Molecular Plant • Volume 7 • Number 3 • Pages 472–480 • March 2014 REVIEW ARTICLE

Covering Your Bases: Inheritance of DNA Methylation in Plant Genomes

Chad E. Niederhuth and **Robert J. Schmitz**[1]

Department of Genetics, University of Georgia, Athens, GA 30602, USA

ABSTRACT Cytosine methylation is an important base modification that is inherited across mitotic and meiotic cell divisions in plant genomes. Heritable methylation variants can contribute to within-species phenotypic variation. Few methylation variants were known until recently, making it possible to begin to address major unanswered questions: the extent of natural methylation variation within plant genomes, its effects on phenotypic variation, its degree of dependence on genotype, and how it fits into an evolutionary context. Techniques like whole-genome bisulfite sequencing (WGBS) make it possible to determine cytosine methylation states at single-base resolution across entire genomes and populations. Application of this method to natural and novel experimental populations is revealing answers to these long-standing questions about the role of DNA methylation in plant genomes.

Key words: DNA methylation; epigenetics; epiallele; whole-genome bisulfite sequencing.

INTRODUCTION

Methylation of cytosine bases in plant genomes is one mechanism associated with variation in gene expression (Finnegan et al., 1996). DNA methylation is covalently attached to cytosines and, as a result, is inherited through mitosis and/or meiosis (Finnegan et al., 1998; Calarco et al., 2012). There is limited evidence that the methylation status of genes is altered and inherited to the next generation as a result of the environment, but there is strong evidence that spontaneous methylation variants arise and do not always adhere to Mendel's law. Because of their interesting patterns of inheritance and the potential for passing on altered gene expression states to the next generation, as a result of environmental perturbation, there is much interest and even controversy in this field (Weigel and Colot, 2012). Until recently, only a limited number of methylation variants had been identified. This has left much unknown about the inheritance patterns of these variants, the extent of natural variation, and the degree of dependence on genetic variation (Schmitz and Ecker, 2012). Answering these questions will be of consequence to understanding plant developmental and evolutionary processes and may have practical applications in plant breeding.

Next-generation sequencing methods, like whole-genome bisulfite sequencing (WGBS), allow the state of methylation to be determined at single-base resolution across the entire genome. This was first applied to the model plant *Arabidopsis thaliana* (Cokus et al., 2008; Lister et al., 2008) and, since then, has been applied to many plant species as well as non-plant species, including humans (Lister et al., 2009). Population epigenomic approaches (Schmitz and Zhang, 2011) that integrate these data with expression data from RNA-seq and small-RNA sequencing or genomic data from whole-genome resequencing makes it possible to infer the relationship between methylation, gene expression, and genetic variation. Application of these techniques to natural and novel experimental populations will enable answers to the long-standing questions regarding the role and effects of DNA methylation in creating phenotypic variation.

Cytosine residues are methylated in three sequence contexts, CG, CHG, and CHH (where H = A, T, or C) (Law and Jacobsen, 2010). The molecular pathways that induce and maintain them further differentiate each of these three types of methylation. For example, CG methylation is maintained by the MET1 DNA methyltransferase (Finnegan et al., 1996; Ronemus et al., 1996), CHG methylation is maintained in *Arabidopsis* by the methyltransferase CMT3 (Du et al., 2012), whereas CHH is targeted by the activities of either DRM1/DRM2 or CMT2 methyltransferases (Chan et al., 2006; Zemach et al., 2013).

[1] To whom correspondence should be addressed. E-mail schmitz@uga.edu, fax 706 542 3910, tel. 706-5421882

© The Author 2013. Published by the Molecular Plant Shanghai Editorial Office in association with Oxford University Press on behalf of CSPB and IPPE, SIBS, CAS.

doi:10.1093/mp/sst165, Advance Access publication 22 November 2013

Received 8 October 2013; accepted 11 November 2013

Transposons, repetitive elements, and some genes are silenced when methylated in all three sequence contexts. These are often targets of the RNA-directed DNA Methylation pathway (RdDM) guided by 24 nucleotide (nt) short interfering RNAs (siRNAs) (Law and Jacobsen, 2010). In some cases, methylation spreads out from the transposon, leading to silencing of nearby loci (Hollister and Gaut, 2009; Ahmed et al., 2011; Eichten et al., 2012). Interestingly, this phenomenon in maize appears to be specific to certain transposon families (Eichten et al., 2012). Methylation does not always lead to silencing. In fact, methylation of only CG residues and the lack of 24-nt siRNAs in gene bodies is associated with moderately high levels of expression (Zhang et al., 2006; Zilberman et al., 2006; Coleman-Derr and Zilberman, 2012).

The maintenance of DNA methylation is an essential process to maintain plant genome integrity. When CG methylation is lost in *met1* mutants, expression of DNA demethylases is reduced in addition to retargeting of H3K9 methylation to heterochromatin in a methylation-independent manner (Mathieu et al., 2007). RNAi mechanisms also appear to play a role in restoring and correcting errors in DNA methylation induced by loss of *ddm1* (Teixeira et al., 2009). A recent extensive study of 86 silencing mutants involved in DNA methylation has helped reveal both the specificity of each type of methylation as well as the degree of interaction amongst pathways (Stroud et al., 2013b). As much interest as there is in the potential for the environment to influence DNA methylation states, there is much stronger evidence that failure to maintain methylation states is a major source of DNA methylation variation.

In this review, we define epigenetic as mitotically and/or meiotically heritable variation in phenotype that is independent of genotype. Methylation levels at certain regions of the genome vary between individuals and are often referred to as epigenetic alleles or 'epialleles'. DNA methylation also shows varying dependence upon genetic variation, often resulting in confusion with regard to the epigenetic nature of these alleles. To clarify this conundrum, epialleles are classified into three groups based on the relative dependence or independence of genotype (Richards, 2006). At the highest level of dependence, obligate epialleles are directly determined by genetic variants and co-segregate with these methylation variants (Bender and Fink, 1995; Liu et al., 2004; Woo et al., 2007). For example, methylation of a gene may be dependent upon the presence or absence of a nearby transposon. Facilitated epialleles, although linked to and even caused by a genetic variant, are not fully dependent upon it and exhibit greater instability than obligate epialleles. In the example of methylation spreading into a gene after the insertion of a neighboring transposon, methylation of the gene is maintained across generations even after the facilitating transposon is excised or segregated away. At the other end of the spectrum are pure epialleles. As the name suggests, these are completely independent of genetic variation. Although these definitions serve to clarify mechanisms of epiallelic variation, it is often difficult to ascertain the distinguishing feature of each group, blurring the lines in their definition. Making such distinctions and determining the prevalence of each kind of epiallele will require individual investigation.

