–7.
66. Santos AP, Ferreira L, Maroco J, Oliveira MM. Abiotic stress and induced DNA hypomethylation cause interphase chromatin structural changes in rice rDNA loci. Cytogenet Genome Res. 2011;132(4):297–303.
67. Lysak MA, Fransz PF, Ali HBM, Schubert I. Chromosome painting in *Arabidopsis thaliana*. Plant J. 2001;28(6):689–97.
68. Pecinka A, Schubert V, Meister A, Kreth G, Klatte M, Lysak MA, Fuchs J, Schubert I. Chromosome territory arrangement and homologous pairing in nuclei of *Arabidopsis thaliana* are predominantly random except for NOR-bearing chromosomes. Chromosoma. 2004;113(5):258–69.
69. Fransz P, de Jong JH, Lysak M, Castiglione MR, Schubert I. Interphase chromosomes in *Arabidopsis* are organized as well defined chromocenters from which euchromatin loops emanate. Proc Natl Acad Sci U S A. 2002;99(22):14584–9.
70. Armstrong SJ, Franklin FC, Jones GH. Nucleolus-associated telomere clustering and pairing precede meiotic chromosome synapsis in *Arabidopsis thaliana*. J Cell Sci. 2001;114(23):4207–17.
71. Fransz P, Soppe W, Schubert I. Heterochromatin in interphase nuclei of *Arabidopsis thaliana*. Chromosome Res. 2003;11(3):227–40.
72. Berr A, Schubert I. Interphase chromosome arrangement in *Arabidopsis thaliana* is similar in differentiated and meristematic tissues and shows a transient mirror symmetry after nuclear division. Genetics. 2007;176(2):853–63.
73. Dong F, Jiang J. Non-Rabl patterns of centromere and telomere distribution in the interphase nuclei of plant cells. Chromosome Res. 1998;6(7):551–8.
74. Funabiki H, Hagan I, Uzawa S, Yanagida M. Cell cycle-dependent specific positioning and clustering of centromeres and telomeres in fission yeast. J Cell Biol. 1993;121(5):961–76.
75. Gilson E, Laroche T, Gasser S. Telomeres and the functional architecture of the nucleus. Trends Cell Biol. 1993;3(4):128–34.
76. Palladino F, Laroche T, Gilson E, Axelrod A, Pillus L, Gasser SM. SIR3 and SIR4 proteins are required for the positioning and integrity of yeast telomeres. Cell. 1993;75(3):543–55.
77. Manuelidis L. Individual interphase chromosome domains revealed by in situ hybridization. Hum Genet. 1985;71(4):288–93.

78. Manuelidis L, Borden J. Reproducible compartmentalization of individual chromosome domains in human CNS cells revealed by in situ hybridization and three-dimensional reconstruction. Chromosoma. 1988;96(6):397–410.
79. Vourc'h C, Taruscio D, Boyle AL, Ward DC. Cell cycle-dependent distribution of telomeres, centromeres and chromosome-specific subsatellite domains in the interphase nucleus of mouse lymphocytes. Exp Cell Res. 1993;205(1):142–51.
80. Croft JA, Bridger JM, Boyle S, Perry P, Teague P, Bickmore WA. Differences in the localization and morphology of chromosomes in the human nucleus. J Cell Biol. 1999;145(6):1119–31.
81. Bridger JM, Boyle S, Kill IR. Re-modelling of nuclear architecture in quiescent and senescent human fibroblasts. Curr Biol. 2000;10(3):149–52.
82. Manders EMM, Kimura H, Cook PR. Direct imaging of DNA in living cells reveals the dynamics of chromosome formation. J Cell Biol. 1999;144(5):813–22.
83. Fussel CP. Rabl distribution of interphase and prophase telomeres in *Allium cepa* is not altered by colchicines and/or ultracentrifugation. Am J Bot. 1992;79(7):771–7.
84. Wilson KL, Foisner R. Lamin-binding proteins. Cold Spring Harb Perspect Biol. 2010;2(4):000554.
85. Fajkus J, Kovarík A, Královics R, Bezděk M. Organization of telomeric and subtelomeric chromatin in the higher plant *Nicotiana tabacum*. Mol Gen Genet. 1995;247(5):633–8.
86. Vershinin AV, Heslop-Harrison JS. Comparative analysis of the nucleosomal structure of rye, wheat and their relatives. Plant Mol Biol. 1998;36(1):149–61.
87. Ingber DE, Jamieson JD. Cells as tensegrity structures: architectural regulation of histodifferentiation by physical forces transduced over basement membrane. In: Andersson LC, Gahmberg CG, Ekblom P, editors. In gene expression during normal and malignant differentiation. Orlando: Academic; 1985. p. 13–32.
88. Pienta KJ, Getzenberg RH, Coffey DS. Cell structure and DNA organization. Crit Rev Eukaryot Gene Express. 1991;1(4):355–85.
89. Ingber DE. Cellular tensegrity: defining new rules of biological design that govern the cytoskeleton. J Cell Sci. 1993;104(3):613–27.
90. Gerlitz G, Reiner O, Bustin M. Microtubule dynamics alter the interphase nucleus. Cell Mol Life Sci. 2013;70(7):1255–68.
91. Panstruga R, Büschges R, Piffanelli P, Schulze-Lefert P. A contiguous 60 kb genomic stretch from barley reveals molecular evidence for gene islands in a monocot genome. Nucleic Acids Res. 1998;26(4):1056–62.
92. Kim JM, To TK, Nishioka T, Seki M. Chromatin regulation functions in plant abiotic stress responses. Plant Cell Environ. 2010;33(4):604–11.
93. Luo M, Liu X, Singh P, Cui Y, Zimmerli L, Wu K. Chromatin modifications and remodeling in plant abiotic stress responses. Biochim Biophys Acta. 2012;1819(2):129–36.
94. Chinnusamy V, Zhu JK. Epigenetic regulation of stress responses in plants. Curr Opin Plant Biol. 2009;12(2):133–9.
95. Santos AP, Serra T, Figueiredo DD, Barros P, Lourenco T, Chander S, Oliveira MM, Saibo NJM. Transcription regulation of abiotic stress responses in rice: a combined action of transcription factors and epigenetic mechanisms. Omics: J Integr Biol. 2011;15(12):839–57.
96. McClintock B. The significance of responses of the genome to challenge. Science. 1984;226(4676):792–801.
97. Grandbastien MA. Activation of plant transposons under stress conditions. Trends Plant Sci. 1998;3(5):181–7.
98. Madlung A, Comai L. The effect of stress on genome regulation and structure. Ann Bot. 2004;94(4):481–95.
99. Arnholdt-Schmitt B. Stress-induced cell reprogramming. A role for global genome regulation? Plant Phys. 2004;136(1):2579–86.

100. Tessadori F, van Zanten M, Pavlova P, Clifton R, Pontvianne F, Snoek LB, Millenaar FF, Schulkes RK, van Driel R, Voesenek LA, Spillane C, Pikaard CS, Fransz P, Peeters AJ. Phytochrome B and histone deacetylase 6 control light-induced chromatin compaction in *Arabidopsis thaliana*. PLoS Genet. 2009;5(9):e1000638.
101. van Zanten M, Tessadori F, Bossen L, Peeters AJM, Fransz P. Large-scale chromatin decompaction induced by low light is not accompanied by nucleosomal displacement. Plant Signal Behav. 2010;5(12):1677–8.
102. Pecinka A, Dinh HQ, Baubec T, Rosa M, Lettner N, Scheid OM. Epigenetic regulation of repetitive elements is attenuated by prolonged heat stress in *Arabidopsis*. Plant Cell. 2010;22(9):3118–29.
103. Tessadori F, Chupeau MC, Chupeau Y, Knip M, Germann S, van Driel R, Fransz P, Gaudin V. Large-scale dissociation and sequential reassembly of pericentric heterochromatin in dedifferentiated *Arabidopsis* cells. J Cell Sci. 2007;120(7):1200–8.
104. Tessadori F, Schulkes RK, van Driel R, Fransz P. Light-regulated large-scale reorganization of chromatin during the floral transition in *Arabidopsis*. Plant J. 2007;50(5):848–57.
105. Dekker J, Rippe K, Dekker M, Kleckner N. Capturing chromosome conformation. Science. 2002;295(5558):1306–11.
106. Louwers M, Bader R, Haring M, van Driel R, de Laat W, Stam M. Tissue and expression level-specific chromatin looping at maize b1 epialleles. Plant Cell. 2009;21(3):832–42.
107. Crevillén P, Sonmez C, Wu Z, Dean C. A gene loop containing the floral repressor FLC is disrupted in the early phase of vernalization. EMBO J. 2013;32(1):140–8.
108. Belmont AS, Straight AF. In vivo visualization of chromosomes using lac operator–repressor binding. Trends Cell Biol. 1998;8(3):121–4.
109. Robinett C, Straight A, Li G, Willhelm C, Sudlow G, Murray A, Belmont AS. In vivo localization of DNA sequences and visualization of large-scale chromatin organization using lac operator/repressor recognition. J Cell Biol. 1996;135(6):1685–700.
110. Belmont AS, Li G, Sudlow G, Robinett C. Visualization of large-scale chromatin structure and dynamics using the lac operator/lac repressor reporter system. Methods Cell Biol. 1999;58:203–22.
111. Michaelis C, Ciosk R, Nasmyth K. Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. Cell. 1997;91:35–45.
112. Fuchs J, Lorenz A, Loidl J. Chromosome associations in budding yeast caused by integrated tandemly repeated transgenes. J Cell Sci. 2002;115:1213–20.
113. Kato N, Lam E. Detection of chromosomes tagged with green fluorescent protein in live *Arabidopsis thaliana* plants. Genome Biol. 2001;2(11):research 0045.1–research 0045.10.
114. Kato N, Lam E. Chromatin of endoreduplicated pavement cells has greater range of movement than that of diploid guard cells in *Arabidopsis thaliana*. J Cell Sci. 2003;116(11):2195–201.
115. Rosin FM, Watanabe N, Cacas JL, Kato N, Arroyo JM, Fang Y, May B, Vaughn M, Simorowski J, Ramu U. Genome-wide transposon tagging reveals location-dependent effects on transcription and chromatin organization in *Arabidopsis*. Plant J. 2008;55(3):514–25.
116. Matzke AJ, Huettel B, van der Winden J, Matzke M. Use of two-color fluorescence-tagged transgenes to study interphase chromosomes in living plants. Plant Physiol. 2005;139(4):1586–96.
117. Matzke AJ, Watanabe K, van der Winden J, Naumann U, Matzke M. High frequency, cell type-specific visualization of fluorescent-tagged genomic sites in interphase and mitotic cells of living *Arabidopsis* plants. Plant Methods. 2010;6:2.
118. Rosa S, de Lucia F, Mylne JS, Zhu D, Ohmido N, Pendle A, Kato N, Shaw P, Dean C. Physical clustering of FLC alleles during polycomb-mediated epigenetic silencing in vernalization. Genes Dev. 2013;27(17):1845–50.
119. Pecinka A, Kato N, Meister A, Probst AV, Schubert I, Lam E. Tandem repetitive transgenes and fluorescent chromatin tags alter local interphase chromosome arrangement in *Arabidopsis thaliana*. J Cell Sci. 2005;118(16):3751–8.

120. Bertrand E, Chartrand P, Schaefer M, Shenoy SM, Singer RH, Long RM. Localization of ASH1 mRNA particles in living yeast. Mol Cell. 1998;2(4):437–45.
121. Schonberger J, Hammes UZ, Dresselhaus T. In vivo visualization of RNA in plants cells using the kN22 system and a GATEWAY-compatible vector series for candidate RNAs. Plant J. 2012;71(1):173–81.
122. Tilsner J, Linnik O, Christensen NM, Bell K, Roberts IM, Lacomme C, Oparka KJ. Live-cell imaging of viral RNA genomes using a Pumilio-based reporter. Plant J. 2009;57(4):758–70.
123. Paige JS, Wu KY, Jaffrey SR. RNA mimics of green fluorescent protein. Science. 2011;333(6042):642–6.
124. Levsky JM, Singer RH. Fluorescence in situ hybridization: past, present and future. J Cell Sci. 2003;116(14):2833–8.
125. Markaki Y, Smeets D, Fiedler S, Schmid VJ, Schermelleh L, Cremer T, Cremer M. The potential of 3D-FISH and super-resolution structured illumination microscopy for studies of 3D nuclear architecture: 3D structured illumination microscopy of defined chromosomal structures visualized by 3D (immuno)-FISH opens new perspectives for studies of nuclear architecture. Bioessays. 2012;34(5):412–26.
126. Hübner B, Cremer T, Neumann J. Correlative microscopy of individual cells: sequential application of microscopic systems with increasing resolution to study the nuclear landscape. Methods Mol Biol. 2013;1042:299–336.

Chapter 5
Role of Epigenetic Modifications in Plant Responses to Environmental Stresses

Xuncheng Liu, Ming Luo, Songguang Yang and Keqiang Wu

Introduction

In eukaryotes, genomic DNA is packaged with histones to form chromatin. The fundamental unit of chromatin is the nucleosome, which is composed of ~146 base pairs of DNA wrapped on a histone octamer containing two copies of histone H2A, H2B, H3 and H4. Chromatin structure influences the accessibility of transcription factors and cofactor for DNA-tempered processes. The structure and function of chromatin are regulated by multiple epigenetic mechanisms, including DNA methylation, histone modifications, adenosine triphosphate (ATP)-dependent chromatin remodelling, placement of histone variants and regulation by noncoding RNA [1]. Among the multiple epigenetic mechanisms that alter chromatin to regulate gene expression, histone modifications are the most versatile. The N-terminal tails of the histone proteins can undergo a variety of posttranslational modifications, including acetylation, methylation, ubiquitination, phosphorylation and sumoylation. All histone modifications are removable, which may therefore provide a flexible way for regulation of gene expression in plant response to environmental changes. The large number of possible histone modification patterns allows highly specific and complex signalling mechanisms and particular histone modification patterns may be associated with specific transcriptional effects. Generally, lysine acetylation correlates with transcriptional activation, whereas lysine methylation leads to either transcriptional activation or repression depending on

K. Wu (✉)
Institute of Plant Biology, College of Life Science, National Taiwan University, Taipei, Taiwan
e-mail: kewu@ntu.edu.tw

X. Liu · S. Yang
South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, Guangdong, China

M. Luo
Key Laboratory of Plant Resources Conservation and Sustainable Utilization, South China Botanical Garden, Chinese Academy of Sciences, Guangzhou 510650, Guangdong, China

O. Pontes, H. Jin (eds.), *Nuclear Functions in Plant Transcription, Signaling and Development,* DOI 10.1007/978-1-4939-2386-1_5

which residues are modified and the type of modification present. In particular, methylation of lysine 4 and lysine 36 on histone H3 are associated with gene activation, whereas methylation of lysine 9 and lysine 27 are associated with gene silencing. Recent reports have shown that different environmental stresses led to altered methylation status of DNA, histone modifications and chromatin remodelling [2–4].

Epigenetic Changes in Plant Responses to Environmental Stresses

Environmental stresses can alter gene expression through modulating DNA methylation in plants. An increase in cytosine methylation at CG sites was observed in rapeseed (*Brassica napus*) under salt stress [5]. In tobacco (*Nicotiana tabacum*), antisense silence of the DNA methyltransferase gene *NtMET1* released a subset of stress-responsive genes, including a glycerophosphodiesterase-like gene, *NtGPDL* [6]. Hypomethylation of the coding sequence of *NtGPDL* was observed under salt and cold stresses, suggesting that the expression of stress-responsive genes might be correlated with DNA methylation. Furthermore, chromium (potassium dichromate) stress increased the cytosine methylation of radish (*Raphanus sativus* L.), and the hypermethylation pattern was correlated with the stress dosage, suggesting that chromium stress could induce de novo cytosine methylation [7].

Changes in histone modifications in plant responses to environmental stresses have also been documented. Tobacco and *Arabidopsis* cells show dynamic changes in histone modifications in response to high salinity and cold stress, manifested by transient up-regulation of H3 phosphoacetylation and histone H4 acetylation [8]. Enrichment of histone acetylation at H3K23 and H3K27 occurs in response to drought stress on the coding regions of drought stress-responsive genes, *RD29B, RD20* and *At2g20880* [9]. Drought stress causes dynamic genome-wide changes of histone H3K4me1, H3K4me2 and H3K4me3 in *Arabidopsis* [10]. Furthermore, the plant hormone abscisic acid (ABA) was reported to regulate stress-inducible gene expression by affecting histone H3 acetylation and methylation [3]. Both ABA and salt stress can induce histone H3K9K14 acetylation and H3K4me3 but decrease H3K9me2 of some stress-responsive genes, suggesting that functionally related genes are regulated coordinately through histone modifications in response to stresses in plant cells.

In rice seedlings, hypoxia leads to the acetylation of histone H3 and to the conversion of H3K4me2 to H3K4me3 in stress-responsive genes *ADH1* and *PDC1* [11]. Both modifications, which are associated with active transcription, were reversible after the removal of the stress. In addition, 4837 genes are differentially H3K4me3 modified in rice seedlings under drought stress [12]. In maize seedlings, a genome-wide demethylation was observed in root tissues upon cold exposure [13]. Cold treatment also induces the expression of the genes encoding histone deacetylases (HDACs), leading to global histones H3 and H4 deacetylation in maize [14]. Taken

together, these studies suggest that posttranslational histone modifications play an important role in modulating the gene activity in plant responses to environmental stresses.

Histone Acetyltransferases

Specific lysine residues on core histone are acetylated and deacetylated by histone acetyltransferases (HATs) and HDACs, respectively. HATs are classified into four different families: GCN5-related N-terminal acetyltransferase (GNAT), MOZ, Ybf2/Sas3, Sas2 and Tip60 (MYST), CREB-binding protein (CBP) and TATA-binding protein-associated factors (TAFII250) family, with a total of 12 members in *Arabidopsis* [15, 16]. The general control non-de-repressible 5 (GCN5) protein is the catalytic subunit of several multi-protein HAT complexes. Alteration/deficiency in activation 2 (ADA2) and SGF29 are the components of GCN5-containing complexes. Two ADA2-related factors (ADA2a and ADA2b) and two SGF29-like proteins (SGF29a and SGF29b) have been identified in *Arabidopsis* [17, 18]. *ada2b* mutants were hypersensitive, whereas *sgf29a* mutants were more resistant to salt stress than their wild-type counterparts, suggesting that ABA2B and SGF29a may play an antagonistic role in plant abiotic stress responses [18].

The loss of function of *GCN5/HAG1* and *ADA2* affects the expression of several cold-regulated (*COR*) genes and tolerance to freezing temperatures [19]. Further analyses indicated that *Arabidopsis* AtGCN5 and ADA2 proteins interact with the transcription factor C-repeat/DRE-binding factor 1 (CBF1) that is responsible for cold-induced gene expression. Cold acclimation resulted in H3 acetylation increases and nucleosome occupancy decreases at *COR* promoters [20]. Furthermore, a constitutive increase in H3 acetylation and decrease in nucleosome occupancy were also observed in transgenic plants overexpressing *CBF1*. These data indicate that *CBF1* might stimulate transcription through the recruitment of HAT complexes to the promoters of its target genes. Both *ada2b* and *gcn5* mutant plants show diminished expression of *COR* genes during cold acclimation. H3 acetylation at *COR* promoters was stimulated upon cold acclimation in *ada2b* and *gcn5* plants as in wild-type plants, but the decrease in nucleosome occupancy was diminished, suggesting GCN5 and ADA2b are not required for cold-stimulated histone acetylation at *COR* gene promoters, but they are required for changes in nucleosome occupancy during cold acclimation.

In yeast, the Elongator complex is a HAT complex consisting of six proteins, elongation protein (Elp)1 to Elp6 [21]. Elp3 has motifs characteristic of the GCN5 family of HATs and is capable of acetylating the core histones in vitro, whereas the other components of the complex are necessary for in vivo histone acetylation and confer specificity towards histone H3 and H4. ABA-overly sensitive 1 (ABO1)/ELO2 is an *Arabidopsis* Elp1 homolog [22]. The *abo1* mutant is hypersensitive to ABA in seed germination and seedling growth with increased ABA-induced stomatal closure and drought resistance, suggesting a role for ABO1/ELO2 in ABA

signalling [22]. Furthermore, *Arabidopsis elp1/abo1/elo2, elp2, elp4/elo1* and *elp6* mutants all display narrow leaves, reduced root growth, ABA hypersensitivity and an increased accumulation of anthocyanins [22, 23]. In addition, *elp1, elp2, elp4* and *elp6* mutants were all more resistant to oxidative stress and caesium chloride (CsCl) than the wild type. These results suggest that Elongator plays crucial roles in regulating plant responses to abiotic stress in *Arabidopsis*.

Histone Deacetylases

HDACs are classified into three major families, namely RPD3/HDA1, silent information regulator 2 (SIR2) and HD2 family, with total 16 members in *Arabidopsis* [15, 24]. Mutations in a *RPD3/HDA1* family *HDAC, HDA6*, affected gene expression, DNA methylation and plant development in *Arabidopsis* [25–31]. The loss of function of another *RPD3/HDA1 HDAC, HDA19*, affects a range of plant development processes [32–37]. HDA19-overexpressing plants showed an increased resistance to the pathogen *Alternaria brassicicola* [36]. Further analysis indicated that HDA19 can interact with WRKY transcription factors to regulate plant basal defence responses [38]. HDA15 and HDA18 are required for the chlorophyll biosynthesis and root cellular patterning in *Arabidopsis*, respectively [39, 40].

Both HDA6 and HDA19 are involved in gene regulation in ABA and salt stress responses in *Arabidopsis* [3, 41, 42]. *hda6* and *hda19* mutant plants are more sensitive to ABA and salt stress compared with wild-type plants. In addition, the expression of ABA and abiotic stress-responsive genes was decreased in *hda6* and *hda19* plants. A recent study indicated that HDA6 is also required for freezing tolerance in *Arabidopsis* [43]. In *Arabidopsis*, HDA19 may be functionally associated with the ethylene response factor (ERF) transcription repressors, AtERF3, AtERF4 and AtERF7, in gene regulation [44, 45]. AtERF3, AtERF4 and AtERF7 are EAR-motif-containing transcriptional repressors that are involved in regulating ABA and abiotic stress responses in *Arabidopsis* [46]. It was found that AtERF7 interacts with the *Arabidopsis* homologue of a human global corepressor, AtSin3, which in turn may interact with HDA19, suggesting that AtERF7, AtSin3 and HDA19 may form a transcriptional repressor complex [44].On the other hand, AtERF3 and AtERF4 can interact with AtSAP18, an orthologue of human SAP18 that is a subunit of the human HDAC complex [45]. Taken together, ERF repressors such as AtERF3, AtERF4 and AtERF7 may form a transcription complex with AtSAP18-, AtSin3- and HDA19-involved gene regulation in ABA and abiotic responses. The observations that *hda6* and *hda19* mutants responded similarly to ABA and salt stress as well as the similar gene expression profiles affected in the two mutants suggest that HDA6 and HDA19 may play a redundant role in ABA and abiotic stress response [41]. It remains to be determined whether HDA6 can also form a transcriptional complex with ERF repressors to regulate gene expression in *Arabidopsis*.

Histone deacetylase 2C (HD2C) is a HD2-type HDAC involved in ABA and salt stress response in *Arabidopsis*. *HD2C* overexpression conferred an ABA-insensitive

phenotype and enhanced tolerance to salt and drought stresses, whereas the loss-of-function mutants of *HD2C* displayed increased sensitivity to ABA and decreased tolerance to salt stress [42, 47]. Furthermore, HD2C interacts physically with the RPD3-type histone deacetylases, HDA6 and HDA19, suggesting that HD2C may functionally associate with HDA6 and HDA19. Therefore, HD2C is a part of the HDAC complexes to regulate gene expression involved in stress responses.

High expression of osmotically responsive genes 15 (HOS15) is a WD40-repeat protein crucial for repression of genes in cold stress tolerance through histone deacetylation in *Arabidopsis* [48]. HOS15 protein shares a high-sequence similarity with the human protein transducin beta-like 1 (TBL1), a component of the human silencing mediator for retinoid and thyroid receptors/nuclear receptor co-repressor 1 (SMRT/N-CoR) repressor complex that is involved in modification of chromatin structure by association with HDACs [49]. The level of acetylated histone H4 is higher in the *hos15* mutant than that in wild-type plants, suggesting that HOS1 is a component of the HDAC complex involved in histone deacetylation. Further research is required to identify the specific HDAC associated with HOS15.

Histone Lysine Methyltransferases

Histone lysine methyltransferases (HKMTs) are SET domain proteins that are classified into five classes in plants [50]. Class I, II and IV HKMTs are homologues of enhancer of zeste (EZ), Trithorax and SUVH (SU(VAR)3–9 of *Drosophila*, respectively. *Arabidopsis* homolog of trithorax (ATX1) is a class III HKMT involved in drought stress response [51]. A link between phosphatidylinositol 5-phosphate (PtdIns5P) and the activity of ATX1 in response to dehydration stress was proposed [52]. The ATX1-dependent gene, *WRKY70*, is down-regulated during dehydration accompanied by a drastic reduction in H3K4me3 of its nucleosomes and an increase in cellular PtdIns5P in *Arabidopsis*. During dehydration stress, the binding of ATX1 to the *WRKY70* locus was diminished and ATX1 was retained in the cytoplasm when PtdIns5P was elevated. These findings indicate a mechanistic link between ATX1 and a lipid (PtdIns5P) synthesis in a signalling pathway that ultimately results in decreased expression of ATX1-dependent genes in response to dehydration stress.

The WD40 protein, multicopy suppressor of IRA1 (MSI1), is a subunit of the Polycomb group (PcG) complex that has an H3K27me3 activity [53]. *Arabidopsis* has five genes encoding MSI1-like proteins, MSI1–MSI5. Strong reduction of *MSI1* in transgenic co-suppression lines (msi1-cs), which contain only about 5% of the wild-type *MSI1* level, results in pleiotropic phenotypes [54]. Transcriptional profiling analysis revealed the up-regulation of a subset of ABA-responsive genes in the transgenic *MSI1* co-suppression plants. Furthermore, MSI1 can bind to the chromatin of *RD20*, a drought-inducible gene. In addition, MSI4/FVE is a negative regulator of *COR* genes containing the C-repeat/dehydration-responsive element (C/DRE) [55]. *msi4* mutants display enhanced expression of *COR* genes and greater

freezing tolerance than wild-type plants when cold acclimated. These results demonstrate the involvement of MSI1 and MSI4 in abiotic stress responses in *Arabidopsis*.

Arginine methylation mainly occurs at Arg2 (R2), Arg8 (R8), Arg17 (R17), Arg26 (R26) of histone H3 and Arg3 (R3) of histone H4 and is catalysed by a small group of protein arginine methyltransferases (PRMTs) [56]. An *Arabidopsis* PRMT, SKB1, also named protein arginine methyltransferase 5 (PRMT5), is involved in abiotic stress response. *skb1* mutants are hypersensitive to salt and defective in splicing [57]. SKB1 associates with chromatin and increases the level of histone H4 arginine 3 (H4R3) symmetric dimethylation (H4R3sme2) to suppress the transcription of a number of stress-responsive genes. Under salt stress, the H4R3sme2 level is reduced, as a result of SKB1 disassociation from chromatin and consequently the induction of stress-responsive genes. These data suggest that SKB1 mediates the salt response by altering the methylation status of H4R3sme2.

ATP-Dependent Chromatin Remodelling Complexes

ATP-dependent chromatin remodelling uses the energy of ATP hydrolysis to alter the structure of chromatin [58, 59]. There are three main classes of ATP-dependent chromatin remodelling complexes: the SWI/SNF adenosine triphosphatases (ATPases), the imitation switch (ISWI) ATPases, and the chromodomain and helicase-like domain (CHD) ATPases. The role of an *Arabidopsis* SNF2/Brahma (BRM)-type chromatin-remodelling protein, AtCHR12, in plant growth response to adverse environmental conditions was investigated [60]. Exposing an *AtCHR12*-overexpressing mutant to stress conditions such as drought, heat and salinity leads to growth arrest of normally active primary buds, as well as to reduced growth of the primary stem. In contrast, the *AtCHR12* knockout mutant shows less growth arrest than the wild type when exposed to stress. Modulation of *AtCHR12* expression correlates with changes in the expression of dormancy-associated genes. These results suggest that AtCHR12 plays a vital role in mediating the temporary growth arrest of *Arabidopsis* under stress conditions. More recently, the SWI2/SNF2 chromatin remodelling ATPase, BRM (BRAHMA), was also shown to play an essential role in response of stresses in *Arabidopsis*. *brm* mutants display increased drought tolerance [61]. The loss-of-function of *BRM* led to destabilization of nucleosomes and repression of *ABI5* transcription, indicating that BRM regulates stress response through the regulation of nucleosome stability of *ABI5*.

The *Arabidopsis* SWI3 subunit of SWI/SNF complexes, SWI3B, was found to act as a positive regulator in ABA-mediated inhibition of seed germination and growth [62]. In addition, SWI3B can interact with hypersensitive to ABA1 (HAB1), a protein phosphatase type 2C acting as a negative regulator of ABA signalling. *swi3b* mutants showed a reduced sensitivity to ABA, and the expression of the ABA-responsive genes, *RAB18* and *RD29B*, was decreased, indicating that SWI3B is a positive regulator of ABA signalling. Further, chromatin immunoprecipitation

experiments showed that the association of HAB1 with the promoters of *RD29B* and *RAB18* was abolished by ABA, indicating the involvement of HAB1 in the regulation of ABA-induced transcription. Taken together, these results suggest that SWIB may act with HAB1 to modulate ABA response.

PICKLE (PKL) is a CHD3 ATPase involved in the repression of *ABI3* and *ABI5* during seed germination in response to ABA [63]. *pkl* mutants displayed a high expression of *ABI3* and *ABI5* upon ABA stimulation, which may lead to hypersensitive germination responses to ABA in *pkl* seeds. ABA-treated *pkl* mutant seeds had lower H3K9 and H3K27 methylation levels at *ABI3* and *ABI5* promoters, suggesting that PKL affect the expression of *ABI* genes by reducing repressive histone marks. These data indicate that PKL is necessary to maintain *ABI3* and *ABI5* chromatin in a repressed state during germination.

Crosstalk of DNA Methylation and Histone Modifications

In many organisms, silent genes are not only DNA methylated but also deacetylated at histones H3 and H4 [64]. In mammalian cells, crosstalk between DNA methylation and histone deacetylation is well supported on the level of protein interactions through the direct interaction between HDACs and DNA methyltransferases (DMTs) or the recruitment of HDACs to methylated DNA via methyl CpG-binding proteins (MBPs) [65]. More recently, it was found that the *Arabidopsis* DMTs MET1 and DNMT2 associate with the histone deacetylases HDA6 and HD2, respectively [66, 67]. The involvement of HDA6 and HD2C in ABA and abiotic stress responses in *Arabidopsis* have been reported [3, 42]. The interaction between DMTSs and HDACs suggests a possible interplay among DMTs and histone deacetylase enzymes in abiotic stress response.

The crosstalk between DNA methylation and histone methylation was also reported [68]. In *Arabidopsis*, CG methylation is propagated during DNA replication by DNA methyltransferase 1 (MET1), which robustly copies methylation patterns on newly synthesized DNA strands. The maintenance of asymmetrical CHH methylation is mostly ensured by domains rearranged methyltransferase 2 (DRM2) in a process known as RNA-directed DNA methylation (RdDM), which involves the RNA polymerases IV and V [68]. The perpetuation of CHG methylation patterns is largely ensured by the plant-specific chromomethylase CMT3, and genetic analyses suggest that targeting of CMT3 to chromatin relies on H3K9me2, indicating that H3K9me2 acts upstream of CHG methylation [69, 70]. More recently, it was found that CMT3 associates with H3K9me2-containing nucleosomes through dual binding of its bromo-adjacent homology (BAH) and chromo domains to H3K9me2 in order to target DNA methylation [71]. Furthermore, the JmjC-domain histone demethylase IBM1 represses gene expression by H3K9me2 demethylating and DNA methylation [72].

ABA and salt stress can enrich the gene activation marks, H3K9K14, acetylation and H3K4me3 but decrease the gene repression mark, H3K9m2, of ABA and salt-

responsive genes [3, 6]. Enrichment of acetylation of histone H3K23 and H3K27 in response to drought stress on the coding regions of *RD29B* and *RD20* was also reported [9]. These results suggest that the induction of ABA and abiotic stress-responsive genes is associated with changes in histone acetylation and methylation. Compared with wild type, ABA and salt-induced H3K4me3 were decreased in the *hda6* mutant, suggesting that the histone deacetylases HDA6 is required for H3K4 trimethylation associated with gene activation [3]. Taken together, these observations suggest a synergistic interplay between histone methylation and acetylation enzymes. Crosstalks between histone deacetylation and demethylation have previously been reported to modulate gene expression in mammalian cells [73, 74]. The mammalian histone demethylase, lysine-specific demethylase 1 (LSD1), is an integral component of histone deacetylase corepressor complexes in which HDACs and LSD1 may cooperate to remove activating acetyl and methyl histone modifications. HDAC inhibitors can diminish histone demethylation activity, whereas the abrogation of LSD1 activity by mutations can decrease the deacetylation activity [73], suggesting that the enzymatic activities of HDACs and LSD1 are closely linked. More recently, it was found that *Arabidopsis* HDA6 physically associates with the LSD1 histone demethylase FLD in vitro and in vivo, indicating that these two proteins act in the same protein complex [31]. Taken together, these results suggest that epigenetic regulation involves the integration of multiple chromatin-modifying activities such as HDACs and histone demethylases (HDMs) acting in a coordinated fashion. Further research is required to study the crosstalk among DNA methylation and histone modifications in plant responses to environmental stresses.

Outlook

Recent studies demonstrate the importance of histone modifications and chromatin remodelling in plant stress responses. However, the function for most proteins involved in epigenetic modifications remains largely unknown in plants. It would be intriguing to identify additional histone modifiers and chromatin factors involved in plant responses to environmental stresses. Further studies on how epigenetic modifications are involved in plant stress responses will contribute significantly to our understanding of the molecular mechanisms underlying plant epigenetic regulation.

References

1. Berger SL. The complex language of chromatin regulation during transcription. Nature. 2007;447:407–12.
2. Boyko A, Kovalchuk I. Epigenetic control of plant stress response. Environ Mol Mutagen. 2008;49:61–72.
3. Chen L, Luo M, Wang Y, Wu K. Involvement of *Arabidopsis* histone deacetylase HDA6 in ABA and salt stress response. J Exp Bot. 2010;61:3345–53.

4. Kim JM, To TK, Nishioka T, Seki M. Chromatin regulation functions in plant abiotic stress responses. Plant Cell Environ. 2010;33:604–11.
5. Labra M, Gianazza E, Waitt R, Eberini I, Sozzi A, Regondi S, Grassi F, Agradi E. Genetic and DNA-methylation changes induced by potassium dichromate in *Brassica napus* L. Chemosphere. 2004;54:1049–58.
6. Wada Y, Miyamoto K, Kusano T, Sano H. Association between up-regulation of stress-responsive genes and hypomethylation of genomic DNA in tobacco plants. Mol Genet Genom. 2004;271:658–66.
7. Labra M, Ghiani A, Citterio S, Sgorbati S, Sala F, Vannini C, Ruffini-Castiglione M, Bracale M. Analysis of cytosine methylation pattern in response to water deficit in pea root tips. Plant Biol. 2002;4:694–9.
8. Sokol A, Kwiatkowska A, Jerzmanowski A, Prymakowska-Bosak M. Up-regulation of stress-inducible genes in tobacco and *Arabidopsis* cells in response to abiotic stresses and ABA treatment correlates with dynamic changes in histone H3 and H4 modifications. Planta. 2007;227:245–54.
9. Kim JM, To TK, Ishida J, Morosawa T, Kawashima M, Matsui A, Toyoda T, Kimura H, Shinozaki K, Seki M. Alterations of lysine modifications on the histone H3N-tail under drought stress conditions in *Arabidopsis thaliana*. Plant Cell Physiol. 2008;49:1580–8.
10. van Dijk K, Ding Y, Malkaram S, Riethoven JJ, Liu R, Yang J, Laczko P, Chen H, Xia Y, Ladunga I, Avramova Z, Fromm M. Dynamic changes in genome-wide histone H3 lysine 4 methylation patterns in response to dehydration stress in *Arabidopsis thaliana*. BMC Plant Biol. 2010;10:238. doi:10.1186/1471-2229-10-238.
11. Tsuji H, Saika H, Tsutsumi N, Hirai A, Nakazono M. Dynamic and reversible changes in histone H3-Lys4 methylation and H3 acetylation occurring at submergence-inducible genes in rice. Plant Cell Physiol. 2006;47:995–1003.
12. Zong W, Zhong X, You J, Xiong L. Genome-wide profiling of histone H3K4-tri-methylation and gene expression in rice under drought stress. Plant Mol Biol. 2012;81:175–88.
13. Steward N, Ito M, Yamaguchi Y, Koizumi N, Sano H. Periodic DNA methylation in maize nucleosomes and demethylation by environmental stress. J Biol Chem. 2002;277:37741–6.
14. Hu Y, Zhang L, Zhao L, Li J, He S, Zhou K, Yang F, Huang M, Jiang L, Li L. Trichostatin A selectively suppresses the cold-induced transcription of the *ZmDREB1* gene in maize. PLoS ONE. 2011;6:e22132.
15. Pandey R, Muller A, Napoli CA, Selinger DA, Pikaard CS, Richards EJ, Bender J, Mount DW, Jorgensen RA. Analysis of histone acetyltransferase and histone deacetylase families of *Arabidopsis thaliana* suggests functional diversification of chromatin modification among multicellular eukaryotes. Nucleic Acids Res. 2002:30:5036–55.
16. Liu X, Luo M, Zhang W, Zhao J, Zhang J, Wu K, Tian L, Duan J. Histone acetyltransferases in rice (*Oryza sativa* L.): phylogenetic analysis, subcellular localization and expression. BMC Plant Biol. 2012;12:145. doi:10.1186/1471-2229-12-145.
17. Stockinger EJ, Mao Y, Regier MK, Triezenberg SJ, Thomashow MF. Transcriptional adaptor and histone acetyltransferase proteins in *Arabidopsis* and their interactions with CBF1, a transcriptional activator involved in cold-regulated gene expression. Nucleic Acids Res. 2001;29:1524–33.
18. Kaldis A, Tsementzi D, Tanriverdi O, Vlachonasios KE. *Arabidopsis thaliana* transcriptional co-activators ADA2b and SGF29a are implicated in salt stress responses. Planta. 2011;233:749–62.
19. Vlachonasios KE, Thomashow MF, Triezenberg SJ. Disruption mutations of ADA2b and GCN5 transcriptional adaptor genes dramatically affect *Arabidopsis* growth, development, and gene expression. Plant Cell. 2003;15:626–38.
20. Pavangadkar K, Thomashow MF, Triezenberg SJ. Histone dynamics and roles of histone acetyltransferases during cold-induced gene regulation in *Arabidopsis*. Plant Mol Biol. 2010;74:183–200.
21. Versées W, De Groeve S, Van Lijsebettens M. Elongator, a conserved multitasking complex? Mol Microbiol. 2010;76:1065–9.

