Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Plant Biology

The causes and consequences of DNA methylome variation in plants

Danelle K Seymour[1] and Claude Becker[2,3]

Epigenetic variation – polymorphisms at the level of DNA methylation or histone modifications – modulates chromatin accessibility, which can perturb transcriptional activity and spur phenotypic variation. Determining the origin, frequency spectrum, and consequences of epigenetic variants is key to understanding the role of this variation in generating stable phenotypic variation in plants. Here we review recent literature on DNA methylation variation in both model and crop plant species with a focus on the link between genotype, epigenotype, and transcription. We highlight population epigenomics studies that explore the relationship between epigenetic variants and genetic diversity. Moreover, we provide an overview of relevant studies that together advocate a minor, albeit significant role for epigenetic variation in directing specific transcriptional changes.

Addresses
[1] Department of Ecology and Evolutionary Biology, University of California, Irvine, Irvine, CA 92697, United States
[2] Department of Molecular Biology, Max Planck Institute for Developmental Biology, 72072 Tübingen, Germany
[3] Gregor Mendel Institute of Molecular Plant Biology, Austrian Academy of Sciences, 1030 Vienna, Austria

Corresponding authors: Seymour, Danelle K (dkseym@uci.edu), Becker, Claude (claude.becker@gmi.oeaw.ac.at)

Current Opinion in Plant Biology 2017, **36**:56–63

This review comes from a themed issue on **Genome studies and molecular genetics**

Edited by **Ian Henderson** and **Korbinian Schneeberger**

For a complete overview see the Issue and the Editorial

Available online 20th February 2017

http://dx.doi.org/10.1016/j.pbi.2017.01.005

Introduction

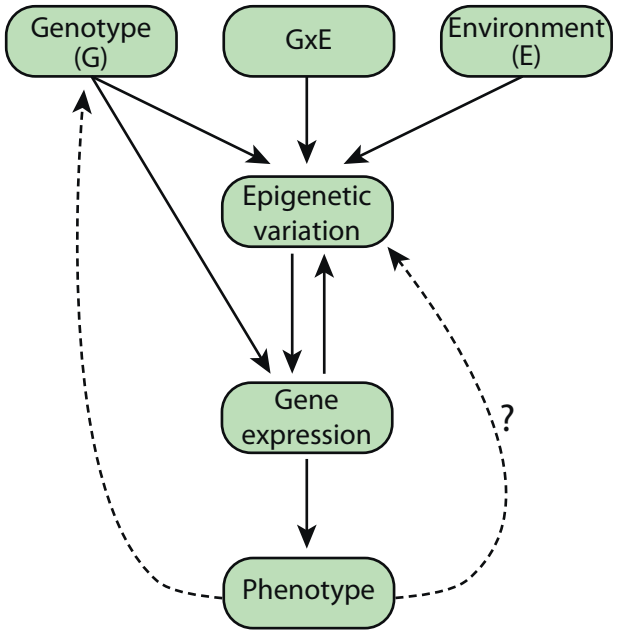
The study of epigenetics is concerned with dissecting the contribution of non-genetic variation to phenotypes. Epigenetic marks include a variety of chemical modifications to both DNA and histone proteins. However, efforts to tie epigenetic variation to phenotypic variation in plants have almost exclusively focused on DNA methylation, largely because there is a mechanism of inheritance, and because this mark is comparatively easy to investigate at a genome-wide level. For simplicity, here we use the term 'epigenetic' specifically to refer to DNA methylation.

The primary questions regarding epigenetic variation in plants are: (1) How do new variants arise? and (2) Can these variants directly alter phenotypic traits? The first question aims at evaluating whether epigenetic variants arise mostly in response to alterations of the underlying DNA sequence, or if such variants can also emerge spontaneously, independent of genetic differences (Figure 1). Such 'pure epialleles' can provide organisms with a genetically independent route to quickly respond to environmental changes [1,2]. Molecular dissection of known epialleles has demonstrated that, in some cases, unstable epigenetic silencing of transposable elements (TEs) was the source of the observed phenotypic variation [3–7]. In support of these early observations, recent work has shown that many, if not most, epigenetic variants are linked to genetic changes, particularly TE insertions.