NATURALLY OCCURRING METHYLATION VARIANTS

Naturally occurring methylation variants have been studied in *Linaria vulgaris*, *Zea mays* (maize), *Cucumis melo* (melon), *Solanum lycopersicum* (tomato), and *Arabidopsis* (Patterson et al., 1993; Bender and Fink, 1995; Cubas et al., 1999; Manning et al., 2006; Martin et al., 2009; Durand et al., 2012; Silveira et al., 2013). The list of known epialleles was limited prior to genome-wide methods and many have only been discovered in recent years. Oftentimes, these were observed only through easily recognized phenotypes that showed unusual instability in their inheritance. It is slightly ironic then that one of the earliest described mutants should turn out to be an epiallele. Carl Linnaeus described the *peloric* mutants of *Linaria* as far back as 1749 (Cubas et al., 1999). Flowers of these mutants are radially symmetric whereas wild-type *Linaria* has dorsoventral asymmetry. This is similar to mutants of the *cycloidea* gene of *Antirrhinum majus* (snapdragon), which was mapped to *Lcyc*, a homolog in *Linaria.* Sequencing failed to reveal any genetic mutations, but differences in methylation of *Lcyc* were detected, suggesting that the *peloric* phenotype was due to methylation variation.

Maize has long been a source of interesting epigenetic phenomena. A particularly interesting example is paramutation, which was first discovered in maize back in the 1950s (Brink, 1956). Paramutation involves the interaction of two alleles, in which the inducing paramutagenic allele causes heritable changes to a susceptible paramutable allele, violating the rules of Mendelian inheritance. Several examples of paramutation are known, including the *r*, *b*, *p1*, and *pl* loci of maize (Brink, 1956; Coe, 1968; Hollick et al., 1995; Sidorenko and Peterson, 2001). The exact mechanisms of each these paramutations is not fully understood, but sequencing shows that these loci are very complex, often containing repeats and inverted copies. Reports vary as to the role of methylation in paramutation. For example, with the *r* gene, crossing the paramutagenic *R-st* allele to the paramutable *R-r* allele results in heritable reductions in *r* expression and a loss of kernel pigmentation. This is associated with changes in the state of methylation at the *r* locus (Walker, 1998). In contrast, no detectable changes in methylation have been observed in the *b* locus, although conversion of the intensely pigmented *B-I* allele by the weakly pigmented *B'* allele is associated with reduced expression of the *b* gene (Patterson et al., 1993).

In melon, the sexual identity of flowers is determined by two loci. The andromonoecious (*a*) locus controls the development of stamen and the gynoecious (*g*) locus affects carpel development. The identity of the *g* locus was mapped to a transposon insertion upstream of a transcription factor,

CmWIP1 (Martin et al., 2009). Methylation spread from the transposon to *CmWIP1* leading to its silencing, whereas, in rare cases of partially reverted flowers, methylation of *CmWIP1* had been lost. This leads to plants whose flowers are hermaphroditic or gynoecious. In a screen of 497 accessions of melon, this insertion was found only in the hermaphroditic and gynoecious accessions. This example from melon clearly supports the strong relationship between genetic variants targeted by DNA methylation and the potential for DNA methylation to spread into neighboring sequences to influence gene expression states.

The *PAI* genes from *Arabidopsis* are excellent examples of obligate epialleles. The Columbia (Col) wild-type accession carries three copies of *PAI* of high sequence identity at the nucleotide level (Bender and Fink, 1995). Whereas *PAI1* and *PAI2* are active in Col, the *PAI3* locus has low expression and little apparent effect. In the Wassilewskija (Ws) accession there is a fourth gene, *PAI4*, an inverted duplication adjacent to the 3′ end of *PAI1.* As a result, the inverted duplication facilitates the methylation and silencing of the entire gene family. This was demonstrated by introduction of the *PAI4* locus into Col through genetic crossing, which induced *de novo* methylation of the three *PAI* Col loci (Luff et al., 1999). Although methylation in symmetrical contexts (CG and CHG) persisted for several generations in Ws upon removal of *PAI4*, it did not persist in asymmetrical contexts (CHH). At a population level, several other wild *Arabidopsis* accessions also contain the *PAI4* inverted duplication and methylation of the *PAI* gene family. However, most accessions studied contain only the three unmethylated copies found in Col (Melquist et al., 1999).

A more recent example from *Arabidopsis* is the *FOLT1*/*FOLT2* genes (Durand et al., 2012). *FOLT1* encodes a folate transporter and is located at the K5 locus on chromosome 5. A second full-length copy, *FOLT2*, along with two truncated copies, is located on the K4 locus on chromosome 4 in the Shahdara (Sha) and the C24 accessions. These copies are absent in Col. In recombinant inbred line (RIL) populations of Col and Sha, genetic incompatibility was found between the Col K4 locus and Sha K5 locus, resulting in distorted segregation ratios. Similar genetic incompatibility is also found in Col/C24 crosses and with other accessions carrying *FOLT2.* Further investigation into this incompatibility revealed that the *FOLT1* alleles in Sha were methylated and silenced by RdDM and that the truncated *FOLT2* copies on K4 were potential sources of siRNAs. Silencing of *FOLT1* was maintained for multiple generations even after the inducing locus was segregated away, making this an excellent example of a facilitated epiallele.