22. Chen Z, Zhang H, Jablonowski D, Zhou X, Ren X, Hong X, Schaffrath R, Zhu JK, Gong Z. Mutations in ABO1/ELO2, a subunit of holo-Elongator, increase abscisic acid sensitivity and drought tolerance in *Arabidopsis thaliana*. Mol Cell Biol. 2006;26:6902–12.
23. Zhou X, Chen Z, Hua D, Zhou Z, Gong Z. Elongator mediates ABA responses, oxidative stress resistance, and anthocyanin biosynthesis in *Arabidopsis*. Plant J. 2009;60:79–90.
24. Alinsug MV, Yu C-W, Wu K. Phylogenetic analysis, subcellular localization, and expression patterns of RPD3/HDA1 family histone deacetylases in plants. BMC Plant Biol. 2009;9:37. doi:10.1186/1471-2229-9-37.
25. Aufsatz W, Mette WF, van der Winden J, Matzke AJM. HDA6, a putative histone deacetylase needed to enhance DNA methylation induced by double-stranded RNA. EMBO J. 2002;21:6832–41.
26. Earely K, Lawrence RJ, Pontes O, Reuther R, Enciso AJ, Silva M, Neves N, Gross M, Viegas W, Pikaard CS. Erasure of histone acetylation by *Arabidopsis* HDA6 mediates large-scale gene silencing in nucleolar dominance. Genes Dev. 2006;20:1283–93.
27. Luo M, Yu CW, Chen F, Zhao L, Tian G, X Liu, Cui Y, Yang JY, Wu K. Histone deacetylase HDA6 is functionally associated with AS1 in repression of *KNOX* genes in *Arabidopsis*. PLoS Genet. 2012;8:e1003114. doi:10.1371/journal.pgen.1003114.
28. Murfett J, Wang X, Hagen G, Guilfoyle TJ. Identification of *Arabidopsis* histone deacetylase HDA6 mutants that affect transgene expression. Plant Cell. 2001;13:1047–61.
29. Probst AV, Fagard M, Proux F, Mourrain P, Boutet S, Earley K, Lawrence RJ, Pikaard CS, Murfett J, Furner I, Vaucheret H, Scheid OM. *Arabidopsis* histone deacetylase HDA6 is required for maintenance of transcriptional gene silencing and determines nuclear organization of rDNA repeats. Plant Cell. 2004;16:1021–34.
30. Wu K, Zhang L, Zhou C, Yu C, Chaikam V. HDA6 is required for jasmonate response, senescence and flowering in *Arabidopsis*. J Exp Bot. 2008;59:225–34.
31. Yu CW, Liu X, Luo M, Chen C, Lin X, Tian G, Lu Q, Cui Y, Wu K. HDA6 interacts with FLD and regulates flowering in *Arabidopsis*. Plant Physiol. 2011;156:173–84.
32. Long JA, Ohno C, Smith ZR, Meyerowitz EM. TOPLESS regulates apical embryonic fate in *Arabidopsis*. Science. 2006;312:1520–3.
33. Tian L, Chen ZJ. Blocking histone deacetylation in *Arabidopsis* induces pleiotropic effects on plant gene regulation and development. Proc Natl Acad Sci U S A. 2001;98:200–5.
34. Tian L, Wang J, Fong MP, Chen M, Cao H, Gelvin SB, Chen ZJ. Genetic control of developmental changes induced by disruption of *Arabidopsis* histone deacetylase 1 (AtHD1) expression. Genetics. 2003;165:399–409.
35. Wu K, Malik K, Tian L, Brown D, Miki B. Functional analysis of a RPD3 histone deacetylase homolog in *Arabidopsis thaliana*. Plant Mol Biol. 2000;44:167–76.
36. Zhou C, Zhang L, Duan J, Miki B, Wu K. Histone deacetylase 19 is involved in jasmonic acid and ethylene signaling of pathogen response in *Arabidopsis*. Plant Cell. 2005;17:1196–204.
37. Zhou Y, Tan B, Luo M, Lin Y, Liu C, Chen C, Yu C-Wu, Yang S, Dong S, Ruan J, Yuan L, Zhang Z, Zhao L, Li C, Chen H, Cui Y, Wu K, Huang S. Histone deacetylase 19 interacts with HSL1 and participates in the repression of seed maturation genes in *Arabidopsis* seedlings. Plant Cell. 2013;25,134–48.
38. Kim K-C., Lai Z, Fan B, Chen Z. *Arabidopsis* WRKY38 and WRKY62 transcription factors interact with histone deacetylase 19 in basal defense. Plant Cell. 2008;20:2357–71.
39. Liu X, Chen C, Wang K, Luo M, Tai R, Yuan L, Zhao M, Yang S, Tian G, Cui Y, Hsieh H, Wu K. Phytochrome interacting factor 3 associates with the histone deacetylase HDA15 in repression of chlorophyll biosynthesis and photosynthesis in etiolated *Arabidopsis* seedlings. Plant Cell. 2013;25:1258–73.
40. Xu CR, Liu C, Wang YL, Li LC, Chen WQ, Xu ZH, Bai SN. Histone acetylation affects expression of cellular patterning genes in the *Arabidopsis* root epidermis. Proc Natl Acad Sci U S A. 2005;102:14469–74.
41. Chen L, Wu K. Role of histone deacetylases HDA6 and HDA19 in ABA and abiotic stress response. Plant Signal Behav 2010;5:1318–20.

42. Luo M, Wang YY, Liu X, Yang S, Lu Q, Cui Y, Wu K. HD2C interacts with HDA6 and is involved in ABA and salt stress response in *Arabidopsis*. J Exp Bot. 2012;63:3297–306.
43. To TK, Nakaminami K, Kim JM, Morosawa T, Ishida J, Tanaka M, Yokoyama S, Shinozaki K, Seki M. *Arabidopsis* HDA6 is required for freezing tolerance. Biochem Biophys Res Commun. 2011;406:414–9.
44. Song CP, Agarwal M, Ohta M, Guo Y, Halfter U, Wang P, Zhu JK. Role of an *Arabidopsis* AP2/EREBP-type transcriptional repressor in abscisic acid and drought stress responses. Plant Cell. 2005;17:2384–96.
45. Song CP, Galbraith DW. AtSAP18, an orthologue of human SAP18, is involved in the regulation of salt stress and mediates transcriptional repression in *Arabidopsis*. Plant Mol Biol. 2006;60:241–57.
46. Yang Z, Tian L, Latoszek-Green M, Brown D, Wu K. *Arabidopsis* ERF4 is a transcriptional repressor capable of modulating ethylene and abscisic acid responses. Plant Mol Biol. 2005;58:585–96.
47. Sridha S, Wu K. Identification of *AtHD2C* as a novel regulator of abscisic acid responses in *Arabidopsis*. Plant J. 2006;46:124–33.
48. Zhu J, Jeong JC, Zhu Y, Sokolchik I, Miyazaki S, Zhu JK, Hasegawa PM, Bohnert HJ, Shi H, Yun DJ, Bressan RA. Involvement of *Arabidopsis* HOS15 in histone deacetylation and cold tolerance. Proc Natl Acad Sci U S A. 2008;105:4945–50.
49. Yoon HG, Chan DW, Huang ZQ, Li J, Fondell JD, Qin J, Wong J. Purification and functional characterization of the human N-CoR complex: the roles of HDAC3, TBL1 and TBLR1. EMBO J. 2003;2:1336–46.
50. Thorstensen T, Grini PE, Aalen RB. SET domain proteins in plant development. Biochim Biophys Acta. 2011;1809:407–720.
51. Ding Y, Lapko H, Ndamukong I, Xia Y, Al-Abdallat A, Lalithambika S, Sadder M, Saleh A, Fromm M, Riethoven JJ, Lu G, Avramova Z. The *Arabidopsis* chromatin modifier ATX1, the myotubularin-like AtMTM and the response to drought. Plant Signal Behav. 2009;4:1049–58.
52. Ndamukong I, Jones DR, Lapko H, Divecha N, Avramova Z. Phosphatidylinositol 5-phosphate links dehydration stress to the activity of *Arabidopsis* trithorax-like factor ATX1. PLoS ONE. 2010;5:e13396. doi:10.1371/journal.pone.0013396.
53. Hennig L, Bouveret R, Gruissem W. MSI1-like proteins: an escort service for chromatin assembly and remodeling complexes. Trends Cell Biol. 2005;15:295–302.
54. Hennig L, Taranto P, Walser M, Schonrock N, Gruissem W. *Arabidopsis* MSI1 is required for epigenetic maintenance of reproductive development. Development. 2003;130:2555–65.
55. Kim H J, Hyun Y, Park JY, Park MJ, Park MK, Kim MD, Kim HJ, Lee MH, Moon J, Lee I, Kim J. A genetic link between cold responses and flowering time through FVE in *Arabidopsis thaliana*. Nat Genet. 2004;36:167–71.
56. Liu C, Lu F, Cui X, Cao X. Histone methylation in higher plants. Annu Rev Plant Biol. 2010;61:395–420.
57. Zhang Z, Zhang S, Zhang Y, Wang X, Li D, Li Q, Yue M, Li Q, Zhang YE, Xu Y, Xue Y, Chong K, Bao S. *Arabidopsis* floral Initiator SKB1 confers high salt tolerance by regulating transcription and pre-mRNA splicing through altering histone H4R3 and small nuclear ribonucleoprotein LSM4 methylation. Plant Cell. 2011;3:369–411.
58. Jerzmanowski A. SWI/SNF chromatin remodeling and linker histones in plants. Biochim Biophys Acta. 2007;1769:330–45.
59. Racki LR, Narlikar GJ. ATP-dependent chromatin remodeling enzymes: two heads are not better, just different. Curr Opin Genet Dev. 2008;18:137–44.
60. Mlynárová L, Nap JP, Bisseling T. The SWI/SNF chromatin-remodeling gene *AtCHR12* mediates temporary growth arrest in *Arabidopsis thaliana* upon perceiving environmental stress. Plant J. 2007;51:874–85.
61. Han SK, Sang Y, Rodrigues A, Wu MF, Rodriguez PL, Wagner D. The SWI2/SNF2 chromatin remodeling ATPase BRAHMA represses abscisic acid responses in the absence of the stress stimulus in *Arabidopsis*. Plant Cell. 2012;24:4892–906.

62. Saez A, Rodrigues A, Santiago J, Rubio S, Rodriguez PL. HAB1–SWI3B interaction reveals a link between abscisic acid signaling and putative SWI/SNF chromatin-remodeling complexes in *Arabidopsis*. Plant Cell. 2008;20:2972–88.
63. Perruc E, Kinoshita N, Lopez-Molina L. The role of chromatin remodeling factor PKL in balancing osmotic stress responses during *Arabidopsis* seed germination. Plant J. 2007;52:927–36.
64. Dobosy JR, Selker EU. Emerging connections between DNA methylation and histone acetylation. Cell Mol Life Sci. 2001;58:721–7.
65. Datta J, Ghoshal K, Sharma SM, Tajima S, Jacob ST. Biochemical fractionation reveals association of DNA methyltransferase (Dnmt) 3b with Dnmt1 and that of Dnmt 3a with a histone methyltransferase and Hdac1. J Cell Biochem. 2003;88:855–64.
66. Liu X, Yu CW, Dun J, Luo M, Wang K, Tian G, Cui Y, Wu K. HDA6 directly interacts with DNA methyltransferase MET1 and maintains transposable elements silencing in *Arabidopsis*. Plant Physiol. 2012;158:119–29.
67. Song Y, Wu K, Dhaubhade S, An L, Tian L. *Arabidopsis* DNA methyltransferase AtDNMT2 associates with histone deacetylase AtHD2s activity. Biochem Biophys Res Commun. 2010;396:187–92.
68. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet. 2010;11:204–20.
69. Feng S, Jacobsen SE. Epigenetic modifications in plants: an evolutionary perspective. Curr Opin Plant Biol. 2011;14:179–86.
70. Lindroth AM, Shultis D, Jasencakova Z, Fuchs J, Johnson L, Schubert D, Patnaik D, Pradhan S, Goodrich J, Schubert I. Dual histone H3 methylation marks at lysines 9 and 27 required for interaction with chromomethylase 3. EMBO J. 2004;23:4146–55.
71. Du J, Zhong X, Bernatavichute YV, Stroud H, Feng S, Caro E, Vashisht AA, Terragni J, Chin HG, Tu A, Hetzel J, Wohlschlegel JA, Pradhan S, Patel DJ, Jacobsen SE. Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. Cell. 2012;151:167–80.
72. Fan D, Dai Y, Wang X, Wang Z, He H, Yang H, Cao Y, Deng XW, Ma L. IBM1, a JmjC domain-containing histone demethylase, is involved in the regulation of RNA-directed DNA methylation through the epigenetic control of RDR2 and DCL3 expression in *Arabidopsis*. Nucleic Acids Res. 2012;40:8905–16.
73. Lee MG, Wynder C, Bochar DA, Hakimi MA, Cooch N, Shiekhattar R. Functional interplay between histone demethylase and deacetylase enzymes. Mol Cell Biol. 2006;26:6395–402.
74. Shi YJ, Matson C, Lan F, Iwase S, Baba T, Shi Y. Regulation of LSD1 histone demethylase activity by its associated factors. Mol Cell. 2005;19: 857–64.

Chapter 6
Setting the Stage for the Next Generation: Epigenetic Reprogramming During Sexual Plant Reproduction

Leonor C. Boavida, Marcela Hernandez-Coronado and Jörg D. Becker

Introduction

Every multicellular organism consists of groups of similar cell types (tissues) that are specialized to carry out distinct functions. Cell differentiation is a product of a gradual commitment of a totipotent cell which is capable of developing into all the specialized cells that make up the adult organism in the course of development. The union of a sperm and an egg, two highly differentiated cells, generates a self-contained entity, a totipotent zygote, that has the ability to produce lineages of pluripotent cells that are predestined to a particular cell fate (cell fate specification) which in turn will become highly differentiated cell types (cell fate determination).

As every cell contains the same genomic information, the process of cell differentiation must reflect highly controlled modifications of gene expression. Over the past few years, we have begun to understand how cell-type-specific expression patterns and the process of cell specification seem to be determined or regulated by reversible epigenetic changes which are gradually imposed on the genome during development. These epigenetic modifications can be accomplished in different ways (DNA methylation, histone modification, chromatin remodelling and using the small RNA machinery), be inherited across generations and determine parent-of-origin-specific patterns of inheritance (genomic imprinting) [1].

J. D. Becker (✉) · L. C. Boavida · M. Hernandez-Coronado
Instituto Gulbenkian de Ciencia, Rua da Quinta Grande 6, 2780–156 Oeiras, Portugal
e-mail: jbecker@igc.gulbenkian.pt

L. C. Boavida
e-mail: lboavida@igc.gulbenkian.pt

M. Hernandez-Coronado
e-mail: mcoronado@igc.gulbenkian.pt

O. Pontes, H. Jin (eds.), *Nuclear Functions in Plant Transcription, Signaling and Development,* DOI 10.1007/978-1-4939-2386-1_6

In plants, every cell has the ability to reversibly dedifferentiate and become totipotent, contrasting with animal cells that become committed earlier during embryogenesis at 2–4 cell stage [2]. Several examples occurring in the plant adult stage are illustrative of the remarkable plasticity of plant cells. The plant life cycle in flowering plants is marked by multiple developmental transitions [3] involving the determination of new cell fates: (A) reprogramming of the apical meristem from a vegetative to a reproductive fate; (B) reprogramming of floral somatic cells to a germ cell lineage (pluripotency); (C) differentiation of specialized female and male gametes and (D) fertilization and restoration of a totipotent zygote.

While the epigenetic state of plant cells is thought to be relatively stable during development, epigenetic modifications occurring during germline specification and early embryogenesis are essential for gamete differentiation, re-establishment of pluripotency in the embryo and configuration of parent-specific epigenetic states. Epigenetic reprogramming through selective maintenance and erasure of epigenetic marks in the germline also plays a fundamental role determining other aspects such as inheritance of induced phenotypic traits and maintenance of genomic stability. The sessile nature of plants makes them to face multiple environmental perturbations during their life cycle that can cause physiological and developmental alterations or have consequences in genomic stability. Inducible phenotypical responses have been shown to affect profoundly the expression of genes potentially regulated by epigenetic marks [4]. Epigenetic modifications induced by stress are mitotically stable and if not reset they can be transmitted to the next generations through the germline (transgenerational epigenetic inheritance) [5–8]. Reduction of DNA methylation in *Arabidopsis thaliana* was shown to cause increased mortality, stunted growth, delayed flowering and lower seed set, supporting that epigenetic modifications contribute to increased plant fitness [9], but the contribution of induced epigenetic modifications in the evolution of natural populations is still a matter of debate [10]. Environmental stresses are also accompanied by an increase in transposon activity [4] and their mobilization can have adverse effects, generating deletions, genomic rearrangements and causing gene misregulation, ultimately compromising genome integrity and stability [11]. More rarely, mutations arising from transposon activity can generate genetic variation that may allow plants to adapt to adverse environments [12]. A strict control of transposon activity is thus imperative to protect the genome from potential deleterious effects across generations, especially in plants, where the germline is specified late during development. An understanding of how plant cells reprogramme themselves, the regulatory circuitry that maintains their ability to become totipotent or pluripotent and the events that commit progenitor cells to particular differentiation states are of paramount importance. In this chapter, we summarize and discuss recent studies into the epigenetic reprogramming occurring during one of the most dramatic phase transitions in the plant life cycle, the transition from a diploid to a haploid phase (sporophytic-to-gametophytic phase transition) which leads to germline specification and the renewal of the plant life cycle through fertilization and re-establishment of pluripotency in the embryo.

Epigenetic Control of Transition to Reproductive Development

Flowering plants spend most of their life cycle in a vegetative phase (diploid sporophytic generation) and the transition to a reproductive phase occurs when the plant reaches maturity. The timing when this developmental transition takes place is of utmost importance to ensure reproductive success, since flowering should occur in favourable physiological and environmental conditions to assure the completion of the fertilization process and dispersal of seeds.

The transition to flowering is a significant developmental change in the plant life cycle and is dependent on complex genetic pathways that integrate information from endogenous factors (hormone and nutrients) and environmental cues such as day length (photoperiod) and temperature (vernalization) [13]. In recent years, it has become clear that the genetic pathways controlling important phase transitions, namely juvenile-to-adult phase and the transition to a reproductive phase, share some common regulatory factors [14, 15]. Moreover, the expression of many of the different flowering genes involved in these phase transitions seems to be regulated by epigenetic modifications, alternative splicing, antisense RNA and chromatin silencing [3, 16–18]. In *Arabidopsis*, the expression of *FLOWERING LOCUS C* (*FLC*), a *MADS*-box transcription factor, is responsible for the direct repression of flowering pathway integrators *FLOWERING LOCUS T* (*FT*) and *SUPPRESSOR OF OVEREXPRESSION OF CONSTANS 1* (*SOC1*). A number of chromatin-modifying components seem to be involved in the repression of *FLC* expression [19]. In response to cold, the upstream component of the vernalization pathway, a chromatin remodelling plant homeodomain (PHD) finger protein, VERNALIZATION INSENSITIVE 3 (VIN3), binds to FLC chromatin [20] interacting with components of the polycomb-group repressive complex 2 (PRC2) to mediate *FLC* silencing through histone H3K27 tri-methylation (H3K27me3) [21, 22], which is essential to reinforce and stabilize a stable epigenetic memory of vernalization through mitotic divisions [23]. Moreover, *FLC* antisense transcripts negatively regulate *FLC* sense transcription by triggering localized histone H3K4 demethylation [24, 25]. In addition, two evolutionary highly conserved microRNAs, *miR156* and *miR172*, and their respective target genes (*SQUAMOSA PROMOTER BINDING PROTEIN-LIKE* and *APETALA 2* transcription factors) were also implicated in the regulation of flower transition (reviewed in 3).

In order for plants to be able to respond to vernalization and prevent early flowering, some type of reprogramming process is likely needed to erase the effects of repressive modifications in FLC chromatin leading to flowering. Independent from the epigenetic state of the maternal plants, *FLC* seems to be temporarily reactivated in male and female reproductive tissues (anther tapetum and ovule integuments) but is repressed in both male and female gametophytes. After fertilization, the paternal derived *FLC* copy is reactivated in the zygote and the maternal derived *FLC* copy is first expressed in the early multicellular embryo [26, 27], while no expression is detected in endosperm. This resetting could thus result from reprogramming associated

with histone turnover as is shown in the zygote (see below) or from local recruitment of H3K27me3 demethylases or specific chromatin remodelers after fertilization. The resetting of *FLC* expression during early embryogenesis thus seems to be a prerequisite for the repressive effect of FLC in flowering in the following generation.

In the plant apex, the integration of signals promoting flowering triggers the activation of flower meristem identity genes such as *LEAFY, APETALA 1* (*AP1*) and *CAULIFLOWER,* transforming the apical meristem into an inflorescence [28]. The activation of the five homeotic classes (ABCDE) of flower identity genes in different regions of the meristem and their combinatorial interaction determine the positioning of four different types of flower organs [29]: in the outer most whorl the sepals, then petals, the male reproductive organs (stamens and anthers, collectively known as androecium) and in the central whorl the female reproductive organ, consisting of one or more carpels that make up the gynoecium or pistil.

Specification of Gametophytic Cell Fate

In flowering plants, differentiation of plant gametes is an event occurring late during flower development. In contrast to animals where gametes are directly specified from meiotic products, plant gametes result from a post-meiotic developmental process that results in the formation of multicellular structures, the gametophytes, where gametes are coupled to accessory cells that facilitate gamete function [30]. A series of proliferative and cell specification events determine the formation of specialized reproductive organs that culminate with the differentiation of tissues and cells with reproductive and nonreproductive functions; however, at this stage no traceable germ cell lineage is yet specified.

Within the anther, an inner cell layer of secretory cells called tapetum supports microsporogenesis that initiates in a central core of cells, the sporogeneous pollen mother cells (PMC). The PMCs define the male reproductive cell lineage leading to the development of a male gametophyte, the pollen grain. At this stage in the anther, the PMCs are connected by enlarged cytomitic channels which create a cytoplasmic continuum thought to promote effective synchronization during meiosis. Two meiotic divisions transform each PMC into a tetrad of haploid and unicellular microspores (UNM); the cytomitic channels disappear, establishing an individual cytoplasm in each microspore, each with its own callose envelope and all encased in the callose wall of the tetrad. Pollen wall deposition initiates and an exine layer develops around each microspore. After the haploid microspores are released from the tetrads, important cytoplasmic and gene expression changes are believed to reset the sporophytic programme and launch a gametophytic programme (reviewed by [31]). Gametogenesis initiates with a reorganization of the microspore cytoplasm, whereby small vacuoles coalesce into a single vacuole that polarizes the nucleus to one side of the microspore. This cell polarization is determinant to establish the first asymmetric division called pollen mitosis I (PMI), in which the two daughter cells acquire different cell fates: the large vegetative cell (VC), the pollen grain,

enclosing a small generative cell (GC) [32] (Fig. 6.1a). Thus, only at the bicellular pollen (BCP) stage the germline, in the form of a generative cell, is clearly specified. It was proposed that the different cell fates are caused by polarized distribution of gametophytic regulatory factors during the asymmetric division, resulting in the repression of vegetative cell-specific genes in the generative cell [33]; however, the processes determining commitment to different cell fates are still poorly understood. The GC undergoes a second mitosis, pollen mitosis II (PMII) to produce two functional twin male gametes, the sperm cells (SC; Fig. 6.1a). Depending on the species, PMII can either take place before dehiscence or during pollen tube growth within the female tissues. The "cell within a cell" structural organization of the male gametophyte in angiosperms relies on an intimate association of the male gametes (sperm cells) with the vegetative nucleus (VN), called the male germ unit (MGU) [34] (Fig. 6.2). In this structural association, a membrane of vegetative origin encloses both sperm cells and extends from one of the sperm cells through a projection to the vegetative nucleus. Within the MGU, both sperm cells are transported as a unit by the leading VN until pollen tubes deliver the sperm cells in the vicinity of female gametes. While this endomembrane of vegetative origin may provide protection, isolating sperm cells from direct contact with the cytoplasmic environment of the pollen tube, it also establishes a possible communication route between the vegetative nucleus and sperm cells [35]. Moreover, the physical adherence of both sperm cells revealed by two sperm-specific tetraspanins (*TETs*) localizing in a membrane microdomain in the interface between both sperm cells may provide another level of communication [36] (Fig. 6.2). Intercellular interactions occurring within female tissues during pollen tube growth are essential for pollen tubes to acquire competence to sense female guidance signals [37], but they also seem to impact sperm cell behaviour as in some plant species sperm cells must undergo a process of maturation before they acquire fertilization competence [38]. These physical connections between the vegetative nucleus and sperm cells or between sperm cells may have functional implications in maintaining an equivalent fertilization competence of both sperm cells [39] and/or in maintaining germline identity within the pollen vegetative cytoplasmic environment. Ultrastructural studies will be essential to reveal the nature of these cellular connections.

Development of the female gametophyte takes place in the carpel. Here, rows of ovule primordia arise from the placental tissue along the margins of the site of carpel fusion, the septum. In each ovule primordium, within a supportive and nourishing tissue called nucellus, megasporogenesis occurs with the differentiation of the reproductive female lineage, the archeosporial cell. This cell differentiates into the megaspore mother cell (MMC), which undergoes meiotic reduction to give rise to four haploid megaspores (Fig. 6.1b). After meiosis, four haploid cells are surrounded by a very thin cell wall, with plasmodesmata connecting the chalazal megaspore to neighbouring nucellar cells. A positional signal is thought to promote cell death of three of the four megaspores. Megagametogenesis begins when the surviving functional megaspore goes through three rounds of mitosis to form first a two-nucleated, four-nucleated and subsequently eight-nucleated syncytium gametophyte. Different patterns of gametophyte development have been described

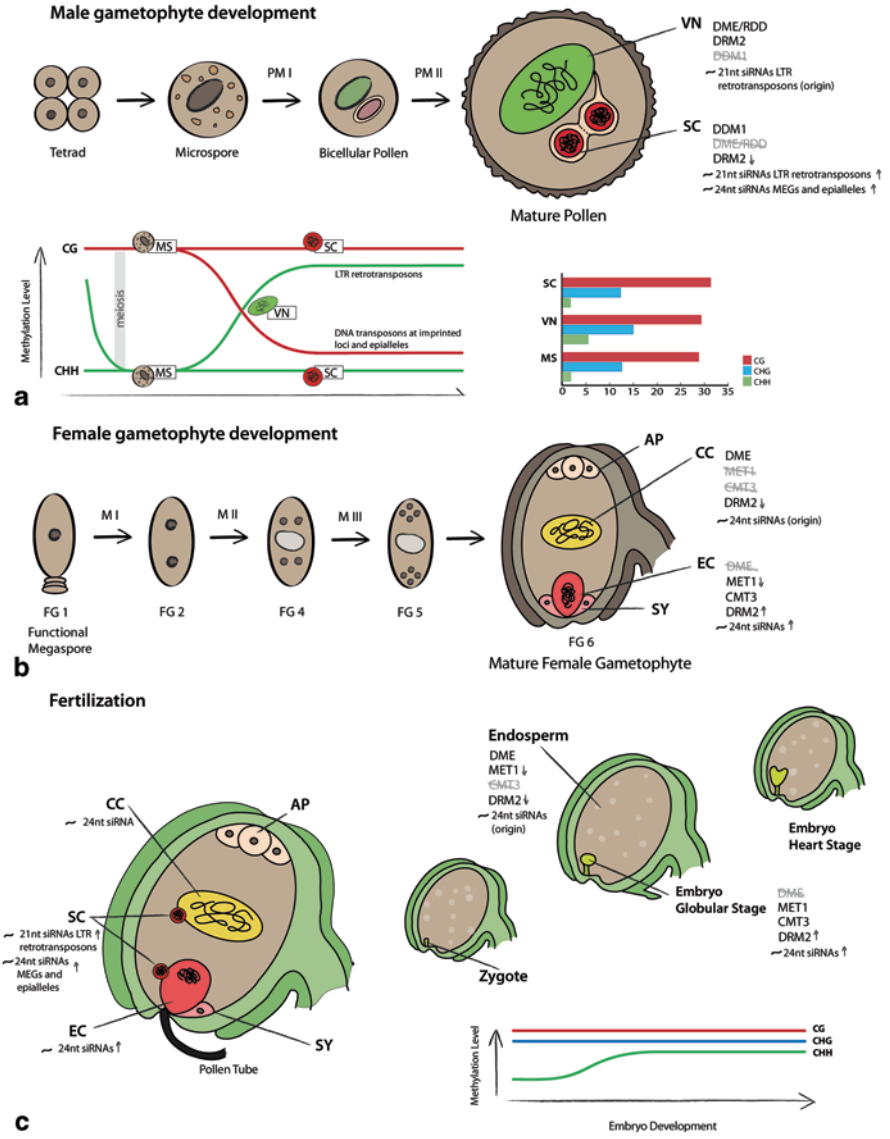

Fig. 6.1 Epigenetic reprogramming during gametophyte development, fertilization and embryogenesis. Sexual reproduction in *Arabidopsis thaliana* involves extensive epigenetic programming with far-reaching potential consequences for genome stability, imprinting and epigenetic inheritance. **a** During male gametophyte development, haploid microspores (MS) divide asymmetrically (PM I) to give rise to bicellular pollen consisting of a vegetative cell harbouring a smaller generative cell. The latter will undergo a second mitosis (PM II), originating two sperm cells (SC) that stay connected to the vegetative nucleus (VN) via a membrane projection. Microgametogenesis is characterized by a loss of CHH methylation from LTR retrotransposons in microspores and sperm cells due to reduced expression of *DRM2*, while it is restored in the vegetative nucleus. Missing expression of the chromatin remodeler *DDM1* in the vegetative nucleus however leads to the

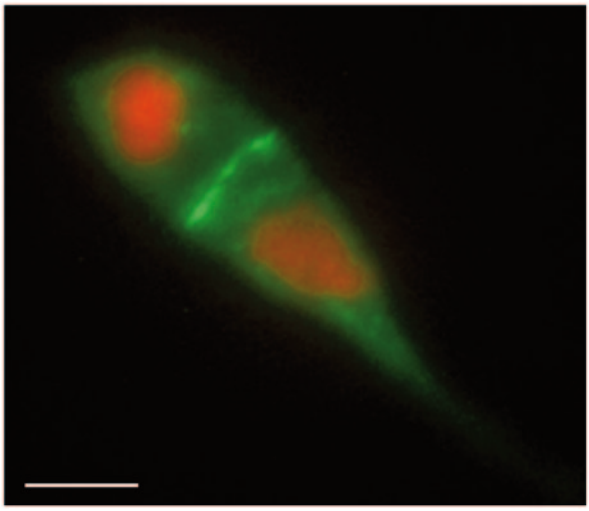

Fig. 6.2 Sperm cell connection in the male germ unit. *TET12p*:TET12-GFP and *HTR10p*:HTR10-mRFP protein fusions expressed in sperm cells within an *in vitro* germinated pollen tube. GFP fluorescence highlights the membrane microdomain connection between sperm cells within the male germ unit; mRFP fluorescence corresponds to nuclear localization of the male germline specific histone H3.3 variant (HTR10). Scale bar: 2 μm

in several species, differing from each other by variations on cytokinesis during meiosis, on the number of mitotic divisions and in the cellularization pattern. Cellularization and the final differentiation of a *Polygonum*-type female gametophyte

activation of *ATHILA* retrotransposons and accumulation of correspondent 21-nt siRNAs in the sperm cells. Moreover, the loss of expression of the DNA glycosylases *DME/RDD* in the vegetative nucleus affects CG methylation of transposons neighbouring imprinted loci and epialleles, with correspondent 24-nt siRNAs accumulating in the sperm cells. PM I pollen mitosis I, PM II pollen mitosis II, MEG maternally expressed gene, ↓ down-regulation, ↑ up-regulation, ~~Gene~~ not expressed. **b** During female gametogenesis, the functional megaspore undergo three consecutive mitosis (M I–III) to generate an eight nucleated cell (FG5). Subsequent cellularization and differentiation results in a seven-celled embryo sac (FG6) consisting of two female gametes, the diploid central cell (CC) and the haploid egg cell (EC) with two types of accessory cells, three antipodals (AP) and two synergids (SY). In the CC, the loss of expression of the methyltransferases *MET1* and *CMT3* and the downregulation of *DRM2* together with high expression of *DME* generate a hypomethylated state. The resulting transcriptional activation of transposable elements leads to the production of 24-nt siRNAs that accumulate in the adjacent egg cell, where RNA-directed DNA methylation (RdDM) pathways are active (*DMR2* expression). FG (female gametophyte) stages according to the classification described in [128]. **c** Upon double fertilization, one sperm cell fuses with the egg cell and the second with the central cell, giving rise to the diploid embryo and the triploid endosperm, respectively. Each gamete carries genetic and epigenetic information, e.g. hypomethylated transposon sequences and siRNAs that create a condition of epigenetic reconciliation upon fusion. The epigenetic state of the endosperm resembles the one of the central cell before fusion with low levels of methylation and production of 24-nt siRNAs. The accumulation of those 24-nt siRNAs in the embryo will participate, possibly together with 24-nt siRNAs delivered by the sperm cells, in *DRM2*-driven RdDM to restore CHH marks during embryo development. SN (sperm nucleus)

involve the specification of four cell types with distinct functions in a seven-celled embryo sac (ES): two gametes, consisting of a haploid egg cell (EC) and a diploid central cell (CC), and two accessory cell types consisting of two synergids (Sy) and three antipodals (AP) ([31] and references therein) (Fig. 6.1b).

Methylation Reprogramming During Pollen Development in *Arabidopsis*

Much of our current understanding of epigenetic modifications during germline specification in plants comes from the terminal developmental stages of gametophytes, but there are growing evidences that epigenetic modifications might be relevant during somatic-to-reproductive transition. These epigenetic changes can have a significant impact in the next generation, as the transition from a mitotic to a meiotic cell cycle programme sets the stage whereby a group of cells acquire a sporogenous reproductive cell fate. The differentiation of PMCs is intimately linked with the differentiation of surrounding tissues in the anther. Modifications in microRNA activity, chromatin remodelling and DNA methylation play a critical role in the differentiation of anther tissues and in tapetum-programmed cell death [40–42]. How these epigenetic signals contribute to differentiation of microspores is still poorly understood, but evidences showing that *trans*-acting 24-nucleotide (nt) short interfering RNAs (siRNAs) produced in anther tapetum can impact male gametophyte development support this hypothesis [43].

The nature of the switch that defines the mitotic to meiotic transition is still unknown, but DNA methylation, histone modifications and siRNAs seem to be involved in this cell fate transition. DNA METHYLTRANSFERASE 1 (MET1) and DECREASE IN DNA METHYLATION 1 (DDM1) are important for meiotic progression and pattern the recombination frequency along chromosomes [44–47]. In rice, the germline specific *MEIOSIS ARRESTED AT LEPTOTENE 1* (*MEL1*), a homologue of *Arabidopsis ARGONAUTE 5* (*AGO5*), is required for mitosis in sporogeneous PMCs and meiotic progression, but does not affect specification of the reproductive cell lineage [48]. However, the *Arabidopsis ago5* knock-out does not show any discernible phenotype [49], suggesting that in *Arabidopsis*, *AGO5* may act redundantly with other *AGO* members.

Transcriptomic analyses of plant meiocytes are relative recent due to the challenge of isolating enriched populations from surrounding anther tissues [50]. The expression profile of *Arabidopsis* meiocytes indicates a partial reactivation of silenced transposable elements (TE) associated to *Copia*-like and *Gypsy*-like long terminal repeat (*LTR*) elements with a preferential enrichment of one of the non-*LTR* retrotransposon families, the short interspersed nuclear elements (*SINE*) [51]. TE activation in meiocytes was correlated with co-expression of neighbouring genes and with localized changes in chromatin structure, thought to facilitate meiotic progression [51]. Transient TE element activation in the last diploid stage before meiosis could thus represent an opportunity to introduce genomic variability or to

generate siRNAs that, if not deleterious for meiosis or pollen development, could function in germline specification and/or be transmitted to the next generation.

The differentiation of the male gametophyte (post-PMI) involves chromatin modifications and changes in methylation patterns that lead similarly to the female gametophytes (see below) to the production of two epigenetic dimorphic cell types. During pollen development, the ubiquitous centromeric histone *HTR12* is first detected in unicellular microspores (UNM), after PMI and PMII *HTR12* remain expressed in the generative nucleus and sperm cells, respectively, but is not detected in the vegetative nucleus. Moreover, upon PMII, the vegetative nucleus and sperm cells present distinct histone-based signatures with *HTR5*, *HTR8* and *HTR14* detected in the vegetative nucleus, while sperm cells express *HTR5* and the sperm-cell specific H3.3 variant *HTR10* (also known as *MALE GAMETE SPECIFIC HISTONE H3*, *MGH3*). Thus, sperm cell chromatin becomes distinct from the non-gametic lineage (vegetative nucleus) during pollen development [52, 53].