Regardless of the source of new epigenetic variants, the second question seeks to address how often new variants perturb phenotypic traits (Figure 1). The initial advances in this field came from the molecular dissection of epialleles linked to traits of interest [8,9]. This work established that there is a direct link between epigenetic and phenotypic variation, typically via induced shifts of gene expression levels, and spurred enormous excitement in dissecting the role of DNA methylation in regulating gene expression.

Although such phenotype-first approaches have been instrumental in developing this field, the work is both time-intensive and labor-intensive. The advent of new sequencing technologies, coupled to steady decreases in the cost of sequencing, led to a flurry of genome-wide surveys of DNA methylation in both model-plant and crop species. However, the relationship between DNA methylation and gene expression is still surprisingly muddled and, as a result, the relative contribution of DNA methylation to phenotypic variation is unknown. Despite recent progress with respect to both genomic and epigenomic analysis, there are a number of major technological and statistical limitations both in determining the link between genetic and epigenetic variants and in quantifying the impact of such variants on gene expression. In this review we will discuss these limitations, as well as our current understanding of the relative contribution of genetic and non-genetic variants to

Figure 1



The interplay between genetic, epigenetic, and phenotypic variation. Genotypic (G) and environmental (E) differences can give rise to epigenetic variants. Additionally, epigenetic changes may only result under specific environmental conditions in a subset of genotypes (GxE). Genetic and epigenetic variation, both independently and together, are able to control gene expression, leading to phenotypic changes. In some cases, gene transcription may even alter the epigenetic state. The process of natural selection acting on preferred phenotypes modulates the frequency of genetic variants in a population, but there is no evidence to date that epigenetic variants are specifically subject to selection in natural populations.

shaping the epigenetic landscape and their role in regulating gene expression in plants (Figure 1).

Typically, three general types of experimental systems are used to probe the link between genotype, epigenotype, and phenotype: epigenetic recombinant inbred lines (epiRILs), generated from crosses between wild-type plants and mutants in key regulators of DNA methylation (reviewed in Ref. [10]), F_1 hybrids derived from genetically distinct strains (reviewed in Refs. [11,12]), and natural populations. Because of space limitations, we will focus on this last type of study, as large-scale surveys of natural populations in a number of plant species have recently provided some definitive answers to our two main questions of interest.

The genetic source of local epigenetic variation

To date, only a limited number of species have been investigated at single-nucleotide resolution at sufficient genomic depth to catalogue the majority of genetic and epigenetic variants present in the population. The model species *Arabidopsis thaliana* remains by far the best-studied plant species with regard to the genome-wide distribution of DNA methylation. The 1001 Epigenomes Project recently released high-quality, whole-genome bisulfite sequencing (WGBS) data together with transcriptomic data for more than 1100 *A. thaliana* accessions [13•]. This data set, combined with a parallel report on genetic variation in the same accessions [14], constitutes one of the largest repositories of high-resolution genomic, epigenomic, and transcriptomic information for a eukaryotic species to date. A main finding of this large survey was that protein-coding genes harboring methylation polymorphisms between accessions were enriched for immunity and defense-response functions, similar to what had been observed in a smaller representative collection of *A. thaliana* accessions [15]. That epigenetic polymorphisms are enriched in these gene categories, which are among the fastest evolving in the genome [16–18], suggests that genetic and epigenetic variation are linked.

This correlation of genetic and epigenetic diversity at defense response genes reflects a general pattern observed in the majority of population epigenomics studies in plants – that genetic and epigenetic variation are frequently associated with one another, although the magnitude of this association can vary by species. The first species-wide epigenomic genome-wide association scan performed in *A. thaliana* [15] found that approximately 35% of regional methylation polymorphisms, often referred to as differentially methylation regions (DMRs), were associated with a DNA sequence polymorphism. Fewer DNA methylation variants, only 18%, were associated with a genetic variant in a large Swedish collection of *A. thaliana* accessions [19••]. A much higher degree of genotype–epigenotype association was observed in a natural population of closely related individuals from North America [20•], with nearly half of the observed variation in DNA methylation being explained by the genetic structure of the population.