Colorless non-ripening (*Cnr*) is a naturally occurring mutant in tomato first identified in a population of commercial varieties (Thompson et al., 1999). This mutant plant is dominant and results in fruit lacking color and with decreased cell adhesion. Genetic mapping of the *Cnr* locus and sequencing failed to reveal any changes in DNA sequence between mutants and wild-type. However, differential expression of an SBP-box transcription factor called *LeSPL–CNR* was detected, suggesting the possibility of silencing by DNA methylation. Supporting this was the observation of rare reversions of *Cnr* to wild-type. Bisulfite sequencing of *LeSPL–CNR* showed significantly higher levels of methylation in the promoter region of *LeSPL–CNR* in mutant plants, suggesting that *Cnr* may be a rare example of a pure epiallele (Manning et al., 2006).

The *QUA-QUINE STARCH* (*QQS*) gene is a recently evolved *de novo* gene in *Arabidopsis* and shows extensive variation in DNA methylation in natural accessions and in the lab (Silveira et al., 2013). *QQS* is of particular interest because it is a clear example of a pure epiallele. In 36 different natural accessions, *QQS* was methylated in 29 and unmethylated in seven. Although *QQS* is surrounded by transposable elements, *QQS* alleles were consistently methylated in all lines studied. Nor was there any correlation between neighboring genetic variants and the state of methylation. In RIL lines of Cvi (Cape Verde Islands) and Col, the hypomethylated Cvi copy of *QQS* was stably inherited and unaffected by variants in other regions of the genome. The combined evidence suggests that methylation variants of *QQS* in natural accessions are independent of the genetic background.

INSIGHTS FROM epiRILS

A twist to the traditional method of making RILs in the mapping of quantitative trait loci was recently applied to analyzing the effects of variation in methylation on phenotypes (Johannes et al., 2009; Reinders et al., 2009). These epigenetic recombinant inbred lines (epiRILs) exploit the extensive genome-wide reductions in methylation in mutants of *decreased dna methylation 1* (*ddm1*) (Vongs et al., 1993) or *methyltransferase 1* (*met1*) (Finnegan et al., 1996; Ronemus et al., 1996). By crossing *ddm1* or *met1* mutants in the Col accession to wild-type Col, subsequent segregating populations can be generated that include individuals that are homozygous for their respective wild-type alleles. This creates a unique situation where the methylation status of the progeny is disrupted even as, in large part, the genetic background is controlled for. Carrying these lines through multiple rounds of selfing and single-seed descent results in a population of plants varying in their methylation state, but essentially identical genetically. As these methylation variants are induced through genetic perturbation, they do not necessarily reveal natural variation, but rather the hidden or 'cryptic' variation in plant genomes that is normally silenced.

These epiRIL populations show widespread phenotypic variance for both responses to biotic and abiotic stressors and developmental traits like flowering time and plant height (Johannes et al., 2009; Reinders et al., 2009). The heritability of developmental traits in *DDM1* epiRILs was estimated to be ~25%–30% (Roux et al., 2011), which is similar to observations made in natural accessions. These lines offer the opportunity for the mapping of quantitative traits to methylation variation. However, this is somewhat complicated by several

sources of variation within the epiRILs themselves. In *met1* and *ddm1* epiRILs, remobilization of transposable elements leads to genetic variation, but with careful analysis this can be controlled for using large populations (Richards, 2009).

The epiRIL populations have already spawned a number of reports providing further insight into basic mechanisms requiring DNA methylation, such as the aforementioned *QQS* epiallele. They have also been used to study the effects of DNA methylation on recombination (Johannes et al., 2009; Reinders et al., 2009). Along with comparisons between wild-type and *met1* mutants, these results indicate that DNA methylation affects the distribution of cross-over events, but does not affect the rate at which it happens (Colome-Tatche et al., 2012; Mirouze et al., 2012). In *met1* mutants, it has been observed that many transposons remain silenced despite transcriptional activation upon the loss of CG methylation, suggesting that posttranscriptional processes may suppress them. In *MET1* epiRILs, the retrotransposon *Evade* (*EVD*) was shown to be an exception to this, causing noticeable disruptive mutations in genes (Mirouze et al., 2012). Taking advantage of epiRIL individuals where only a single copy of *EVD* was reactivated, the nature and course of cellular responses to increasing transposition and copy number were tracked (Mari-Ordonez et al., 2013). At low copy numbers, *EVD* remained unmethylated, but was associated with an abundance of 21–22-nt siRNAs. Analysis of the *EVD* transcript showed that it possesses a GAG domain that appears to protect the transcript from degradation by RNAi. However, a threshold level appears to exist at ~40 copies, at which the presence of 24-nt siRNAs could be detected, followed by methylation and silencing of *EVD* by the RdDM pathway. These results were recapitulated using transgenic approaches and provide new insight into the course of events leading to methylation and silencing of transposons.

POPULATION STUDIES OF DNA METHYLATION

Aforementioned studies of individual methylation variants have raised more questions than answers on the extent, stability, and nature of DNA methylation in plant genomes. Although examples of obligate, facilitative, and pure epialleles are known, how common these are in natural populations and how they are inherited are relatively unknown. Answering these questions will be essential in understanding the importance of methylation variation in not only an evolutionary context, but also to basic plant biology and for practical applications such as crop breeding. Whole-genome approaches that leverage techniques from both genomics and population genetics are now being used to answer these questions.

One of the first studies to take this approach used the methylation-dependent restriction enzyme MrcBC and tiling arrays to compare differences in DNA methylation between the Col and Landsberg *erecta* (Ler) accessions of *Arabidopsis* (Vaughn et al., 2007). Gene-body methylation for two loci in Col x Ler crosses was followed in the F_1 and F_2 generations, showing that it was inherited primarily with the parental genotype. However, it was unstable and loss of methylation was observed in some lines in the F_1 generation. This was followed by a wider examination of 96 different *Arabidopsis* accessions, where widespread variation in DNA methylation was found. An early maize epigenomic study also revealed that there is widespread natural variation of DNA methylation (Eichten et al., 2011). Additionally, using identity by descent analysis, evidence for the existence of rare pure epialleles was identified, which were stably inherited upon examination of near-isogenic lines.