Significant advances in fluorescence-activated cell sorting (FACS) of *Arabidopsis thaliana* pollen, of its individual cellular components (vegetative nucleus and sperm cells) and of its progenitor cell, the unicellular microspore, were the prerequisite to decipher transcriptomic and epigenetic changes occurring during pollen development. In a first approach, sorted male gametes expressing GFP under a sperm cell specific promoter were used to characterize their transcriptome (mRNA and small RNAs) and the methylation pattern of specific transposons [35, 54, 55]. In order to analyse epigenetic changes during pollen development on a whole-genome scale, this method was further improved to allow simultaneous sorting of the vegetative nucleus and sperm cells as well as FACS isolation of microspores [56, 57]. In a similar fashion, but using DNA dyes instead of fluorescent protein labels, the epigenome of the male germ unit of wild type and mutant pollen grains was analysed [58–60]. Together, these efforts led to genome-wide maps of cytosine methylation in all three sequence contexts (symmetric CG, CHG and asymmetric CHH methylation, where H stands for any residue except G) for sorted microspores, vegetative nuclei and sperm cells. The in-depth analysis of these data contributed significantly to our current understanding of epigenetic reprogramming in the male gametophyte [57, 58] (Fig. 6.3). Through mapping of differentially methylated regions (DMRs) in their genomic context (genic, intergenic or transposable elements) and combination with small RNA abundance, it became apparent that the large majority of these DMRs affect transposon and repetitive element sequences. Their reactivation leads to the production of siRNAs capable of controlling their activity either by transcriptional gene silencing (TGS) or post-transcriptional gene silencing (PTGS). A relaxation of control over potentially harmful transposable elements, particularly in the gametophyte harbouring the germline, seems contradictory. However, the emerging picture is more that of an epigenetic pathway that has been co-opted for inheritance of epigenetic marks (epialleles and imprinted genes) and control of transposable elements during gametophyte development and embryogenesis. The first indication for epigenetic activation of transposable elements in the male gametophyte of *Arabidopsis* came from a study by Slotkin et al. [35]. These retrotransposons get activated in the vegetative nucleus but not in the sperm

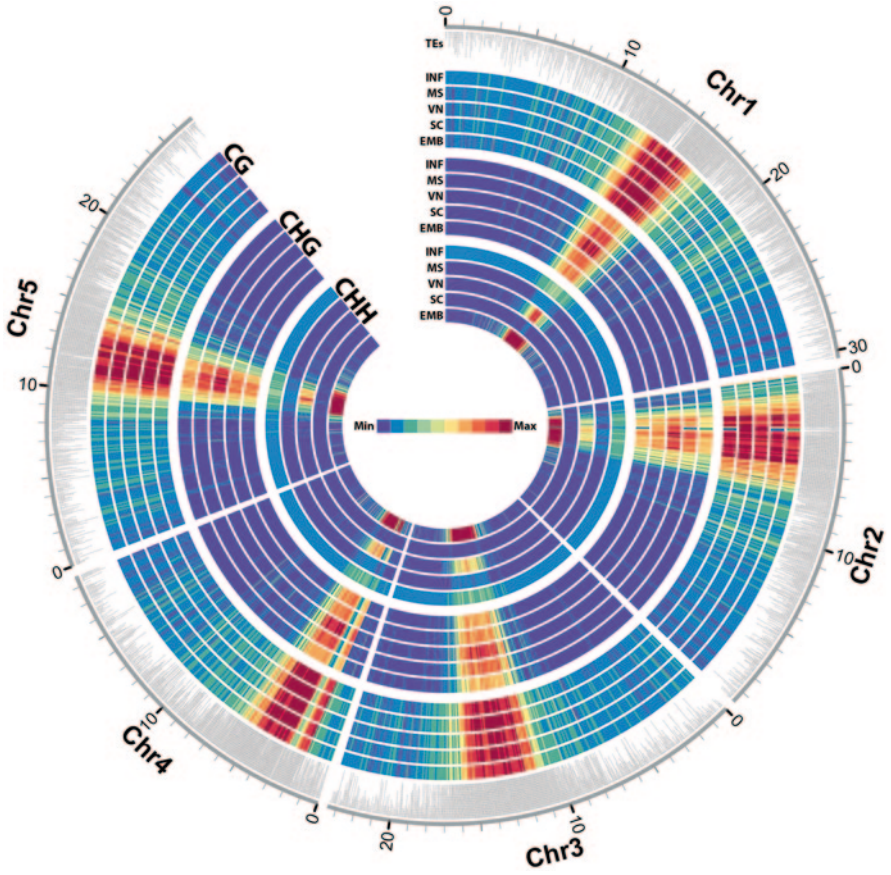

Fig. 6.3 Whole-genome representation of DNA methylation dynamics during pollen development. The heat map is based on bisulphite sequencing data of genomic DNA extracted from inflorescence and embryo as well as distinct pollen cell types and nuclei. Methylation density is shown in 10 kb blocks in the three sequence contexts (CG, CHG and CHH) and along the five chromosomes (low in *blue*; high in *red*). The outer track in grey represents transposable element density. Loss of CHH methylation in microspore and sperm cells in comparison with the vegetative nucleus is evident in the pericentromeric regions. INF inflorescence, MS microspore, VN vegetative nucleus, SC sperm cell, EMB embryo. (Adapted from [57] with permission from Elsevier)

cells, coinciding with differential expression of the chromatin-remodeler *DDM1*, whose function is crucial to maintain DNA methylation of transposable elements. It does so by facilitating access of DNA methyltransferases to linker histone H1-containing heterochromatin [61] and complete loss of *DDM1* leads to increasing developmental abnormalities in inbred lines caused by TE activation [11, 62]. First indicated by transcriptional profiling of *Arabidopsis* sperm cells [54] it was confirmed by translational fusion with GFP that expression of *DDM1* in the male germ unit is restricted to the sperm cells [35]. The transcriptional activation of *Athila* retrotransposons in the vegetative nucleus leads to the production of 21-nt siRNAs in

pollen, which surprisingly accumulate to higher levels in the sperm cells than in the vegetative nucleus [35]. In the sperm, they are thought to reinforce retrotransposon silencing post-transcriptionally (Fig. 6.1a), but they could also play a role during embryogenesis, if delivered upon fertilization (discussed below).

Interestingly, the transcriptional activation of *Athila6A* family retrotransposons in the VN also leads to the production of *Athila6A*-derived 21-nt siRNA854. In pollen, this siRNA acts *in trans* via incorporation into AGO1 and results in post-transcriptional cleavage of *UBP1b* mRNA [63]. In the sporophyte, it has been shown that UBP1b protein is localized to stress granules and is involved in repressing Athila6 GAG capsid protein production, if the transposable element is transcriptionally activated upon stress conditions and the RNAi pathway is non-functional. Here, siRNA854 acts as a suppressor of host transposable element silencing via translational inhibition of UBP1b [64].

The most obvious differences in methylation profiles between VN, SC and microspores were observed for asymmetric CHH methylation in the pericentromeric regions, with the microspore and SCs showing hypomethylation in comparison to the VN [57, 58] (Figs. 6.1a and 6.3). These CHH DMRs map mostly to class I *LTR/Gypsy* retrotransposons and the sperm cells contain 21-nt siRNAs matching the hypomethylated retrotransposons. These siRNAs are likely to be involved in PTGS, thus preventing hazardous transposition of these elements in the male germline. The observed loss of CHH methylation in the male gametes, but not in the VN, is best explained by a passive loss of CHH methylation during microgametogenesis due to reduced maintenance of CHH methylation in the germline. Supporting this hypothesis, expression of the important RdDM component *DOMAINS REARRANGED METHYLTRANSFERASE 2* (*DRM2*) is restricted to the VN during pollen development, with very low levels in the microspore, generative cell and sperm cells [57] (Fig. 6.1a).

But the vegetative nucleus also shows hypomethylated regions, although in a CG sequence context and to a much lesser extent [57, 58]. CG and CHG symmetric methylation is maintained during cell divisions by DDM1 and the DNA methyltransferases MET1 and CHROMOMETHYLASE 3 (CMT3). *MET1* and *CMT3* are expressed during pollen mitosis I and II [65] and levels of symmetric methylation levels in the vegetative cell and sperm cells are largely unaltered when compared to those in somatic cells. Therefore, the observed partial loss of CG methylation in the VN must be achieved through active de-methylation by the DNA GLYCOSYLASE DEMETER (DME) and/or its homologous REPRESSOR OF SILENCING 1 (ROS1), DEMETER-LIKE 2 (DML2) and DML3 [59]. And indeed, *DME* and *ROS1/DML2/DML3* (*RDD*) are not expressed in the sperm cells, but they are expressed in the vegetative nucleus and lead to CG hypomethylation of class II DNA transposons of the *MuDR* and *RC/Helitron* class [57, 59]. It has been shown that imprinted genes are often neighboured by such DNA transposons [66, 67]. In fact, a number of TEs flanking imprinted genes lose CG methylation in the vegetative nucleus relative to microspores and sperm cells. But mainly those flanking genes that are maternally expressed in the endosperm (MEGs) and not those close to paternally expressed genes (PEGs) show high levels of CHH methylation in microspores and

sperm cells. This observed preservation of CHH methylation at TEs neighbouring MEGs depends at least partially on the activity of DME at the corresponding loci in the vegetative nucleus, as indicated by *dme*/+ sperm cells showing reduced CHH methylation when compared with wild type sperm cells [58]. In addition, 24-nt siRNAs corresponding to MEGs accumulate preferentially in the sperm cells [57], supposedly promoting MEG-specific CHH methylation (Fig. 6.1a).

Methylation of recently discovered hypervariable epialleles might be affected in a similar way. Through analysis of leaf methylation profiles after 30 generations of inbreeding by single-seed descent, it could be shown that more than 100 loci are prone to gain methylation sporadically and recurrently [68, 69], many of those being targets of RDD. More than half of the 100 loci are hypermethylated in sperm cells, readily explained by the low levels of *DME/RDD* expression in microspores and high expression in the vegetative nucleus [57]. This hypermethylation in sperm cells supports the idea that hypervariable epialleles already methylated in the inflorescence might be heritable through appropriate epigenetic reprogramming in the male germline [68, 69].

As outlined above, epigenetic reprogramming during male gametogenesis in *Arabidopsis* results in two cell types with distinct epigenetic features (Fig. 6.1a). Most notably, the vegetative nucleus undergoes extensive reprogramming involving loss of pericentromeric heterochromatin, loss of CG methylation and production of 21-nt siRNAs from activated retrotransposons. In this context, it is important to remember that the vegetative nucleus does not contribute genetic material to the next generation. In relation to the sperm cells the vegetative nucleus functions as a nurse cell, similar in that sense to the role the central cell plays for the egg cell (see below). Thus, partly due to the reprogramming of its companion cell, but also due to alterations in its own epigenetic make-up, the male gametes represent a unique epigenetic state upon anthesis: (A) *LTR/Gypsy* retrotransposons are hypomethylated in a CHH context, (B) 24-nt siRNAs matching TEs that flank imprinted, maternally expressed genes are abundant, (C) 21-nt siRNAs matching VN-activated retrotransposons accumulate.

Exactly how and when this sequestration of 21- and 24-nt siRNAs into the male gametes is occurring is still a matter of debate. Theoretically, these siRNAs could move from the vegetative nucleus through the cytoplasmic extension of the male germ unit or via the pollen cytoplasm. And indeed, an artificial microRNA (*amiR*) expressed in the vegetative nucleus that targets a sperm cell expressed mRNA encoding GFP led to a reduction in GFP signal [35]. It has been argued however that the *LAT52* promoter used to drive expression in the vegetative nucleus is already active at the microspore stage and therefore the *amiR* could have been carried over to the precursor of the sperm cells, the generative cell, already during pollen mitosis I. Support for this hypothesis comes from a study using a promoter specific to the vegetative cell (post PMI, late bicelullar pollen) to drive *amiR*GFP expression, in which case germline-specific GFP expression could not be silenced [70]. Regardless of the origin of siRNAs accumulating in sperm cells, their role could be to reinforce silencing of maternally expressed genes as well as a subset of retrotransposons in the male gamete. An alternative function however could only come to bare after their delivery to the egg cell and during early embryogenesis (Fig. 6.1c).

Epigenetic Regulation in the Female Gametophyte

The specification of MMC from surrounding somatic cells in ovule primordia seems to be accompanied by extensive chromatin reprogramming important to establish an epigenetic and transcriptional landscape distinct from the surrounding somatic tissues. Chromatin decondensation associated with depletion of linker histones, reprogramming of histone variants and establishment of histone marks characteristic of a more permissive chromatin state, contribute to an epigenetic landscape supporting pluripotency and competence for a postmeiotic development of the female gametophyte [71].

The specification of a single functional MMC in a nucellus primordium seems to be epigenetically regulated through the action of small non-coding RNAs [48, 72]. Maize *AGO104,* closely related to *Arabidopsis AGO9,* is required for meiosis II during megasporogenesis [73]. Maize *ago104* mutants develop normal MMCs, but fail to undergo meiosis and instead undergo a mitotic-like division followed by megagametogenesis to produce functional unreduced megaspores [73]. In *Arabidopsis ago9* mutants, the somatic cells adjacent to MMCs acquire a reproductive cell fate bypassing meiosis and generating multiple unreduced megaspores [72, 74]. Thus *AGO9* seems to act by repressing germ cell fate in somatic tissues, while *AGO104* acts to repress somatic fate in reproductive cells. These processes most resemble diplosporic and aposporic development in apomictic plants, suggesting that the regulation of these two gametophytic pathways could be interconnected [72, 73]. AGO5 and AGO9 are known to preferentially associate with 24-nt small interfering RNAs (siRNAs) derived from transposons and repetitive elements but with different specificities to the 5′-terminal nucleotide. The specific accumulation of AGO104 and AGO9 in somatic nucellar cells, but their absence from MMCs raised the hypothesis whereby mobile siRNAs derived from transposable elements could move from adjacent somatic cells to MMC to restrict reproductive development to the functional megaspore [73]. *MEL1*, the orthologue of *Arabidopsis AGO5,* is initially expressed in the sub-epidermal cells in the ovule primordia during MMC differentiation, but later the expression becomes restricted to the MMC, disappearing during meiosis [48]. In rice *mel1/ago5* mutants the specification of the reproductive cell lineage is not affected, but MMCs fail to undergo meiosis leading to the absence or arrest of female gametophyte development. In *Arabidopsis, ago5* mutants do not show any visible phenotype, but a semidominant form of AGO5, *ago5-4,* presents defects in the initiation of megagametogenesis [49]. The *ago5-4* truncated form lacking the MID domain and catalytic PIWI domain may compromise sRNA binding efficiency, reminiscent of viral suppressor proteins which sequester siRNAS [75]. Consistently, the expression of the viral RNAi suppressor of 24-nt siRNAs, *P1/Hc-Pro*, in somatic nucellus cells produced a similar phenotype, supporting that RNA-directed DNA methylation pathways act in somatic nucellar cells to initiate megagametogenesis [49]. While a possible redundancy of *AGO5* cannot be excluded, these studies support the existence of two RdDM independent pathways, one acting through *AGO9* restricting the reproductive potential to the functional megaspore [72, 73] and an independent pathway initiating megagame-

togenesis [49]. Also supporting this hypothesis, the loss of function *of dmt102* and *dmt103* in maize, homologous to *Arabidopsis CMT3* and *DRM2,* produces unreduced gametes and ectopic embryo sac formation from supernumerary MMCs [76], suggesting that RNA-directed DNA methylation pathways are associated to transcriptional repressive states that could determine the distinction between an apomictic and a sexual reproductive development [76]. Interestingly, *AGO5* and *AGO9* also show enriched expression in the sperm cells [54, 55], but no specific function has been identified in the male gametes so far.

In contrast to male gametogenesis where two identical sperm cells are formed with the same potential to fertilize any of the female gametes [39], female gametogenesis gives rise to two dimorphic female gametes, the egg and the central cell, which contribute to distinct fates in post-fertilization products, the embryo and the endosperm (Fig. 6.1b, c). As all female gametophytic cells are genetically identical, it was proposed that female gamete dimorphism could reflect individual epigenetic states and these should be established in the syncytium stage, prior to embryo sac cellularization and differentiation [77]. Epigenetic reprogramming of female gametes by DNA methylation, histone modifications and replacement of histone variants contribute to post-fertilization dimorphic epigenetic states regulating transposon activity with functional implications in acquisition of zygotic totipotency and initiation of embryogenesis [78–80]. In particular, the composition and incorporation of specific histone variants contributes to establish specific chromatin epigenetic states in the egg and central cell [77]. The central cell expresses several *H3.1* isoforms and shows enrichment for two specific *H3.3* variants (*HTR8* and the unusual *HTR14*), while the mature egg cell is depleted of *H3.1* variants, expressing a single *H3.3* variant (*HTR5*) [52]. Moreover, egg cells exhibit a hypermethylated quiescent status correlating with low levels of transcriptional activity of PolII and enrichment of repressive chromatin histone marks associated predominantly with silenced states in euchromatic and heterochromatic regions (*H3K9me2* and *LHP1*). In contrast, the central cell is hypomethylated and transcriptionally active, correlating with a more permissive chromatin configuration. The dimorphic *H3K9me2* chromatin marks seem to be regulated by CMT3 in the egg cell and by DEMETER-LIKE (DML) activity in the central cell [77, 81]. Consistently, female gametes present an asymmetry of DNA methylation patterns, largely associated to the reduced activity of DME in the central cell [82, 83] and transcriptional repression of DNA methyltransferases like *MET1*, mediated by the retinoblastoma pathway [84, 85]. Passive loss of DNA methylation in the egg cell appears to be counteracted by *de novo* DNA methyltransferases DRM1 and DRM2 while in the central cell only low levels of *DRM2* were detected [86]. The requirement of CMT3-induced egg cell silencing in both transposon and euchromatic regions led to the hypothesis that siRNAs produced in transcriptional active cells (such as central cell or somatic cells) could target the egg cell [81]. DRM2 activity relies on RNA polymerases PolIV and PolV that have evolved an exclusive function in siRNA biogenesis and transcriptional silencing, respectively [87]. PolIV-dependent 24-nt siRNAs, many of which generated from TEs, were shown to be maternally specific and accumulate predominantly in the central cell [58, 88]. The finding that AGO9 associates with

24-nt siRNAs derived from transposable elements and is expressed in somatic ovule cells but not in female gametes raised the possibility that somatic cell-derived siRNAs could act as a non-cell-autonomous signal travelling to female gametophytic cells to restrict TE activity. This hypothesis was also consistent with reactivation of TEs in the egg and synergid cells of *ago9* mutants [72, 74].

Post-fertilization Epigenetic Reprogramming

In two parallel events, sperm cells fertilize the egg cell and the central cell and after karyogamy the distinct cell fate of both fertilization products, the zygote and the endosperm, is reflected in distinct epigenetic profiles (Fig. 6.1c). In the endosperm, which nurtures embryogenesis during seed development [89], low levels of DNA methylation are achieved through active DME-driven demethylation [83, 90] and continued repression of the main methyltransferases *MET1, CMT3* and *DRM2* [82, 86, 91] set a profile that most resembles the epigenetic make-up of the transcriptionally active central cell before fertilization. In the embryo, a crucial epigenetic reprogramming reinforces maintenance of DNA methylation in CG and CHG contexts by *MET1* and *CMT3* expression and RdDM activity with high expression levels of *DRM2* assures the gradual *de novo* methylation in CHH contexts (Fig. 6.4), which may contribute to maintain stable inheritance of epialleles across generations, reset silenced imprinted genes in the embryo or reinforce TE silencing contributing to genomic stability.

In the embryo methylation of the hypomethylated *LTR/Gypsy* retrotransposons coming in from the paternal side is restored, possibly through CMT2-dependent re-methylation pathways [61], although 21-nt siRNA might also be involved at the post-transcriptional level. These siRNAs could have their origin in the central cell or endosperm. That such movement of siRNAs from the female companion cell (central cell) to the egg cell is possible was shown in an experiment analogous to that in pollen, in which an artificial miRNA expressed in the central cell led to the reduction of a GFP signal in the egg cell [58].

In *Arabidopsis,* delivery of paternal 21-nt siRNAs targeting maternal *LTR* retrotransposons post-transcriptionally and simultaneous transcriptional targeting of CHH hypomethylated paternal retrotransposons by 24-nt siRNA of maternal origin could potentially create a mix and match situation of hypomethylated TEs and their targeting siRNAs upon fertilization (Fig. 6.1c). Such interaction between the 21-nt post-transcriptional and 24-nt transcriptional pathways has been predicted [87] and if occurring during reproduction it could have far-reaching consequences for the embryo and endosperm [92]. Two scenarios are conceivable: (A) Transposon sequences of the two genomes involved in an interspecific cross differ to an extent that the siRNA sequences derived from one parent do not match the transposable elements in the other parent. (B) Interploidy crosses lead to a genomic dosage disequilibrium characterized by the siRNA pool of one parent being insufficient to silence all copies of transposable elements of the other parent. These mechanisms

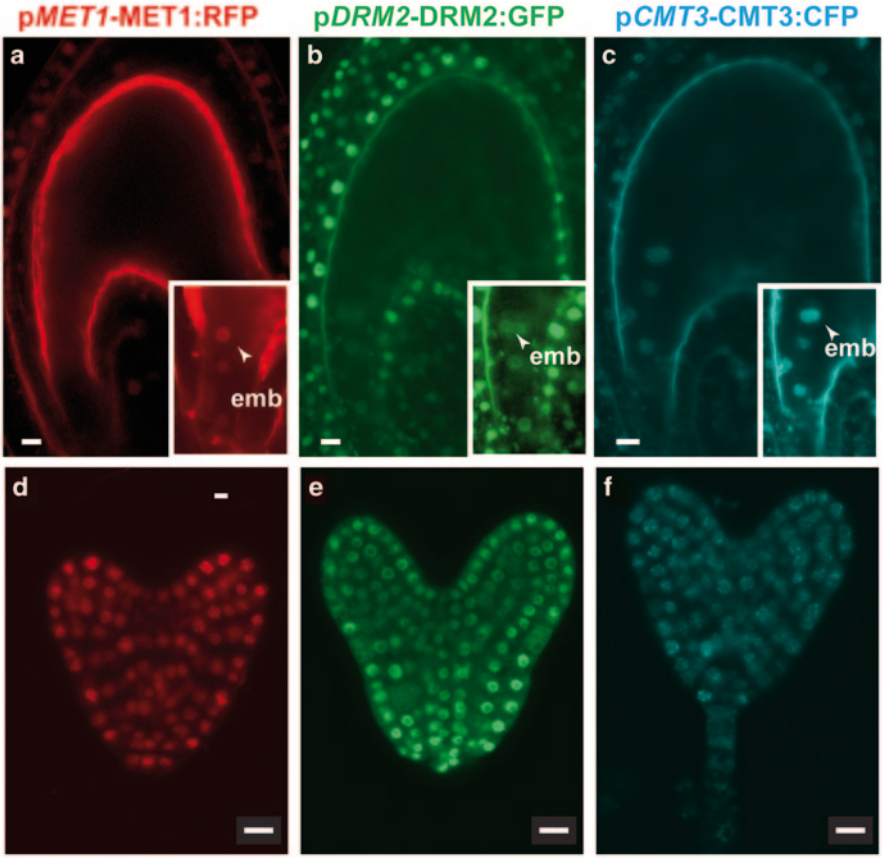

Fig. 6.4 Expression pattern of DNA methyltransferases during seed development. During seed development in *Arabidopsis,* the DNA methyltransferases *MET1*, *CMT3* and *DRM2* are differentially expressed in the developing embryo and its nourishing endosperm, as visualized using fluorescent fusion proteins under the control of the respective native promoter. During early embryogenesis (A–C), expression of the DNA methyltransferases is detected in the nuclei of the one cell (*MET1*) and two cell embryo (*CMT3* and *DRM2*), respectively (arrowheads in inset). The endosperm does not show any detectable expression levels. This divergent expression pattern is maintained during embryogenesis as exemplified by exclusive expression in the embryo at the heart stage (D–F). Scale bar: 20 μm. Thus, DNA methylation in all sequence contexts is highly active during embryogenesis, while it is strongly reduced in the endosperm. (Adapted from [86] with permission from Elsevier)

may present an effective way to assess parental compatibility at fertilization, in case of hybrid incompatibility leading to failure to suppress transposon activity (and deregulation of siRNA/TE controlled genes in their vicinity), ultimately resulting in seed abortion [92–94]. Such seed abortion has been observed for instance in interspecific crosses between *Arabidopsis thaliana* and *Arabidopsis arenosa,* involving upregulation of a normally silenced paternal *ATHILA* retrotransposon in a genome dosage-sensitive way [95, 96]. Underlining a general role of small RNAs

for hybridization barriers, hybrid dysgenesis in *Drosophila melanogaster* involves the reactivation of paternally inherited transposable elements in the progeny, caused by the lack of maternal Piwi-interacting RNAs targeting these TEs [97].

Delivery of small RNAs by the male gametes to the egg cell or central cell has not been proven in plants, but further support for such hypothesis comes from a delivered mRNA. *Short Suspensor* (*SSP*) mRNA accumulates in sperm cells of mature pollen and is delivered to the egg cell at fertilization, where its translation in the zygote triggers the activation of the YODA MITOGEN-ACTIVATED PROTEIN MAP kinase cascade responsible for the regulation of the first asymmetric division of the *Arabidopsis* zygote [98]. Similar to plants, human sperms contain a complex repertoire of coding and non-coding RNAs, but the role the delivery of this paternal RNA pool might play upon fertilization remains to be shown [99, 100]. In contrast, the picture is much clearer in *C. elegans*. Here, compelling evidence indicates that sperm carries ARGONAUTE/small RNA complexes transmitting a transgenerational small RNA memory of paternal gene expression [101].

The selective erasure and maintenance of epigenetic modifications at fertilization may also be critical for zygotic genomic activation (ZGA) which marks the initiation of *de novo* transcription from the zygotic genome. ZGA is essential for the transition from a quiescent transcriptional stage to a pluripotent state in early embryo enabling them to activate any pathway required for organism development [102]. In *Arabidopsis*, maternal transcripts predominate in a 2–4 cell embryonic stage, showing a gradual transition to increased paternal contribution at the globular stage [103]. Consistent with these observations *Arabidopsis* embryo development proceeds up to a globular stage even with low levels of active PolII, suggesting that early stages of embryo development rely on stored maternal transcripts [104, 105]. Specific maternal epigenetic marks were associated with the initial quiescent zygotic stage and with zygotic activation of paternal alleles which should result from a balance between chromatin-based repressive mechanisms and the establishment of a permissive chromatin state for transcriptional activation [103]. Consistently, in *Arabidopsis*, H3.3 variants from both gametes are quickly removed from the zygote after karyogamy, and somatic H3 variants are restored by *de novo* synthesis in the embryo, while in the endosperm there is a progressive dilution of the H3.3 variants through successive nuclear divisions [52, 53]. This epigenetic reprogramming in the egg cell seems to be mediated by CAF1 [103], ultimately limiting the inheritance of epigenetic information carried by H3.3 variants to the next generation [52]. In a process not well understood, H3K9 methylation driven by the maternal allele of *KRYPTONITE* (*KYP*) and *CMT3* as well as siRNAs produced by RdDM pathways that predominantly target TEs [106] seem to target also coding genes, which is believed to set the stage for zygotic activation and embryo development [103].

Genomic imprinting in plants occurs after fertilization whereby specific loci in endosperm are expressed during seed development according with their parental origin. Most imprinted genes are associated with differentially methylated regions (DMRs) that are methylated in a CG context in one of the two parental alleles. This CG methylation is important for epigenetic inheritance through gametogenesis and involved with imprinting in the endosperm [84, 107–109]. Several examples

of genomic imprinting in plants highlight the role of two distinct and interdependent mechanisms, DNA methylation and polycomb group (PcG) complex-mediated H3K27me3, in setting parental-specific epigenetic marks that are established in gametes and erased post-fertilization in endosperm and embryo. *FWA* and *MEDEA* maternal imprinting in endosperm are a good example of this epigenetic regulatory control: during fertilization, DME removes CG methylation in the *FWA* maternal allele, sustaining the maternal monoallellic expression in endosperm while the low activity of *MET1* in the endosperm is sufficient to maintain CG methylation of the *FWA* paternal allele [82, 86, 91]. On the other hand, *MEDEA* uses a more complex mechanism in which DME in the central cell demethylates CG sites flanking *MEDEA*, promoting *MEDEA* maternal expression in endosperm. MEDEA then binds to the Fertilization Independent Seed (FIS)-PcG complex. After fertilization in the endosperm the PcG complex represses the expression of the paternal allele through enrichment of repressive histone marks (H3K27me3). Moreover, the mechanism by which *PHERES1 (PHE1)* is paternally imprinted in endosperm involves a DMR in the 3′ end of *PHE1* and recruitment of PcG complex. In the central cell, the DMR in *PHE1* 3′ end is hypomethylated, allowing recruitment of FIS-PcG complex to *PHE1* promoter and resulting in stable maternal *PHE1* repression in the endosperm. The FIS complex is absent in sperm cells and the methylated DMR prevents the silencing activity, causing the paternal allele to be active in the endosperm ([110] and references therein).

Until recently, few genes were known to be imprinted in the embryo and for this reason the regulatory control was poorly understood [111]. The study of *MATERNALLY EXPRESSED IN EMBRYO 1* (*MEE1*) gene of maize, imprinted in both the embryo and endosperm, has shown that imprinted alleles can acquire DNA methylation after fertilization. The demethylated state of maternal *MEE1* in the central cell and the methylated paternal allele are consistent with the exclusive maternal expression of *MEE1* in endosperm. In the egg cell, the maternal *MEE1* is methylated but upon fertilization is rapidly de-methylated, establishing differential parental epigenetic states in the embryo. During embryo development the maternal allele is then remethylated resetting the maternal imprinting marks later in embryogenesis [111]. This embryonic *MEE1* maternal imprinting is consistent with an active demethy-lation in the zygote; however, the mechanism of regulation raised some intriguing questions implying the existence of an allele-specific recognition mechanism that could distinguish between two equal methylated parental alleles [78, 111]. One possible hypothesis is that 24-nt siRNAs accumulating in the endosperm could be transported to the embryo [58, 90] to guide demetylation of the maternal *MEE1* allele in the embryo [112]. Evidences show that siRNAs interacting with ROS3 are sufficient to guide sequence-specific demethylation involving ROS1 [113], supporting the possibility that siRNAs can also guide DNA demethylation. However, it still remains to be shown that strand-specific differences exist between parental alleles. Such possibility could be achieved by differential histone marks in parental alleles. The recent identification of several genes imprinted in embryo do not exclude this hypothesis and support that DNA methylation is not a primary imprinting mark in embryos. The regulation of embryo imprinting seems to be partially imposed by H3K27me3 mediated by the PRC2 complex but independent of MET1

[114]. While genomic imprinting in the endosperm does not raise major implications for the following generation, evidences indicate that imprinting in the embryo is maintained during embryogenesis and early seedling development, but PEG and MEG imprinting marks need to be eventually erased and reset before reproductive development in the next generation [114].

The fact that a subset of maternal and paternal imprinted expressed genes are associated with DMRs, probably deriving from accidental silencing of repetitive elements neighbouring these genes, led to the hypothesis that genomic imprinting could have evolved as a by-product of TE silencing [115]. Recent evidences support that RdDM pathways acting in plant gametes can regulate parental-specific genomic imprinting at specific loci in endosperm [116]. In early endosperm development, low RdDM activity together with low activity of *MET1* and other methyltransferases [86, 91, 117] is expected to contribute to maintenance of the demethylated state of the maternal allele and activation of the maternal imprinted genes (MEGs). Silencing of the paternal alleles seems however to be dependent on RdDM pathways, whereby NRPD2A-dependent siRNAs generated potentially from activity of transposon elements in diploid somatic paternal tissues seem to be sufficient to target *de novo* DNA methylation and silencing prior to gametogenesis. The methylated status of paternal alleles is maintained by MET1 during sporophytic pollen development, sustained in the haploid gametes and inherited after fertilization in endosperm [116].

The observation that most imprinted genes in plants and animals are expressed in the placenta and endosperm, respectively, tissues allocated to nutrition of the developing embryo gave rise to the prediction that the function of imprinted genes in endosperm could have played a major role in the evolution and selection of genomic imprinting. The parental conflict theory predicts that nutrient allocation in the offspring is resolved in favour of the parents reproductive interests, i.e. maternally expressed imprinted genes should negatively regulate endosperm growth, while paternal imprinted genes positively regulate endosperm growth [118]. While this is the case for many of the identified imprinted genes, the *MATERNALLY EXPRESSED GENE 1* (*MEG1*) in maize has opposite effects promoting seed growth [119]. In addition, the lack of obvious functions and the fact that ectopic expression of many of these imprinted genes in vegetative tissues do not cause deleterious effects in plant development [120, 121], have led to a new hypothesis in which, genomic imprinting might have evolved from positive selection of beneficial mutations in vegetative tissues that were then retained and co-opted for positive functions in embryo or endosperm [120].

The epigenetic reprogramming observed in the gametes and their products has implications that go beyond their alleged roles for transposon and imprinting control during the reproductive phase. This could be the case for the observed variability of epialleles across generations [68, 69], possibly driven by gains and losses in CHH methylation during gametogenesis and early embryo development [122]. In this respect it should also be noted that both 21- and 24-nt siRNAs can be transported over long distances between shoot and root (and *vice versa*) and that these mobile 24-nt siRNAs are able to direct RdDM in meristematic root stem cells [123–125]. Assum-

ing that transport could also be effective to the shoot meristem and eventually to the gametes, these siRNAs could act in a pathway underlying epigenetic inheritance. This could explain transgenerational memory of stress as observed in plants [126]. Support for this hypothesis comes from salicylic acid stress in the sporophyte leading to the production of TE-associated 21-nt siRNAs that are similar to those that are epigenetically activated in the vegetative nucleus of pollen [127].

Perspectives

In recent years, we have seen significant advances in experimental methods allowing the isolation of specific reproductive cell types or cell-type components. Through combination with genome-scale approaches like bisulphite sequencing and RNAseq, these have led to a better comprehension of the epigenetic landscape associated to regulation of a number of developmental transitions, namely how plants determine a sporogenous reproductive cell lineage from somatic cells, how dimorphic epigenetic states may contribute to male germline differentiation in a multicellular gametophyte or how dimorphic epigenetic states in female gametes contribute to establish distinct developmental programs post-fertilization. However, the epigenetic profiles of several specific developmental transitions are still incomplete due to the challenge of isolating particular cell types (e.g. meiocytes or female gametophytic cells) from their neighbouring somatic cells or in specific developmental stages (e.g. during pollen tube growth), information that will be crucial to improve resolution of epigenetic states in specific cell types and define the precise timing of epigenetic reprogramming during gametogenesis. These reprogramming events are also a prerequisite to establish temporary transcriptional repressive states during cell fate transitions in order to restore totipotency and allow cell-type differentiation programmes.

Moreover, it will be of fundamental importance to understand the regulatory mechanisms that specifically operate during these phase transitions, enabling plants to retain, erase or reset specific epigenetic information acquired during the plant life cycle as well as parent specific allelic expression patterns that can contribute to the next generation. Histone modifications and small RNAs seem to be involved in this transgenerational epigenetic inheritance. It will be of particular interest to understand if mobile siRNAs, either from companion cells or inherited, can modify the epigenetic landscape of the zygote or function as critical regulators of gene expression during embryo development. An exciting development has been the association of epigenetic mechanisms with the potential regulation of apomictic development, raising a profound interest in the epigenetic principles that distinguish sexual reproduction from apomixis. Given the emerging impact of epigenetic processes on several aspects of sexual reproduction in angiosperms future insights hold the promise to yield novel experimental tools, ultimately opening new paths to improve crop species.

References

1. Grant-Downton RT, Dickinson HG. Epigenetics and its implications for plant biology 2. The 'epigenetic epiphany': epigenetics, evolution and beyond. Ann Bot. 2006;97(1):11–27.
2. Pereira PA, Navarro-Costa P, Martinho RG, Becker JD. Evolutionarily-conserved mechanisms of male germline development in flowering plants and animals. Biochem Soc Trans. 2014;42:377–82.
3. Huijser P, Schmid M. The control of developmental phase transitions in plants. Development. 2011;138(19):4117–29.
4. Madlung A, Comai L. The effect of stress on genome regulation and structure. Ann Bot. 2004;94(4):481–95.
5. Boyko A, Kovalchuk I. Transgenerational response to stress in *Arabidopsis thaliana*. Plant Signal Behav. 2010;5(8):995–8.
6. Saze H. Epigenetic memory transmission through mitosis and meiosis in plants. Semin Cell Dev Biol. 2008;19(6):527–36.
7. Eichten S, Borevitz J. Epigenomics: methylation's mark on inheritance. Nature. 2013;495(7440):181–2.
8. Jacobsen SE, Meyerowitz EM. Hypermethylated SUPERMAN epigenetic alleles in Arabidopsis. Science. 1997;277(5329):1100–3.
9. Finnegan EJ, Peacock WJ, Dennis ES. Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. Proc Natl Acad Sci U S A. 1996;93(16):8449–54.
10. Hirsch S, Baumberger R, Grossniklaus U. Epigenetic variation, inheritance, and selection in plant populations. Cold Spring Harb Symp Quant Biol. 2012;77:97–104.
11. Miura A, Yonebayashi S, Watanabe K, Toyama T, Shimada H, Kakutani T. Mobilization of transposons by a mutation abolishing full DNA methylation in Arabidopsis. Nature. 2001;411(6834):212–4.
12. Oliver KR, McComb JA, Greene WK. Transposable elements: powerful contributors to angiosperm evolution and diversity. Genome Biol Evol. 2013;5(10):1886–901.
13. Amasino RM, Michaels SD. The timing of flowering. Plant Physiol. 2010;154(2):516–20.
14. Deng W, Ying H, Helliwell CA, Taylor JM, Peacock WJ, Dennis ES. *FLOWERING LOCUS C* (*FLC*) regulates development pathways throughout the life cycle of Arabidopsis. Proc Natl Acad Sci U S A. 2011;108(16):6680–5.
15. Willmann MR, Poethig RS. The effect of the floral repressor *FLC* on the timing and progression of vegetative phase change in Arabidopsis. Development. 2011;138(4):677–85.
16. He Y, Amasino RM. Role of chromatin modification in flowering-time control. Trends Plant Sci. 2005;10(1):30–5.
17. Bastow R, Mylne JS, Lister C, Lippman Z, Martienssen RA, Dean C. Vernalization requires epigenetic silencing of *FLC* by histone methylation. Nature. 2004;427(6970):164–7.
18. He Y, Michaels SD, Amasino RM. Regulation of flowering time by histone acetylation in Arabidopsis. Science. 2003;302(5651):1751–4.
19. He Y. Control of the transition to flowering by chromatin modifications. Mol Plant. 2009;2(4):554–64.
20. Sung S, Amasino RM. Vernalization in *Arabidopsis thaliana* is mediated by the PHD finger protein VIN3. Nature. 2004;427(6970):159–64.
21. De Lucia F, Crevillen P, Jones AM, Greb T, Dean C. A PHD-polycomb repressive complex 2 triggers the epigenetic silencing of *FLC* during vernalization. Proc Natl Acad Sci U S A. 2008;105(44):16831–6.
22. Schubert D, Primavesi L, Bishopp A, Roberts G, Doonan J, Jenuwein T, et al. Silencing by plant polycomb-group genes requires dispersed trimethylation of histone H3 at lysine 27. EMBO J. 2006;25(19):4638–49.
23. Angel A, Song J, Dean C, Howard M. A polycomb-based switch underlying quantitative epigenetic memory. Nature. 2011;476(7358):105–8.