Intriguingly, the situation seems quite different outside of *A. thaliana*. Work in a number of grass species suggests that a larger fraction of methylation polymorphisms between strains are linked to genetic variants. Schmitz *et al.* [21] reported that the vast majority (91%) of DMRs detected among soybean recombinant inbred lines (RILs) were associated with a methylQTL. Similarly, in maize RILs, more than half of all DMRs were associated with a local genetic variant [22]. Relative to *A. thaliana*, these grass species have much larger genome sizes, caused by an increased repertoire of TEs. One possibility is that unstable control of the increased TE content in these species can give rise to the observed increase in genetically dependent epialleles. In contrast to what has been observed in most grass species, we should note that a small survey of 6 strains of *B. distachyon* suggested that TE polymorphisms across the strains were not enriched for

methylation variants, although genetic variation in general was associated with increased methylation [23].

Despite mounting evidence that genetic variants underlie a large fraction of population-level methylation polymorphisms, there is some evidence for recurring variation at non-genetic, or 'pure' epialleles. The best evidence comes from epigenetic variation in a set of isogenic *A. thaliana* lines [24,25]. Interestingly, there is a high degree of overlap between these 'pure' variants and epigenetic variants segregating in a set of North American *A. thaliana* accessions [20•], as well as in the set of variants induced by hyperosmotic stress conditions [26••]. This non-random overlap suggests that there is indeed a set of loci that undergo changes in their methylation state, independent of genetic background or environmental history, and may constitute 'pure' epigenetic variants. That these regions appear to be intrinsically labile indicates that 'pure' variants may be challenging for natural selection to target, and until we know the function of these changes it is difficult to speculate further on their role.