The extent of methylation variation at single-base resolution in a population was recently examined by sequencing the methylomes of 152 naturally occurring *Arabidopsis* accessions (Schmitz et al., 2013b). The effects on gene expression and interaction with the genome were determined by integrating these data with the transcriptomes, as determined by RNA-seq, of 144 lines and genome resequencing of 217 lines. Extensive variation was found in DNA methylation between accessions. Differentially Methylated Regions (DMRs) can occur in the CG context (CG-DMRs) or the CG, CHG, and CHH contexts (C-DMRs). By treating DMRs as a phenotype, it is possible to identify genetic loci, methylQTL, that may explain the differences in methylation using association mapping. In fact, using a similar approach in maize revealed that slightly over half (~51%) of the identified DMRs were associated with local genetic variants (Eichten et al., 2013).

RILs are an ideal population for studying the inheritance of methylation as the segregation and linkage of epialleles and genetic variants can be traced and used to map methylQTL because allele frequencies are typically balanced as opposed to natural populations. In *Glycine max* (soybeans), the methylomes of 83 RILs and their parental lines were determined by WGBS (Schmitz et al., 2013a). The majority (~91%) of C-DMRs identified were associated with a methylQTL. Of these methylQTL, ~97% were localized to the same chromosome as the C-DMR whereas only ~3% mapped to another chromosome. The remaining ~9% of C-DMRs that were not associated with a region of the genome likely represent possible facilitated and pure epialleles. These data support the conclusion that a subset of these epialleles is due to genetic variants, similar to results from association mapping in *Arabidopsis*. However, it is not possible to rule out that some of these may be pure epialleles that are stably inherited over the limited number of generations used to create the RILs. Long-term studies would be needed to address this likely alternative explanation.

The methylomes of maize B73, Mo17, and 9 RILs from a B73 x Mo17 cross were determined by WGBS (Regulski et al., 2013). Similarly, this study found that CG methylation largely segregated with the parental genotype, further supporting the observation that most epialleles are directly associated with genetic variants. However, 1772 DMRs were observed to switch genotypes and it was estimated that 10% or more of

DMRs in maize might undergo paramutation-like changes in methylation, although contrasting reports were revealed in a separate maize RIL (Eichten et al., 2013). In the lines analyzed (Regulski et al., 2013), it was far more common for the DMRs to be hypermethylated as opposed to hypomethylated and also more common for the hypermethylation to come from Mo17. Interestingly, enrichment of CHG methylation was observed around the acceptor sites of exon–intron junctions in some genes. The authors suggest that this methylation may affect the efficiency of splicing or be involved in alternative splicing. RNA-Seq data seemed to support this hypothesis, as transcripts associated with CHG methylation at acceptor sites were less efficiently spliced. This hints that there may be additional roles for methylation in plants.

Both association mapping and the analysis of RILs were recently applied to methylation in maize and teosinte (Eichten et al., 2013). The methylation of 20 inbred maize lines was determined using meDIP-chip profiling. This was then extended to a set of 11 teosinte inbred lines and an additional 20 maize inbred lines. From this, 1966 common DMRs found in three or more lines and 1754 rare DMRs found in only one or two lines were identified. The levels of DMRs seemed to be fairly constant across genotypes and clustering of lines by DMRs showed similar results to that of clustering by SNPs. This is similar to observations made in *Arabidopsis* (Schmitz et al., 2013b). Between maize and teosinte, 172 DMRs were identified that formed distinctive clusters separating the two (Eichten et al., 2013). Only ~10% of these, however, were found in regions of known selective sweeps in maize, indicating that selection during domestication alone could not explain the differences in methylation.

Most of the epigenomic studies in plant species thus far have primarily focused on variation in DNA methylation, but analysis of regions of the epigenome that are invariably methylated in the *Arabidopsis* data set described previously revealed that targets of RdDM were specifically reactivated in pollen (Schmitz et al., 2013b). This corroborates previous observations that show that RdDM is repressed in pollen, leading to activation of transposons (Slotkin et al., 2009; Calarco et al., 2012; Ibarra et al., 2012). Interestingly, this reactivation is spatially restricted to the vegetative nucleus of pollen (Slotkin et al., 2009). Reactivation of some of these targets leads to the production of 24-nt siRNAs that may reinforce RdDM silencing in the germ line. It is proposed that this reactivation may be a mechanism for faithfully transmitting methylation of RdDM targets to the next generation (Slotkin et al., 2009). There also appears to be an additional role, as many of the genes reactivated are also important for pollen tube growth and development (Schmitz et al., 2013b). How ubiquitous this mechanism is outside of *Arabidopsis* has yet to be determined.

As these studies demonstrate, high-throughput methods are revolutionizing the study of methylation by identifying thousands of new epialleles, just as such technologies are revealing the extent of genetic variants. This is further making it possible to apply the methods of population genetics to methylation with the ultimate goal of understanding the extent of methylation variation and how it co-evolves with genetic variation to give rise to phenotypic variation.

METHYLATION VARIANTS IN EVOLUTION

Analogous to genetic mutations, it is known that spontaneous conversions from one state of methylation to another do occur (metastability), as in the case of the *peloric* and *CNR* epialleles (Cubas et al., 1999; Manning et al., 1999). Mutation accumulation (MA) lines in *Arabidopsis* were studied for their phenotypic divergence (Shaw et al., 2000) and used to estimate the rate and scope of spontaneous mutations (Ossowski et al., 2010). In two separate studies, WGBS was applied to the ancestors and descendants of these lines spanning 30 generations (Becker et al., 2011; Schmitz et al., 2011). These results show that, at the whole-genome level, methylation is remarkably stable and heritable, with methylation of transposable elements being the most stable and consistent across lines. The rates of C-DMR formation were similar to the rate of spontaneous mutations occurring at about one event per generation per line. Methylation of CG residues in gene bodies, however, was less stable and this likely reflects the maintenance of CG methylation in gene bodies by the maintenance methyltransferase MET1. These types of epimutations accumulate over time at a rate of 4.46×10^{-4} methylation polymorphisms per CG site per generation (Schmitz et al., 2011), around four orders higher than the previously estimated genetic mutation rate in the same population (Ossowski et al., 2010). It is important to note, however, that these epimutations do not accumulate linearly and their functional consequences in plant genomes have not yet been established. The instability of some epialleles means that reversion is common and can cycle over short time scales (Becker et al., 2011) with uncertain implications for evolution.