24. Liu F, Marquardt S, Lister C, Swiezewski S, Dean C. Targeted 3' processing of antisense transcripts triggers Arabidopsis *FLC* chromatin silencing. Science. 2010;327(5961):94–7.
25. Swiezewski S, Liu F, Magusin A, Dean C. Cold-induced silencing by long antisense transcripts of an Arabidopsis polycomb target. Nature. 2009;462(7274):799–802.
26. Sheldon CC, Hills MJ, Lister C, Dean C, Dennis ES, Peacock WJ. Resetting of *FLOWERING LOCUS C* expression after epigenetic repression by vernalization. Proc Natl Acad Sci U S A. 2008;105(6):2214–9.
27. Choi J, Hyun Y, Kang MJ, In Yun H, Yun JY, Lister C, et al. Resetting and regulation of *FLOWERING LOCUS C* expression during Arabidopsis reproductive development. Plant J. 2009;57(5):918–31.
28. Turck F, Fornara F, Coupland G. Regulation and identity of florigen: *FLOWERING LOCUS T* moves center stage. Annu Rev Plant Biol. 2008;59:573–94.
29. Smaczniak C, Immink RG, Muino JM, Blanvillain R, Busscher M, Busscher-Lange J, et al. Characterization of MADS-domain transcription factor complexes in Arabidopsis flower development. Proc Natl Acad Sci U S A. 2012;109(5):1560–5.
30. Feng X, Zilberman D, Dickinson H. A conversation across generations: soma-germ cell crosstalk in plants. Dev Cell. 2013;24(3):215–25.
31. Boavida LC, Becker JD, Feijo JA. The making of gametes in higher plants. Int J Dev Biol. 2005;49(5–6):595–614.
32. Horvitz HR, Herskowitz I. Mechanisms of asymmetric cell division: two Bs or not two Bs, that is the question. Cell. 1992;68(2):237–55.
33. Eady C, Lindsey K, Twell D. The significance of microspore division and division symmetry for vegetative cell-specific transcription and generative cell differentiation. Plant Cell. 1995;7(1):65–74.
34. McCue AD, Cresti M, Feijo JA, Slotkin RK. Cytoplasmic connection of sperm cells to the pollen vegetative cell nucleus: potential roles of the male germ unit revisited. J Exp Bot. 2011;62(5):1621–31.
35. Slotkin RK, Vaughn M, Borges F, Tanurdzic M, Becker JD, Feijo JA, et al. Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell. 2009;136(3):461–72.
36. Boavida LC, Qin P, Broz M, Becker JD, McCormick S. Arabidopsis tetraspanins are confined to discrete expression domains and cell types in reproductive tissues and form homo- and heterodimers when expressed in yeast. Plant Physiol. 2013;163(2):696–712.
37. Palanivelu R, Preuss D. Distinct short-range ovule signals attract or repel *Arabidopsis thaliana* pollen tubes *in vitro*. BMC Plant Biol. 2006;6:7.
38. Friedman WE. Expression of the cell cycle in sperm of Arabidopsis: implications for understanding patterns of gametogenesis and fertilization in plants and other eukaryotes. Development. 1999;126(5):1065–75.
39. Ingouff M, Sakata T, Li J, Sprunck S, Dresselhaus T, Berger F. The two male gametes share equal ability to fertilize the egg cell in *Arabidopsis thaliana*. Curr Biol. 2009;19(1):R19–20.
40. Solis MT, Chakrabarti N, Corredor E, Cortes-Eslava J, Rodriguez-Serrano M, Biggiogera M, et al. Epigenetic changes accompany developmental programmed cell death in tapetum cells. Plant Cell Physiol. 2014;55(1):16–29.
41. Millar AA, Gubler F. The Arabidopsis *GAMYB*-like genes, *MYB33* and *MYB65*, are microRNA-regulated genes that redundantly facilitate anther development. Plant Cell. 2005;17(3):705–21.
42. Grini PE, Thorstensen T, Alm V, Vizcay-Barrena G, Windju SS, Jorstad TS, et al. The ASH1 HOMOLOG 2 (ASHH2) histone H3 methyltransferase is required for ovule and anther development in Arabidopsis. PloS One. 2009;4(11):e7817.
43. Tarutani Y, Shiba H, Iwano M, Kakizaki T, Suzuki G, Watanabe M, et al. Trans-acting small RNA determines dominance relationships in Brassica self-incompatibility. Nature. 2010;466(7309):983–6.
44. Mirouze M, Lieberman-Lazarovich M, Aversano R, Bucher E, Nicolet J, Reinders J, et al. Loss of DNA methylation affects the recombination landscape in Arabidopsis. Proc Natl Acad Sci U S A. 2012;109(15):5880–5.

45. Yelina NE, Choi K, Chelysheva L, Macaulay M, de Snoo B, Wijnker E, et al. Epigenetic remodeling of meiotic crossover frequency in Arabidopsis thaliana DNA methyltransferase mutants. PLoS Genet. 2012;8(8):e1002844.
46. Perrella G, Consiglio MF, Aiese-Cigliano R, Cremona G, Sanchez-Moran E, Barra L, et al. Histone hyperacetylation affects meiotic recombination and chromosome segregation in Arabidopsis. Plant J. 2010;62(5):796–806.
47. Melamed-Bessudo C, Levy AA. Deficiency in DNA methylation increases meiotic crossover rates in euchromatic but not in heterochromatic regions in Arabidopsis. Proc Natl Acad Sci U S A. 2012;109(16):E981–8.
48. Nonomura K, Morohoshi A, Nakano M, Eiguchi M, Miyao A, Hirochika H, et al. A germ cell specific gene of the *ARGONAUTE* family is essential for the progression of premeiotic mitosis and meiosis during sporogenesis in rice. Plant Cell. 2007;19(8):2583–94.
49. Tucker MR, Okada T, Hu Y, Scholefield A, Taylor JM, Koltunow AM. Somatic small RNA pathways promote the mitotic events of megagametogenesis during female reproductive development in Arabidopsis. Development. 2012;139(8):1399–404.
50. Chen C, Retzel EF. Analyzing the meiotic transcriptome using isolated meiocytes of *Arabidopsis thaliana*. Methods Mol Biol. 2013;990:203–13.
51. Yang H, Lu P, Wang Y, Ma H. The transcriptome landscape of Arabidopsis male meiocytes from high-throughput sequencing: the complexity and evolution of the meiotic process. Plant J. 2011;65(4):503–16.
52. Ingouff M, Rademacher S, Holec S, Soljic L, Xin N, Readshaw A, et al. Zygotic resetting of the HISTONE 3 variant repertoire participates in epigenetic reprogramming in Arabidopsis. Curr Biol. 2010;20(23):2137–43.
53. Ingouff M, Hamamura Y, Gourgues M, Higashiyama T, Berger F. Distinct dynamics of HISTONE3 variants between the two fertilization products in plants. Curr Biol. 2007;17(12):1032–7.
54. Borges F, Gomes G, Gardner R, Moreno N, McCormick S, Feijo JA, et al. Comparative transcriptomics of Arabidopsis sperm cells. Plant Physiol. 2008;148(2):1168–81.
55. Borges F, Pereira PA, Slotkin RK, Martienssen RA, Becker JD. MicroRNA activity in the Arabidopsis male germline. J Exp Bot. 2011;62(5):1611–20.
56. Borges F, Gardner R, Lopes T, Calarco JP, Boavida LC, Slotkin RK, et al. FACS-based purification of Arabidopsis microspores, sperm cells and vegetative nuclei. Plant Methods. 2012;8(1):44.
57. Calarco JP, Borges F, Donoghue MT, Van Ex F, Jullien PE, Lopes T, et al. Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. Cell. 2012;151(1):194–205.
58. Ibarra CA, Feng X, Schoft VK, Hsieh TF, Uzawa R, Rodrigues JA, et al. Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. Science. 2012;337(6100):1360–4.
59. Schoft VK, Chumak N, Choi Y, Hannon M, Garcia-Aguilar M, Machlicova A, et al. Function of the *DEMETER* DNA glycosylase in the Arabidopsis thaliana male gametophyte. Proc Natl Acad Sci U S A. 2011;108(19):8042–7.
60. Schoft VK, Chumak N, Mosiolek M, Slusarz L, Komnenovic V, Brownfield L, et al. Induction of RNA-directed DNA methylation upon decondensation of constitutive heterochromatin. EMBO Rep. 2009;10(9):1015–21.
61. Zemach A, Kim MY, Hsieh PH, Coleman-Derr D, Eshed-Williams L, Thao K, et al. The Arabidopsis nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. Cell. 2013;153(1):193–205.
62. Kakutani T, Jeddeloh JA, Flowers SK, Munakata K, Richards EJ. Developmental abnormalities and epimutations associated with DNA hypomethylation mutations. Proc Natl Acad Sci U S A. 1996;93(22):12406–11.
63. McCue AD, Nuthikattu S, Reeder SH, Slotkin RK. Gene expression and stress response mediated by the epigenetic regulation of a transposable element small RNA. PLoS Genet. 2012;8(2):e1002474.

64. McCue AD, Nuthikattu S, Slotkin RK. Genome-wide identification of genes regulated in *trans* by transposable element small interfering RNAs. RNA Biol. 2013;10(8):1379–95.
65. Honys D, Twell D. Transcriptome analysis of haploid male gametophyte development in Arabidopsis. Genome Biol. 2004;5(11):R85.
66. Gehring M, Bubb KL, Henikoff S. Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science. 2009;324(5933):1447–51.
67. Wolff P, Weinhofer I, Seguin J, Roszak P, Beisel C, Donoghue MT, et al. High-resolution analysis of parent-of-origin allelic expression in the Arabidopsis endosperm. PLoS Genet. 2011;7(6):e1002126.
68. Becker C, Hagmann J, Muller J, Koenig D, Stegle O, Borgwardt K, et al. Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. Nature. 2011;480(7376):245–9.
69. Schmitz RJ, Schultz MD, Lewsey MG, O'Malley RC, Urich MA, Libiger O, et al. Transgenerational epigenetic instability is a source of novel methylation variants. Science. 2011;334(6054):369–73.
70. Grant-Downton R, Kourmpetli S, Hafidh S, Khatab H, Le Trionnaire G, Dickinson H, et al. Artificial microRNAs reveal cell-specific differences in small RNA activity in pollen. Curr Biol. 2013;23(14):R599–601.
71. She W, Grimanelli D, Rutowicz K, Whitehead MW, Puzio M, Kotlinski M, et al. Chromatin reprogramming during the somatic-to-reproductive cell fate transition in plants. Development. 2013;140(19):4008–19.
72. Olmedo-Monfil V, Duran-Figueroa N, Arteaga-Vazquez M, Demesa-Arevalo E, Autran D, Grimanelli D, et al. Control of female gamete formation by a small RNA pathway in Arabidopsis. Nature. 2010;464(7288):628–32.
73. Singh M, Goel S, Meeley RB, Dantec C, Parrinello H, Michaud C, et al. Production of viable gametes without meiosis in maize deficient for an ARGONAUTE protein. Plant Cell. 2011;23(2):443–58.
74. Duran-Figueroa N, Vielle-Calzada JP. *ARGONAUTE9*-dependent silencing of transposable elements in pericentromeric regions of Arabidopsis. Plant Signal Behav. 2010;5(11):1476–9.
75. Dunoyer P, Lecellier CH, Parizotto EA, Himber C, Voinnet O. Probing the microRNA and small interfering RNA pathways with virus-encoded suppressors of RNA silencing. Plant Cell. 2004;16(5):1235–50.
76. Garcia-Aguilar M, Michaud C, Leblanc O, Grimanelli D. Inactivation of a DNA methylation pathway in maize reproductive organs results in apomixis-like phenotypes. Plant Cell. 2010;22(10):3249–67.
77. Pillot M, Baroux C, Vazquez MA, Autran D, Leblanc O, Vielle-Calzada JP, et al. Embryo and endosperm inherit distinct chromatin and transcriptional states from the female gametes in Arabidopsis. Plant Cell. 2010;22(2):307–20.
78. Gutierrez-Marcos JF, Costa LM, Dal Pra M, Scholten S, Kranz E, Perez P, et al. Epigenetic asymmetry of imprinted genes in plant gametes. Nat Genet. 2006;38(8):876–8.
79. Calarco JP, Martienssen RA. Imprinting: DNA methyltransferases illuminate reprogramming. Curr Biol. 2012;22(21):R929–31.
80. Ikeda Y. Plant imprinted genes identified by genome-wide approaches and their regulatory mechanisms. Plant Cell Physiol. 2012;53(5):809–16.
81. Pillot M, Autran D, Leblanc O, Grimanelli D. A role for CHROMOMETHYLASE3 in mediating transposon and euchromatin silencing during egg cell reprogramming in Arabidopsis. Plant Signal Behav. 2010;5(10):1167–70.
82. Choi Y, Gehring M, Johnson L, Hannon M, Harada JJ, Goldberg RB, et al. DEMETER, a DNA glycosylase domain protein, is required for endosperm gene imprinting and seed viability in Arabidopsis. Cell. 2002;110(1):33–42.
83. Gehring M, Huh JH, Hsieh TF, Penterman J, Choi Y, Harada JJ, et al. DEMETER DNA glycosylase establishes *MEDEA* polycomb gene self-imprinting by allele-specific demethylation. Cell. 2006;124(3):495–506.
84. Jullien PE, Kinoshita T, Ohad N, Berger F. Maintenance of DNA methylation during the Arabidopsis life cycle is essential for parental imprinting. Plant Cell. 2006;18(6):1360–72.

85. Costa LM, Gutierrez-Marcos JF. Retinoblastoma makes its mark on imprinting in plants. PLoS Biol. 2008;6(8):e212.
86. Jullien PE, Susaki D, Yelagandula R, Higashiyama T, Berger F. DNA methylation dynamics during sexual reproduction in *Arabidopsis thaliana*. Curr Biol. 2012;22(19):1825–30.
87. Eamens A, Vaistij FE, Jones L. NRPD1a and NRPD1b are required to maintain post-transcriptional RNA silencing and RNA-directed DNA methylation in Arabidopsis. Plant J. 2008;55(4):596–606.
88. Mosher RA, Melnyk CW, Kelly KA, Dunn RM, Studholme DJ, Baulcombe DC. Uniparental expression of PolIV-dependent siRNAs in developing endosperm of Arabidopsis. Nature. 2009;460(7252):283–6.
89. Berger F, Grini PE, Schnittger A. Endosperm: an integrator of seed growth and development. Curr Opin Plant Biol. 2006;9(6):664–70.
90. Hsieh TF, Ibarra CA, Silva P, Zemach A, Eshed-Williams L, Fischer RL, et al. Genome-wide demethylation of Arabidopsis endosperm. Science. 2009;324(5933):1451–4.
91. Jullien PE, Mosquna A, Ingouff M, Sakata T, Ohad N, Berger F. Retinoblastoma and its binding partner MSI1 control imprinting in Arabidopsis. PLoS Biol. 2008;6(8):e194.
92. Martienssen RA. Heterochromatin, small RNA and post-fertilization dysgenesis in allopolyploid and interploid hybrids of Arabidopsis. New Phytol. 2010;186(1):46–53.
93. Bourc'his D, Voinnet O. A small-RNA perspective on gametogenesis, fertilization, and early zygotic development. Science. 2010;330(6004):617–22.
94. Ng DW, Lu J, Chen ZJ. Big roles for small RNAs in polyploidy, hybrid vigor, and hybrid incompatibility. Curr Opin Plant Biol. 2012;15(2):154–61.
95. Josefsson C, Dilkes B, Comai L. Parent-dependent loss of gene silencing during interspecies hybridization. Curr Biol. 2006;16(13):1322–8.
96. Ha M, Lu J, Tian L, Ramachandran V, Kasschau KD, Chapman EJ, et al. Small RNAs serve as a genetic buffer against genomic shock in Arabidopsis interspecific hybrids and allopolyploids. Proc Natl Acad Sci U S A. 2009;106(42):17835–40.
97. Brennecke J, Malone CD, Aravin AA, Sachidanandam R, Stark A, Hannon GJ. An epigenetic role for maternally inherited piRNAs in transposon silencing. Science. 2008;322(5906):1387–92.
98. Bayer M, Nawy T, Giglione C, Galli M, Meinnel T, Lukowitz W. Paternal control of embryonic patterning in *Arabidopsis thaliana*. Science. 2009;323(5920):1485–8.
99. Krawetz SA, Kruger A, Lalancette C, Tagett R, Anton E, Draghici S, et al. A survey of small RNAs in human sperm. Hum Reprod. 2011;26(12):3401–12.
100. Sendler E, Johnson GD, Mao S, Goodrich RJ, Diamond MP, Hauser R, et al. Stability, delivery and functions of human sperm RNAs at fertilization. Nuc Acids Res. 2013;41(7):4104–17.
101. Conine CC, Moresco JJ, Gu W, Shirayama M, Conte D, Jr., Yates JR, 3rd, et al. Argonautes promote male fertility and provide a paternal memory of germline gene expression in C. elegans. Cell. 2013;155(7):1532–44.
102. Baroux C, Autran D, Gillmor CS, Grimanelli D, Grossniklaus U. The maternal to zygotic transition in animals and plants. Cold Spring Harb Symp Quant Biol. 2008;73:89–100.
103. Autran D, Baroux C, Raissig MT, Lenormand T, Wittig M, Grob S, et al. Maternal epigenetic pathways control parental contributions to Arabidopsis early embryogenesis. Cell. 2011;145(5):707–19.
104. Vielle-Calzada JP, Baskar R, Grossniklaus U. Delayed activation of the paternal genome during seed development. Nature. 2000;404(6773):91–4.
105. Grimanelli D, Perotti E, Ramirez J, Leblanc O. Timing of the maternal-to-zygotic transition during early seed development in maize. Plant Cell. 2005;17(4):1061–72.
106. Feng S, Jacobsen SE, Reik W. Epigenetic reprogramming in plant and animal development. Science. 2010;330(6004):622–7.
107. Saze H, Mittelsten Scheid O, Paszkowski J. Maintenance of CpG methylation is essential for epigenetic inheritance during plant gametogenesis. Nat Genet. 2003;34(1):65–9.

108. Meyer S, Scholten S. Equivalent parental contribution to early plant zygotic development. Curr Biol. 2007;17(19):1686–91.
109. Scott RJ, Spielman M. Genomic imprinting in plants and mammals: how life history constrains convergence. Cytogenet Genome Res. 2006;113(1–4):53–67.
110. Kohler C, Wolff P, Spillane C. Epigenetic mechanisms underlying genomic imprinting in plants. Annu Rev Plant Biol. 2012;63:331–52.
111. Jahnke S, Scholten S. Epigenetic resetting of a gene imprinted in plant embryos. Curr Biol. 2009;19(19):1677–81.
112. Jiang H, Kohler C. Evolution, function, and regulation of genomic imprinting in plant seed development. J Exp Bot. 2012;63(13):4713–22.
113. Zheng X, Pontes O, Zhu J, Miki D, Zhang F, Li WX, et al. ROS3 is an RNA-binding protein required for DNA demethylation in Arabidopsis. Nature. 2008;455(7217):1259–62.
114. Raissig MT, Bemer M, Baroux C, Grossniklaus U. Genomic imprinting in the Arabidopsis embryo is partly regulated by PRC2. PLoS Genet. 2013;9(12):e1003862.
115. Kohler C, Weinhofer-Molisch I. Mechanisms and evolution of genomic imprinting in plants. Heredity. 2010;105(1):57–63.
116. Vu TM, Nakamura M, Calarco JP, Susaki D, Lim PQ, Kinoshita T, et al. RNA-directed DNA methylation regulates parental genomic imprinting at several loci in Arabidopsis. Development. 2013;140(14):2953–60.
117. Belmonte MF, Kirkbride RC, Stone SL, Pelletier JM, Bui AQ, Yeung EC, et al. Comprehensive developmental profiles of gene activity in regions and subregions of the Arabidopsis seed. Proc Natl Acad Sci U S A. 2013;110(5):E435–44.
118. Zhang H, Chaudhury A, Wu X. Imprinting in plants and its underlying mechanisms. J Genet Genomics. 2013;40(5):239–47.
119. Costa LM, Yuan J, Rouster J, Paul W, Dickinson H, Gutierrez-Marcos JF. Maternal control of nutrient allocation in plant seeds by genomic imprinting. Curr Biol. 2012;22(2):160–5.
120. Berger F, Vu TM, Li J, Chen B. Hypothesis: selection of imprinted genes is driven by silencing deleterious gene activity in somatic tissues. Cold Spring Harb Symp Quant Biol. 2012;77:23–9.
121. Ikeda Y, Kobayashi Y, Yamaguchi A, Abe M, Araki T. Molecular basis of late-flowering phenotype caused by dominant epi-alleles of the *FWA* locus in Arabidopsis. Plant Cell Physiol. 2007;48(2):205–20.
122. Borges F, Martienssen RA. Establishing epigenetic variation during genome reprogramming. RNA Biol. 2013;10(4):490–4.
123. Dunoyer P, Brosnan CA, Schott G, Wang Y, Jay F, Alioua A, et al. An endogenous, systemic RNAi pathway in plants. EMBO J. 2010;29(10):1699–712.
124. Molnar A, Melnyk CW, Bassett A, Hardcastle TJ, Dunn R, Baulcombe DC. Small silencing RNAs in plants are mobile and direct epigenetic modification in recipient cells. Science. 2010;328(5980):872–5.
125. Dunoyer P, Schott G, Himber C, Meyer D, Takeda A, Carrington JC, et al. Small RNA duplexes function as mobile silencing signals between plant cells. Science. 2010;328(5980):912–6.
126. Molinier J, Ries G, Zipfel C, Hohn B. Transgeneration memory of stress in plants. Nature. 2006;442(7106):1046–9.
127. Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, et al. Widespread dynamic DNA methylation in response to biotic stress. Proc Natl Acad Sci U S A. 2012;109(32):E2183–91.
128. Christensen CA, King EJ, Jordan JR, Drews GN. Megagametogenesis in Arabidopsis wild type and the Gf mutant. Sex Plant Reprod. 1997;10:49–64.

Chapter 7
Epigenetic Modifications at Developmental Transitions in *Arabidopsis*

Heike Wollmann and Frédéric Berger

Chromatin Features and Transcriptional Activity

Our current understanding of chromatin dynamics during developmental transitions in plants is to a large extent based on correlations between transcription rates and profiles of corresponding chromatin environments. Many chromatin features show distinct enrichment patterns over gene bodies, some of which positively or negatively correlate with transcription rates.

DNA Methylation

DNA methylation occurs at three distinct sequence contexts, the symmetric CG and CHG and the asymmetric CHH. CG methylation is the most common type of methylation, whereas methylated CHG and CHH sites are less abundant [1]. The predominant role of DNA methylation is maintaining genome integrity by silencing repetitive sequences and transposable elements (TEs) present in pericentromeric regions and DNA methylation over these regions is usually enriched in the three sequence contexts [1–4]. Aside from its predominant enrichment over silent chromatin, DNA methylation in CG context is enriched over gene bodies [1, 2, 5–7]. In contrast to its role over silent heterochromatin, the function of euchromatic or genic enrichment of DNA methylation is less clear [8, 9]. CG methylation is predomi-

F. Berger (✉) · H. Wollmann
Temasek Lifesciences Laboratory, 1 Research Link, National University of Singapore, Singapore 117604, Republic of Singapore
e-mail: fred@tll.org.sg

Department of Biological Sciences, National University of Singapore, 14 Science Drive 4, Singapore 117543, Republic of Singapore

H. Wollmann
Genome Institute of Singapore, 60 Biopolis Street, Genome, #02-01, Singapore 138672, Republic of Singapore

O. Pontes, H. Jin (eds.), *Nuclear Functions in Plant Transcription, Signaling and Development,* DOI 10.1007/978-1-4939-2386-1_7

nantly catalyzed by the DNA methyltransferase MET1 [1, 4]. DNA methylation enrichment over gene bodies increases toward the 3′ end of genes and with transcription rates, reaching highest values over genes that show moderate expression rates [1, 4, 7, 10]. Genes with highest expression levels display less gene body methylation than genes with moderate expression levels.

Histone Modifications

Aside from the crucial structural role of histone proteins within the chromatin, covalent modifications at the N-terminal end of histone proteins are associated with particular chromatin states. The enrichment of subsets of specific histone modifications indicates the state of the associated chromatin as active, repressed, or permanently silent. Amongst a wide variety of histone modifications, the best characterized of these modifications occur on the N-terminal tail of histone H3 and comprise acetylation and methylation of lysine residues and methylation of arginine residues. Two of these marks, trimethylation of the lysine residues number 4 and 27 (H3K4me3 and H3K27me3, respectively), are tightly associated with gene expression and developmental transitions. The variety of methyl marks is complex, as mono-, di-, or trimethylation of specific lysines can have very distinct genomic locations and therefore functions in the chromatin landscape. In plants, a good example is H3K27me1, which is enriched over constitutive heterochromatin, while H3K27me3 is enriched in euchromatic regions and associated with (developmentally) repressed gene loci [11]. Despite the large number of histone marks known and associated with transcriptional and chromatin activity, they tend to occur in a limited number of combinations associated with specific chromatin functional status [11]. Enrichment profiles assessed through genome-wide analysis of chromatin immunoprecipitation (ChIP) led to the classification of as few as four major chromatin states (CS1–CS4) on *Arabidopsis* chromosome 4 [11]. Most active genes can be found in CS1, a chromatin state enriched for H3K4me3, H3K9me3, H3K36me3, and H3K56ac, marks characteristically associated with active transcription. CS2 is predominantly enriched for H3K27me3, a typical mark found in gene bodies of repressed genes [6]. CS3 and CS4 regions are largely devoid of genes and while CS3 includes constitutive heterochromatin, mostly enriched for H3K9me2, H3K27me1/2, and H4K20me1, there is no particular enrichment of any of the investigated marks in CS4 [11].

Similarly, in differentiated animal cells, the two marks H3K4me3 and H3K27me3 are associated with distinct chromatin states and their enrichment domains are typically exclusive. However, in embryonic stems cells, dual appearance of both marks (bivalence) has been described on a small subset of developmental genes, marking a specific chromatin state “poised” for gene activation [12, 13]. In plants, there is little evidence for bivalent marks and we do not know whether they have a function similar to that described in animals.

Histone Variants

Nucleosomes are octamers built from two histones of the same histone family, H2A, H2B, H3, and H4, while H1 interacts with the linker DNA between nucleosomes. To the exception of histone H4, other histone families comprise variants in amino acid sequence. Some variants differ only by a few key amino acid residues, while others comprise additional N- or C-terminal domains. With an increasing number of histone variants being characterized, their crucial role in chromatin regulation is emerging [14, 15]. In plants, histone families underwent at least two duplications and a remarkable diversification, leading to a substantial number of gene loci encoding identical and distinct histone variants from the different families [14, 16]. Different variants from the H2A and H3 families have been studied in detail and show profiles with distinct genomic locations. These include the variants H3.3 and H2A.Z, which are involved in transcriptional dynamics during developmental transitions.

Only recently, a comprehensive analysis using ChIP-seq characterized the genomic profile and potential function of H2A.Z and H3.3 in *Arabidopsis*. The variant H2A.Z is particularly enriched over euchromatic regions and has mostly been associated with gene regulatory functions during development and adaptation to abiotic stress. Over gene bodies, H2A.Z is predominantly enriched over 5′ gene ends, at gene promoters/transcription start sites (TSS) [17]. This pattern, though overlapping with that of covalent histone modifications like H3K4me3, is less correlated with transcription rates and complete loss of H2A.Z function mostly affects a subset of genes with predominant gene body enrichment of H2A.Z [18]. At the same time, the enrichment of H2A.Z at 5′ gene ends has been shown to be important for transcriptional regulation as well [19].

The plant H3 family consists of 15 *HISTONE THREE REATED* (*HTR*) genes, encoding for a total of eight different H3 variants [16]. The most abundant euchromatic H3 variants during plant development are H3.1 and H3.3, encoded by five and three independent gene loci, respectively [20]. H3.1 is also referred to as the canonical H3 and is distributed largely evenly across the genome and over gene bodies [21, 22]. H3.1 is highly abundant in cells undergoing active cell divisions, which has been shown particularly in meristems [22, 23]. With its transcription being coupled to cell division, H3.1 is the predominant replicative H3, with expression peaks coinciding with DNA replication [24] (www.cyclebase.org). In contrast, H3.3 is expressed throughout the cell cycle and can therefore act as a replacement H3, being incorporated upon the loss of nucleosomes by, e.g., transcription [24, 25]. The two H3 types differ by only a few amino acids, but still are placed into the chromatin by distinct chaperone complexes [26]. In contrast to H3.1, H3.3 is enriched over active chromatin, predominantly toward the 3′ end of gene bodies close to transcription termination sites (TTS), where its levels correlate with transcription rates and RNA polymerase II enrichment sites in plants and animals [21, 22, 27–29]. The expression of most other plant H3 variants has either been extremely low to undetectable or highly specific, like the male-gamete-specific H3.10 (formerly MGH3 or

HTR10), which is likely involved in chromatin remodeling during spermatogenesis [20, 23, 30]. Lastly, the centromeric CenH3 (also known as CENH3, CENP-A, or HTR12) is highly enriched over centromeric chromatin and largely absent from euchromatic regions [31]. Taken together, only two of the many *Arabidopsis* H3 variants have thus been associated with transcription modulation or developmental transitions, H3.3 and H3.10.

Regulation of Responsiveness

Transcriptional changes allow cell differentiation and enable developmental transitions. Similarly, adaptive processes to, e.g., environmental changes require modifications of gene expression levels. Recent evidence suggests that the responsiveness of a gene, i.e., its competence for transcriptional changes is at least partially encoded in its chromatin environment [8, 32].

The large number of transcriptome studies available for different tissue types or responses to various environmental or hormonal stimuli enabled a comprehensive analysis of the responsiveness of each *Arabidopsis* gene, by comparison of expression patterns in these different conditions [33]. As a result, the response factor (RF) of each gene provides the means to characterize gene responsiveness to developmental or to stress-induced stimuli. Genes with low RF comprise housekeeping genes, while extremely high overall RF values enabled the identification of hypervariable genes [33]. Hypervariable genes are characteristically short, with few introns and notably low DNA methylation levels, compared to highly methylated housekeeping genes [33]. These initial observations indicate that in comparison, housekeeping and responsive genes show distinct chromatin structures. This link is further supported by the tight link between gene responsiveness and gene body enrichment of the histone variant H2A.Z [18].

Mutations in ARP6 and other components of the H2A.Z chaperone complex disturb the regulation of the vegetative-to-reproductive transition, due to misregulation of temperature responsive genes [19, 34–38]. *arp6* mutants flower early in normal temperature, similar to wild-type plants in elevated temperature conditions, highlighting the tight link between environmental stimuli and developmental timing. The transcriptome of *arp6* mutants in normal temperature resembles that of warm grown plants and H2A.Z nucleosome dynamics at promoters of temperature responsive genes, which suggested that the histone variant fulfills a thermosensory role [19]. However, it is likely that the role of H2A.Z is not specific to temperature sensing. Responsive genes in general show a preferential enrichment of H2A.Z over their gene bodies and a large number of these genes are misregulated in *h2a.z* mutants [18]. In agreement with this, the loss-of-function mutants in two of the three *H2A.Z* loci or H2A.Z chaperone component *PIE1* show misregulation of immune response genes and enhanced resistance to pathogens [39]. Furthermore, phosphate starvation response genes are constitutively derepressed in *arp6* mutants, and this activation is accompanied by a depletion of H2A.Z over a subset of the misregulated genes [40]. These studies provide a paradigm for the tight link

between development and adaptation, as plants rely on environmental cues to time developmental transitions. Besides, H2A.Z acts antagonistic to DNA methylation, providing a potential basis for a distinct chromatin environment over highly regulated and constitutively expressed genes [8].

Developmental Dynamics of Chromatin Features

During the plant life cycle, several transitions occur that mark the end of one developmental phase and the start of another. These transitions are crucial for plant life and reproduction and therefore highly regulated on the transcriptional and chromatin level (Fig. 7.1). All plants undergo a switch between haploid and diploid generations, the gametophytic and sporophytic cycles, respectively. In flowering plants, the sporophytic cycle is marked by a series of distinct phases, starting with embryo and seed development and proceeding over germination to juvenile-to-adult and vegetative-to-reproductive transition, initiating flowering, which leads to gametogenesis and a new gametophytic phase. Chromatin features have also been shown to dynamically accompany the transition from undifferentiated meristematic to differentiated cell states, e.g., to enable faithful silencing of meristematic genes after cell differentiation or allow transcription of developmental genes. During reproductive development, silenced TEs become derepressed, a potential mechanism to reinforce the silenced state in the next generation. Many of the known chromatin marks participate in more than one of these transitions and loss-of-function mutants can have devastating effects on plant growth and development. Below, we highlight the tight link between chromatin features and developmental transitions by describing a few examples.

Dynamics of DNA Methylation During Reproduction and Embryogenesis

Derepression of TEs can cause their transposition, jeopardizing genome integrity. TEs are therefore strictly silenced, correlating with the enrichment of several chromatin features, including DNA methylation. However, during reproduction and seed development, otherwise stable DNA methylation patterns become dynamic in a cell-type and parent-of-origin-specific manner, crucial for this developmental transition [41]. Reproduction requires extensive reprogramming; the removal of parental epigenetic information to allow developmental programs in the embryo, while at the same time, the silencing of TEs should be maintained. An attractive model based on the analysis of differential DNA methylation patterns suggests that germ-cell- and embryo-associated cells that accompany, but not contribute to, cells forming the next generation, undergo global DNA demethylation, causing the activation of TEs [42–45]. TE expression leads to production of small RNAs, which in turn reinforce silencing through small RNA-mediated DNA methylation in the germ cells and the developing embryo.

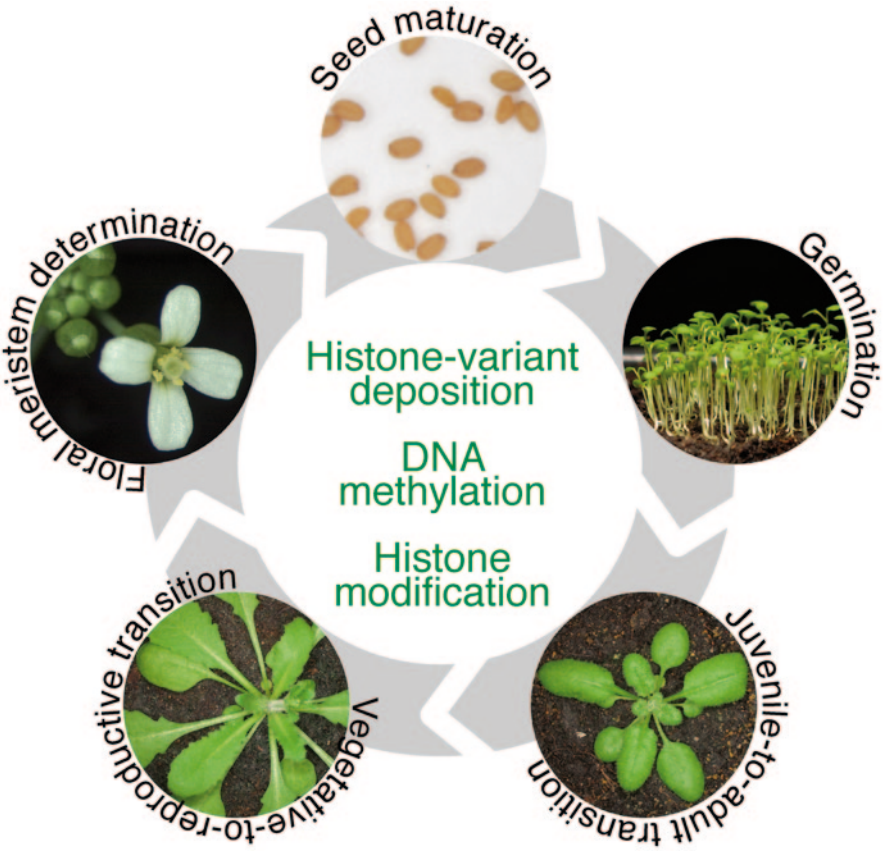

Fig. 7.1 Chromatin features dynamical change during developmental transitions. Several key developmental transitions mark the plant life cycle, some of which are depicted in this diagram. Most of these transitions have been associated with dynamic changes in the chromatin landscape, at least in one of the three types of chromatin features, DNA methylation, covalent histone modification, and histone variant replacement

In plants, reprogramming of DNA methylation during gametogenesis also affects gene expression in a male- or female-specific manner, leading to parent-of-origin-specific expression after fertilization. Such genes are referred to as imprinted genes. Imprinted genes are associated with differentially methylated regions (DMRs) [46–49]. Imprinting in plants is almost exclusively confined to the endosperm, the tissue nurturing the embryo and results from differential methylation of male and female gametes at the time of fertilization [50].

PcG and H3K27me3 in Developmental Transitions

H3K27me3 is associated with multiple developmental transitions in plants, including seed maturation, flowering, and differentiation. Although it is clear that

H3K9me3 is propagated through cell divisions and allows faithful transmission of patterns of silenced gene expression [51], this remains controversial in the case of H3K27me3 [52, 53]. H3K27me3 is associated with gene repression and a large number of genes are targeted by this mark in *Arabidopsis* [6]. The developmental roles of H3K27me3 became obvious by the severe phenotypes of mutants in polycomb repressive complex 2 (PRC2) members [54]. PRC2 is highly conserved and catalyzes H3K27me3. Mutants in PRC2 components affect various aspects of plant growth and development.

Comparison of H3K27me3 patterns in undifferentiated meristem cells and differentiated leave tissue revealed the dynamic nature of this mark during cell differentiation [55]. Enrichment and depletion follow transcriptional remodeling affecting known cell-type or tissue-specific regulatory genes, but also signaling pathways. From the early studies of the *FERTILIZATION INDEPENDENT SEED* (*FIS*) mutants affected by the lack of PRC2 activity in endosperm, it appeared that H3K27 methylation controls the transition from early to late developmental programs that separate the early syncytial from the late cellular developmental phase of the endosperm [56]. Similar roles were later reported for other transitions. *FERTILIZATION INDEPENDENT ENDOSPERM* (*FIE*) is a component of the PRC2 complex. In *fie* mutants, H3K27me3 levels are greatly reduced and seeds display enhanced dormancy and germination defects [57, 58]. The phenotypic defects are accompanied by upregulation of reproductive genes, indicating that the loss of the repressive function of H3K27me3 prevents the embryo-to-seedling transition in these plants. Germination requires silencing of seed developmental genes and activation of genes essential for vegetative growth. Only recently it has been proposed that a chromatin state switch from H3K4me3 to H3K27me3 is required to silence seed developmental genes and that the lack of this switch retards germination [59].

H3K27me3 plays also an important role in regulating floral transition via the complex regulation of the floral repressor FLC, a topic that has recently been covered in great detail elsewhere [60].

Furthermore, H3K27me3 has been shown to be a crucial mark for the termination of stem cell fate in flower development [61–63]. In contrast to the indeterminate shoot apical meristem, floral meristems are determinate, producing a defined number of organs before termination. The transition from indeterminate to determinate fate is regulated by the floral homeotic factor AGAMOUS (AG), inducing expression of KNUCKLES (KNU), a zinc finger protein that represses WUSCHEL (WUS), the major factor maintaining stem cell identity and indeterminacy of meristems [61–63]. Notably, the induction of KNU by AG requires about 2 days, a time delay that is regulated by H3K27me3 turnover. AG binding sites at the *KNU* upstream region reside within Polycomb response elements (PREs) and AG binding prevents access of Polycomb group (PcG) proteins, causing H3K27me3 depletion and eventually KNU expression. Importantly, this epigenetic timing mechanism is tightly linked to the cell cycle, which is required for H3K27me3 turnover. This mechanism could be envisaged as a broadly applicable way of timing PcG-regulated gene expression during developmental transitions (Fig. 7.2). Vice versa, the acquisition of pluripotency during the transition of differentiated leaf cells into callus cells also requires genome-wide reprogramming of H3K27me3 and mutants in PRC2 components lack the potential to form callus from leaf tissue [64].

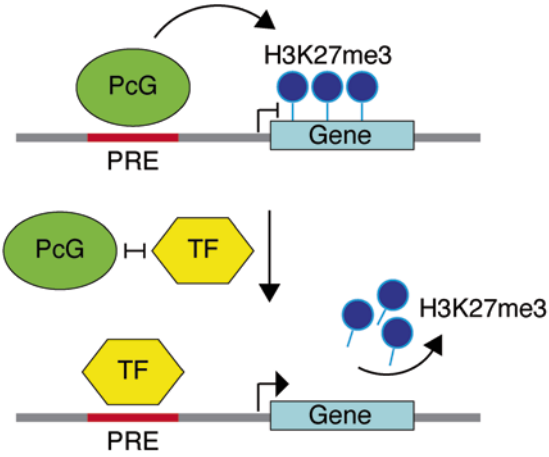

Fig. 7.2 Developmental timing through cell-cycle-dependent turnover of histone marks. Simplified model showing potential transcriptional timing during developmental transitions through turnover of histone marks. Polycomb group (PcG) proteins bind at polycomb responsive elements (PREs) in the promoter region of PcG-regulated developmental genes and maintain transcriptional repression through deposition of H3K27me3. Upon transition, competitive binding of transcription factors with co-localized binding sites in the promoter prevents access of the *PcG* complex, leading to a loss of H3K27me3 during the next rounds of cell division and activation of gene expression. To date, only a single case of this type of regulation is known, however, the general applicability of this model is plausible

The impact of PRC2 on developmental transitions is likely very ancient as it is clear that PRC2 activity is conserved in the moss *Physcomitrella patens* and regulates the transition between the haploid gametophyte and the diploid sporophyte [65, 66]. After mosses, the plant sporophyte became the dominant life phase and PRC2 components duplicated to reach for example three major sets in *Arabidopsis*, where clearly the gametophytic genes *MEDEA* and *FIS2* still play a role distinct from the genes *EMBRYONIC FLOWER 2, CURLY LEAF*, and *SWINGER* that are active primarily in the sporophyte. Sporophytic PRC2 complexes have specialized in regulating the major developmental transitions in the life cycle of higher plants. How exactly this specialization is conserved across flowering plants remains unknown.