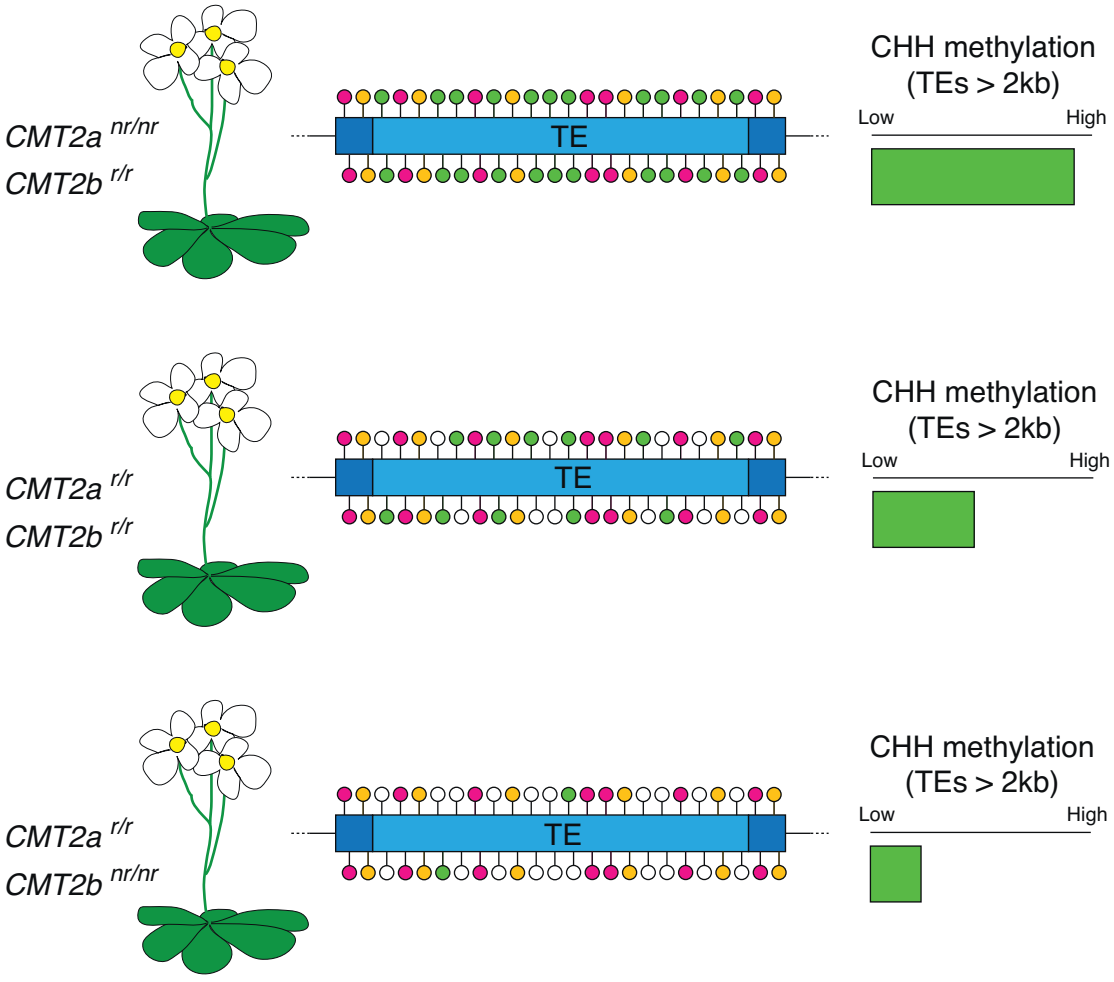
Estimates of the degree of genetic control of DNA methylation variation differ by study, highlighting some of the technical limitations of population level surveys. First, and foremost, most genome-wide studies are based on a single reference genome. As a result, not all genetic variants are known. Lineage-specific TE insertions, or other structural variants, are the most likely to be ignored when using a reference-based strategy [27]. Consequently, epialleles may appear 'pure' when, in fact, they are genetically linked, diminishing the degree of association between genetic and epigenetic variants. The higher degree of genotype–epigenotype association in the *A. thaliana* North American population [20•] compared to other populations [15,19••] may result from alleviating this so-called reference bias, as a specific reference sequence was generated for the analysis of the North American population [20•]. The limitations of genome-wide association techniques can also conceal associations between genetic and epigenetic variants. For example, recent evidence suggests that allelic heterogeneity, or cases where independent genetic changes generate the same methylation polymorphism, are likely responsible for epivariants that appear to be 'pure' [28••]. Finally, estimates of genotype–epigenotype associations are typically based on local associations and exclude the contribution of large-effect *trans* associations. In sum, although the genetic contribution to epigenetic variation varies by species, and even by population, all studies agree that a significant degree of variation in DNA methylation is genetically controlled.

Single-gene mutations with large-effect epigenetic consequences

One noteworthy pattern in population epigenomic studies is that the genetic variants reported to be associated with methylation polymorphisms are enriched for key components of the DNA methylation machinery. Mutations in core machinery genes are expected to vastly perturb the epigenetic landscape, and this has been observed in a number of cases [29]. In an *A. thaliana* mutation accumulation experiment that spanned 30 generations, an increased accumulation of epimutations occurred in one lineage relative to the rest [24,25]. Although these lines are mostly isogenic, a few single nucleotide polymorphisms (SNPs) did accrue over the course of the experiment (~30 per line). The lineage in question contained one SNP that was located in *MATERNAL EFFECT EMBRYO ARREST 57* (*MEE57*), a *DNA METHYLTRANSFERASE 2* (*MET2)* homolog, indicating that a single SNP in a key DNA methylation maintenance gene likely accelerated the genome-wide epimutation rate in this lineage [24]. More conclusive evidence for the role of large-effect mutations in shaping the DNA methylation landscape comes from the collection of Swedish *A. thaliana* accessions [19••]. In this study, two SNPs (*CMT2a* and *CMT2b*) located in proximity to *CHROMOMETHYLASE 2* (*CMT2*) were associated with higher CHH methylation levels in TEs (Figure 2). Not only is *CMT2* a gene with a known role in regulating TE methylation [30,31], but the authors provide further support for the association by demonstrating that TE methylation levels co-segregate with the expected low- and high-methylation alleles in an independent population [19••]. The *CMT2a* allele alone accounted for a whopping 21% of all significant associations between genetic and epigenetic variants. In another study, a third large-effect mutation, this time located at the *MET2a* locus, was associated with increased transposition of specific TE families in *A. thaliana* [32]. This locus is linked to TE mobilization, not DNA methylation per se, but it suggests that mutations in DNA methyltransferases destabilize TE methylation levels, increasing their propensity to transpose.