Understanding the potential role of methylation in evolution has been studied by cross-species comparisons of methylomes. Two studies have examined deep comparisons across eukaryotes, including multiple plant species (Feng et al., 2010; Zemach et al., 2010), whereas a third compared the methylomes of the model grass *Brachypodium distachyon* and *Oryza sativa* ssp. *japonica* (rice) (Takuno and Gaut, 2013). Methylation in all three contexts was found in all plants and green algae, as was methylation of gene bodies and transposons. Although CG-gene-body methylation was abundant in angiosperms, very little of it was found in two species that diverged early from the angiosperms: *Selaginella moellendorffii* and *Physcomitrella patens* (Zemach et al., 2010). These results show that methylation and silencing in plants have very ancient origins. In contrast, non-plant species predominantly had only CG methylation (Feng et al., 2010; Zemach et al., 2010).

Comparison of Brachypodium and rice revealed a great deal of conservation in gene-body methylation amongst orthologs (Takuno and Gaut, 2013). In particular, there is a clear bias towards genes with a low ratio of cytosines in the CG context and genes that tended to evolve at much slower rates. Despite the reported methylation polymorphism of *Arabidopsis*, a reanalysis of *Arabidopsis* methylomes and comparisons to orthologs in *Arabidopsis lyrata* revealed similar trends. In contrast, there was much more variance in CHH methylation between Brachypodium and rice. The authors suggest that SMPs may have little evolutionary or functional consequence, in comparison to methylation levels within a region, and vary greatly as long as regional methylation levels remain above a certain threshold. Although the methylation states of orthologs are highly conserved, evidence from analysis of the soybean methylome revealed possible roles for methylation in variation of gene expression of paralogs, as substantial differential targeting by RdDM was observed (Schmitz et al., 2013a). Soybeans are known to have undergone two recent whole-genome duplications in their evolution (Schmutz et al., 2010). Compared to the *Arabidopsis* methylome, a much greater percentage of cytosines were methylated in soybean and far more protein-coding genes were targeted by RdDM (Schmitz et al., 2013a). Further analysis showed that more recent gene copies were preferentially methylated, suggesting that methylation and targeting by RdDM may be a mechanism for coping with the effects of whole-genome duplication, silencing new genes until they have a change to either undergo sub- or neo-functionalization or are purged from the genome.

ENVIRONMENTALLY INDUCED VARIATION OF DNA METHYLATION

Rare examples of DNA methylation states that are affected by the environment and are associated with the inheritance of an environmentally induced trait have been observed, often with controversy (Weigel and Colot, 2012). Perhaps the clearest example comes from rice tissue cultures, where extensive and stochastically induced hypomethylation was observed between individual plants (Stroud et al., 2013a). This was found even in individuals derived from the same parental plant. A more direct relationship between DNA methylation states and environmental cues was recently observed in tomato fruit ripening and upon bacterial infection in *Arabidopsis* (Dowen et al., 2012; Zhong et al., 2013). In both reports, no evidence was presented for the effects of these environmentally induced alterations on DNA methylation in subsequent generations. Furthermore, in both studies, the effects on DNA methylation were widespread throughout regions of the genome targeted by RdDM, indicating that some responses to the environment may be tightly linked to opposing the activity of this pathway. This area of investigation will clearly grow in the coming years, but meticulous efforts must be taken when designing and analyzing the potential for the environment to create heritable phenotypes via altering DNA methylation states.

CONCLUSIONS AND FUTURE PROSPECTS

The often-convoluted connections between genetic variation, the environment, and phenotype are central to the questions of biology. Sitting at the boundary of these in plant genomes is DNA methylation. Past studies of DNA methylation have revealed a significant amount regarding the underlying mechanisms, but were limited by the number of known epialleles. This has left questions of the extent of methylation variation, its links to genetic variation, and its evolutionary consequences largely unaddressed. The development of genome-wide technologies capable of capturing methylation states at single-base resolution such as WGBS (Lister and Ecker, 2009) makes it possible to see the extent of methylation variation as never before. Application of this technology to natural populations, previously unstudied species, and experimental populations is beginning to answer these unresolved questions.

Several trends have begun to emerge from these new data. Methylation variation is extensive in plant populations (Vaughn et al., 2007; Eichten et al., 2013; Schmitz et al., 2013b) and the rate at which epimutations occur is several-fold higher than the normal rate of mutation (Becker et al., 2011; Schmitz et al., 2011). However, a subset of epialleles are closely associated with genetic variation, both locally (*cis*) and distantly (*trans*), segregating with this variation in a Mendelian fashion (Vaughn et al., 2007; Eichten et al., 2013; Regulski et al., 2013; Schmitz et al., 2013a, 2013b). This challenges popular conceptions of DNA methylation variation being independent of the genome. That being said, a significant number of epialleles appear to not be associated with genotype and may be examples of pure epialleles (Regulski et al., 2013; Schmitz et al., 2013a, 2013b). The use of epiRIL lines is helping to reveal the extent of cryptic variation on quantitative traits and, at the same time, contributing to further understanding of the molecular mechanisms leading to methylation variation, as in the case of *de novo* methylation of transposons (Johannes et al., 2009; Reinders et al., 2009; Mari-Ordonez et al., 2013). Comparison between species shows surprising conservation of methylation amongst slow-evolving orthologs and that methylation and silencing of transposons are an ancient process in plants (Feng et al., 2010; Zemach et al., 2010; Takuno and Gaut, 2013). The role of DNA methylation variants and their contribution to the evolution of plant species are still unknown but, with the ever-growing list of DNA methylation variants, this will be a major area of investigation in the coming years.