Conclusion and Hypotheses

The turnover of histone modifications plays an obvious role in developmental transitions through the interplay with transcription factors. However, it remains unclear how the process of deposition and removal of histone marks is regulated dynamically. Although Jumonji (JMJ) domain proteins with a H3K27 demethylase activity have been identified [67], they do not show an impact on developmental transitions comparable to that described for PRC2 members for example. In addition, histone marks not only affect the promoter region where transcription factors reside but

they also extend in the gene body and regulation as well as impact of this expansion on transcription rates remain unclear. Gene bodies are also marked by differential enrichment of H2A.Z, H3.3, and DNA methylation and we propose below that these three chromatin features influence nucleosome dynamics that would impact the turnover of histone marks and enable chromatin dynamics during developmental transitions.

H3.3 Dynamics during Differentiation

In differentiated tissue, cell division rates are generally low and highly dividing regions are typically undifferentiated. Cell division and DNA replication allow dynamic exchanges of chromatin features; however, after the termination of cell division, changes in chromatin environment require active turnover of chromatin features (Fig. 7.3). The deposition of histone variants is critical, as some of them are dependently expressed cell cycle and therefore not available in large quantities in differentiated cell types. In the case of the H3 variants H3.1 and H3.3, H3.1 expression is largely limited to dividing cells and levels drop distinctly in nondividing cells, leaving H3.3 to be the most abundant H3 variant in differentiated, nondividing tissue [22]. Notably, H3.3 levels over genes are highly dynamic and accompany changes in the transcriptome that occur during cell differentiation. Developmentally

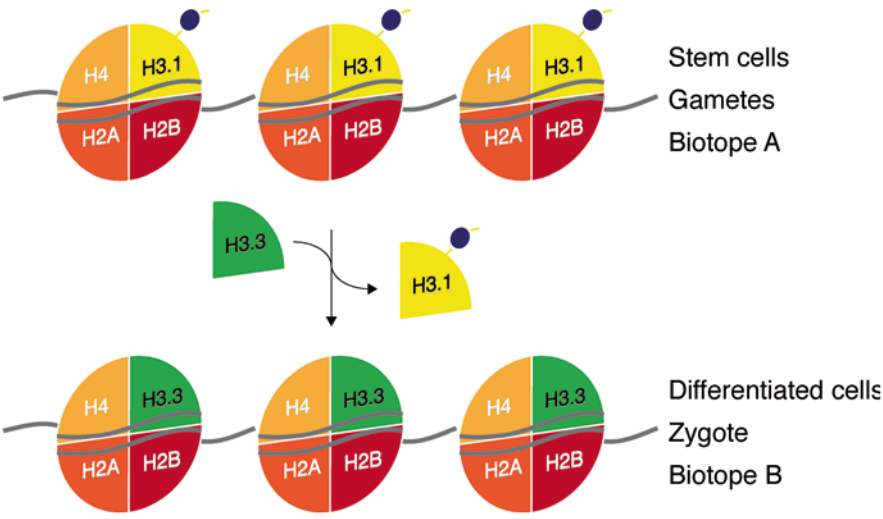

Fig. 7.3 Replacement of histone marks through exchange of histone variants. Dynamic histone variant exchange as in the case of H3.1/H3.3 allows the replacement of, e.g., repressive histone marks in the plant chromatin, possibly playing a crucial role during developmental transitions. H3.3 is expressed throughout the cell cycle and is therefore available to replace marked H3 in dividing and nondividing cells, independent of DNA replication, potentially supporting flexible transitions of transcriptional regimes

repressed genes lose H3.3 enrichment, while active genes show H3.3 enrichment in differentiated cells, compared to undifferentiated ones [22]. Therefore, H3.3 is a likely candidate to be involved in transcriptional remodeling during differentiation.

DNA methylation enrichment particularly marks housekeeping genes, with largely stable expression patterns throughout plant development. Housekeeping genes by default are required to remain unresponsive to developmental or environmental stimuli, to maintain the basic cellular functions during phase transitions or adaptive processes. Enrichment of DNA methylation prevents H2A.Z deposition. H2A.Z is associated with regulation of gene responsiveness during developmental transitions and adaptation to different environmental stress conditions. Consequently, DNA methylation has been suggested to be required for the establishment of constitutive expression patterns, by preventing H2A.Z incorporation over gene bodies [8]. In this scenario, DNA methylation could be required to stabilize transcription rates of housekeeping genes, preventing transcriptional fluctuations. In agreement with this, body methylated genes were found to be more constitutively expressed [7].

The potential role of DNA methylation in regulating or stabilizing transcription rates of genes during developmental transitions is still unclear. Similarly, the function of histone variants in their respective chromatin domains and their dynamics during developmental transitions are only now emerging. Future investigations will show whether or not histone variant replacement simply accompanies transitions or rather enables transitions to occur in plants. The link between histone variants, histone marks, DNA methylation, and associated noncoding RNAs remains to be studied in detail to elucidate the epigenetic basis of developmental transitions.

References

1. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, et al. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. Cell. 2008;133(3):523–36.
2. Cokus SJ, Feng S, Zhang X, Chen Z, Merriman B, Haudenschild CD, et al. Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. Nature. 2008;452(7184):215–9.
3. Stroud H, Do T, Du J, Zhong X, Feng S, Johnson L, et al. Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis*. Nat Struct Mol Biol. 2014;21(1):64–72.
4. Stroud H, Greenberg MV, Feng S, Bernatavichute YV, Jacobsen SE. Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. Cell. 2013;152(1–2):352–64.
5. Tran RK, Henikoff JG, Zilberman D, Ditt RF, Jacobsen SE, Henikoff S. DNA methylation profiling identifies CG methylation clusters in Arabidopsis genes. Curr Biol. 2005;15(2):154–9.
6. Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, et al. Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. PLoS Biol. 2007;5(5):e129.
7. Zhang X, Yazaki J, Sundaresan A, Cokus S, Chan SW, Chen H, et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. Cell. 2006;126(6):1189–201.
8. Coleman-Derr D, Zilberman D. DNA methylation, H2A.Z, and the regulation of constitutive expression. Cold Spring Harb Symp Quant Biol. 2012;77:147–54.

9. Saze H, Kakutani T. Differentiation of epigenetic modifications between transposons and genes. Curr Opin Plant Biol. 2011;14(1):81–7.
10. Zilberman D, Gehring M, Tran RK, Ballinger T, Henikoff S. Genome-wide analysis of Arabidopsis thaliana DNA methylation uncovers an interdependence between methylation and transcription. Nat Genet. 2007;39(1):61–9.
11. Roudier F, Ahmed I, Berard C, Sarazin A, Mary-Huard T, Cortijo S, et al. Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. EMBO J. 2011;30(10):1928–38.
12. Bernstein BE, Mikkelsen TS, Xie X, Kamal M, Huebert DJ, Cuff J, et al. A bivalent chromatin structure marks key developmental genes in embryonic stem cells. Cell. 2006;125(2):315–26.
13. Azuara V, Perry P, Sauer S, Spivakov M, Jorgensen HF, John RM, et al. Chromatin signatures of pluripotent cell lines. Nat Cell Biol. 2006;8(5):532–8.
14. Talbert PB, Ahmad K, Almouzni G, Ausio J, Berger F, Bhalla PL, et al. A unified phylogeny-based nomenclature for histone variants. Epigenetics Chromatin. 2012;5:7.
15. Volle C, Dalal Y. Histone variants: the tricksters of the chromatin world. Curr Opin Genet Dev. 2014;25:8–14.
16. Ingouff M, Berger F. Histone3 variants in plants. Chromosoma. 2010;119(1):27–33.
17. Zilberman D, Coleman-Derr D, Ballinger T, Henikoff S. Histone H2A.Z and DNA methylation are mutually antagonistic chromatin marks. Nature. 2008;456(7218):125–9.
18. Coleman-Derr D, Zilberman D. Deposition of histone variant H2A.Z within gene bodies regulates responsive genes. PLoS Genet. 2012;8(10):e1002988.
19. Kumar SV, Wigge PA. H2A.Z-containing nucleosomes mediate the thermosensory response in Arabidopsis. Cell. 2010;140(1):136–47.
20. Okada T, Endo M, Singh MB, Bhalla PL. Analysis of the histone H3 gene family in Arabidopsis and identification of the male-gamete-specific variant AtMGH3. Plant J. 2005;44(4):557–68.
21. Stroud H, Otero S, Desvoyes B, Ramirez-Parra E, Jacobsen SE, Gutierrez C. Genome-wide analysis of histone H3.1 and H3.3 variants in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A. 2012;109(14):5370–5.
22. Wollmann H, Holec S, Alden K, Clarke ND, Jacques PE, Berger F. Dynamic deposition of histone variant H3.3 accompanies developmental remodeling of the Arabidopsis transcriptome. PLoS Genet. 2012;8(5):e1002658.
23. Ingouff M, Rademacher S, Holec S, Soljic L, Xin N, Readshaw A, et al. Zygotic resetting of the HISTONE 3 variant repertoire participates in epigenetic reprogramming in Arabidopsis. Curr Biol. 2010;20(23):2137–43.
24. Ahmad K, Henikoff S. Histone H3 variants specify modes of chromatin assembly. Proc Natl Acad Sci U S A. 2002;99 (Suppl 4):16477–84.
25. Schwartz BE, Ahmad K. Transcriptional activation triggers deposition and removal of the histone variant H3.3. Genes Dev. 2005;19(7):804–14.
26. Tagami H, Ray-Gallet D, Almouzni G, Nakatani Y. Histone H3.1 and H3.3 complexes mediate nucleosome assembly pathways dependent or independent of DNA synthesis. Cell. 2004;116(1):51–61.
27. Ooi SL, Priess JR, Henikoff S. Histone H3.3 variant dynamics in the germline of *Caenorhabditis elegans*. PLoS Genet. 2006;2(6):e97.
28. Goldberg AD, Banaszynski LA, Noh KM, Lewis PW, Elsaesser SJ, Stadler S, et al. Distinct factors control histone variant H3.3 localization at specific genomic regions. Cell. 2010;140(5):678–91.
29. Ray-Gallet D, Woolfe A, Vassias I, Pellentz C, Lacoste N, Puri A, et al. Dynamics of histone H3 deposition in vivo reveal a nucleosome gap-filling mechanism for H3.3 to maintain chromatin integrity. Mol Cell. 2011;44(6):928–41.
30. Ingouff M, Hamamura Y, Gourgues M, Higashiyama T, Berger F. Distinct dynamics of HISTONE3 variants between the two fertilization products in plants. Curr Biol. 2007;17(12):1032–7.

31. Talbert PB, Masuelli R, Tyagi AP, Comai L, Henikoff S. Centromeric localization and adaptive evolution of an Arabidopsis histone H3 variant. Plant Cell. 2002;14(5):1053–66.
32. To TK, Kim JM. Epigenetic regulation of gene responsiveness in Arabidopsis. Front Plant Sci. 2014;4:548.
33. Aceituno FF, Moseyko N, Rhee SY, Gutierrez RA. The rules of gene expression in plants: organ identity and gene body methylation are key factors for regulation of gene expression in *Arabidopsis thaliana*. BMC Genomics. 2008;9:438.
34. March-Diaz R, Reyes JC. The beauty of being a variant: H2A.Z and the SWR1 complex in plants. Mol Plant. 2009;2(4):565–77.
35. Deal RB, Kandasamy MK, McKinney EC, Meagher RB. 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. 2005;17(10):2633–46.
36. Deal RB, Topp CN, McKinney EC, Meagher RB. Repression of flowering in Arabidopsis requires activation of FLOWERING LOCUS C expression by the histone variant H2A.Z. Plant Cell. 2007;19(1):74–83.
37. March-Diaz R, Garcia-Dominguez M, Florencio FJ, Reyes JC. SEF, a new protein required for flowering repression in Arabidopsis, interacts with PIE1 and ARP6. Plant Physiol. 2007;143(2):893–901.
38. Kumar SV, Lucyshyn D, Jaeger KE, Alos E, Alvey E, Harberd NP, et al. Transcription factor PIF4 controls the thermosensory activation of flowering. Nature. 2012;484(7393):242–5.
39. March-Diaz R, Garcia-Dominguez M, Lozano-Juste J, Leon J, Florencio FJ, Reyes JC. Histone H2A.Z and homologues of components of the SWR1 complex are required to control immunity in Arabidopsis. Plant J. 2008;53(3):475–87.
40. Smith AP, Jain A, Deal RB, Nagarajan VK, Poling MD, Raghothama KG, et al. Histone H2A.Z regulates the expression of several classes of phosphate starvation response genes but not as a transcriptional activator. Plant Physiol. 2010;152(1):217–25.
41. Wollmann H, Berger F. Epigenetic reprogramming during plant reproduction and seed development. Curr Opin Plant Biol. 2012;15(1):63–9.
42. Gehring M, Bubb KL, Henikoff S. Extensive demethylation of repetitive elements during seed development underlies gene imprinting. Science. 2009;324(5933):1447–51.
43. Hsieh TF, Ibarra CA, Silva P, Zemach A, Eshed-Williams L, Fischer RL, et al. Genome-wide demethylation of Arabidopsis endosperm. Science. 2009;324(5933):1451–4.
44. Slotkin RK, Vaughn M, Borges F, Tanurdzic M, Becker JD, Feijo JA, et al. Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell. 2009;136(3):461–72.
45. Zemach A, Kim MY, Silva P, Rodrigues JA, Dotson B, Brooks MD, et al. Local DNA hypomethylation activates genes in rice endosperm. Proc Natl Acad Sci U S A. 2010;107(43):18729–34.
46. Gehring M, Missirian V, Henikoff S. Genomic analysis of parent-of-origin allelic expression in *Arabidopsis thaliana* seeds. PLoS One. 2011;6(8):e23687.
47. Hsieh TF, Shin J, Uzawa R, Silva P, Cohen S, Bauer MJ, et al. Regulation of imprinted gene expression in Arabidopsis endosperm. Proc Natl Acad Sci U S A. 2011;108(5):1755–62.
48. Luo M, Taylor JM, Spriggs A, Zhang H, Wu X, Russell S, et al. A genome-wide survey of imprinted genes in rice seeds reveals imprinting primarily occurs in the endosperm. PLoS Genet. 2011;7(6):e1002125.
49. Wolff P, Weinhofer I, Seguin J, Roszak P, Beisel C, Donoghue MT, et al. High-resolution analysis of parent-of-origin allelic expression in the Arabidopsis Endosperm. PLoS Genet. 2011;7(6):e1002126.
50. Li J, Berger F. Endosperm: food for humankind and fodder for scientific discoveries. New Phytol. 2012;195(2):290–305.
51. Hathaway NA, Bell O, Hodges C, Miller EL, Neel DS, Crabtree GR. Dynamics and memory of heterochromatin in living cells. Cell. 2012;149(7):1447–60.

52. Zheng B, Chen X. Dynamics of histone H3 lysine 27 trimethylation in plant development. Curr Opin Plant Biol. 2011;14(2):123–9.
53. Simon JA, Kingston RE. Occupying chromatin: polycomb mechanisms for getting to genomic targets, stopping transcriptional traffic, and staying put. Mol Cell. 2013;49(5):808–24.
54. Holec S, Berger F. Polycomb group complexes mediate developmental transitions in plants. Plant Physiol. 2012;158(1):35–43.
55. Lafos M, Kroll P, Hohenstatt ML, Thorpe FL, Clarenz O, Schubert D. Dynamic regulation of H3K27 trimethylation during Arabidopsis differentiation. PLoS Genet. 2011;7(4):e1002040.
56. Ingouff M, Haseloff J, Berger F. Polycomb group genes control developmental timing of endosperm. Plant J. 2005;42(5):663–74.
57. Muller K, Bouyer D, Schnittger A, Kermode AR. Evolutionarily conserved histone methylation dynamics during seed life-cycle transitions. PLoS ONE. 2012;7(12):e51532.
58. Bouyer D, Roudier F, Heese M, Andersen ED, Gey D, Nowack MK, et al. Polycomb repressive complex 2 controls the embryo-to-seedling phase transition. PLoS Genet. 2011;7(3):e1002014.
59. Molitor AM, Bu Z, Yu Y, Shen WH. Arabidopsis AL PHD-PRC1 complexes promote seed germination through H3K4me3-to-H3K27me3 chromatin state switch in repression of seed developmental genes. PLoS Genet. 2014;10(1):e1004091.
60. Song J, Angel A, Howard M, Dean C. Vernalization–a cold-induced epigenetic switch. J Cell Sci. 2012;125(Pt 16):3723–31.
61. Sun B, Ito T. Floral stem cells: from dynamic balance towards termination. Biochem Soc Trans. 2010;38(2):613–6.
62. Sun B, Looi LS, Guo S, He Z, Gan ES, Huang J, et al. Timing mechanism dependent on cell division is invoked by polycomb eviction in plant stem cells. Science. 2014;343(6170):1248559.
63. Sun B, Xu Y, Ng KH, Ito T. A timing mechanism for stem cell maintenance and differentiation in the Arabidopsis floral meristem. Genes Dev. 2009;23(15):1791–804.
64. He C, Chen X, Huang H, Xu L. Reprogramming of H3K27me3 is critical for acquisition of pluripotency from cultured Arabidopsis tissues. PLoS Genet. 2012;8(8):e1002911.
65. Okano Y, Aono N, Hiwatashi Y, Murata T, Nishiyama T, Ishikawa T, et al. A polycomb repressive complex 2 gene regulates apogamy and gives evolutionary insights into early land plant evolution. Proc Natl Acad Sci U S A. 2009;106(38):16321–6.
66. Mosquna A, Katz A, Decker EL, Rensing SA, Reski R, Ohad N. Regulation of stem cell maintenance by the polycomb protein FIE has been conserved during land plant evolution. Development. 2009;136(14):2433–44.
67. Lu F, Cui X, Zhang S, Jenuwein T, Cao X. Arabidopsis REF6 is a histone H3 lysine 27 demethylase. Nat Genet. 2011;43(7):715–9.

Chapter 8
Mechanisms of Transposable Element Evolution in Plants and Their Effects on Gene Expression

Lisa M. Smith

An Introduction to Transposable Elements in Plants

Transposable elements (TEs), first described in the 1940s by Barbara McClintock, are DNA elements of variable sequence that can (or were previously able to) move within a genome. TEs account for almost half of the human genome and the majority of the genomes of many economically important plants [1–3], with ~80% of the wheat genome comprising TEs [4]. Flowering plants have highly variable haploid genome sizes, differing by more than three orders of magnitude [5, 6], with variance in genome size largely attributable to the relative prevalence of TEs [3, 7–10]. While most plants have comparably small genomes, those with large genomes may be more restricted in the range of their environmental niches [11]. At opposite extremes of genome size, *Fritillaria* species have some of the largest known genomes due to the expansion of many diverse TE families (up to 127 Gb [12]), while bladderwort has a relatively small genome and very few TEs (80 Mb [13]). Compared to the angiosperms, gymnosperm genomes have relatively uniform numbers of TEs [14].

There are two major categories of TEs; those than move via a "copy-and-paste" mechanism with an RNA intermediate that is reverse transcribed prior to reintegration (class I retrotransposons), and those that use a "cut-and-paste" mechanism with a DNA intermediate and have terminal inverted repeats (class II transposons; Fig. 8.1). Class II TEs also include *helitrons* which are thought to replicate by a rolling circle mechanism [15]. The most common family of TEs in plants is the long terminal repeat (LTR) retrotransposons, with the LTRs defining the transcriptional direction and boundaries of the TE [16–18]. Other class I retrotransposons include autonomous long interspersed elements (LINEs) and nonautonomous short interspersed elements (SINEs) that are derived from transfer RNAs (tRNAs) [19].

L. M. Smith (✉)
Department of Animal and Plant Sciences, University of Sheffield, Alfred Denny Building, Western Bank, Sheffield S10 2TN, UK
e-mail: lisa.m.smith@sheffield.ac.uk

O. Pontes, H. Jin (eds.), *Nuclear Functions in Plant Transcription, Signaling and Development,* DOI 10.1007/978-1-4939-2386-1_8

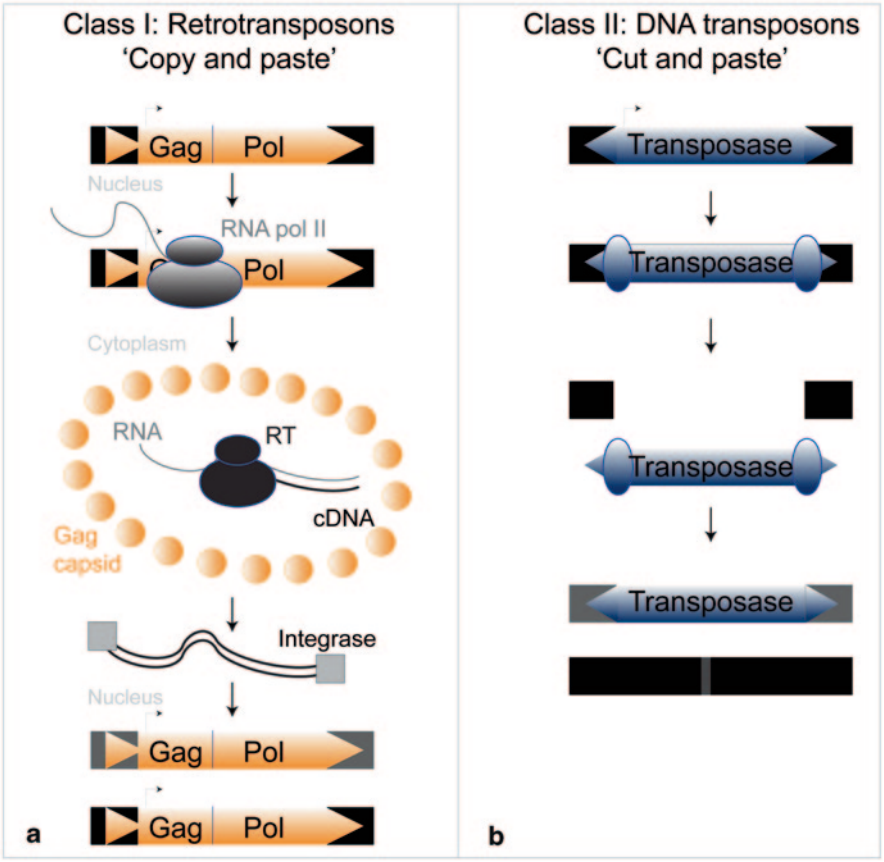

Fig. 8.1 Major classes of transposable elements. **a** Retrotransposons, as illustrated here by the long terminal repeat (*LTR*) retrotransposons, generally encode a single open reading from which a viral capsid-like protein (Gag) is transcribed by RNA polymerase II, along with polyprotein (Pol) components (RNAse H, integrase and reverse transcriptase). In the cytoplasm, the translated proteins associate with the retrotransposon RNA in virus-like particles, allowing reverse transcription and formation of integratable units that are transported back into the nucleus to allow integration into new genomic locations (*gray region*). **b** DNA transposons encode a transposase gene, which, once translated, is transported back to the nucleus where it excises the transposable element and facilitates its integration elsewhere in the genome.

In plants, TEs are most abundant within the pericentromeric regions, with reduced representation in the gene-rich chromosome arms. However, various TE superfamilies and families have different insertion preferences and hence different genomic distributions [20]. Low-copy-number TE families are more likely to be located within euchromatic gene-rich regions, while members of high-copy-number families are often found embedded within other TEs [20].

The TE complement of a genome can be defined as the sum of autonomous TEs (TEs that encode transposases and can facilitate their own transposition), non-

autonomous TEs (TEs that can no longer produce transposase but can transpose using transposases encoded by other TEs), and degenerate relics of TEs (which are no longer able to transpose). Only a small proportion of TEs within a genome are generally autonomous. To be autonomous, a TE must be transcribed, although transcription itself is not sufficient for transposition. Most transcribed TEs will be unable to transpose due to accumulated mutations within their coding sequences [21]. A recent study indicates that 3% of TEs in the model plant *A. thaliana* are transcribed while in the close relative *Arabidopsis lyrata,* which contains approximately three times are many TEs, 8% of TEs show evidence of transcription [22]. This is consistent with an analysis of *Mutator*-like TEs (MULEs) with captured gene fragments in rice, where 5% were found to be transcriptionally active [23], and with a second study that found that 9% of TEs in the maize shoot apical meristem are transcribed [24]. Transcriptional activation of TEs can be tissue specific [25]. In particular, most TEs are up-regulated in the endosperm, pollen vegetative nucleus, and in the central cell, as discussed further below [26–28].

The TE complement in plants can expand and contract rapidly, indicating the importance of TEs in genome evolution. For example, differences in TE content account for much of the difference in genome size between the model plant *A. thaliana* and *A. lyrata* (120 vs. 225 Mb genome size; 5–10 million years divergence [29]). Similarly, maize and its relative *Zea luxurians* diverged 140,000 years ago, but the maize genome is one third smaller (2.9 vs. 4.4 Gb [30]). These rapid changes in genome size may be due to more efficient deletion of TEs in the former of each pair and/or a recent expansion of TEs in the latter, and indicate differences in the control of TEs between closely related plant species [30–33]. Therefore, the control of TE transposition and mechanisms by which TEs are removed from the genome can be considered key factors in the evolution of genomes. TE transcriptional activity and the selection coefficient against TEs are influenced by multiple factors including mating system, stress, heterochromatic modifications, and targeting by epigenetic modifications. The dynamics of TE multiplication, degradation, and control, and the effect of TEs on the protein-coding fraction of plant genomes form the focus of this book chapter. Each facet of TE dynamics should not however be considered in isolation, as there are many interactions: e.g., TE control mechanisms have been linked to TE degradation [34] and repression of nearby genes [22, 35], and TE copy number can determine the control mechanism [36, 37].

Epigenetic Silencing of TEs

Before we contemplate the dynamics of TE births and deaths, let us begin by considering how the plant genome controls its TE complement. In plants, most TEs are epigenetically inactivated or “silenced”; clusters of small interfering RNAs (siRNAs) associate with most TEs and are involved in a “double-lock” mechanism of (si)RNA-directed DNA methylation (RdDM), with a reinforcement loop between the epigenetic marks of DNA methylation, histone methylation, and siRNAs [38–47].

Transposition is thereby controlled via transcriptional gene silencing (TGS). Targets of chromatin modifications such as DNA methylation and histone methylation include both highly repetitive and low-copy-number TEs [38, 39, 48, 49].

The silencing of new TE insertions by siRNAs occurs in multiple phases. Transcripts of young, low-copy-number TEs initially generate 21-nt siRNAs via a posttranscriptional gene silencing (PTGS) pathway; the second strand of the RNA is produced by RNA-DEPENDENT RNA POLYMERASE 6 (RDR6), processed into siRNAs by the RIBONUCLEASE (RNAse) III-like enzymes DICER-LIKE 2 (DCL2) and DCL4, and upon incorporation into a slicing complex with ARGONAUTE 1 (AGO1) or AGO2 these siRNAs target further TE transcripts for cleavage and degradation [37]. Once a critical number of new TE insertions is reached, the double-stranded RNA (dsRNA) levels are sufficient to trigger processing by DCL3 into 24-nt siRNAs that are then used by AGO4 to initiate *de novo* methylation at the TE locus [37]. Over the next few generations, methylation is reinforced and spreads as TGS takes over from PTGS [37]. The siRNAs associated with most TEs are therefore predominantly 24 nt in length.

The progression to TE silencing by TGS can be accelerated when TEs transpose in an inverted repeat formation. Transcription of a partial or full inverted repeat of closely related TEs provides the dsRNA template required by DCL3 for siRNA generation, with the produced siRNAs capable of silencing any family members with close homology. A key example of this is the *trans* silencing of *MuDR* elements by the *Mu killer* (*Muk*) inverted repeat in maize [50]. *Muk* encodes a long RNA hairpin that triggers siRNA generation, resulting in DNA methylation at *MuDR* TEs, which is then normally stably inherited [51]. One *MuDR* insertion has however been identified where DNA methylation is consistently reversed once *Muk* is segregated away, indicating that particular genomic environments can influence the transpositional ability of TEs through active DNA demethylation (see further discussion of DNA methylation and demethylation below [51]).

Under stable TGS, TEs and other repeats are transcribed by the plant-specific RNA polymerase IV (Pol IV, [52–54]), before dsRNA is produced from these noncoding transcripts by RDR2 [48]. As detailed above, the dsRNAs are then processed to 24-nt-long siRNAs by DCL3 [48]. There is, however, partial functional redundancy between the DCL2, DCL3, and DCL4 proteins in plants [48, 55]. Most, if not all, siRNAs in plants are then methylated on the 2' hydroxyl group of the 3' terminal nucleotide by the methyltransferase HUA ENHANCER 1 (HEN1), which protects the siRNAs from uridylation and degradation [56–58]. Once the ds-siRNAs are incorporated into the AGO4 complex, the strand with lower 5' end stability is generally retained as the guide strand and the other strand is degraded [59, 60]. AGO4 complexes associate with a second plant-specific RNA polymerase, Pol V [52, 61]. This AGO4–Pol V complex produces short nonpolyadenylated transcripts at loci with high homology to the siRNA and so guides the deposition and reinforcement of further epigenetic modifications such as asymmetric methylation to siRNA-producing loci [62, 63]. The subcellular localizations of Pol IV and Pol V are distinct, with Pol IV associated with euchromatic nuclear domains, and Pol V found in the ribonucleoprotein processing Cajal bodies [64, 65]. Although the core components

of the RdDM pathway have been discussed here, many further cofactors are known and have recently been reviewed by Law and Jacobsen [66].

Chromatin modifications are an integral component of TGS and include DNA methylation and histone modifications. In plants, methylation of DNA by incorporation of 5-methylcytosine occurs in all sequence contexts: symmetric (CG and CHG di- and trinucleotides, where H is A, T, or G and methylation is present on both DNA strands), and asymmetric cytosine contexts (CHH, where methylation will occur on only one DNA strand). Symmetric methylation is maintained through cell divisions by CHROMOMETHYLTRANSFERASE 3 (CMT3) and METHYLTRANSFERASE 1 (MET1 [39, 67, 68]). The maintenance of CHG methylation by CMT3 is dependent on a second type of chromatin modification, the dimethylation of histone 3 at lysine 9 (H3K9me2) by histone methyltransferase SUVH4/KRYPTONITE (KYP; with SUVH5 and SUVH6 acting redundantly with SUVH4 at some loci [69–72]). Asymmetric CHH methylation through RdDM and *de novo* symmetric methylation are mediated predominantly through DOMAINS-REARRANGED METHYLTRANSFERASE 2 (DRM2 [39]). Patterns of DNA methylation are maintained through the antagonistic actions of DNA methyltransferases and DNA demethylases including REPRESSOR OF SILENCING 1 (ROS1), DEMETER-LIKE 1 (DML1), and DML2 [73, 74].

The majority of TEs have all three types of DNA methylation (CG, CHG, and CHH); however, a quarter contain no methylation and a further 15% have atypical methylation patterns [75]. TEs with absent or atypical methylation are predominantly short, degenerate relics located in the euchromatin, which are depauperate of CG dinucleotides [75]. These methylation patterns indicate evolution of TEs from active and targeted by siRNA-mediated DNA methylation, to inactivated through deletions and deamination of cytosine residues and no longer targeted by siRNA-mediated DNA methylation (mechanisms discussed further below [75]). Although DNA methylation and siRNA-targeting patterns are not perfectly correlated [41, 75], siRNAs are a robust proxy for DNA methylation at TEs [35, 76], with the vast majority of unmethylated TEs also not matching 24-nt siRNAs [75].

RdDM reinforces silencing of TEs during gametogenesis and embryogenesis through TE reactivation and siRNA production in the pollen vegetative nucleus, and in the central cell and endosperm of the developing seed [26–28]. In rice, in contrast to *A. thaliana,* hypomethylation occurs predominantly at a few maternal loci, and although siRNAs are expressed from both parental genomes, their up-regulation is restricted to relatively few strongly expressed loci [74]. Through siRNA transport via an unknown mechanism, TE silencing is then reinforced in an siRNA-mediated manner in the developing embryo.

Given the importance of DNA methylation in the suppression of TEs, it is perhaps unsurprising that TEs have evolved ways to counter RdDM. TnpA, encoded by the maize Suppressor-mutator (Spm) TE, binds to hemi-methylated or unmethylated TEs and actively contributes to DNA demethylation [77].

In addition to DNA methylation, chromatin can be regulated through deposition of histone variants, and methylation, phosphorylation, acetylation, and other covalent modifications on the histones (H2A, H2B, H3 and H4). The most common modifications in plants are H3K9me2 as mentioned above, and methylation at

lysines 4 and 27 of H3 (H3K4me1, H3K4me2, H3K4me3, and H3K27me3 [78]). H3K4 methylation is found predominantly at transcribed chromatin and is absent from TEs [79], while H3K27me3 is a repressive chromatin mark associated with developmental regulation of tissue-specific genes [78, 80]. H3K9me2 is the primary histone modification associated with heterochromatic and euchromatic TEs in plants.

Chromatin is generally classified as euchromatic, regions which are transcriptionally active and less densely packed during interphase, or heterochromatic. There are two types of heterochromatin: constitutive heterochromatin, which occurs at the centromeres and telomeres, and facultative chromatin, which is present in other regions of the genome and is developmentally regulated and reversible. One theory is that the DCL3 pathway, including DRD1 and Pol V, is required to regulate small TEs that are inserted in euchromatic regions and cannot be regulated through heterochromatin formation [63, 81]. Recombination rate is suppressed in heterochromatin, which may help to prevent ectopic recombination between the many repeats in these regions [82, 83]. In *Pol IV/V* and *Pol V* mutants, facultative heterochromatin loses co-localization with the chromocenters, indicating that Pol IV may facilitate assembly of facultative heterochromatin [54, 63]. Additionally, the Pol V pathway may regulate LTRs of retrotransposons within euchromatic and heterochromatic regions to prevent their utilization as promoters, and reduce ectopic expression or antisense transcription of neighboring coding genes [63].

TE Dynamics

TE Origins

The origins of most TEs remain unclear, although their ubiquity in plant and animal genomes indicates that they probably arose very early in the evolution of eukaryotes, or in a pre-eukaryotic ancestor, before the advent of the epigenetic mechanisms that now regulate them [5]. Separate events likely led to the creation of each class of TEs [84]. New families of TEs have probably arisen through derivation from earlier TE families via rearrangements and sequence acquisition from elsewhere in the genome [84].

TE Proliferation

Which TEs Proliferate

The TE complement of a plant genome consists mostly of degenerate TEs that are no longer capable of transposition. The most reliable ways to assess which TEs are competent for transposition are through comparison of TE complements of closely

related species or strains within a species, and through the examination of TE transposition in mutants permissive of TE movement. More often, transcription of TEs is reported as an indicator of potential for TE mobilization; however, increased transcription does not necessarily lead to increased transposition if PTGS is active, or if mutations within the TE have rendered it incapable of mobilization.

Only a small proportion of the TEs in a genome are transcriptionally active under normal conditions, and this varies according to species. In *A. thaliana,* 3 % of TEs are transcribed compared to 8 % in the close relative *A. lyrata* [22]. A study of interspecific F1 hybrids between these two species verified the higher rate of TE transcription in outcrossing *A. lyrata,* and confirmed that this is mostly due to *cis*-regulatory differences between the two species [85]. *Copia* and *LINE* class I TEs, and faster evolving TEs, are most likely to be transcribed [85].

Recent activity of TEs in many plant species has been noted: e.g., *Bs1* in maize [86, 87] and *Tnt1* in tobacco [88]. A number of TEs are known to have been mobilized in the *A. thaliana* lineage, such as *Athila* [89], or are activated in plants defective in epigenetic silencing and RdDM (discussed further below [46, 90–95]). Likewise, rare examples of TEs currently undergoing large of bursts transposition are known. In the best characterized example, the *mPing* nonautonomous TE family is increasing by approximately 40 copies per plant in rice strain EG4 [96]. In a second case, the maize autonomous *TED* TE excises at high frequency during mitotic division in the gametes and transposes to mostly unlinked sites [97].

Comparative studies indicate that over evolutionary timescales TE insertions do not occur at a constant rate but rather occur in short bursts where thousands of new insertions may occur within a short time frame, followed by longer periods of relative inactivity [31, 32]. Modeling of TE dynamics also supports that transposition dynamics vary in a cyclical manner [98, 99]. Taken together, this implies that TE may cause fluctuations of plant genome size over time, where bursts of activity contribute to a larger bloated genome, followed by a longer period of decreased activity during which TE deletion processes dominate and reduce genome size [33].

Defects in epigenetic silencing and RdDM often lead to mobilization of TEs that are otherwise suppressed under normal plant growth conditions. Loss of DNA methylation at CG and non-CG sites can lead to increased TE transcription and transposition [40, 47, 100]. For example, *ago4, rdr2,* and *dcl3* mutants in *A. thaliana* lose siRNAs, DNA methylation, and H3K9me2 at repeat loci that are associated with formation of facultative heterochromatin, resulting in increased transcript levels of TEs such as AtSN1 [48, 101, 102]. Similarly, DNA methylation and H3K9me2 are lost in mutants of the DNA helicase chromatin remodeler, DECREASE IN DNA METHYLATION 1 (DDM1), resulting in TE mobilization [40, 45, 46]. Half of TEs tested in the *ddm1* mutant are strongly up-regulated at a transcriptional level, with younger TEs overrepresented and TEs disrupted by nested insertions underrepresented [40]. Likewise, mutations in *CMT3* or *MET1* result in increased transcription of CACTA TEs; however, transposition only increases in the double mutant, with the exception of one element, *EVADE,* which mobilizes in a *met1* epigenetic recombinant inbred line [47, 103].

While the effects of transcriptional up-regulation are hard to predict, so too is the direction of transcriptional misregulation; in maize, the loss of the RDR2 homologue, MEDIATOR OF PARAMUTATION 1 (MOP1), results in widespread changes in TE and gene expression [24]. Many class II TEs are up-regulated; however, the majority (about two thirds) of differentially regulated class I TEs are down-regulated [24]. A similar proportion of differentially regulated genes are also down-regulated [24]. Together these results indicate the difficulty of predicting the effects of a loss of RdDM on TE regulation, and the complexity of TE regulation in plants due to differences between TE classes and families, and plant species.

Where Do They Go: Insertion Preferences

As for most TE characteristics, TEs exhibit (super)family-level variation in terms of insertional preference. Euchromatic TEs are located preferentially close to genes, although methylated 5' TEs are underrepresented, possibly due to methylated TEs more strongly suppressing proximal gene expression leading to stronger selection against them [75] or due to insertion site preferences. The *mPing* family in rice allows a unique view into insertion site preferences, as the newly inserted copies of *mPing* will not have undergone significant selection, which can otherwise confound the analysis of insertion site preferences. A comprehensive study of nascent *mPing* insertions in rice shows a strong preference for promoter regions of genes and against exon insertions [104]. This may be due to preferential insertion in DNA at boundaries between transcriptionally active and inactive regions.