Outside of *A. thaliana*, drastic effects of single-gene variants on the level and distribution of genome-wide DNA methylation are found in a number of Brassicaceae species. *Arabis alpina* has lost symmetrical DNA methylation at CG nucleotides, possibly caused by a reduced number of methyltransferase genes in the species [33]. Two other Brassicaceae species, *Eutrema salsugineum* and *Conringia planisiliqua*, have independently lost copies of a DNA methyltransferase homologous to *CMT3* in *A. thaliana*. Presumably, loss of this gene has led to the reduction of genome-wide levels of CHG methylation. Interestingly, global levels of gene expression are not affected in these species [34]. That functional mutations in core members of the DNA methylation machinery segregate in natural strains of *A. thaliana*, and that members appear to be completely dispensable in some species, suggests that a significant portion of the DNA methylation landscape is controlled by only a few mutations with extensive epigenetic consequences.

Figure 2



Single genetic mutations have large-scale epigenetic effects.
In a population of *A. thaliana* from north and south Sweden, Dubin *et al*. [19••] detected genetic variants at two loci, *CMT2a* and *CMT2b*. These variants reside near the locus encoding the methyltransferase *CMT2* and are found in three distinct haplotypes. Plants carrying the non-reference (nr) allele for *CMT2a* and the reference (r) allele for *CMT2b* are linked to high levels of CHH methylation in long TEs (>2kb; represented by the blue rectangular box). Their reverse allelic combination (*CMT2a*$^{r/r}$; *CMT2b*$^{nr/nr}$) leads to low levels of CHH TE methylation. The recombinant plants (*CMT2a*$^{r/r}$; *CMT2b*$^{r/r}$) display intermediate methylation levels of TEs. Methylated cytosines are color-coded by sequence context: CG = red, CHG = yellow, CHH = green; unmethylated = white.

The minor, but significant, contribution of DNA methylation to variation in gene expression

Not only is proper function of the core molecular machinery responsible for initiating and maintaining DNA methylation in plant genomes, but it is also required for normal plant growth and development. It was surprising then that null mutations in key genes only marginally perturb gene expression profiles, with only 2% of genes being differentially expressed in *met1* (DNA METHYLTRANSFERASE 1) mutants [35,36], with similar results in other species [37,38]. Recent work has shown that the severe developmental defects displayed by a key epigenetic regulator mutant can be rescued by restoring the function of a single gene [39], suggesting that DNA methylation may only be required to control endogenous gene expression at a small number of loci.

In support of this observation, surveys of natural variation in DNA methylation have repeatedly demonstrated that the correlation of methylation variation with differential gene expression is low, albeit significant [15,22,40,41]. Interpreting the results from these large-scale studies is complicated by at least three factors. First, as we explain above, separating the effects of genetic and epigenetic variation in natural populations is a challenge. Second, the relationship between DNA methylation and gene expression differs by both genomic feature and location. For example, TE methylation is typically associated with

dampened gene expression, but the opposite relationship has also been reported [42,43••]. The last complication is that, unlike the original work on epialleles, genome-wide studies are correlative, and it is not clear whether a causal relationship between DNA methylation and gene expression holds globally. We will discuss each of these challenges and how recent work has provided significant insight into the extent that gene regulation is mediated by DNA methylation.