Potential exists for the application of these findings to the breeding and improvement of crop species, as epialleles could be a potential novel source of variation that plant breeders could work with (Springer, 2013). Hints of this are seen in

the phenotypic variation observed in epiRIL lines that show altered response to biotic and abiotic stresses. Better understanding of the role of DNA methylation in heterosis could also be exploited in breeding programs. As demonstrated in rice cultures, the tissue culture process can alter methylation levels with obvious implications for the development of transgenics (Stroud et al., 2013b). A recent example of the potential of targeting DNA methylation in breeding comes from *Brassica napus*. Here, the genetic background was made isogenic by using doubled haploid lines. Plants were then pushed to two extremes of energy use efficiency over the course of several generations (Hauben et al., 2009). Initial results show extensive changes in methylation and little evidence that the observed differences were genetic. Further technological and methodological advances may make it possible to exploit natural and cryptic variation in DNA methylation for crop improvement.

FUNDING

Funding from the University of Georgia Research Foundation, the National Institutes of Health (R00GM100000), and the National Science Foundation (IOS-1339194) to R.J.S. supported this work.

ACKNOWLEDGMENTS

We would like to thank the editors for the opportunity to write this review, Nathan Springer for discussion and comments on this review, and all those who have contributed to reviewing it. Their suggestions have been invaluable. No conflict of interest declared.

REFERENCES

Ahmed, I., Sarazin, A., Bowler, C., Colot, V., and Quesneville, H. (2011). Genome-wide evidence for local DNA methylation spreading from small RNA-targeted sequences in *Arabidopsis*. Nucleic Acids Res. **39**, 6919–6931.

Becker, C., Hagmann, J., Muller, J., Koenig, D., Stegle, O., Borgwardt, K., and Weigel, D. (2011). Spontaneous epigenetic variation in the *Arabidopsis thaliana* methylome. Nature. **480**, 245–249.

Bender, J., and Fink, G.R. (1995). Epigenetic control of an endogenous gene family is revealed by a novel blue fluorescent mutant of *Arabidopsis*. Cell. **83**, 725–734.

Brink, R.A. (1956). A genetic change associated with the R locus in maize which is directed and potentially reversible. Genetics. **41**, 872–889.

Calarco, J.P., Borges, F., Donoghue, M.T., Van Ex, F., Jullien, P.E., Lopes, T., Gardner, R., Berger, F., Feijo, J.A., Becker, J.D., et al. (2012). Reprogramming of DNA methylation in pollen guides epigenetic inheritance via small RNA. Cell. **151**, 194–205.

Chan, S.W., Zhang, X., Bernatavichute, Y.V., and Jacobsen, S.E. (2006). Two-step recruitment of RNA-directed DNA methylation to tandem repeats. PLoS Biol. **4**, e363.

Coe, E.H., Jr. (1968). Heritable repression due to paramutation in maize. Science. **162**, 925.

Cokus, S.J., Feng, S., Zhang, X., Chen, Z., Merriman, B., Haudenschild, C.D., Pradhan, S., Nelson, S.F., Pellegrini, M., and Jacobsen, S.E. (2008). Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning. Nature. **452**, 215–219.

Coleman-Derr, D., and Zilberman, D. (2012). Deposition of histone variant H2A.Z within gene bodies regulates responsive genes. PLoS Genet. **8**, e1002988.

Colome-Tatche, M., Cortijo, S., Wardenaar, R., Morgado, L., Lahouze, B., Sarazin, A., Etcheverry, M., Martin, A., Feng, S., Duvernois-Berthet, E., et al. (2012). Features of the *Arabidopsis* recombination landscape resulting from the combined loss of sequence variation and DNA methylation. Proc. Natl Acad. Sci. U S A. **109**, 16240–16245.

Cubas, P., Vincent, C., and Coen, E. (1999). An epigenetic mutation responsible for natural variation in floral symmetry. Nature. **401**, 157–161.

Dowen, R.H., Pelizzola, M., Schmitz, R.J., Lister, R., Dowen, J.M., Nery, J.R., Dixon, J.E., and Ecker, J.R. (2012). Widespread dynamic DNA methylation in response to biotic stress. Proc. Natl Acad. Sci. U S A.

Du, J., Zhong, X., Bernatavichute, Y.V., Stroud, H., Feng, S., Caro, E., Vashisht, A.A., Terragni, J., Chin, H.G., Tu, A., et al. (2012). Dual binding of chromomethylase domains to H3K9me2-containing nucleosomes directs DNA methylation in plants. Cell. **151**, 167–180.

Durand, S., Bouche, N., Perez Strand, E., Loudet, O., and Camilleri, C. (2012). Rapid establishment of genetic incompatibility through natural epigenetic variation. Curr. Biol. **22**, 326–331.

Eichten, S.R., Briskine, R., Song, J., Li, Q., Swanson-Wagner, R., Hermanson, P.J., Waters, A.J., Starr, E., West, P.T., Tiffin, P., et al. (2013). Epigenetic and genetic influences on DNA methylation variation in maize populations. Plant Cell. **25**, 2783–2797.

Eichten, S.R., Ellis, N.A., Makarevitch, I., Yeh, C.T., Gent, J.I., Guo, L., McGinnis, K.M., Zhang, X., Schnable, P.S., Vaughn, M.W., et al. (2012). Spreading of heterochromatin is limited to specific families of maize retrotransposons. PLoS Genet. **8**, e1003127.

Eichten, S.R., Swanson-Wagner, R.A., Schnable, J.C., Waters, A.J., Hermanson, P.J., Liu, S., Yeh, C.T., Jia, Y., Gendler, K., Freeling, M., et al. (2011). Heritable epigenetic variation among maize inbreds. PLoS Genet. **7**, e1002372.

Feng, S., Cokus, S.J., Zhang, X., Chen, P.Y., Bostick, M., Goll, M.G., Hetzel, J., Jain, J., Strauss, S.H., Halpern, M.E., et al. (2010). Conservation and divergence of methylation patterning in plants and animals. Proc. Natl Acad. Sci. U S A. **107**, 8689–8694.

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

Finnegan, E.J., Peacock, W.J., and Dennis, E.S. (1996). Reduced DNA methylation in *Arabidopsis thaliana* results in abnormal plant development. Proc. Natl Acad. Sci. U S A. **93**, 8449–8454.