There is an increased level of genic polymorphism in TE proximal genes in *A. thaliana* [34], suggesting that TEs may be more likely to integrate into more dynamic genomic regions; or some genomic regions may be more prone to change both within coding genes and nearby TEs [34]. TE insertional polymorphisms have been noted to favor high single-nucleotide polymorphism (SNP) regions in rice, supporting the former hypothesis [105]. Likewise in humans and chimpanzees, TEs were found to insert close to genes with more diverged expression levels due to TE insertion site preference rather than any effect of the TEs on proximal genes [106].

Families of TEs are more likely to be mutagenic if they have an insertion bias towards euchromatic regions. In general, class I TEs tend to be distributed farther from genes than are class II TEs [35]. Although some LTR retrotransposons show no insertional preference [107], TE families with a preference towards euchromatic regions include *m*iniature *i*nverted repea*t* *e*lements (MITEs), SINEs, helitrons, CACTAs, and MULEs [19, 96, 108–111]. Likewise, the Ds/Spm family preferentially inserts in GC-rich regions around translational initiation sites [112], with a similar insertion pattern noted for the *BraSto* MITE family in *Brassica* species [113]. With higher mutagenic potential, the *Tos17* and *Ac/Ds* TE families in rice show an insertion preference towards exons [114–116]. In contrast, in *A. lyrata,* the high-copy-number *Tal1* COPIA element inserts preferentially into the centromeres [117]. Therefore, TEs have a wide range of insertion site preferences, and these specificities will affect the dynamics of new copies of the TE (e.g., through selec-

tion pressure and likelihood of methylation) and will affect the ways in which TEs interact with protein-coding genes.

TE Deletions

Measures that counteract TE proliferation include inactivation and removal. SNPs and deletions of greater than 3 bp account for the majority of variation between TE family members [34, 118], with both modifications potentially inactivating TEs, and deletions actively reducing the fraction of the genome composed of TEs. Some studies suggest that small deletions in TEs and intergenic regions are the main mechanisms by which genomes undergo cycles of shrinkage [29], while other studies indicate that recombination within chromosomes is the major driving force, at least for TE removal from plant genomes [20, 119, 120]. This discrepancy may indicate that the size of the plant genome and TE complement, among other factors, influence the relative contribution of small DNA deletions vs. recombination in TE removal.

Ectopic recombination may not be a driving force in the reduction of TE numbers in all genomes, as TE density and recombination rate are not correlated in self-compatible *A. thaliana* [121]. However, *A. thaliana* is relatively TE-poor, with TEs accounting for only ~21 % of the genome [29, 75, 122], and ectopic recombination is lower in homozygous genomes [123]. Therefore, TE dynamics in *A. thaliana* may differ from those in TE-rich genomes. Comparison of *A. thaliana* to its close relative *A. lyrata,* which contains approximately three times as many TEs [22], has shown that deletion of TEs in *A. thaliana* is likely an ongoing, active process. In accordance with this hypothesis, recent genome resequencing efforts have highlighted that intraspecific polymorphisms and deletions in *A. thaliana* are disproportionately located within TEs and, to a lesser extent, intergenic regions [124–126]. With current resequencing efforts in many plant species comes the ability to compare highly similar plant genomes and accurately characterize deletions and polymorphisms that affect repetitive sequences such as TEs. Such comparisons of genome evolution from different species with a range of TE complements, genome sizes, and differing life-cycle traits will allow a greater understanding of the mechanisms that drive TE deletions.

Deletions can occur through two types of ectopic recombination: illegitimate recombination (IR), which utilizes only short regions of homology and does not require RecA, and unequal homologous recombination (UHR) between direct repeats with high homology [90, 118, 119, 127–129]. In the case of LTR retrotransposons, the deletion of one LTR and the rest of the TE sequences between the LTRs of a retrotransposon via UHR leaves a single LTR called a solo-LTR (Fig. 8.2ai [9]). The rate of TE removal and solo-LTR formation differs between genomic regions, species, individuals, and chromosomal regions, as does the rate of smaller insertions and deletions that may inactivate a TE [16, 130]. As an example, the average half-life of an LTR TE in rice is approximately two million years [131]. If UHR occurs between different LTR-containing TEs, this can remove more than one TE,

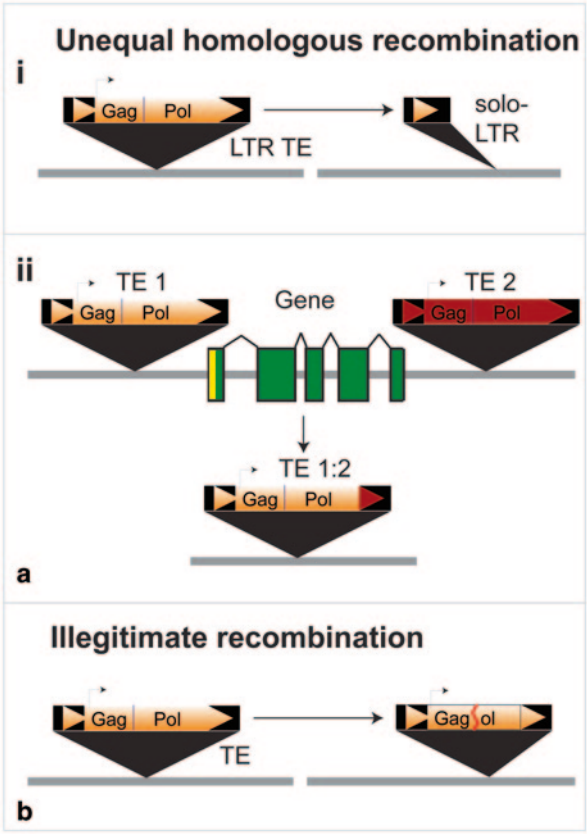

Fig. 8.2 Ectopic recombination of transposable elements. **a** Unequal homologous recombination occurs between regions of extended high homology, such as the long terminal repeats (*LTRs*) flanking many class I TEs. (i) Recombination between *LTRs* of a single transposon can lead to solo-LTR formation, while (ii) recombination between *LTRs* different TEs can also remove intervening sequences or genes. **b** Illegitimate recombination requires only short regions of homology and can lead to small deletions within TEs. *TEs* transposable elements

and any intervening protein-coding genes, leading to faster genome contraction, but with higher potential for deleterious deletions (Fig. 8.2aii [9]). Given that deletion rates are correlated with meiotic crossover events, it is unsurprising that LTR TEs in euchromatic regions are more likely to be truncated than full length and tend to be shorter than LTRs residing in heterochromatic regions [120, 132]. IR tends to remove shorter genome fragments compared to UHR but can also truncate TEs (Fig. 8.2b [133]).

Rearrangements

TEs can initiate more complex DNA rearrangements through recombination than those rearrangements already discussed above. Closely related directly oriented TEs can initiate tandem direct duplication of one of the TEs and its flanking regions. This type of rearrangement occurs when the 3' end of one TE and 5' end of a second TE are targeted by transposase, and the cleaved ends then insert into a flanking region on the sister chromatid [134]. Such transpositions involving two linked,

related TEs can also cause deletions and inversions. Some TEs, such as the class II Ac/Ds TEs in maize, preferentially insert into genes, therefore are more likely to facilitate genome rearrangements that result in gene fusions and novel expression patterns.

It has been proposed that microRNAs (miRNAs) may have evolved from TEs through rearrangements and deletions. miRNAs are short RNAs that are related to siRNAs, but which regulate protein-coding genes though a series of proteins closely related to those in the RdDM pathway. MiRNAs may have evolved from *MITEs* as a sub-functionalization of the siRNA-generating pathway that controls TEs [135]. Indeed, a number of miRNAs in *A. thaliana* and rice show homology to *MITEs,* with TE-derived miRNAs in rice potentially regulating thousands of genes [135–137]. Thus, TE rearrangements can be utilized by plants to produce functional variation and new regulatory networks via miRNAs.

Horizontal Gene Transfer

Most cases of horizontal gene transfer in eukaryotes involve TEs, with more than 200 cases described [138], including some in the angiosperms [139–142]. Horizontal gene transfer can occur via rare wide crosses or through TE transfer by RNA and DNA viruses with broad host ranges; however, in the latter case, viruses will only infrequently cross into the gametes and produce heritable changes to the genome. For instance, RNA viruses have been proposed as a source of LTR retrotransposons in maize [84]. Horizontal gene transfer may be more likely to occur in plant species that are pollinated by insects and other animals; however, transfer could also occur through plant–microbe interactions [14]. Artificial horizontal gene transfer experiments introducing a *Nicotiana tabacum* LTR retrotransposon into *A. thaliana* demonstrate that naïve TE families are silenced in a copy-number-dependent mechanism [36], similar to that of recently mobilized elements [37].

Conditions Affecting Rates of TE Proliferation, Retention, and Deletion

While insertions and deletions affect TE complements within individuals, evolutionary processes at the level of populations are important for determining the fate of TEs within species. The evolution of TE complements over large timescales will be affected by a number of factors including: selection pressure against the TEs as individuals and families (as determined by effective population size and the selection coefficient); mating system of the host species; and transposition rate (as determined by the autonomy of the element and its propensity for transposition [143]). Selection pressures are closely linked to the genomic location of a TE, whether it is heterochromatic or euchromatic, and its proximity to and effect on protein-coding genes [35, 120]. Selected factors affecting TE dynamics are discussed in more detail below.

Mating System

Whether a plant reproduces through outcrossing or self-fertilization should have a profound effect on the population dynamics of its TE complement; however, the direction of the effect remains unclear. Modeling of TE dynamics has suggested that TEs should be lost from the genomes of selfing species and that TE invasions should be prevented (with the exception of invasions initiated via horizontal gene transfer [144]). Self-compatible species may have higher selection coefficients on TEs due to their higher rates of homozygosity, leading over many generations to lower TE abundance through selection against TEs of minor disadvantageous effect [145, 146]. However, from studies of selfing model species with compact genomes like *A. thaliana,* it remains unclear whether selection on TEs and DNA removal rates are higher, or whether TE activity is reduced [29, 147]. Recent reports indicate that, at least for *A. thaliana,* DNA removal is a major contributing factor as deletions are overrepresented in TEs [29, 34].

Self-repression of TEs may be favored in self-compatible species that have a high level of genomic homozygosity as deleterious effects of TEs are expected to be more pronounced [148]. Evidence for self-repression of TEs comes from rice, where one study found that nonautonomous *Stowaway* MITEs can harbor transposition-enhancing elements while the autonomous *Osmar* TEs that facilitate mobilization of *Stowaway* contain motifs repressing their own transposition [149].

TEs of outcrossing *A. lyrata* are on average older than those found in the genome of self-fertilizing *A. thaliana* [150]. While the average TE age in *A. lyrata* is exponential, with a decreasing frequency of older TEs, in *A. thaliana* there is a paucity of TEs younger than 0.5 million years, congruous with the evolution of selfing and subsequent TE repression in this species [150].

Other studies have however indicated that selection pressure against TEs is lower in selfing *Arabidopsis* and tomato species than those that outcross, demonstrating the difficulty of predicting and measuring effects of the mating system on TE mobility [151–153]. Effective population size is reduced in selfing species, which should reduce selection efficiency and slow the removal of TEs from selfing genomes [150], leading to the hypothesis that TEs should increase in population frequency when a plant species becomes self-compatible [151, 154]. Therefore, as selfing increases the rate of homozygous TEs, ectopic recombination (which often results in deleterious deletions and removes TEs from the genome) should decrease [155]. In summary, whether recombination and other effects should result in an increase or decrease in TEs in selfing species remains disputed, with conflicting evidence in both the theoretical and empirical fields [156, 157].

Stress Responses

Biotic and abiotic stresses can affect TE proliferation rates. For example, higher TE transposition and/or transcription rates have been noted after heat shock (e.g., [95, 158]), cold stress (e.g., [159–161]), pathogen/microbe infection (e.g., [162,

163]), high-ultraviolet (UV) exposure (e.g., [164]), and wounding (e.g., [36, 159]). Regeneration of plants from callus tissue can also induce the transposition of TEs, as demonstrated by the LORE1a LTR retrotransposon in *Lotus japonicus* [107], while in rice at least 13 TE families are transpositionally activated in *in vitro* culture [129]. In some cases, such as the tobacco *Tto1* class I TE under pathogen stress, activation of the TE may result from convergent evolution within the TE promoter resulting in transcription factor-binding sites that are utilized under particular conditions [18, 163]. Evolution of stress-inducible promoters in TEs could be advantageous, as this would restrict transposition to circumstances when rapid TE-induced variability in the progeny could allow faster adaptation to a changing environment [18, 165]. This has been substantiated by one study where transcription of a TE was induced upon heat stress in mutants impaired in siRNA biogenesis [95]. Transposition events detected in the subsequent generation conferred heat responsiveness to proximal genes [95].

Hybridization and Polyploidy

Intra- and interspecific hybridization occurs frequently in plants, and is often followed by genome doubling to form allopolyploids or autopolyploids due to the increased production of unreduced gametes in hybrids [14]. TE bursts may occur upon the genomic shock of hybridization or polyploid formation (e.g., [166–172]), but are not predictable, and TE transcriptional up-regulation is not necessarily followed by a TE burst [169, 173]. Indeed, expression of an *A. arenosa Athila* TE can increase upon hybridization with *A. thaliana* [174], but even for the same TE in the same cross from the same laboratory, there are conflicting reports [175].

A number of factors may influence TEs upon hybridization or polyploidy formation. Evolutionary dynamics vary among TE families within a species, by demography and also by mating system [152]. Given that siRNAs that silence TEs in the embryo are maternally derived, the direction of a cross, maternal dosage, and divergence between the TE complements in the two species may be key factors in determining which TEs may become transpositionally competent in a hybrid [27]. Also, in the case of polyploidization, the doubling of the genome causes functional redundancy for all genes such that selection against TE insertions in and near genes will be reduced [171]. Simultaneously, duplication of all TEs may reduce silencing efficacy [22].

There are examples of expansion of specific TE families driving genome size increases in hybrids [31, 32]. In a synthetic allopolyploid of *A. thaliana* and *A. arenosa* the Sunfish TE displays increased transposition [176], while the MITE *mPing,* class I *Tos17,* and class II *hAT* are all more mobile in a hybrid of rice with a wild relative [166, 167]. In wheat allotetraploid hybrids, transcription of a family of LTR retrotransposons has been noted to increase at an early stage of allopolyploid formation [177, 178]. Although these are experimental hybrids, natural hybrids in the sunflower family demonstrate that observations of TE bursts are more widely applicable. Three natural *Helianthus* hybrids that have arisen 60–200,000 years ago

have genomes approximately 50% larger than either parent, largely due to the expansion of one family of class I elements [172].

In the absence of a TE burst, TEs may none the less have a major effect during polyploidy formation through facilitating major structural changes and genome reorganizations [179]. In the allotetraploid *Nicotiana tabacum,* genome compaction through loss of TEs and other repeats, in particular from the paternal genome, has been observed [180]. This is in contrast to the two progenitor species; maternal parent *N. sylvestris* has undergone a recent TE burst or homogenization, while paternal parent *N. tomentosiformis* appears to have a stable genome size [180].

Population Size

Reduction in effective population size should reduce the efficacy of selection against TEs and increase the effects of drift, leading to a higher rate of TE accumulation and an increase in genome size [181, 182]. However, a recent survey of more than 200 plant genomes found no correlation between genome size and effective population size [183]. Especially in polyploidy formation, the population bottleneck is extreme and plants are reproductively isolated; therefore, the chances of TEs families becoming fixed or extinct should be greatly increased [14, 31].

Larger TE family population sizes may allow more rapid genomic structural changes and genome shrinkage as recombination between TEs is more likely [14, 184]. Thus, cycles of TE population amplification and reduction probably contribute immensely to genome plasticity and adaptation.

TE Effects on Gene Expression

TEs can affect gene expression through a number of mechanisms such as introduction of regulatory elements and induction of methylation in promoter regions, new insertions contributing to coding regions, duplication of genes, and shuffling of genomic regions. Major routes by which TEs affect gene expression are considered below.

Insertional Inactivation

Inactivation of a gene though insertion of a TE is perhaps the most obvious way in which a TE can affect gene expression (Fig. 8.3a). Possibly the most famous example is that of an *Ac/Ds* TE insertion in the *rugosa* gene of peas, which inactivates the locus, resulting in Mendel's wrinkled peas [185]. Likewise, *mutator* insertions at the *kn1* locus in maize cause leaf developmental defects [186], while a Stowaway MITE insertion in a gene encoding a flavonoid hydroxylase in the anthocyanin biosynthesis pathway results in pink skinned potatoes [187].

According to research in *A. thaliana,* insertion of methylated TEs into genes is relatively rare; however, poorly methylated TEs are found in genes (introns, exons,

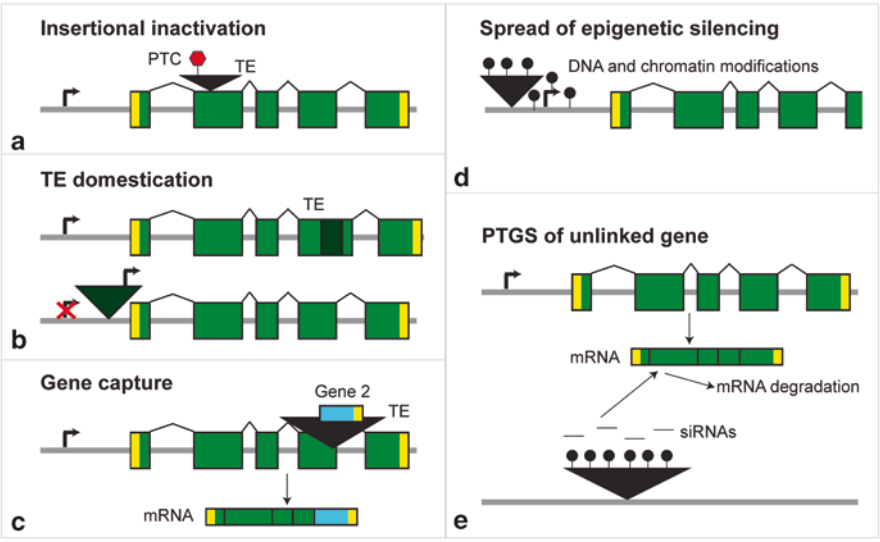

Fig. 8.3 Selected examples of transposable element effects on protein-coding genes. **a** Mobilization of a TE into a protein-coding gene can lead to insertional inactivation of the gene through introduction of a premature termination codon (*PTC*) into the transcribed sequence. **b** Some TEs can be domesticated by contributing encoded regions of the protein, or through introduction of new regulatory sequences. **c** Through integration of coding genes into TEs, gene capture can occur, leading to potential gene shuffling and formation of chimeric mRNAs. **d** Spread of epigenetic modifications from silenced TEs into neighboring genes can cause down-regulation of the protein-coding gene. **e** More complicated rearrangements have lead to the generation of miRNAs from TEs, which can regulate protein-coding genes through post-transcriptional gene silencing (*PTGS*)

and untranslated regions (UTRs)) at approximately the same frequency as if insertion sites were unbiased and random [35]. A number of factors may contribute to this distribution: Poorly or nonmethylated TEs are generally much shorter [75]; therefore, when inserted in introns and UTRs may have less impact on gene expression; or these TEs may reflect TE domestication events (see below); or alternatively, some of these TEs may be actively demethylated to retain gene function [35, 73].

Prior to mobilization, transposition of class II TEs by a "cut-and-paste" mechanism requires that the TE is first removed from its original location. Excision of class II TEs typically leaves behind a footprint of an imperfect target site duplication, resulting in sequence variation that can affect the function of the host gene [188].

Gene Capture and TE Domestication

TEs can contribute to the coding sequences or regulatory elements of genes, in which case they are referred to as domesticated (Fig. 8.3b). Approximately 8–10% of *A. thaliana* and rice genes harbor TE fragments, with *En/Spm*-like and *Copia*-like TEs and kinase genes overrepresented in *A. thaliana* [189, 190]. Some TE-gene chimeras retain DNA-binding ability in the form of transcription factors [19]. Given

the specificity of transposase for binding sites within TEs, this means that domesticated TE transcription factors could be targeted to transposed TEs in such a way that both the DNA sequence and the encoded DNA-binding protein are effectively "co-domesticated," forming new regulatory networks [19].

Perhaps the best-studied example of TE domestication in plants is that of *FLOWERING WAGENINGEN* (*FWA*), a homeobox transcription factor where a SINE element contributes the promoter and the first two exons of the gene. When RdDM is perturbed in *ddm1, drm1 drm2, dcl3, nrpd1a* (*polIV*), *ago4,* or *rdr2* mutants, silencing of FWA in vegetative tissues is released, leading to late flowering [40, 48, 191]. TE domestication has also occurred in the *MUSTANG* family in *A. thaliana* and rice with domesticated TEs termed *MULEs* [192, 193], *hAT* domesticated in the *DAYSLEEPER* transcription factor [194], and *mutator,* which has contributed to the *FAR1* and *FHY3* genes involved in light response and therefore adaptation to different light environments in *A. thaliana* [195].

In a further case of domestication, a *COPIA* element has inserted in the first intron of the disease resistance gene *RPP7* [196]. High H3K9me2 at the TE, deposited by the SUVH histone methyltransferases, is then required to generate functional transcripts from the *RPP7* gene. Upon infection, H3K9me2 levels decrease, and alternative splicing and polyadenylation results in a decrease in the proportion of full-length transcripts encoding RPP7. Thus, the COPIA TE regulates the alternative polyadenylation of *RPP7,* leading to regulation of RPP7 through balancing of coding and noncoding isoforms of the mRNA [196]. Similarly, in *Nicotiana glutinosa,* a *MITE* has been domesticated to provide an alternatively spliced exon in the *N* gene, the inclusion of which is required for tobacco mosaic virus resistance [197].

TEs and TE-derived sequences can contribute new regulatory sequences upstream of coding regions, leading to novel regulatory networks [177]. For example, TE-derived sequences provide stress-inducible promoters/enhancers to nearby protein-coding genes [104, 162, 198], such as an LTR retroelement that provides a cold-responsive promoter in blood oranges [199]. Further examples arise from TE insertions that cause up-regulation of genes that enhance aluminum tolerance in barley, wheat, and sorghum [200–202].

Rearrangements within proximal domesticated TEs can lead to different regulatory effects. One example of TE domestication where rearrangement affects gene expression is the *Tam3* insertion in the promoter of the *niv* gene of *Antirrhinum*. Depending on the state of the insertion, the *niv* gene is either down-regulated or has a new tissue specificity [203]. A second case of TE domestication within a promoter has been described in maize where insertions and rearrangements of a *Mutator* TE in *Adh1* affect the expression pattern [204, 205]. For an in-depth review of TE domestication, see [206].

In cases where a TE is not fixed within a species, TE domestication can contribute to functional phenotypic variation. For instance, *FLOWERING LOCUS C* (*FLC*) demonstrates natural variation through TE effects on the gene; in some strains or accessions, a MULE in the first intron is silenced by RdDM, down-regulating the gene and leading to early flowering [207]. In a second example, an *MITE* insertion close to a cytochrome P450-like gene in maize affects flowering time, either directly or through the *trans* effects of an siRNA [208].

Alternatively, a gene fragment can be "captured" by a TE, leading to its mobility and multiplication. This can result in shuffling of protein domains, the formation of new chimeric genes, and neofunctionalization (Fig. 8.3c). Gene capture has been documented for a group of *MULE* TEs (called *Pack-MULEs*) in rice, *helitrons* in maize, and *CACTA* TEs in soybean and sorghum [23, 209–214]. In rice, only about 1% of *Pack-MULEs* are translated, with those containing chimeric genes more likely to be expressed, and thereby contribute novel coding capacity [213]. In grasses, a significant proportion of miRNA genes have probably been relocated through TE capture [215]. For a full list of angiosperm TE domestication and gene capture events, in addition to gene disruptions, see [216].

Effects of TEs on flanking genes are partly family specific. While many TE insertions, particularly those into exons, are likely to be detrimental to the fitness of the host, the *mPings* of rice, which preferentially insert in the promoters of genes, usually have either no effect or a positive effect on the transcription of adjacent genes [104]. Where they have a positive effect, they can confer stress-inducible expression on the adjacent genes, thereby contributing to new gene-regulatory networks [104]. Similarly, the *BraSto* elements in *Brassica* species may provide new transcription factor-binding sites, thereby regulating gene expression, particularly under stress [113]. Through high insertion rates in heterochromatic regions, TEs may be also be functionally domesticated to provide centromeres (e.g., in potatoes [217, 218]) or matrix attachment regions (e.g., in rice and sorghum [219]). These cases demonstrate that TEs contribute to protein-coding genes in plants and confer adaptive fitness benefits [193].

TE Effects on Nearby Genes

TEs have long been regarded as genomic parasites due to the detrimental effects of insertional inactivation of genes and ectopic recombination of DNA [190]. Even without direct insertional inactivation, TEs can affect the expression level of proximal genes through a number of mechanisms. Multiple studies have demonstrated that TEs on average suppress the expression of proximal coding genes [22, 35, 40, 220]. In *Arabidopsis* species, the transcription level of genes is negatively correlated with the number of siRNA-targeted, methylated TEs in the vicinity, and positively correlated with the distance to the nearest TE [34, 35]; however, the distance threshold for TE proximity tolerance varies between species and with siRNA targeting [22]. Gene polymorphism, TE presence, and TE variance all act additively on the variability of orthologous gene expression [22, 34]. A similar pattern of lower expression of TE-proximal genes and higher variation in orthologous gene expression has been observed for MITEs in rice [221].

Epigenetic silencing can spread, leading to transcriptional repression, as is discussed in more detail in the following section [75]. Methylated TEs are correlated with greater suppression of proximal genes in *A. thaliana* than are unmethylated TEs, regardless of insertion upstream or downstream of the coding region [35]. Purifying selection is only significant for methylated TEs proximal to genes [35].

Longer TEs (which are more likely to be autonomous) may be under stronger purifying selection because they are more frequently methylated, rather than because they facilitate ectopic recombination [35, 75].

An illustrative example of domestication of TE repression of a neighboring gene has recently emerged. As part of the induced immune response in *A. thaliana,* a number of components of RdDM are temporarily suppressed and ROS1-dependent DNA demethylation is triggered, resulting in up-regulation of TEs and a disease resistance gene *Resistance Methylated Gene 1* (*RMG1*: [222]). Derepression of *RMG1* probably occurs due to two AtREP helitron TEs in the promoter region, and thus the induced immune response indirectly leads to TE-mediated induction of a disease resistance gene [222].

Production of antisense transcripts can result in a similar outcome of proximal gene suppression through post-transcriptional repression of neighboring genes [177]. Indeed, antisense transcripts and methylation can combine in the differential regulation of TEs and nearby genes. In rice, tissue- and species-specific expression patterns of genes adjacent to *Dasheng* LTR retrotransposons result from epiallelic variation [223]. When *Dasheng* TEs are methylated in a given tissue or rice strain, the adjacent genes are expressed, while a lack of methylation in other tissues and strains results in antisense transcription and post-transcriptional repression of flanking genes [223].

Spread of Epigenetic Modifications

One way in which TEs can affect the expression of proximal genes is through spread of epigenetic modifications from the TE into nearby coding or regulatory regions (Fig. 8.3d). Some predominantly heterochromatic TEs that are not targeted by siRNAs are nonetheless methylated [75]. In these cases, methylation is proposed to spread into the nontargeted TEs from proximal siRNA-targeted TEs [75]. In the TE-rich heterochromatin of *A. thaliana,* methylation spreads about 500 bp from methylated TEs into neighboring untargeted TEs, while methylation in euchromatic TEs spreads on average 200 bp beyond the siRNA-targeted region [75]. Given that siRNA targeting and TE methylation are highly correlated [35, 75, 76], this is consistent with the siRNA+/− effect on proximal gene suppression dissipating by 400 bp [22]; however, decreased gene expression due to a proximal TE (regardless of siRNA targeting) is evident for about 2.5 kb [22]. This extended zone of influence indicates that chromatin properties other than DNA methylation may influence proximal gene expression [22]. These chromatin properties could include histone modifications or the compactness of the chromatin; however, species differ in their tolerance of TEs in the proximity of genes [22]. Regardless of the level of TE tolerance, TEs that are siRNA targeted and methylated are, on average, located further away from expressed genes than TEs that are neither strongly methylated nor associated with siRNAs [35, 75]. The silencing of TEs via RdDM amplifies their deleterious effect on the expression of nearby genes, indicating an evolutionary conflict for the host between inactivating TEs and maintaining gene expression [35].

Even within a species, the spread of chromatin modifications may be dependent on TE family. A study of maize found that while DNA and histone methylation did not spread more than 200 bp beyond most TEs, a third of TE families initiated chromatin modifications that spread for 0.8 kb or more, with distance correlated with family [224]. TE families that initiate spreading chromatin modifications tend to be younger, of higher copy number, and longer in length than average [224].

To counteract and limit the spread of DNA methylation, in *A. thaliana* and presumably also other plant species, the H3K9 demethylase INCREASE IN BONSAI METHYLATION 1 (IBM1) acts in the exclusion of H3K9me2 and DNA methylation from genic regions [225–227]. The *IBM1* gene also provides a model for how plants can overcome the effects of epigenetic spread from an intragenic TE; an RNA-binding protein called IBM2 binds to CHG methylation caused by an intronic TE at *IMB1,* and facilitates generation of full-length mRNAs despite the intragenic heterochromatin [228].

TE Effects on Unlinked Genes

TEs can also affect gene expression through silencing of unlinked genes (Fig. 8.3e). During the initial phases of epigenetic TE silencing, 21-nt- and sometimes 22-nt-long siRNAs are produced. If these siRNAs are incorporated into AGO1 rather than AGO4, and if the siRNAs have sufficient sequence complementarity to the transcripts of other genes, this may result in *trans* regulation of off-target genes. This has been demonstrated in the case of *siRNA854,* which is derived from an *Athila* TE in *A. thaliana*. The presence of *siRNA854* leads to repression of *oligouridylate-binding protein 1b* (*UBP1b*), which encodes an RNA-binding protein involved in stress granule formation [229]. Hence, derepression of the *Athila* TE can result in down-regulation of UBP1b, and induce a stress-sensitive phenotype [229]. Additionally, two siRNAs from *Stowaway* MITEs in rice positively regulate abiotic stress tolerance and ABA signaling [230].

In a third example of *trans* effects on gene expression, *CACTA* TEs in rice encode an miRNA [231]; miR820 targets an mRNA encoding a component of the RdDM pathway, *OsDRM2,* leading to down-regulation of RdDM and up-regulation of TE transcripts [231]. Interestingly, the miRNA and its target have coevolved in terms of sequence, retaining targeting in a range of *Oryza* species and demonstrating regulation of host–TE interactions via the TE partner [231]. The *trans* regulation of genes in plants by TEs may, however, be limited due to spatial separation of AGO proteins involved in the epigenetic control of TEs and AGO proteins that regulate protein-coding gene expression [232].

Methylation and TE Evolution

Given that most TEs are targets of RdDM and are highly methylated [38–47, 81, 233–235], the increased frequency of intra- and interspecific SNPs in TEs may

indicate a higher rate of mutations caused by methylcytosine deamination [105, 124, 125, 234]. Recent methylation analyses of TEs support the hypothesis that high levels of TE methylation boost the mutation rate of TEs [75], with C:G to A:T mutation rates higher than expected by chance and highest at sites targeted by CHG methylation [124, 236]. Ahmed et al. also found that shorter, degenerate relics of TEs have lower CG levels than expected, which is proposed to lead to incomplete perpetuation of methylation at CHG and CHH sites, and therefore lower methylation rates [75].

Methylation patterns may also influence the efficacy of selection on TEs. In *A. thaliana,* TEs are distributed such that older methylated TEs are on average further away from coding genes, indicating that selection strength is likely influenced by methylation and gene proximity [35]. Longer TEs are more often methylated and are also under stronger purifying selection [35, 127].

As a counterbalance to methylation, TE-encoded genes may in some instances prevent DNA methylation or actively demethylate TE sequences [77, 237]. In addition, some TEs proximal to genes are actively demethylated by *A. thaliana,* presumably to relieve strongly deleterious gene suppression [41, 225–227].

TE Evolution through Silencing and Deletions

Under a model detailed by Hollister and Gaut for evolution of TEs of neutral or detrimental effect within euchromatic regions [35], new insertions occur randomly. Mobilization is followed by epigenetic silencing, as siRNAs targeting the original TE insertion now also silence the new copy in *trans*. As the TEs accumulate polymorphisms (accelerated by deamination of methylated cytosines), the siRNAs generated from the original and offspring TEs also diverge in sequence. TEs associated with diverged siRNAs are expressed at lower levels and more likely to be methylated than those targeted only by shared siRNAs [22, 41], so the efficacy of TE silencing increases. The expression level of proximal genes is further reduced and thus selection pressure against the TEs may increase [34]. Purifying selection removes TEs affecting proximal gene expression resulting in a paucity of methylated TEs close to genes [35]. During the slow selection again gene-proximal TEs (most TEs have low selection coefficients [182]), epigenetic silencing helps to inactivate the TEs and prevent further transposition events. A rapid burst of proliferation resulting in many offspring TEs will delay the establishment of unique siRNAs against each TE, with a corresponding impediment to increasing the efficacy of silencing. With decreased epigenetic control of the TEs, a feed-forward process of TE proliferation could result.

siRNA targeting may also facilitate a second method of TE inactivation—through preferential deletion of TE regions targeted by siRNAs [34]. This deletion process could actively be promoted by the siRNAs and other epigenetic marks in a mechanism analogous to the siRNA-guided removal of internal eliminated sequences including TEs in *Tetrahymena* [238, 239]. In support of this hypothesis, small dele-

tions within TEs occur more frequently than ectopic recombination events at LTRs [119, 121]. Alternatively, the selection coefficient against siRNA-targeted TEs may be sufficient for this process to occur passively through IR and unequal intra-strand homologous recombination [34]. Deletions within TEs could reduce the selection coefficient by removing siRNA targets sites, partially release proximal gene repression, and reduce deleterious ectopic recombination events in addition to accelerating the transpositional inactivation of TEs.

TEs and Epigenetic Variation Between Accessions

Surprisingly little research has covered genome-wide natural variation in TEs and epigenetic modification in plants; however, this is expected to change rapidly with high-throughput sequencing, community efforts such as the 1001 genomes project, and the increasing quality of re-sequenced genomes [39, 125, 126]. One of the first efforts to compare genome-wide differences in siRNAs from two *A. thaliana* accessions (Ler and Col-0) found that differences in siRNAs were generally well correlated with DNA methylation variation, and that many differential siRNA loci were located in TE or repeat sequences [240]. A second comparison of the same accessions concluded that fewer than 10% of TEs are differentially methylated, with methylation more variable between genic regions than between TEs, while TEs are more likely to be deleted or show presence/absence polymorphisms [234].

A further study of *A. thaliana* accessions Ler and C24 identified differential expression of siRNAs at up to 30% of clusters, with significant differences at approximately 10% of loci with sufficient sequencing coverage in both accessions [241]. A further study confirmed that between pairs of *A. thaliana* accessions about 6% of TEs are significantly variable [34]. Assembly of the genomes of accessions C24 and Bur-0 led to a 7% increase in mapped siRNAs compared to mapping against the reference Col-0 genome [126]. This increase in mapped siRNAs indicates that a significant number of siRNA-generating loci are diverged between accessions and is consistent with the rate of variable TEs [34]. Therefore, a number of the differentially expressed loci in earlier studies, or studies where siRNAs from different accessions are mapped back to a reference genome, may reflect TE presence/absence polymorphisms or structural variation in the siRNA-targeted regions. This is supported by the observation that siRNAs are enriched in the regions of TEs that are variable between accessions [34].

Variable TEs are equally distributed throughout the *A. thaliana* genome but occur preferentially distal to genes, where IR may be less likely to interfere with coding genes and their regulatory sequences [34]. Whether this pattern of distribution is normal for variant plant TEs will require genome-wide comparisons of TEs from multiple accessions of a variety of species that differ in life history traits and TE content. TEs that are variable between accessions do not, however, differ in the repression of proximal genes compared to invariable TEs [34].

TEs are removed from plant genomes through IR and unequal intra-strand homologous recombination, and may be prevented from reaching fixation within population through negative selection coefficients, especially for gene-proximal methylated TEs [35, 75, 182]. Therefore, it is perhaps unsurprising that TEs are overrepresented in analyses of structural variants among accessions and between species [105, 124, 125, 234]. A recent genome-wide comparison of 80 *A. thaliana* accessions found evidence of structural variation for 80% of TEs [124]. Similarly, Hollister and Gaut found that 44% of TEs were polymorphic in a comparison of more than 600 TEs in 48 accessions [35]. In the closely related but outcrossing *A. lyrata,* TEs are found at intermediate to high population frequencies; however, most are not fixed within the population [182] suggesting that similar TE demographics may indeed be present compared to *A. thaliana* [124]. Taken together, these studies indicate that TEs are not generally fixed within species and that there is enrichment of deletions within TEs.

Conclusions

Considering their diverse effects on genome structure and gene regulation, it is unsurprising that TEs have been implicated as a major driving force in plant genome evolution and speciation [216, 242]. Given the abundance of TEs in angiosperms compared to gymnosperms, TEs are likely to have contributed to the successful radiation of flowering plants and their rapid adaptation to different environmental niches through facilitation of genome plasticity. As TEs are the most variable component of plant genomes, they provide a prolific source of diversity to aid plant genome evolution and adaptation.

In 1984, Barbara McClintock first proposed that genomic changes initiated by TEs under stress conditions could contribute to adaptation and even speciation [165]. With the recognition that a particular form of stress can lead to new regulatory networks that respond to that specific stress [95], perhaps activation of native TEs by stress can be used in agriculture to breed new crop varieties that are more resistant to major stresses.

To fully understand the forces governing TE dynamics will require many further comparisons of TEs from closely related species that vary in terms of TE content, TE families, genome size, mating system, population size, ploidy, and epigenetic control of TEs. With next generation sequencing improvements, this is becoming increasingly feasible, even for species with a high TE content that has previously made genome assembly complex. With an increasing number of genome-wide analyses in different species, hopefully our understanding of TE dynamics will progress by leaps and bounds in the near future.

Acknowledgments I wish to thank my colleague Dr. Jurriaan Ton of the University of Sheffield for his comments and insightful discussions on the manuscript.