Hauben, M., Haesendonckx, B., Standaert, E., Van Der Kelen, K., Azmi, A., Akpo, H., Van Breusegem, F., Guisez, Y., Bots, M.,

Lambert, B., et al. (2009). Energy use efficiency is characterized by an epigenetic component that can be directed through artificial selection to increase yield. Proc. Natl Acad. Sci. U S A. **106**, 20109–20114.

Hollick, J.B., Patterson, G.I., Coe, E.H., Jr., Cone, K.C., and Chandler, V.L. (1995). Allelic interactions heritably alter the activity of a metastable maize pl allele. Genetics. **141**, 709–719.

Hollister, J.D., and Gaut, B.S. (2009). Epigenetic silencing of transposable elements: a trade-off between reduced transposition and deleterious effects on neighboring gene expression. Genome Res. **19**, 1419–1428.

Ibarra, C.A., Feng, X., Schoft, V.K., Hsieh, T.F., Uzawa, R., Rodrigues, J.A., Zemach, A., Chumak, N., Machlicova, A., Nishimura, T., et al. (2012). Active DNA demethylation in plant companion cells reinforces transposon methylation in gametes. Science. **337**, 1360–1364.

Johannes, F., Porcher, E., Teixeira, F.K., Saliba-Colombani, V., Simon, M., Agier, N., Bulski, A., Albuisson, J., Heredia, F., Audigier, P., et al. (2009). Assessing the impact of transgenerational epigenetic variation on complex traits. PLoS Genet. **5**, e1000530.

Law, J.A., and Jacobsen, S.E. (2010). Establishing, maintaining and modifying DNA methylation patterns in plants and animals. Nat. Rev. Genet. **11**, 204–220.

Lister, R., and Ecker, J.R. (2009). Finding the fifth base: genome-wide sequencing of cytosine methylation. Genome Res. **19**, 959–966.

Lister, R., O'Malley, R.C., Tonti-Filippini, J., Gregory, B.D., Berry, C.C., Millar, A.H., and Ecker, J.R. (2008). Highly integrated single-base resolution maps of the epigenome in *Arabidopsis*. Cell. **133**, 523–536.

Lister, R., Pelizzola, M., Dowen, R.H., Hawkins, R.D., Hon, G., Tonti-Filippini, J., Nery, J.R., Lee, L., Ye, Z., Ngo, Q.M., et al. (2009). Human DNA methylomes at base resolution show widespread epigenomic differences. Nature. **462**, 315–322.

Liu, J., He, Y., Amasino, R., and Chen, X. (2004). siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in *Arabidopsis*. Genes Dev. **18**, 2873–2878.

Luff, B., Pawlowski, L., and Bender, J. (1999). An inverted repeat triggers cytosine methylation of identical sequences in *Arabidopsis*. Mol. Cell. **3**, 505–511.

Manning, A.P., Mes-Masson, A.M., Seymour, R.J., Tetrault, M., Provencher, D.M., and Tonin, P.N. (1999). Expression of FHIT in primary cultures of human epithelial ovarian tumors and malignant ovarian ascites. Mol. Carcinog. **24**, 218–225.

Manning, K., Tor, M., Poole, M., Hong, Y., Thompson, A.J., King, G.J., Giovannoni, J.J., and Seymour, G.B. (2006). A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening. Nat. Genet. **38**, 948–952.

Mari-Ordonez, A., Marchais, A., Etcheverry, M., Martin, A., Colot, V., and Voinnet, O. (2013). Reconstructing de novo silencing of an active plant retrotransposon. Nat. Genet. **45**, 1029–1039.

Martin, A., Troadec, C., Boualem, A., Rajab, M., Fernandez, R., Morin, H., Pitrat, M., Dogimont, C., and Bendahmane, A. (2009). A transposon-induced epigenetic change leads to sex determination in melon. Nature. **461**, 1135–1138.

Mathieu, O., Reinders, J., Caikovski, M., Smathajitt, C., and Paszkowski, J. (2007). Transgenerational stability of the *Arabidopsis* epigenome is coordinated by CG methylation. Cell. **130**, 851–862.

Melquist, S., Luff, B., and Bender, J. (1999). *Arabidopsis* PAI gene arrangements, cytosine methylation and expression. Genetics. **153**, 401–413.

Mirouze, M., Lieberman-Lazarovich, M., Aversano, R., Bucher, E., Nicolet, J., Reinders, J., and Paszkowski, J. (2012). Loss of DNA methylation affects the recombination landscape in *Arabidopsis*. Proc. Natl Acad. Sci. U S A. **109**, 5880–5885.

Ossowski, S., Schneeberger, K., Lucas-Lledo, J.I., Warthmann, N., Clark, R.M., Shaw, R.G., Weigel, D., and Lynch, M. (2010). The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana*. Science. **327**, 92–94.

Patterson, G.I., Thorpe, C.J., and Chandler, V.L. (1993). Paramutation, an allelic interaction, is associated with a stable and heritable reduction of transcription of the maize b regulatory gene. Genetics. **135**, 881–894.

Regulski, M., Lu, Z., Kendall, J., Donoghue, M.T., Reinders, J., Llaca, V., Deschamps, S., Smith, A., Levy, D., McCombie, W.R., et al. (2013). The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. Genome Res.

Reinders, J., Wulff, B.B., Mirouze, M., Mari-Ordonez, A., Dapp, M., Rozhon, W., Bucher, E., Theiler, G., and Paszkowski, J. (2009). Compromised stability of DNA methylation and transposon immobilization in mosaic *Arabidopsis* epigenomes. Genes Dev. **23**, 939–950.

Richards, E.J. (2006). Inherited epigenetic variation: revisiting soft inheritance. Nat. Rev. Genet. **7**, 395–401.

Richards, E.J. (2009). Quantitative epigenetics: DNA sequence variation need not apply. Genes Dev. **23**, 1601–1605.

Ronemus, M.J., Galbiati, M., Ticknor, C., Chen, J., and Dellaporta, S.L. (1996). Demethylation-induced developmental pleiotropy in *Arabidopsis*. Science. **273**, 654–657.