References

1. Lander ES, et al. Initial sequencing and analysis of the human genome. Nature. 2001;409(6822):860–921.
2. Messing J, et al. Sequence composition and genome organization of maize. Proc Natl Acad Sci U S A. 2004;101(40):14349–54.
3. SanMiguel PJ, et al. Transposable elements, genes and recombination in a 215-kb contig from wheat chromosome 5A(m). Funct Integr Genomics. 2002;2(1–2):70–80.
4. Brenchley R, et al. Analysis of the bread wheat genome using whole-genome shotgun sequencing. Nature. 2012;491(7426):705–10.
5. Fedoroff NV. Transposable elements, epigenetics, and genome evolution. Science. 2012. 338(6108):758–67.
6. Bennett MD, Leitch IJ. Nuclear DNA amounts in angiosperms: targets, trends and tomorrow. Ann Bot. 2011;107(3):467–590.
7. Bennetzen JL. Transposable elements, gene creation and genome rearrangement in flowering plants. Curr Opin Genet Dev. 2005;15(6):621–7.
8. SanMiguel P, et al. Nested retrotransposons in the intergenic regions of the maize genome. Science. 1996;274(5288):765–8.
9. Bennetzen JL. Mechanisms and rates of genome expansion and contraction in flowering plants. Genetica. 2002;115(1):29–36.
10. Shirasu K, et al. A contiguous 66-kb barley DNA sequence provides evidence for reversible genome expansion. Genome Res. 2000;10(7):908–15.
11. Kelly LJ, Leitch IJ. Exploring giant plant genomes with next-generation sequencing technology. Chromosome Res. 2011;19(7):939–53.
12. Ambrozova K, et al. Diverse retrotransposon families and an AT-rich satellite DNA revealed in giant genomes of *Fritillaria lilies*. Ann Bot. 2011;107(2):255–68.
13. Ibarra-Laclette E, et al. Architecture and evolution of a minute plant genome. Nature. 2013;498(7452):94–8.
14. Oliver KR, McComb JA, Greene WK. Transposable elements: powerful contributors to angiosperm evolution and diversity. Genome Biol Evol. 2013;5(10):1886–901.
15. Kapitonov VV, Jurka J. Rolling-circle transposons in eukaryotes. Proc Natl Acad Sci U S A. 2001;98(15):8714–9.
16. Vitte C, Bennetzen JL. Analysis of retrotransposon structural diversity uncovers properties and propensities in angiosperm genome evolution. Proc Natl Acad Sci U S A. 2006;103(47):17638–43.
17. Du J, et al. Evolutionary conservation, diversity and specificity of LTR-retrotransposons in flowering plants: insights from genome-wide analysis and multi-specific comparison. Plant J. 2010;63(4):584–98.
18. Casacuberta JM, Santiago N. Plant LTR-retrotransposons and MITEs: control of transposition and impact on the evolution of plant genes and genomes. Gene. 2003;311:1–11.
19. Deragon JM, Casacuberta JM, Panaud O. Plant transposable elements. In: Volff J-N, editor. Plant genomes. Basel: Karger; 2008. p. 69–82.
20. Baucom RS, et al. Exceptional diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize genome. PLoS Genet. 2009;5(11):e1000732.
21. Bucher E, Reinders J, Mirouze M. Epigenetic control of transposon transcription and mobility in Arabidopsis. Curr Opin Plant Biol. 2012;15(5):503–10.
22. Hollister JD, et al. Transposable elements and small RNAs contribute to gene expression divergence between *Arabidopsis thaliana* and *Arabidopsis lyrata*. Proc Natl Acad Sci U S A. 2011;108(6):2322–7.
23. Jiang N, et al. Pack-MULE transposable elements mediate gene evolution in plants. Nature. 2004;431(7008):569–73.

24. Jia Y, et al. Loss of RNA-dependent RNA polymerase 2 (RDR2) function causes widespread and unexpected changes in the expression of transposons, genes, and 24-nt small RNAs. PLoS Genet. 2009;5(11):e1000737.
25. Vicient CM. Transcriptional activity of transposable elements in maize. BMC Genomics. 2010;11:601.
26. Mosher R, et al. Uniparental expression of PolIV-dependent siRNAs in developing endosperm of Arabidopsis. Nature. 2009;460(7252):283–6.
27. Slotkin R, et al. Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell. 2009;136(3):461–72.
28. Hsieh TF, et al. Genome-wide demethylation of Arabidopsis endosperm. Science. 2009; 324(5933):1451–4.
29. Hu TT, et al. The *Arabidopsis lyrata* genome sequence and the basis of rapid genome size change. Nat Genet. 2011;43(5):476–81.
30. Tenaillon MI, et al. Genome size and transposable element content as determined by high-throughput sequencing in maize and *Zea luxurians*. Genome Biol Evol. 2011;3:219–29.
31. Hawkins JS, et al. Differential lineage-specific amplification of transposable elements is responsible for genome size variation in Gossypium. Genome Res. 2006;16(10):1252–61.
32. Piegu B, et al. Doubling genome size without polyploidization: dynamics of retrotransposition-driven genomic expansions in *Oryza australiensis*, a wild relative of rice. Genome Res. 2006;16(10):1262–9.
33. Vitte C, Panaud O. LTR retrotransposons and flowering plant genome size: emergence of the increase/decrease model. Cytogenet Genome Res. 2005;110(1–4):91–107.
34. Wang X, Weigel D, Smith LM. Transposon variants and their effects on gene expression in Arabidopsis. PLoS Genet. 2013;9(2):e1003255.
35. Hollister JD, Gaut BS. Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome Res. 2009;19:1419–28.
36. Perez-Hormaeche J, et al. Invasion of the Arabidopsis genome by the tobacco retrotransposon Tnt1 is controlled by reversible transcriptional gene silencing. Plant Physiol. 2008;147(3):1264–78.
37. Mari-Ordonez A, et al. Reconstructing de novo silencing of an active plant retrotransposon. Nat Genet. 2013;45(9):1029–39.
38. Lippman Z, et al. Distinct mechanisms determine transposon inheritance and methylation via small interfering RNA and histone modification. PLoS Biol. 2003;1(3):420–8.
39. Cao X, et al. Role of the DRM and CMT3 methyltransferases in RNA-directed DNA methylation. Curr Biol. 2003;13(24):2212–7.
40. Lippman Z, et al. Role of transposable elements in heterochromatin and epigenetic control. Nature. 2004;430:471–6.
41. Lister R, et al. Highly integrated single-base resolution maps of the epigenome in Arabidopsis. Cell. 2008;133(3):523–36.
42. Matzke MA, Mette MF, Matzke AJM. Transgene silencing by the host genome defense: implications for the evolution of epigenetic control mechanisms in plants and vertebrates. Plant Mol Biol. 2000;43(2):401–15.
43. Zilberman D, et al. Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. Nat Genet. 2007;39(1):61–9.
44. Kasschau KD, et al. Genome-wide profiling and analysis of Arabidopsis siRNAs. PLoS Biol. 2007;5(3):e57.
45. Singer T, Yordan C, Martienssen RA. Robertson's mutator transposons in *A. thaliana* are regulated by the chromatin-remodeling gene decrease in DNA methylation (DDM1). Genes Dev. 2001;15(5):591–602.
46. Miura A, et al. Mobilization of transposons by a mutation abolishing full DNA methylation in Arabidopsis. Nature. 2001;411(6834):212–4.
47. Kato M, et al. Role of CG and non-CG methylation in immobilization of transposons in Arabidopsis. Curr Biol. 2003;13(5):421–6.

48. Xie Z, et al. Genetic and functional diversification of small RNA pathways in plants. PLoS Biol. 2004;2(5):642–52.
49. Zilberman D, et al. Role of Arabidopsis ARGONAUTE4 in RNA-directed DNA methylation triggered by inverted repeats. Curr Biol. 2004;14:1214–20.
50. Slotkin RK, Freeling M, Lisch D. Mu killer causes the heritable inactivation of the mutator family of transposable elements in *Zea mays*. Genetics. 2003;165(2):781–97.
51. Singh J, Freeling M, Lisch D. A position effect on the heritability of epigenetic silencing. PLoS Genet. 2008;4(10):e1000216.
52. Wierzbicki AT, Haag JR, Pikaard CS. Noncoding transcription by RNA polymerase Pol IVb/Pol V mediates transcriptional silencing of overlapping and adjacent genes. Cell. 2008;135(4):635–48.
53. Herr AJ, et al. RNA polymerase IV directs silencing of endogenous DNA. Science. 2005;308(5718):118–20.
54. Onodera Y, et al. Plant nuclear RNA polymerase IV mediates siRNA and DNA methylation-dependent heterochromatin formation. Cell. 2005;120(5):613–22.
55. Gasciolli V, et al. Partially redundant functions of Arabidopsis DICER-like enzymes and a role for DCL4 in producing trans-acting siRNAs. Curr Biol. 2005;15(16):1494–500.
56. Yu B, et al. Methylation as a crucial step in plant microRNA biogenesis. Science. 2005;307:932–5.
57. Li J, et al. Methylation protects miRNAs and siRNAs from a 3'-End uridylation activity in Arabidopsis. Curr Biol. 2005;15:1501–7.
58. Yang Z, et al. HEN1 recognizes 21–24 nt small RNA duplexes and deposits a methyl group onto the 2' OH of the 3' terminal nucleotide. Nucleic Acids Res. 2006. 34(2):667–75.
59. Khvorova A, Reynolds A, Jayasena SD. Functional siRNAs and miRNAs exhibit strand bias. Cell. 2003;115:209–16.
60. Schwarz DS, et al. Asymmetry in the assembly of the RNAi enzyme complex. Cell. 2003;115:199–208.
61. Pontier D, et al. Reinforcement of silencing at transposons and highly repeated sequences requires the concerted action of two distinct RNA polymerases IV in Arabidopsis. Genes Dev. 2005;19:2030–40.
62. Wierzbicki AT, et al. Spatial and functional relationships among Pol V-associated loci, Pol IV-dependent siRNAs, and cytosine methylation in the Arabidopsis epigenome. Genes Dev. 2012;26(16):1825–36.
63. Huettel B, et al. Endogenous targets of RNA-directed DNA methylation and Pol IV in Arabidopsis. EMBO J. 2006;25:2828–36.
64. Li CF, et al. An ARGONAUTE4-containing nuclear processing center colocalized with Cajal bodies in *Arabidopsis thaliana*. Cell. 2006;126:93–106.
65. Pontes O, et al. The Arabidopsis chromatin-modifying nuclear siRNA pathway involves a nucleolar RNA processing center. Cell. 2006;126:79–92.
66. Law JA, Jacobsen SE. Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat Rev Genet. 2010;11(3):204–20.
67. Aufsatz W, et al. The role of MET1 in RNA-directed de novo and maintenance methylation of CG dinucleotides. Plant Mol Biol. 2004;54(6):793–804.
68. Bartee L, Malagnac F, Bender J. Arabidopsis cmt3 chromomethylase mutations block non-CG methylation and silencing of an endogenous gene. Genes Dev. 2001;15:1753–8.
69. Ebbs ML, Bartee L, Bender J. H3 lysine 9 methylation is maintained on a transcribed inverted repeat by combined action of SUVH6 and SUVH4 methyltransferases. Mol Cell Biol. 2005;25(23):10507–15.
70. Ebbs ML, Bender J. Locus-specific control of DNA methylation by the Arabidopsis SUVH5 histone methyltransferase. Plant Cell. 2006;18(5):1166–76.
71. Jackson JP, et al. Control of CpNpG DNA methylation by the KRYPTONITE histone H3 methyltransferase. Nature. 2002;416(6880):556–60.
72. Malagnac F, Bartee L, Bender J. An Arabidopsis SET domain protein required for maintenance but not establishment of DNA methylation. EMBO J. 2002;21(24):6842–52.

73. Penterman J, et al. DNA demethylation in the Arabidopsis genome. Proc Natl Acad Sci U S A. 2007;104(16):6752–7.
74. Rodrigues JA, et al. Imprinted expression of genes and small RNA is associated with localized hypomethylation of the maternal genome in rice endosperm. Proc Natl Acad Sci U S A. 2013;110(19):7934–9.
75. Ahmed I, et al. Genome-wide evidence for local DNA methylation spreading from small RNA-targeted sequences in Arabidopsis. Nucleic Acids Res. 2011;39(16):6919–31.
76. Zhang X, et al. Genome-wide high-resolution mapping and functional analysis of DNA methylation in Arabidopsis. Cell. 2006;126(6):1189–201.
77. Cui H, Fedoroff NV. Inducible DNA demethylation mediated by the maize suppressor-mutator transposon-encoded TnpA protein. Plant Cell. 2002;14(11):2883–99.
78. Feng S, Jacobsen SE. Epigenetic modifications in plants: an evolutionary perspective. Curr Opin Plant Biol. 2011;14(2):179–86.
79. Zhang X, et al. Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in *Arabidopsis thaliana*. Genome Biol. 2009;10(6):R62.
80. Zhang X, et al. Whole-genome analysis of histone H3 lysine 27 trimethylation in Arabidopsis. PLoS Biol. 2007;5(5):e129.
81. Tran RK, et al. Chromatin and siRNA pathways cooperate to maintain DNA methylation of small transposable elements in Arabidopsis. Genome Biol. 2005;6:R90. doi:10.1186/gb-2005–6-11-r90.
82. Blumenstiel JP. Evolutionary dynamics of transposable elements in a small RNA world. Trends Genet. 2011;27(1):23–31.
83. Gaut BS, et al. Recombination: an underappreciated factor in the evolution of plant genomes. Nat Rev Genet. 2007;8(1):77–84.
84. Bennetzen JL. Transposable element contributions to plant gene and genome evolution. Plant Mol Biol. 2000;42(1):251–69.
85. He F, et al. Widespread interspecific divergence in cis-regulation of transposable elements in the Arabidopsis genus. Mol Biol Evol. 2012;29(3):1081–91.
86. Jin YK, Bennetzen JL. Structure and coding properties of Bs1, a maize retrovirus-like transposon. Proc Natl Acad Sci U S A. 1989;86(16):6235–9.
87. Johns MA, et al. An unusually compact retrotransposon in maize. Plant Mol Biol. 1989; 12(6):633–42.
88. Grandbastien MA, Spielmann A, Caboche M. Tnt1, a mobile retroviral-like transposable element of tobacco isolated by plant cell genetics. Nature. 1989;337(6205):376–80.
89. Pelissier T, et al. Athila, a new retroelement from *Arabidopsis thaliana*. Plant Mol Biol. 1995;29(3):441–52.
90. Pereira V. Insertion bias and purifying selection of retrotransposons in the *Arabidopsis thaliana* genome. Genome Biol. 2004;5:R79. doi: 10.1186/gb-2004–5-10-r79.
91. Frank MJ, et al. Tag1 is an autonomous transposable element that shows somatic excision in both Arabidopsis and tobacco. Plant Cell. 1997;9(10):1745–56.
92. Tsay YF, et al. Identification of a mobile endogenous transposon in *Arabidopsis thaliana*. Science. 1993;260(5106):342–4.
93. Lenoir A, et al. The evolutionary origin and genomic organization of SINEs in *Arabidopsis thaliana*. Mol Biol Evol. 2001;18(12):2315–22.
94. Casacuberta E, Pardue ML. Transposon telomeres are widely distributed in the Drosophila genus: TART elements in the virilis group. Proc Natl Acad Sci U S A. 2003;100(6):3363–8.
95. Ito H, et al. An siRNA pathway prevents transgenerational retrotransposition in plants subjected to stress. Nature. 2011;472(7341):115–9.
96. Naito K, et al. Dramatic amplification of a rice transposable element during recent domestication. Proc Natl Acad Sci U S A. 2006;103(47):17620–5.
97. Li Y, Harris L, Dooner HK. TED, an autonomous and rare maize transposon of the mutator superfamily with a high gametophytic excision frequency. Plant Cell. 2013;25(9):3251–65.
98. Le Rouzic A, Capy P. Population genetics models of competition between transposable element subfamilies. Genetics. 2006;174(2):785–93.
99. Le Rouzic A, Boutin TS, Capy P. Long-term evolution of transposable elements. Proc Natl Acad Sci U S A. 2007;104(49):19375–80.

100. Rangwala SH, Richards EJ. Differential epigenetic regulation within an Arabidopsis retroposon family. Genetics. 2007;176(1):151–60.
101. Hamilton AJ, et al. Two classes of short interfering RNA in RNA silencing. EMBO J. 2002;21(17):4671–9.
102. Zilberman D, Cao X, Jacobsen SE. ARGONAUTE4 control of locus specific siRNA accumulation and DNA and histone methylation. Science. 2003;299:716–9.
103. Mirouze M, et al. Selective epigenetic control of retrotransposition in Arabidopsis. Nature. 2009;461(7262):427–30.
104. Naito K, et al. Unexpected consequences of a sudden and massive transposon amplification on rice gene expression. Nature. 2009;461(7267):1130–4.
105. Huang X, et al. Genome-wide analysis of transposon insertion polymorphisms reveals intraspecific variation in cultivated rice. Plant Physiol. 2008;148(1):25–40.
106. Warnefors M, Pereira V, Eyre-Walker A. Transposable elements: insertion pattern and impact on gene expression evolution in hominids. Mol Biol Evol. 2010;27(8):1955–62.
107. Fukai E, et al. Derepression of the plant chromovirus LORE1 induces germline transposition in regenerated plants. PLoS Genet. 2010;6(3):e1000868.
108. Bureau TE, Wessler SR. Tourist: a large family of small inverted repeat elements frequently associated with maize genes. Plant Cell. 1992;4(10):1283–94.
109. Bureau TE, Ronald PC, Wessler SR. A computer-based systematic survey reveals the predominance of small inverted-repeat elements in wild-type rice genes. Proc Natl Acad Sci U S A. 1996;93(16):8524–9.
110. Wei F, et al. Detailed analysis of a contiguous 22-Mb region of the maize genome. PLoS Genet. 2009;5(11):e1000728.
111. Zerjal T, et al. Contrasting evolutionary patterns and target specificities among three Tourist-like MITE families in the maize genome. Plant Mol Biol. 2009;71(1–2):99–114.
112. Pan X, Li Y, Stein L. Site preferences of insertional mutagenesis agents in Arabidopsis. Plant Physiol. 2005;137(1):168–75.
113. Sarilar V, et al. BraSto, a Stowaway MITE from Brassica: recently active copies preferentially accumulate in the gene space. Plant Mol Biol. 2011;77(1–2):59–75.
114. Miyao A, et al. Target site specificity of the Tos17 retrotransposon shows a preference for insertion within genes and against insertion in retrotransposon-rich regions of the genome. Plant Cell. 2003;15(8):1771–80.
115. Piffanelli P, et al. Large-scale characterization of Tos17 insertion sites in a rice T-DNA mutant library. Plant Mol Biol. 2007;65(5):587–601.
116. Kolesnik T, et al. Establishing an efficient Ac/Ds tagging system in rice: large-scale analysis of Ds flanking sequences. Plant J. 2004;37(2):301–14.
117. Tsukahara S, et al. Centromere-targeted de novo integrations of an LTR retrotransposon of *Arabidopsis lyrata*. Genes Dev. 2012;26(7):705–13.
118. Le QH, et al. Transposon diversity in *Arabidopsis thaliana*. Proc Natl Acad Sci U S A. 2000;97(13):7376–81.
119. Devos KM, Brown JK, Bennetzen JL. Genome size reduction through illegitimate recombination counteracts genome expansion in Arabidopsis. Genome Res. 2002;12(7):1075–9.
120. Tian Z, et al. Do genetic recombination and gene density shape the pattern of DNA elimination in rice long terminal repeat retrotransposons? Genome Res. 2009;19(12):2221–30.
121. Wright SI, Agrawal N, Bureau TE. Effects of recombination rate and gene density on transposable element distributions in *Arabidopsis thaliana*. Genome Res. 2003;13(8):1897–903.
122. Buisine N, Quesneville H, Colot V. Improved detection and annotation of transposable elements in sequenced genomes using multiple reference sequence sets. Genomics. 2008;91(5):467–75.
123. Montgomery EA, et al. Chromosome rearrangement by ectopic recombination in *Drosophila melanogaster*: genome structure and evolution. Genetics. 1991;129(4):1085–98.
124. Cao J, et al. Whole-genome sequencing of multiple *Arabidopsis thaliana* populations. Nat Genet. 2011;43(10):956–63.
125. Gan X, et al. Multiple reference genomes and transcriptomes for *Arabidopsis thaliana*. Nature. 2011;477(7365):419–23.

126. Schneeberger K, et al. Reference-guided assembly of four diverse *Arabidopsis thaliana* genomes. Proc Natl Acad Sci U S A. 2011;108(25):10249–54.
127. Hollister JD, Gaut BS. Population and evolutionary dynamics of Helitron transposable elements in *Arabidopsis thaliana*. Mol Biol Evol. 2007;24(11):2515–24.
128. Ma J, Devos KM, Bennetzen JL. Analyses of LTR-retrotransposon structures reveal recent and rapid genomic DNA loss in rice. Genome Res. 2004;14(5):860–9.
129. Sabot F, et al. Transpositional landscape of the rice genome revealed by paired-end mapping of high-throughput re-sequencing data. Plant J. 2011;66(2):241–6.
130. Kalendar R, et al. Genome evolution of wild barley (*Hordeum spontaneum*) by BARE-1 retrotransposon dynamics in response to sharp microclimatic divergence. Proc Natl Acad Sci U S A. 2000;97(12):6603–7.
131. Ma J, Bennetzen JL. Rapid recent growth and divergence of rice nuclear genomes. Proc Natl Acad Sci U S A. 2004;101(34):12404–10.
132. Ma J, Bennetzen JL. Recombination, rearrangement, reshuffling, and divergence in a centromeric region of rice. Proc Natl Acad Sci U S A. 2006;103(2):383–8.
133. Hawkins JS, Grover CE, Wendel JF. Repeated big bangs and the expanding universe: directionality in plant genome size evolution. Plant Sci. 2008;174(6):557–62.
134. Zhang J, Zuo T, Peterson T. Generation of tandem direct duplications by reversed-ends transposition of maize ac elements. PLoS Genet. 2013;9(8):e1003691.
135. Piriyapongsa J, Jordan IK. Dual coding of siRNAs and miRNAs by plant transposable elements. RNA. 2008;14(5):814–21.
136. Li Y, et al. Domestication of transposable elements into MicroRNA genes in plants. PLoS One. 2011;6(5):e19212.
137. Ou-Yang F, et al. Transposable element-associated microRNA hairpins produce 21-nt sRNAs integrated into typical microRNA pathways in rice. Funct Integr Genomics. 2013;13(2):207–16.
138. Schaack S, Gilbert C, Feschotte C. Promiscuous DNA: horizontal transfer of transposable elements and why it matters for eukaryotic evolution. Trends Ecol Evol. 2010;25(9):537–46.
139. Cheng X, et al. A new family of Ty1-copia-like retrotransposons originated in the tomato genome by a recent horizontal transfer event. Genetics. 2009;181(4):1183–93.
140. Diao X, Freeling M, Lisch D. Horizontal transfer of a plant transposon. PLoS Biol. 2006;4(1):e5.
141. Roulin A, et al. Whole genome surveys of rice, maize and sorghum reveal multiple horizontal transfers of the LTR-retrotransposon Route66 in Poaceae. BMC Evol Biol. 2009;9:58.
142. Woodrow P, et al. Ty1-copia group retrotransposons and the evolution of retroelements in several angiosperm plants: evidence of horizontal transmission. Bioinformation. 2012;8(6):267–71.
143. Tenaillon MI, Hollister JD, Gaut BS. A triptych of the evolution of plant transposable elements. Trends Plant Sci. 2010;15(8):471–8.
144. Boutin TS, Le Rouzic A, Capy P. How does selfing affect the dynamics of selfish transposable elements? Mob DNA. 2012;3:5.
145. Dolgin ES, Charlesworth B. The fate of transposable elements in asexual populations. Genetics. 2006;174(2):817–27.
146. Wright SI, Schoen DJ. Transposon dynamics and the breeding system. Genetica. 1999;107(1–3):139–48.
147. Agren JA, Wright SI. Co-evolution between transposable elements and their hosts: a major factor in genome size evolution? Chromosome Res. 2011;19(6):777–86.
148. Charlesworth B, Langley CH. The evolution of self-regulated transposition of transposable elements. Genetics. 1986;112(2):359–83.
149. Yang G, et al. Tuned for transposition: molecular determinants underlying the hyperactivity of a Stowaway MITE. Science. 2009;325(5946):1391–4.
150. de la Chaux N, et al. The predominantly selfing plant Arabidopsis thaliana experienced a recent reduction in transposable element abundance compared to its outcrossing relative *Arabidopsis lyrata*. Mob DNA. 2012;3(1):2.

151. Wright SI, et al. Population dynamics of an Ac-like transposable element in self- and cross-pollinating Arabidopsis. Genetics. 2001;158(3):1279–88.
152. Lockton S, Gaut BS. The evolution of transposable elements in natural populations of self-fertilizing *Arabidopsis thaliana* and its outcrossing relative *Arabidopsis lyrata*. BMC Evol Biol. 2010;10:10. doi:10.1186/1471–2148-10–10.
153. Tam SM, et al. The distribution of copia-type retrotransposons and the evolutionary history of tomato and related wild species. J Evol Biol. 2007;20(3):1056–72.
154. Wright SI, et al. Genomic consequences of outcrossing and selfing in plants. Int J Plant Sci. 2008;169(1):105–118.
155. Charlesworth D, Charlesworth B. Transposable elements in inbreeding and outbreeding populations. Genetics. 1995;140(1):415–7.
156. Morgan MT. Transposable element number in mixed mating populations. Genet Res. 2001;77(3):261–75.
157. Charlesworth D, Wright SI. Breeding systems and genome evolution. Curr Opin Genet Dev. 2001;11(6):685–90.
158. Matsunaga W, et al. The effects of heat induction and the siRNA biogenesis pathway on the transgenerational transposition of ONSEN, a copia-like retrotransposon in *Arabidopsis thaliana*. Plant Cell Physiol. 2012;53(5):824–33.
159. Mhiri C. et al. The promoter of the tobacco Tnt1 retrotransposon is induced by wounding and by abiotic stress. Plant Mol Biol. 1997;33(2):257–66.
160. Ivashuta S, et al. Genotype-dependent transcriptional activation of novel repetitive elements during cold acclimation of alfalfa (*Medicago sativa*). Plant J. 2002;31(5):615–27.
161. Uchiyama T, et al. A pair of transposons coordinately suppresses gene expression, independent of pathways mediated by siRNA in Antirrhinum. New Phytol. 2013;197(2):431–40.
162. Grandbastien MA, et al. Stress activation and genomic impact of Tnt1 retrotransposons in Solanaceae. Cytogenet Genome Res. 2005;110(1–4):229–41.
163. Beguiristain T, et al. Three Tnt1 subfamilies show different stress-associated patterns of expression in tobacco. Consequences for retrotransposon control and evolution in plants. Plant Physiol. 2001;127(1):212–21.
164. Walbot V. UV-B damage amplified by transposons in maize. Nature. 1999;397(6718):398–9.
165. McClintock B. The significance of responses of the genome to challenge. Science. 1984;226(4676):792–801.
166. Shan X, et al. Mobilization of the active MITE transposons mPing and Pong in rice by introgression from wild rice (*Zizania latifolia* Griseb.). Mol Biol Evol. 2005;22(4):976–90.
167. Wang N, et al. Transpositional reactivation of the Dart transposon family in rice lines derived from introgressive hybridization with *Zizania latifolia*. BMC Plant Biol. 2010;10:190.
168. Zou J, et al. De novo genetic variation associated with retrotransposon activation, genomic rearrangements and trait variation in a recombinant inbred line population of *Brassica napus* derived from interspecific hybridization with *Brassica rapa*. Plant J. 2011;68(2):212–24.
169. Kawakami T, et al. Transposable element proliferation and genome expansion are rare in contemporary sunflower hybrid populations despite widespread transcriptional activity of LTR retrotransposons. Genome Biol Evol. 2011;3:156–67.
170. Wang Q, et al. Intergenomic rearrangements after polyploidization of *Kengyilia thoroldiana* (Poaceae: Triticeae) affected by environmental factors. PLoS One. 2012;7(2):e31033.
171. Matzke MA, Matzke AJ. Polyploidy and transposons. Trends Ecol Evol. 1998;13(6):241.
172. Ungerer MC, Strakosh SC, Zhen Y. Genome expansion in three hybrid sunflower species is associated with retrotransposon proliferation. Curr Biol. 2006;16(20):R872–3.
173. Baumel A, et al. Retrotransposons and genomic stability in populations of the young allopolyploid species *Spartina anglica* C.E. Hubbard (Poaceae). Mol Biol Evol. 2002;19(8):1218–27.
174. Josefsson C, Dilkes B, Comai L. Parent-dependent loss of gene silencing during interspecies hybridization. Curr Biol. 2006;16(13):1322–1328.

175. Burkart-Waco D, et al. Early disruption of maternal-zygotic interaction and activation of defense-like responses in Arabidopsis interspecific crosses. Plant Cell. 2013;25(6):2037–55.
176. Madlung A, et al. Genomic changes in synthetic Arabidopsis polyploids. Plant J. 2005;41(2):221–30.
177. Kashkush K, Feldman M, Levy AA. Transcriptional activation of retrotransposons alters the expression of adjacent genes in wheat. Nat Genet. 2003;33(1):102–6.
178. Kashkush K, Feldman M, Levy AA. Gene loss, silencing and activation in a newly synthesized wheat allotetraploid. Genetics. 2002;160(4):1651–9.
179. Parisod C, et al. Rapid structural and epigenetic reorganization near transposable elements in hybrid and allopolyploid genomes in Spartina. New Phytol. 2009;184(4):1003–15.
180. Renny-Byfield S, et al. Next generation sequencing reveals genome downsizing in allotetraploid *Nicotiana tabacum*, predominantly through the elimination of paternally derived repetitive DNAs. Mol Biol Evol. 2011;28(10):2843–54.
181. Lynch M, Conery JS. The origins of genome complexity. Science. 2003;302(5649):1401–4.
182. Lockton S, Ross-Ibarra J, Gaut BS. Demography and weak selection drive patterns of transposable element diversity in natural populations of *Arabidopsis lyrata*. Proc Natl Acad Sci U S A. 2008;105(37):13965–70.
183. Whitney KD, et al. A role for nonadaptive processes in plant genome size evolution? Evolution. 2010;64(7):2097–109.
184. Nystedt B, et al. The Norway spruce genome sequence and conifer genome evolution. Nature. 2013;497(7451):579–84.
185. Bhattacharyya MK, et al. The wrinkled-seed character of pea described by Mendel is caused by a transposon-like insertion in a gene encoding starch-branching enzyme. Cell. 1990;60(1):115–22.
186. Greene B, Walko R, Hake S. Mutator insertions in an intron of the maize knotted1 gene result in dominant suppressible mutations. Genetics. 1994;138(4):1275–85.
187. Momose M, Abe Y, Ozeki Y. Miniature inverted-repeat transposable elements of Stowaway are active in potato. Genetics. 2010;186(1):59–66.
188. Scott L, LaFoe D, Weil CF. Adjacent sequences influence DNA repair accompanying transposon excision in maize. Genetics. 1996;142(1):237–46.
189. Sakai H, Tanaka T, Itoh T. Birth and death of genes promoted by transposable elements in *Oryza sativa*. Gene. 2007;392(1–2):59–63.
190. Lockton S, Gaut BS. The contribution of transposable elements to expressed coding sequence in *Arabidopsis thaliana*. J Mol Evol. 2009;68(1):80–9.
191. Chan SW, et al. RNA silencing genes control de novo DNA methylation. Science. 2004;303(5662):1336.
192. Cowan RK, et al. MUSTANG is a novel family of domesticated transposase genes found in diverse angiosperms. Mol Biol Evol. 2005;22(10):2084–9.
193. Joly-Lopez Z, et al. A gene family derived from transposable elements during early angiosperm evolution has reproductive fitness benefits in *Arabidopsis thaliana*. PLoS Genet. 2012;8(9):e1002931.
194. Bundock P, Hooykaas P. An Arabidopsis hAT-like transposase is essential for plant development. Nature. 2005;436(7048):282–4.
195. Hudson ME, Lisch DR, Quail PH. The FHY3 and FAR1 genes encode transposase-related proteins involved in regulation of gene expression by the phytochrome A-signaling pathway. Plant J. 2003;34(4):453–71.
196. Tsuchiya T, Eulgem T. An alternative polyadenylation mechanism coopted to the Arabidopsis RPP7 gene through intronic retrotransposon domestication. Proc Natl Acad Sci U S A. 2013;110(37):E3535–43.
197. Kuang H, et al. Identification of miniature inverted-repeat transposable elements (MITEs) and biogenesis of their siRNAs in the Solanaceae: new functional implications for MITEs. Genome Res. 2009;19(1):42–56.

198. Sugimoto K, Takeda S, Hirochika H. Transcriptional activation mediated by binding of a plant GATA-type zinc finger protein AGP1 to the AG-motif (AGATCCAA) of the wound-inducible Myb gene NtMyb2. Plant J. 2003;36(4):550–64.
199. Butelli E, et al. Retrotransposons control fruit-specific, cold-dependent accumulation of anthocyanins in blood oranges. Plant Cell. 2012;24(3):1242–55.
200. Fujii M, et al. Acquisition of aluminium tolerance by modification of a single gene in barley. Nat Commun. 2012;3:713.
201. Magalhaes JV, et al. A gene in the multidrug and toxic compound extrusion (MATE) family confers aluminum tolerance in sorghum. Nat Genet. 2007;39(9):1156–61.
202. Delhaize E, Ma JF, Ryan PR. Transcriptional regulation of aluminium tolerance genes. Trends Plant Sci. 2012;17(6):341–8.
203. Lister C, Jackson D, Martin C. Transposon-induced inversion in Antirrhinum modifies nivea gene expression to give a novel flower color pattern under the control of cycloidearadialis. Plant Cell. 1993;5(11):1541–53.
204. Kloeckener-Gruissem B, Vogel JM, Freeling M. The TATA box promoter region of maize Adh1 affects its organ-specific expression. EMBO J. 1992;11(1):157–66.
205. Kloeckener-Gruissem B, Freeling M. Transposon-induced promoter scrambling: a mechanism for the evolution of new alleles. Proc Natl Acad Sci U S A. 1995;92(6):1836–40.
206. Hoen DR, Bureau TE. Transposable element exaptation in plants. In: Grandbastien M-A and Casacuberta JM, editors. Plant transposable elements. Springer: Berlin; 2012.
207. Liu J, et al. siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in Arabidopsis. Genes Dev. 2004;18(23):2873–8.
208. Zerjal T, et al. Maize genetic diversity and association mapping using transposable element insertion polymorphisms. Theor Appl Genet. 2012;124(8):1521–37.
209. Zabala G, Vodkin LO. The wp mutation of glycine max carries a gene-fragment-rich transposon of the CACTA superfamily. Plant Cell. 2005;17(10):2619–32.
210. Morgante M, et al. Gene duplication and exon shuffling by helitron-like transposons generate intraspecies diversity in maize. Nat Genet. 2005;37(9):997–1002.
211. Brunner S, Pea G, Rafalski A. Origins, genetic organization and transcription of a family of non-autonomous helitron elements in maize. Plant J. 2005;43(6):799–810.
212. Lai J, et al. Gene movement by helitron transposons contributes to the haplotype variability of maize. Proc Natl Acad Sci U S A. 2005;102(25):9068–73.
213. Hanada K, et al. The functional role of pack-MULEs in rice inferred from purifying selection and expression profile. Plant Cell. 2009;21(1):25–38.
214. Paterson AH, et al. The Sorghum bicolor genome and the diversification of grasses. Nature. 2009;457(7229):551–6.
215. Abrouk M, et al. Grass microRNA gene paleohistory unveils new insights into gene dosage balance in subgenome partitioning after whole-genome duplication. Plant Cell. 2012;24(5):1776–92.
216. Oliver KR, Greene WK. Transposable elements: powerful facilitators of evolution. Bioessays. 2009;31(7):703–14.
217. Macas J, et al. Hypervariable 3' UTR region of plant LTR-retrotransposons as a source of novel satellite repeats. Gene. 2009;448(2):198–206.
218. Gong Z, et al. Repeatless and repeat-based centromeres in potato: implications for centromere evolution. Plant Cell. 2012;24(9):3559–74.
219. Avramova Z, et al. Matrix attachment regions and structural colinearity in the genomes of two grass species. Nucleic Acids Res. 1998;26(3):761–7.
220. Zhang X, et al. Global analysis of genetic, epigenetic and transcriptional polymorphisms in *Arabidopsis thaliana* using whole genome tiling arrays. PLoS Genet. 2008;4:e1000032. doi:10.1371/journal.pgen.1000032.
221. Lu C, et al. Miniature inverted-repeat transposable elements (MITEs) have been accumulated through amplification bursts and play important roles in gene expression and species diversity in *Oryza sativa*. Mol Biol Evol. 2012;29(3):1005–17.

222. Yu A, et al. Dynamics and biological relevance of DNA demethylation in Arabidopsis antibacterial defense. Proc Natl Acad Sci U S A. 2013;110(6):2389–94.
223. Kashkush K, Khasdan V. Large-scale survey of cytosine methylation of retrotransposons and the impact of readout transcription from long terminal repeats on expression of adjacent rice genes. Genetics. 2007;177(4):1975–85.
224. Eichten SR, et al. Spreading of heterochromatin is limited to specific families of maize retrotransposons. PLoS Genet. 2012;8(12):e1003127.
225. Saze H, et al. Control of genic DNA methylation by a jmjC domain-containing protein in *Arabidopsis thaliana*. Science. 2008;319(5862):462–5.
226. Miura A, et al. An Arabidopsis jmjC domain protein protects transcribed genes from DNA methylation at CHG sites. EMBO J. 2009;28(8):1078–86.
227. Inagaki S, et al. Autocatalytic differentiation of epigenetic modifications within the Arabidopsis genome. EMBO J. 2010;29(20):3496–506.
228. Saze H, et al. Mechanism for full-length RNA processing of Arabidopsis genes containing intragenic heterochromatin. Nat Commun. 2013;4:2301.
229. McCue AD, et al. Gene expression and stress response mediated by the epigenetic regulation of a transposable element small RNA. PLoS Genet. 2012;8(2):e1002474.
230. Yan Y, et al. Small RNAs from MITE-derived stem-loop precursors regulate abscisic acid signaling and abiotic stress responses in rice. Plant J. 2011;65(5):820–8.
231. Nosaka M, et al. Role of transposon-derived small RNAs in the interplay between genomes and parasitic DNA in rice. PLoS Genet. 2012;8(9):e1002953.
232. Schott G, et al. Differential effects of viral silencing suppressors on siRNA and miRNA loading support the existence of two distinct cellular pools of ARGONAUTE1. EMBO J. 2012;31(11):2553–65.
233. Rabinowicz PD, et al. Genes and transposons are differentially methylated in plants, but not in mammals. Genome Res. 2003;13(12):2658–64.
234. Vaughn MW, et al. Epigenetic natural variation in *Arabidopsis thaliana*. PLoS Biol. 2007;5:e174. doi: 10.1371/journal.pbio.0050174.
235. Cokus SJ, et al. Shotgun bisulphite sequencing of the Arabidopsis genome reveals DNA methylation patterning. Nature. 2008;452(7184):215–9.
236. Ossowski S, et al. The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. Science. 2010;327(5961):92–4.
237. Lisch D. Epigenetic regulation of transposable elements in plants. Ann Rev Plant Biol. 2009;60:43–66.
238. Mochizuki K, Gorovsky MA. Conjugation-specific small RNAs in Tetrahymena have predicted properties of scan (scn) RNAs involved in genome rearrangement. Genes Dev. 2004;18(17):2068–73.
239. Mochizuki K, et al. Analysis of a piwi-related gene implicates small RNAs in genome rearrangement in Tetrahymena. Cell. 2002;110(6):689–99.
240. Zhai J, et al. Small RNA-directed epigenetic natural variation in *Arabidopsis thaliana*. PLoS Genet. 2008;4(4):e1000056.
241. Groszmann M, et al. Changes in 24-nt siRNA levels in Arabidopsis hybrids suggest an epigenetic contribution to hybrid vigor. Proc Natl Acad Sci USA. 2011;108(6):2617–22.
242. Zeh DW, Zeh JA, Ishida Y. Transposable elements and an epigenetic basis for punctuated equilibria. Bioessays. 2009;31(7):715–26.