Roux, F., Colome-Tatche, M., Edelist, C., Wardenaar, R., Guerche, P., Hospital, F., Colot, V., Jansen, R.C., and Johannes, F. (2011). Genome-wide epigenetic perturbation jump-starts patterns of heritable variation found in nature. Genetics. **188**, 1015–1017.

Schmitz, R.J., and Ecker, J.R. (2012). Epigenetic and epigenomic variation in *Arabidopsis thaliana*. Trends Plant Sci. **17**, 149–154.

Schmitz, R.J., and Zhang, X. (2011). High-throughput approaches for plant epigenomic studies. Curr. Opin. Plant Biol. **14**, 130–136.

Schmitz, R.J., He, Y., Valdes-Lopez, O., Khan, S.M., Joshi, T., Urich, M.A., Nery, J.R., Diers, B., Xu, D., Stacey, G., et al. (2013a). Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. Genome Res. **23**, 1663–1674.

Schmitz, R.J., Schultz, M.D., Lewsey, M.G., O'Malley, R.C., Urich, M.A., Libiger, O., Schork, N.J., and Ecker, J.R. (2011). Transgenerational epigenetic instability is a source of novel methylation variants. Science. **334**, 369–373.

Schmitz, R.J., Schultz, M.D., Urich, M.A., Nery, J.R., Pelizzola, M., Libiger, O., Alix, A., McCosh, R.B., Chen, H., Schork, N.J., et al. (2013b). Patterns of population epigenomic diversity. Nature. **495**, 193–198.

Schmutz, J., Cannon, S.B., Schlueter, J., Ma, J., Mitros, T., Nelson, W., Hyten, D.L., Song, Q., Thelen, J.J., Cheng, J., et al. (2010).

Genome sequence of the palaeopolyploid soybean. Nature. **463**, 178–183.

Shaw, R.G., Byers, D.L., and Darmo, E. (2000). Spontaneous mutational effects on reproductive traits of *Arabidopsis thaliana*. Genetics. **155**, 369–378.

Sidorenko, L.V., and Peterson, T. (2001). Transgene-induced silencing identifies sequences involved in the establishment of paramutation of the maize p1 gene. Plant Cell. **13**, 319–335.

Silveira, A.B., Trontin, C., Cortijo, S., Barau, J., Del Bem, L.E., Loudet, O., Colot, V., and Vincentz, M. (2013). Extensive natural epigenetic variation at a originated gene. PLoS Genet. **9**, e1003437.

Slotkin, R.K., Vaughn, M., Borges, F., Tanurdzic, M., Becker, J.D., Feijo, J.A., and Martienssen, R.A. (2009). Epigenetic reprogramming and small RNA silencing of transposable elements in pollen. Cell. **136**, 461–472.

Springer, N.M. (2013). Epigenetics and crop improvement. Trends Genet. **29**, 241–247.

Stroud, H., Ding, B., Simon, S.A., Feng, S., Bellizzi, M., Pellegrini, M., Wang, G.L., Meyers, B.C., and Jacobsen, S.E. (2013a). Plants regenerated from tissue culture contain stable epigenome changes in rice. Elife. **2**, e00354.

Stroud, H., Greenberg, M.V., Feng, S., Bernatavichute, Y.V., and Jacobsen, S.E. (2013b). Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome. Cell. **152**, 352–364.

Takuno, S., and Gaut, B.S. (2013). Gene body methylation is conserved between plant orthologs and is of evolutionary consequence. Proc. Natl Acad. Sci. U S A. **110**, 1797–1802.

Teixeira, F.K., Heredia, F., Sarazin, A., Roudier, F., Boccara, M., Ciaudo, C., Cruaud, C., Poulain, J., Berdasco, M., Fraga, M.F., et al. (2009). A role for RNAi in the selective correction of DNA methylation defects. Science. **323**, 1600–1604.

Thompson, A.J., Tor, M., Barry, C.S., Vrebalov, J., Orfila, C., Jarvis, M.C., Giovannoni, J.J., Grierson, D., and Seymour, G.B. (1999). Molecular and genetic characterization of a novel pleiotropic tomato-ripening mutant. Plant Physiol. **120**, 383–390.

Vaughn, M.W., Tanurd Ić, M., Lippman, Z., Jiang, H., Carrasquillo, R., Rabinowicz, P.D., Dedhia, N., McCombie, W.R., Agier, N., Bulski, A., et al. (2007). Epigenetic natural variation in *Arabidopsis thaliana*. PLoS Biol. **5**, e174.

Vongs, A., Kakutani, T., Martienssen, R.A., and Richards, E.J. (1993). *Arabidopsis thaliana* DNA methylation mutants. Science. **260**, 1926–1928.

Walker, E.L. (1998). Paramutation of the r1 locus of maize is associated with increased cytosine methylation. Genetics. **148**, 1973–1981.

Weigel, D., and Colot, V. (2012). Epialleles in plant evolution. Genome Biol. **13**, 249.

Woo, H.R., Pontes, O., Pikaard, C.S., and Richards, E.J. (2007). VIM1, a methylcytosine-binding protein required for centromeric heterochromatinization. Genes Dev. **21**, 267–277.

Zemach, A., Kim, M.Y., Hsieh, P.H., Coleman-Derr, D., Eshed-Williams, L., Thao, K., Harmer, S.L., and Zilberman, D. (2013). The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin. Cell. **153**, 193–205.

Zemach, A., McDaniel, I.E., Silva, P., and Zilberman, D. (2010). Genome-wide evolutionary analysis of eukaryotic DNA methylation. Science. **328**, 916–919.

Zhang, X., Yazaki, J., Sundaresan, A., Cokus, S., Chan, S.W., Chen, H., Henderson, I.R., Shinn, P., Pellegrini, M., Jacobsen, S.E., et al. (2006). Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis*. Cell. **126**, 1189–1201.

Zhong, S., Fei, Z., Chen, Y.R., Zheng, Y., Huang, M., Vrebalov, J., McQuinn, R., Gapper, N., Liu, B., Xiang, J., et al. (2013). Single-base resolution methylomes of tomato fruit development reveal epigenome modifications associated with ripening. Nat. Biotechnol. **31**, 154–159.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2006). Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. Nat. Genet. **39**, 61–69.