Chapter 9
Population Epigenetics

Christy M. Foust, Aaron W. Schrey and Christina L. Richards

Population Epigenetics

Population genetics examines the extent of genetic variation, and changes to genetic variation, in response to evolutionary processes in populations [1]. There is an enormous body of knowledge surrounding population genetics, extending from the 1930s with the introduction of the modern synthesis (MS) to the present. The MS encompasses population genetics questions that include the dynamics of alleles in populations as they pertain to the Hardy–Weinberg equilibrium (HWE; i.e., the assumptions of infinite population size, random mating, no migration, no mutation, and no selection) [2]. In addition to the assumptions of HWE, the MS is also bound by the confines of Mendelian genetics which calls for the random segregation and assortment of alleles [3].

Though the MS has been incredibly informative regarding evolutionary questions, limitations have been identified with advancements in molecular biology [4, 5]. For example, the MS does not consider the importance of non-Mendelian inheritance to evolution [4, 6]. Epigenetic inheritance is one type of non-Mendelian inheritance; which through mechanisms like DNA methylation or histone modification can account for the transmission of phenotypic variation that is independent from genotypic variation (i.e., DNA sequence) [7]. Since the MS is largely a theory of the dynamics of DNA sequence-based transfer through a population, and gene sequences are largely unaffected by environmental factors, environmental effects on differences in phenotype are not considered important factors within the MS [4]. The MS has been a revolutionary advancement to the study of evolution, because

C. M. Foust (✉) · C. L. Richards
Department of Integrative Biology, University of South Florida, 4202 E Fowler Avenue, Tampa, FL 33620, USA
e-mail: christyfoust@mail.usf.edu

A. W. Schrey
Department of Biology, Armstrong State University, 11935 Abercorn Street, Savannah, GA, 31419 USA

O. Pontes, H. Jin (eds.), *Nuclear Functions in Plant Transcription, Signaling and Development,* DOI 10.1007/978-1-4939-2386-1_9

it reconciles Mendelian inheritance with Darwinian gradual change in response to natural selection. Certainly, its ideas are still critical for evolutionary thinking, but the time has come to extend the MS to address many developments in the field.

Population epigenetics addresses some of the limitations of the MS. Similar to population genetics, population epigenetics examines the extent of epigenetic variation present in populations of organisms and the way this variation changes over time, or across the landscape with changing environmental conditions [8, 9]. However, in stark contrast to genetic variation, epigenetic variation may not be inherited via Mendelian processes [10–12], and epigenetic variation can be directly affected by the environment [13, 14]. Though the basic concepts defining population genetics are known, the body of knowledge concerning processes of change and mechanisms of inheritance of epigenetic variation remains largely unknown [15]. Theorists are developing models of epigenetic inheritance [16, 19], and the effect of selection on epigenetic marks [20]. These models provide a theoretical underpinning for understanding the behavior of epigenetic effects and are discussed later in this chapter.

History and Definition of Epigenetics

Conrad Waddington coined the term "epigenetics" while investigating genetic assimilation and the canalization of traits in the 1940s [21, 22]. Waddington [22] recognized that some characteristics that arise in response to environmental conditions (e.g., stressors) could be incorporated into the genome via the process of genetic assimilation. He also recognized that some phenotypes could be highly conserved regardless of genotype (i.e., canalized). His definition of epigenetics was broad and incorporated these processes and all interactions between genes and gene products that lead to an organism's final phenotype [23, 24]. The discovery of molecular mechanisms and processes that regulate gene activity has resulted in changes to Waddington's definition [23].

Epigenetics has more recently been defined as heritable, differential gene expression not based on changes in DNA sequence [8, 25–27], and the inclusion of "heritable" in this definition is debated [26]. Some biologists think that epigenetics should include all processes that result in differential gene expression in the absence of changes to DNA sequences, heritable or not [28]. Others think that heritability is a necessary component of epigenetics, because inheritance raises direct evolutionary questions that are consistent with population genetics expectations [8, 29–30]. Currently, little empirical data exist to address heritability of epigenetic traits. While it is relatively simple to screen for epigenetic variation, it is much harder to determine if this variation is heritable. We have proposed a definition of epigenetics as the study of molecular-level mechanisms that result in changes in phenotype that are not due to changes in DNA sequence and may lead to heritable change in phenotype [31, 32].

Measuring DNA Methylation

Several mechanisms contribute to epigenetic variation, but DNA methylation is the most studied in ecological epigenetics [33]. DNA methylation occurs when a methyl group attaches to a cytosine in several possible sequence motifs (CpG, CpCpG, CpHpHp, and CpNpG) [34]. DNA methylation can have variable effects on gene activity, but methylation of the promoter region often decreases gene activity, because it interferes with transcription enzymes [27, 35, 36]. Epigenetic variation is more labile than genetic variation and changes in DNA methylation can occur in direct response to environmental cues, and independently from DNA sequence [9, 13, 14]. Because of this, environmental conditions can cause differential methylation and ultimately differential gene expression in response to those environmental conditions.

Methylation-sensitive amplified fragment length polymorphism (MS-AFLP) has been the primary technique used to study DNA methylation in population-level epigenetic studies, though next-generation sequencing-based techniques are promising to become more common [33]. There are several benefits to using MS-AFLP to assess DNA methylation in population-level studies [33]. Like standard AFLP, MS-AFLP offers a genome-wide scan of DNA methylation, and patterns in DNA methylation can be compared among individuals within sites, habitats, or other groups of interest, even in nonmodel organisms that lack a sequenced genome. The technique is also relatively economical, which accommodates the large sample sizes necessary for population-level studies. MS-AFLP also utilizes the same equipment and methods as AFLP, which allows the techniques to be used in concert to address ecological questions and directly compare genetic and epigenetic variation.

One major area of research for population-level epigenetic variation is the relationship between genetic and epigenetic variation. How can the epigenetic contribution to an organism's response be teased apart from genetic variation? The correlation between epigenetic and genetic variation may be categorized as obligate, facilitated, or pure (Fig. 9.1) [8, 9]. Obligate epigenetic variation is directly correlated with genetic variation and appears to follow the rules of Mendelian genetics [9]. Facilitated epigenetic variation is partially correlated with genetic variation, and genetic code directs the epigenetic variation but the outcome is context dependent. Pure epigenetic variation is independent from genetic variation. Both facilitated and pure tend to exhibit non-Mendelian segregation and both can be affected by stochasticity in addition to environmental conditions.

Researchers have begun designing elegant studies that control for genetic differences among individuals [14, 37–40]. These studies have suggested that some epigenetic variation is correlated with environment and some environmentally induced epigenetic variation is passed on to the next generation. These studies have also identified correlations between epigenetic markers and environmental conditions (often stressors), even in the absence of genetic variation. Though directly connecting the effect of epigenetic variation on phenotype is generally lacking, these studies suggest that epigenetic effects could be important to evolution.

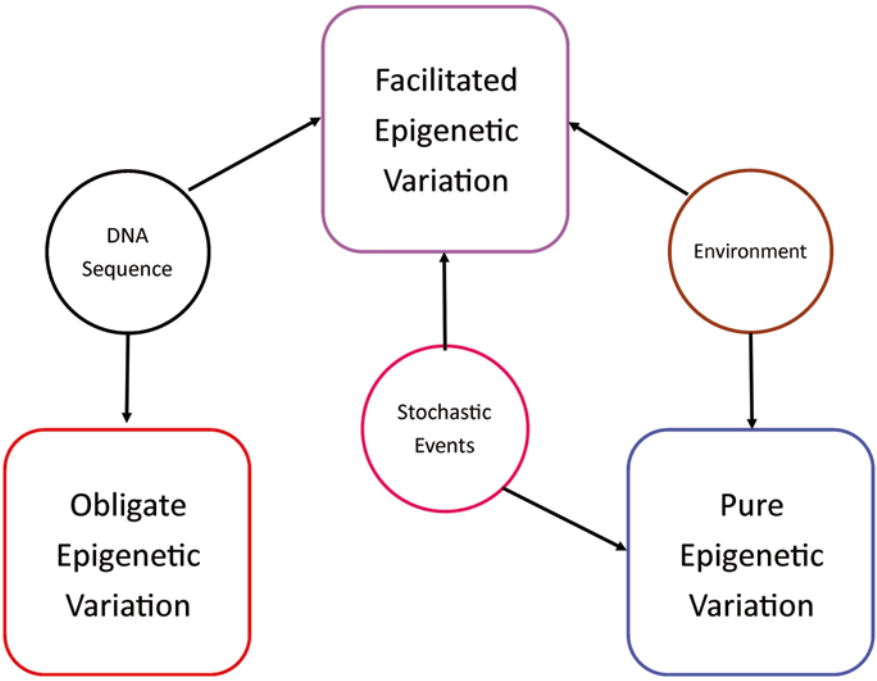

Fig. 9.1 Relationship among obligate, facilitated, and pure epigenetic variation with DNA sequence, stochastic events, and environmental conditions. (Adapted from [8, 9])

Questions We Should Be Asking

Richards [9] laid the foundation for population epigenetics by identifying four fundamental questions, and Bossdorf et al. [27] supplied a similar list of questions to ecologists. To date, with the exception of documenting epigenetic variation in natural populations, these questions remain largely unanswered.

1. How much epigenetic variation is present?
2. How independent is epigenetic variation from genetic variation?
3. What is the extent to which epigenetic states are inherited?
4. What is the importance of epigenetic variation in populations, whether inherited or not?

How Much Epigenetic Variation is Present?

In most epigenetics studies, epigenetic variation (i.e., DNA methylation) exceeds genetic variation [14, 37, 40–43]. This may be related to the much more rapid mutation rate in epigenetic than genetic contexts: replicate "mutation accumulation"

lines of a single *Arabidopsis thaliana* accession exhibited 10^3–10^4 more epigenetic mutations than genetic [44, 45].

Overall, "how much" epigenetic variation is present and the manner that the variation is generated may be critically important to understand the epigenetic response to environmental conditions. To address how much epigenetic variation there is, many ecologists have measured epigenetics in natural populations in the context of several ecological topics. These topics include identifying mechanisms for plasticity, rapid differentiation to habitat, and mechanisms which contribute to invasive species success.

Plasticity

Epigenetic diversity alone may be important for allowing organisms to respond to different environmental conditions. Using a classic phenotypic plasticity design, Bossdorf et al. [46] found that experimental alteration of DNA methylation altered important traits, and the plasticity of those traits, in response to nutrient addition in *A. thaliana*. Among 22 *A. thaliana* genotypes, there was also a difference in the *degree* to which trait means and plasticities were affected by the methylation inhibitor 5-azacytidine. Overall patterns of variability among the genotypes indicated that epigenetic changes can affect not only the short-term environmental responses of plants but also the evolutionary potential of important traits and their plasticities [46].

Epigenetic variation can also play an important role in the ability for clonal yeast (*Metschnikowia reukaufii*) to occupy different nectar habitats. Herrera et al. [37] grew clonal colonies of yeast in experimental media that mimicked naturally occurring differences in sugar combinations and concentrations of nectars to determine if the yeast exhibited epigenetic responses to the nectar variation. The clonal replicates controlled for genetic variation contributing to the response, and demonstrated that plasticity in resource use was correlated with changes in methylation (with MS-AFLP). The probability of an MS-AFLP marker changing from unmethylated to methylated was significantly correlated with sugar concentration, content, and their interaction. They also used 5-azacytidine to show that yeast growth response was significantly inhibited if epigenetic changes were prevented.

More recently, Herrera and Bazaga [39] examined the epigenetic signature between prickly and nonprickly leaves on *Ilex aquifolium,* English holly. *Ilex aquifolium* exhibits more prickly leaves in response to mammalian herbivory [47–49]. Both prickly and nonprickly leaves were sampled from five shrubs along a 450-m transect, and sampling the different leaf types from the same shrubs ensured that the genotypes were identical. Variation among MS-AFLP loci was significantly associated with leaf type, and the probability of methylation declined in nonprickly leaves compared to prickly. They found six MS-AFLP loci that distinguished prickly from nonprickly leaves, suggesting that the genes associated with these MS-AFLP loci could be important in mediating plasticity in leaf morphology in response to herbivory.

Habitat Differentiation

The studies exploring plasticity show that epigenetic diversity can contribute to response to environmental factors, suggesting that population epigenetic variation may be structured spatially based on environmental variables in natural populations. Differentiation of epigenetic marks among habitats could indicate that epigenetic effects allow organisms to adapt to local environmental conditions without a change in genetic code [27].

In white mangroves, both DNA methylation diversity and differentiation could be important for persistence in variable habitats. Lira-Medeiros et al. [50] examined epigenetic and genetic variation in *Laguncularia racemosa* from both riverine and salt marsh habitats in Brazil. Mangroves from both habitats exhibited higher levels of epigenetic variation than genetic variation as identified by comparing MS-AFLP and AFLP results. They also found epigenetic differentiation between mangroves from salt marsh and riverine habitats for *L. racemosa*, but no genetic differentiation. Plant phenotypes significantly differed between habitats for a number of ecologically relevant traits including height, diameter at breast height (DBH), leaf width, and leaf area. Though Lira-Medeiros et al. [50] found epigenetic population differentiation, it was not possible to determine if these differences were generated by stable or induced epigenetic effects, because they collected plant material from the field. Their findings suggest that epigenetics may play an important role in response to local habitat conditions; however, manipulative studies are required to determine the contribution of stable versus environmentally induced epigenetic marks.

Epigenetic diversity and differentiation could also be important in the response to herbivory of *Viola cazorlensis* [41]. Herrera and Bazaga [41] analyzed epigenetic and genetic diversity and differentiation among *V. cazorlensis* plants that had been exposed to 20 years of variable ungulate herbivory. As observed in white mangroves, *V. cazorlensis* also exhibited higher levels of epigenetic variation (e.g., DNA methylation) than genetic variation when comparing MS-AFLP and AFLP results. Additionally, *V. cazorlensis* exhibited both epigenetic and genetic population structure in response to herbivory. The epigenetic differences were correlated with genetic loci implicated with adaptive floral traits. These findings suggest that epigenetics could play an important role in flowering dynamics in the *V. cazorlensis* system that could ultimately result in the adaptive differentiation of populations over evolutionary time [41].

Population Bottlenecks

Epigenetic diversity may provide an additional source of variation to populations that have undergone reductions in genetic variation from demographic changes, like in invasive species. One consequence of population bottlenecks is inbreeding depression. Vergreer et al. [51] showed that epigenetics can play a role in inbreeding depression in the self-pollinating perennial plant *Scabiosa columbaria*. They inbred and outcrossed the plant and found that inbred plants exhibited decreased leaf number, biomass, and photosynthetic efficiency when compared with outcrossed plants.

They also found that genome-wide DNA methylation was 10 % higher in the inbred plants. Additionally, they treated some plants with 5-azacytidine, which decreased genome-wide methylation by 11 %. In the 5-azacytidine treatment, they found that biomass decreased in outcrossed plants and increased in inbred plants. Photosynthetic efficiency and leaf number were not affected in outcrossed plants, but inbred plants exhibited levels consistent with control, outcrossed plants. The biomass of inbred plants treated with 5-azacytidine was also significantly higher than control, inbred plants. The combination of these findings suggests that DNA methylation plays a role in mediating the effects of inbreeding depression for various traits.

These findings were supported in a study on house sparrows (*Passer domesticus*) introduced to Kenya in the 1950s [52]. Schrey et al. [53] showed that Kenyan house sparrows have lower levels of genetic variation than native populations, and Liebl et al. [54] showed that all house sparrows sampled from seven cities in Kenya exhibited different epigenotypes. Though the sparrows exhibited high levels of epigenetic variation, they did not exhibit epigenetic population structure among the sites. However, certain MS-AFLP loci were methylated more often in Kenyan populations when compared to native populations [53], which suggests that these loci could be involved in traits important to living in the new habitat. Liebl et al. [54] also discovered a trend suggesting the epigenetic variation may compensate for decreased genetic variation and increased inbreeding. Individuals exhibiting lower levels of genetic variation had higher inbreeding coefficients and exhibited higher levels of epigenetic variation. Overall, these findings suggest that epigenetic variation may compensate for decreased genetic variation in the early stages of invasion, thereby allowing the sparrows to occupy an expanded range of response that allows them to occupy many different habitats.

Different patterns of epigenetic variation in Japanese knotweed populations further support the hypothesis that epigenetic effects may contribute to response to local habitat. Richards et al. [40] sampled invasive Japanese knotweed (*Fallopia* spp.) individuals from roadside, beach, and salt marsh habitats across eastern New York state. Individuals were grown in a common garden to control for induced environmental effects prior to assessing genetic and epigenetic variation. Genetic diversity in these populations was low, with only four variable AFLP loci detected out of 200 loci scored. Epigenetic variation was significantly higher, with 19 variable loci detected out of 180 loci scored. Both genetic and epigenetic population structure was found among the different habitats, but habitat explained a significant proportion of the structure only for the epigenetic variation and not for genetic variation. Given the low genotypic diversity, these findings suggest that epigenetic effects may play an important role in knotweed's response to variable environmental conditions.

How Independent Is Epigenetic Variation from Genetic Variation?

The relationship between genetic and epigenetic variation can be complex [9], and determining how much epigenetic variation is obligate, facilitated, or pure has been challenging [27, 31]. Currently, the best way to address this question has been to

assess the importance of epigenetic variation in situations where genetic variation is lacking or minimal. Studies in clonal plants support the importance of pure or facilitated epigenetic effects by clonal species (i.e., dandelions, knotweed).

Verhoeven et al. [14] examined DNA methylation in response to five ecologically relevant treatments in replicates of a single dandelion (*Taraxacum officinale*) genotype. DNA methylation variation was significantly higher in response to each of three experimental treatments (i.e., low nutrients, jasmonic acid, and salicylic acid), than control, and the majority of the differences were inherited. Similarly, Richards et al. [40] found replicates of genetically identical individuals showed epigenetic differentiation to beach, roadside or salt marsh habitats. They grew the individuals in a common garden, and these patterns were persistent through clonal reproduction. The findings from these two clonal plant species suggest that epigenetic differences elicited by a single genotype in response to different environmental factors may persist into the next generation, and that the epigenetic effects are more than just a simple readout of the genotype.

Using clonal species is one way to control for the effect of genotype in population-level studies. In populations where it is challenging or impossible to control genetic variation among individuals with experimental design, researchers will have to find other ways to control for the effect of genotype. Adapting statistical tests to control for the genetic component of response is one way to do this. For example, we found both genetic and epigenetic population structure among habitat types in natural *Spartina alterniflora* populations (Foust et al. unpublished data). We used a partial Mantel test to correlate epigenetic distances with habitat (low, intermediate, and high-salt areas) while controlling for genetic distance. The partial Mantel test allowed us to find significant epigenetic population structure while controlling for genetic population structure. This is one example where statistics offered a way to tease apart the correlation between epigenetic and genetic variation.

What Is the Extent to Which Epigenetic States Are Inherited?

To be naturally selected, traits controlled by epigenetic variation must be transmitted to the next generation. Currently, there are few ecological epigenetic studies that directly connect heritable phenotypic traits to epigenetic marks. However, we know that epigenetic changes can be induced by environmental conditions, that epigenetic changes affect phenotype, and that these phenotypic changes can be passed to the next generation [7]. Epigenetic marks can also be highly conserved through transmission, with variability ranging from 1 % to 10^{-6} variations per generation [7, 8, 55], which means that induced epigenetic variation will likely persist in future generations.

Epigenetic recombinant inbred lines (epiRILs) are one valuable resource that have shed more light on this question [11, 56]. Two populations of epiRILs were developed from backcrosses of the methylation mutants *(ddm1* and *met1*) to Col-0 wild type in *A. thaliana*. Johannes et al. [11] showed that both flowering time and

plant height varied among epiRILs, and that these effects persisted for at least eight generations. While these findings were initially discovered in epiRILs created from one single genotype, Cortijo et al. [57] have confirmed that these epigenetic associations with phenotype are found across 138 natural accessions for which genome-wide methylation data and phenotype data were available [45, 58, 59].

Perhaps the most clearly defined example of epigenetic inheritance in a non-model species was discovered in *Linaria vulgaris* (toadflax) [12]. *Linaria vulgaris* exhibits radial flower symmetry when hypermethylation of the *Lcyc* gene, an epimutation, occurs. Individuals without the epimutation exhibit dorsoventral symmetry. This change in phenotype is ecologically important because it could affect pollination and overall fitness of the plant [60]. It could also be evolutionarily important because the epiallele can be directly transmitted to the next generation, but the mechanism is unclear because the epimutation is sometimes reset during somatic development [12]. More research in an ecological context is required to determine if this epimutation affects the evolutionary trajectory of the species.

Epigenetic inheritance also occurs in the perennial herbs, *Helleborus foetidus* [38] and *T. officinale* [14]. In both studies, MS-AFLP was used to identify patterns of epigenetic variation in parent plants and the next generation. *Helleborus foetidus* was sampled from three populations located on a latitudinal gradient in Spain. In addition to the parents, patterns of epigenetic variation were also assessed in the pollen produced by parent plants (i.e., the male gametophyte). Though there was some epigenetic resetting from the sporophyte to gametophyte generation, significant epigenetic population structure persisted among the gametophytes of the three populations, as predicted from the parental generation. Similarly, Verhoeven et al. [14] showed that between 74 and 92 % of epigenetic states present in the parent generation occurred in the offspring, even though the stressors (i.e., low nutrients, salt, jasmonic acid, salicylic acid, and control) were no longer present. These findings suggest that epigenetic population structure persists into future generations, which could help offspring respond to local environmental conditions [27].

What Is the Importance of Epigenetic Variation in Populations, Whether Inherited or Not?

Epigenetic studies of nonmodel organisms and studies at the population level typically do not directly identify the connection between epigenetic state and phenotype. This issue must be addressed to fully understand the importance of epigenetic variation. Theorists have developed models that show epigenetics can affect the evolutionary trajectories of organisms [16–19]. Bonduriansky and Day [18] recognize that epigenetics allows for the decoupling of genotype from phenotype, which releases some of the constraints on evolution as defined by Mendelian genetics. Day and Bonduriansky [61] developed a model that incorporates both genetic and epigenetic inheritance, and allows some genomic elements to accumulate more epigenetic variation than others [45]. Pál [16] and Pál and Miklós [17] found that epi-

genetic inheritance systems (e.g., transmission of DNA methylation patterns) can allow organisms to avoid valleys on an adaptive landscape. Ultimately, epigenetic inheritance can change the adaptive landscape. Thus, an organism that is trapped in an adaptive valley may be rescued by epigenetic variation. It is also possible that genetic inheritance may "take over" after epigenetic mechanisms have facilitated the transition across an adaptive valley (i.e., genetic assimilation [19, 21, 22, 62]).

Klironomos et al. [19] developed a model that showed epigenetics can affect the tempo and the overall outcome of evolution. They describe cases where adaptation occurs by epigenetic mechanisms, and epigenetic inheritance can ensure that the phenotype persists. Epigenetically derived phenotypes can "speed up" evolution and potentially change the trajectory of evolution. Geoghegan and Spencer [20] support this idea and show that incorporating epigenetic marks into selection models can also drastically alter trajectories. This is especially true considering that environmentally mediated epigenetic marks can regenerate depending on the scenario. This means that as the environment changes, epigenetic marks can revert and potentially take organisms back to the fitness peak that was pertinent to past environmental conditions. Each of these models demonstrates the potential importance of epigenetics to evolution. Incorporating empirical data to these theoretical advances will greatly expand our understanding of the importance of epigenetics.

Future Directions

The available epigenetics studies in natural populations and nonmodel organisms present compelling evidence that epigenetic mechanisms are important at the population level. However, there are major areas that need to be addressed by future studies. These include:

1. Understanding the relationship between genetic and epigenetic variation
2. Identifying stable versus environmentally induced epigenetic variation
3. Identifying the importance of inheritance of epigenetic variation

As epigenetic techniques and next-generation sequencing become more cost-effective, researchers should use these tools to address these issues in more detail [33]. Epigenetic sequencing coupled with next-generation sequencing will allow for cross comparisons with genomic information obtained from studies of model organisms on specific genes and gene networks associated with various environmental conditions.

Adapting statistical analyses to control for genetic response will be important in gaining understanding about the relationship between genetic and epigenetic variation. Adaptation of the partial Mantel test has been useful in uncovering epigenetic population structure associated with low, intermediate, and high-salt habitats, that is independent from genetic population structure in natural *S. alterniflora* populations (Foust et al. unpublished data). Other statistical tests could be modified to accomplish the same goal in systems where it is challenging or impossible to control

for genetic variation among individuals by experimental design. Using statistical approaches to control for genetic response will allow research to broaden to include more nonmodel organisms, thereby increasing our overall understanding of the relationship between genotype and epigenotype.

Another consideration that will be important in future studies of population epigenetics is the effect of nonstable epigenetic effects. which can be minimized in some cases with a modified experimental design. Most studies to date have collected samples in the field, thus confounding stable and induced DNA methylation. Richards et al. [40] controlled for these effects by collecting Japanese knotweed rhizomes in the field and growing them in common garden prior to assessing genetic and epigenetic variation. This allowed for any highly labile marks to "normalize," and stable marks that might contribute to habitat response to persist. Reciprocal crosses would help determine epigenetic response to local conditions, and address the potential "chicken or egg" scenario between the association of epigenetic marks and environmental conditions. For example, are the patterns of epigenetic variation present because the plants responded to environmental conditions? Or did plants that possess certain patterns of epigenetic variation thrive in specific habitats?

To address these questions, future studies should build upon current epigenetic surveys and begin including more manipulative experimental designs with nonmodel organisms in natural systems. Using multifaceted experimental designs will help to determine if patterns of epigenetic variation change in response to environmental conditions, which result in population structure, or if patterns of epigenetic variation that are already present enable an organism to live in a certain area [40].

Finally, to determine the importance of epigenetic variation to evolutionary questions, more multigenerational studies need to be performed. Ideally, multigenerational studies will be designed to address the questions brought forth in the current, proposed models. These questions include: what are the implications of phenotype being decoupled from genotype, how do epigenetic changes alter the adaptive landscape or allow populations to avoid adaptive valleys, does epigenetics really speed up evolutionary processes, and how does it change evolutionary trajectories? [16–20, 61]. To do this, researchers need to examine both short- and long-term epigenetic response to environmental conditions [19] and assess both genetic and epigenetic contributions to evolution.

Conclusions

Currently, the importance of epigenetics at the population level is often made via correlation. Manipulative field studies will need to be performed to determine if epigenetic population structure results from habitat response or if organisms with certain epigenetic signatures can more easily live in certain habitats [40]. Since patterns of epigenetic variation can be established in direct response to local environmental conditions, but can also be inherited from parents, it is important to address this causality problem to gain further understanding about how organisms respond

epigenetically to variable conditions. Understanding the different ways populations can respond to environmental conditions via epigenetics is also important, because population genetics ignores the effects of short-term and within-generation environmental variation on evolution.

Researchers have already begun addressing the questions Richards [9] set forth, and each of those papers further support the importance of epigenetics to ecologically-relevant traits in nonmodel and natural populations. The theoretical models presented in this chapter have established a good outline for the future direction of population epigenetics, and future studies should include multigenerational, population-level studies to test these models. Progress is being made in all of these areas; however, there is much that we do not know.

We do know that an extension of the MS that encompasses epigenetic mechanisms and inheritance is necessary [5]. It is becoming more clear, especially as more studies are performed at the population level, that epigenetics is important to short-term environmental response, and that DNA methylation changes in response to environmental conditions can persist into the future [63]. In some cases, these epigenetic responses provide an adaptive advantage to offspring that are subjected to those same environmental conditions [46]. To fully address these questions, laboratory-based studies on model organisms should be coupled with natural and manipulative studies on natural populations and nonmodel organisms to obtain a clearer picture of how epigenetics affects organisms [38].

References

1. Lewontin RC. The genetic basis of evolutionary change. West Sussex: Columbia University Press; 1974.
2. Pigliucci M. Do we need an extended evolutionary synthesis? Evolution 2007;61:2743–49.
3. Hartl DL, Jones EW. Essential genetics. 2nd ed. Sudbury: Jones and Bartlett; 1999.
4. Jablonka E, Lamb MJ. The epigenome in evolution: beyond the modern synthesis. VOGis Herald 2008;12:242–54.
5. Pigliucci M, Müller GB. Evolution – the extended synthesis. Cambridge: MIT Press; 2010.
6. Jablonka E, Lamb MJ. Evolution in four dimensions: genetic, epigenetic, behavioral, and symbolic. Cambridge: MIT Press; 2005.
7. Jablonka E, Raz G. Transgenerational epigenetic inheritance: prevalence, mechanisms, and implications for the study of heredity and evolution. Q Rev Biol. 2009;84:131–76.
8. Richards EJ. Inherited epigenetic variation – revisiting soft inheritance. Nat Rev Genet. 2006;7:395–401.
9. Richards EJ. Population epigenetics. Curr Opin Genet Dev. 2008;18:221–6.
10. Grandjean V, Yaman R, Cuzin F, Rassoulzadegan M. Inheritance of an epigenetic mark: the CpG DNA methyltransferase 1 is required for de novo establishment of complex patterns of non-CpG methylation. PLoS One. 2007;11:e1136.
11. Johannes F, Porcher E, Teixeira FK, Saliba-Combani V, Simon M, Agier N, Bulski A, Albuisson J, Heredia F, Audigier P, Bouchez D, Dillman C, Guerche P, Hospital F, Colot V. Assessing the impact of transgenerational epigenetic variation on complex traits. PLoS Genet. 2009;6:e1000530.
12. Cubas P, Vincent C, Coen E. An epigenetic mutation responsible for natural variation in floral symmetry. Nature. 1999;401:157–61.

13. Bollati V, Baccarelli A. Environmental epigenetics. Heredity. 2010;105:105–12.
14. Verhoeven KJF, Jansen JJ, van Dijk PJ, Biere A. Stress-induced DNA methylation changes and their heritability in asexual dandelions. New Phytol. 2010;185:1108–18.
15. Angers B, Castonguay E, Massicotte R. Environmentally induced phenotypes and DNA methylation: how to deal with unpredictable conditions until the next generation and after. Mol Ecol. 2010;19:1283–95.
16. Pál C. Plasticity, memory, and the adaptive landscape of the genotype. Proc R Soc Lond Biol. 1998;265:1319–23.
17. Pál C, Miklós I. Epigenetic inheritance, genetic assimilation, and speciation. J Theor Biol. 1999;200:19–37.
18. Bonduriansky R, Day T. Non-genetic inheritance and its evolutionary implications. Annu Rev Ecol Evol Syst. 2009;40:103–25.
19. Klironomos FD, Berg J, Collins S. How epigenetic mutations can affect genetic evolution: model and mechanism. BioEssays. 2013;35:571–8.
20. Geoghegan JL, Spencer HG. Population-epigenetic models of selection. Theor Pop Biol. 2012;81:232–42.
21. Waddington CH. Canalization of development and the inheritance of acquired characters. Nature. 1942;150:563–65.
22. Waddington CH. Genetic assimilation of an acquired character. Evolution. 1953;7:118–26.
23. Jablonka E, Lamb MJ. The changing concept of epigenetics. Ann NY Acad Sci. 2002;981:82–96.
24. Holliday R. Epigenetics: a historical review. Epigenetics. 2006;1:76–80.
25. Holliday R. Epigenetics: an overview. Dev Genet. 1994;15:453–7.
26. Bird A. Perceptions of epigenetics. Nature. 2007;447:396–8.
27. Bossdorf O, Richards CL, Pigliucci M. Epigenetics for ecologists. Ecol Lett. 2008;11:106–15.
28. Hallgrímsson B, Hall BK, editors. Epigenetics: linking genotype and phenotype in development and evolution. Berkeley: University of California Press; 2011.
29. Ho DH, Burggren WW. Epigenetics and transgenerational transfer: a physiological perspective. J Exp Biol. 2010;213:3–16.
30. Kovalchuk I. Transgenerational epigenetic inheritance in animals. Front Genet. 2012;3:1–2.
31. Ledón-Rettig CC, Richards CL, Martin LB. Epigenetics for behavioral ecologists. Behav Ecol. 2013;24:311–24.
32. Kilvitis H, Alvarez M, Foust CM, Robertson M, Schrey AW, Richards CL. Ecological epigenetics. In Landry C, Aubin-Horth N, editors. Ecological and evolutionary genomics. Dordrecht: Springer; 2014.
33. Schrey AW, Alvarez M, Foust CM, Kilvitis HJ, Lee JD, Liebl AL, Martin LB, Richards CL, Robertson M. Ecological epigenetics: beyond MS-AFLP. Integr Comp Biol. 2013;53:340–50.
34. Rapp RA, Wendel JF. Epigenetics and plant evolution. New Phytol. 2005;168:81–91.
35. Jaenisch R, Bird A. Epigenetic regulation of gene expression: how the genome integrates intrinsic and environmental signals. Nat Genet. 2003;33:245–54.
36. Jablonka E, Lamb MJ. The evolution of information in the major transitions. J Theor Biol. 2006;239:236–46.
37. Herrera CM, Pozo MI, Bazaga P. Jack of all nectars, master of most: DNA methylation and the epigenetic basis of niche width in a flower-living yeast. Mol Ecol. 2012;21:2602–16.
38. Herrera CM, Medrano M, Bazaga P. Epigenetic differentiation persists after male gametogenesis in natural populations of the perennial herb *Helleborus foetidus* (Ranunculaceae). PLoS One. 2013;7:e70730.
39. Herrera CM, Bazaga P. Epigenetic correlates of plant phenotypic plasticity: DNA methylation differs between prickly and nonprickly leaves in heterophyllous *Ilex aquifolium* (Aquifoliaceae) trees. Bot J Linean Soc. 2013;171:441–52.
40. Richards CL, Schrey AW, Pigliucci M. Invasion of diverse habitats by few Japanese knotweed genotypes is correlated with epigenetic differentiation. Ecol Lett. 2012;15:1016–25.
41. Herrera CM, Bazaga P. Epigenetic differentiation and relationship to adaptive genetic divergence in discrete populations of the violet *Viola cazorlensis*. New Phytol. 2010;187:867–76.

42. Massicotte R, Whitelaw E, Angers B. DNA methylation: a source of random variation in natural populations. Epigenetics. 2011;6:421–7.
43. Schrey AW, Coon CAC, Grispo MT, Awad M, Imboma T, McCoy ED, Mushinsky HR, Richards CL, Martin LB. Epigenetic variation may compensate for decreased genetic variation with introductions: a case study using house sparrows (*Passer domesticus*) on two continents. Genet Res Int. 2012;2012:1–7.
44. Becker C, Hagmann J, Müller J, Koenig D, Stegle O, Borgwardt K, Weigel D. Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. Nature. 2011;480:245–9.
45. Schmitz RJ, Schultz MD, Urich MA, Nery JR, Pelizzola M, Libiger O, Alix A, McCosh RB, Chen H, Schork NJ, Ecker JR. Patterns of population epigenomic diversity. Nature. 2013;495:193–8.
46. Bossdorf O, Richards CL, Pigliucci M. Experimental alteration of DNA methylation affects the phenotypic plasticity of ecologically relevant traits in *Arabidopsis thaliana*. Evol Ecol. 2010;24:541–53.
47. Supnick M. On the function of leaf spines in *Ilex opaca*. Bull Torrey Bot Club. 1983;110:228–30.
48. Potter DA, Kimmerer TW. Do holly leaf spines really deter herbivory? Oecologia 1988;75:216–21.
49. Obeso JR. The induction of spinescence in European holly leaves by browsing ungulates. Plant Ecol 1997;129:149–56.
50. Lira-Medeiros, CF, Parisod C, Fernandes RA, Mata CS, Cardoso MA, Ferreira PCG. Epigenetic variation in mangrove plants occurring in contrasting natural environment. PLoS One. 2010;5:e10326.
51. Vergreer P, Wagemaker NCAM, Ouborg NJ. Evidence for an epigenetic role in inbreeding depression. Biol Lett. 2012;8:798–801.
52. Anderson TR. Biology of the ubiquitous house sparrow: from genes to populations. New York: Oxford University Press; 2006.
53. Schrey AW, Grispo M, Awad M, Cook MB, McCoy ED, Mushinsky HR, Albayrak T, Bensch S, Burke T, Butler LK, Dor R, Fokidis HB, Jensen H, Imboma T, Kessler-Rios MM, Marzal A, Stewart IRK, Westerdahl H, Westneat DF, Zehtindjiev P, Martin LB. Broad-scale latitudinal patterns of genetic diversity among native European and introduced house sparrow (*Passer domesticus*) populations. Mol Ecol. 2011;20:1133–43.
54. Liebl AL, Schrey AW, Richards CL, Martin LB. Patterns of DNA methylation throughout a range expansion of an introduced songbird. Integr Comp Biol. 2013;53:351–8.
55. Genereux DP, Miner BE, Bergstrom CT, Laird CD. A population-epigenetic model to infer site-specific methylation rates from double-stranded DNA methylation patterns. Proc Natl Acad Sci. 2005;16:5802–07.
56. Reinders J, Wulff BH, Mirouze M, Marí-Ordóñez A, Dapp M, Rozhon W, Bucher E, Theiler G, Paszkowski J. Compromised stability of DNA methylation and transposon immobilization in mosaic Arabidopsis epigenomes. Genes Dev. 2009;23:939–50.
57. Cortijo S, Wardenaar R, Tatché MC, Gilly A, Etcheverny M, Labaclie K, Caillieux E, Hospital F, Aury JM, Wincker P, Roudier F, Jansen RC, Colot V, Johannes F. Mapping the transgenerational epigenetic basis for complex traits. Science. 2014;343:1045–1048.
58. Brachi B, Faure N, Horton M, Flahauw E, Vazquez A, Nordborg M, Bergelson J, Cuguen J, Roux F. Linkage and association mapping of *Arabidopsis thaliana* flowering time in nature. PLoS Genet. 2010;6:e1000940.
59. Atwel, S, Huang YS, Vilhja´lmsson BJ, Willems G, Horton M, Li Y, Meng D, Platt A, Tarone AM, Hu TT, Jiang R, Muliyati NW, Zhang X, Amer MA, Baxter I, Brachi B, Chory J, Dean C, Debieu M, de Meaux J, Ecker JR, Faure N, Kniskern JM, Jones JDG, Michael T, Nemri A, Roux F, Salt DE, Tang C, Todesco M, Traw MB, Weigel D, Marjoram P, Borevitz JO, Bergelson J, Nordborg M. Genome-wide association study of 107 phenotypes in *Arabidopsis thaliana* inbred lines. Nature. 2010;465:627–31.

60. Fenster CB, Armbruster WS, Wilson P, Dudash MR, Thomson JD. Pollination syndromes and floral specialization. Annu Rev Ecol Evol Syst. 2004;35:375–403.
61. Day T, Bonduriansky R. A unified approach to the evolutionary consequences of genetic and non-genetic inheritance. Am Nat. 2011;178:E18–36.
62. West-Eberhard MJ. Developmental plasticity and the origin of species differences. Proc Natl Acad Sci. 2005;102:6543–9.
63. Duckworth RA. Epigenetic inheritance systems act as a bridge between ecological and evolutionary timescales. Behav Ecol. 2013;24:327–8.

Index

O. Pontes, H. Jin (eds.), *Nuclear Functions in Plant Transcription, Signaling and Development,* DOI 10.1007/978-1-4939-2386-1