The first challenge in evaluating the impact of DNA methylation variation on gene expression is separating their effects from genetic variants. Recent work in *A. thaliana* by Meng *et al.* has presented a truly quantitative evaluation of the link between changes in DNA methylation and genetic variants [28••]. Highlighting the extent of this challenge, the authors demonstrate that at a genome-wide level, DNA methylation can explain no additional variation in gene expression levels than genetic variants. Despite this limitation, the authors surveyed single genes for associations with *cis*- DNA methylation variants and identified a small subset of genes (212) where gene expression was significantly associated with local methylation variants, with only 64 of those genes also lacking genetic associations. Although previous studies have alluded to this finding [15,20•,22,40,41], we can safely conclude from this work that when DNA methylation is not linked to genetic variants, it plays a minor, but significant, role in regulating gene expression.

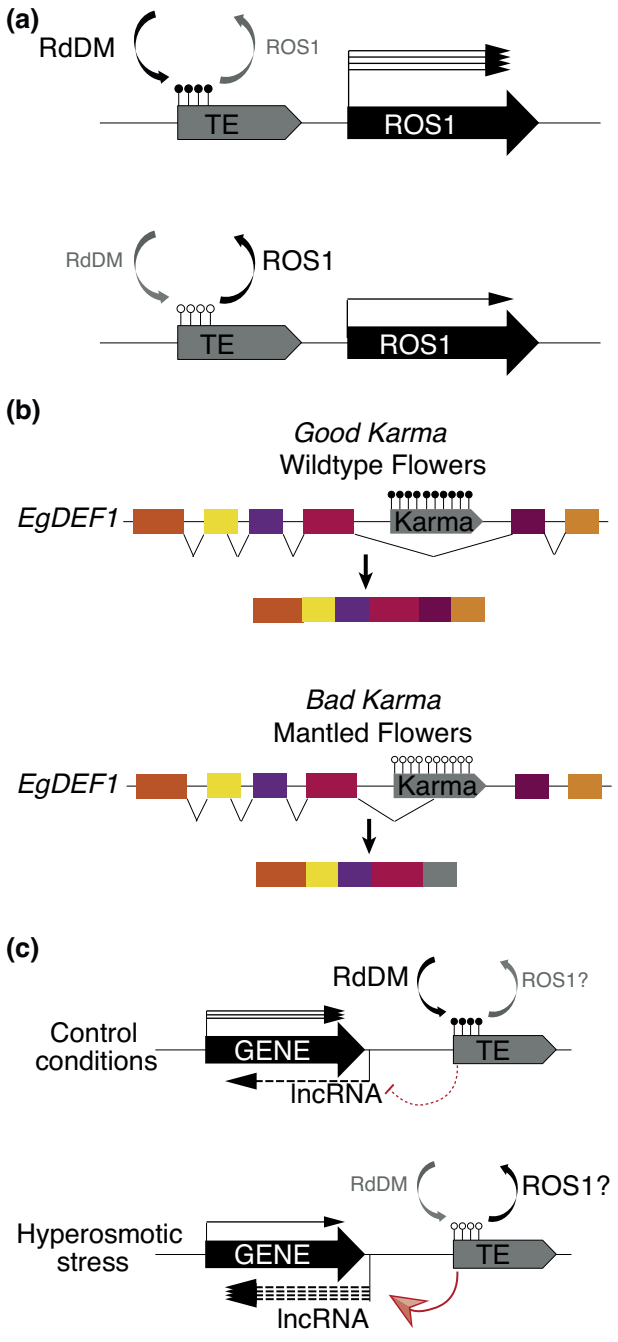
The relationship between the levels of gene expression and DNA methylation is complex. In *A. thaliana*, methylated TEs are strongly linked to reduced expression of flanking genes [28••], with the degree of the suppressive effect dependent on the distance between the two elements [44,45]. In contrast, a weak, but positive correlation has been repeatedly observed between gene expression levels and CG-only methylation of genes, also known as gene body methylation [35,36,46–48]. While these overall trends hold true in most surveyed plant species, there are some notable exceptions. First, although methylation in non-CG contexts is typically linked to TEs and their repressive effect on expression, recent work in maize has shown that non-CG methylation can positively correlate with expression [40,49,50]. In maize, CHH islands that delineate the transitions between chromatin domains have a weak, positive association with expression levels, but the loss of these islands has little impact on transcriptional levels of local genes [40,49,50]. Similarly, CHG methylation of genes is expected to have a negative effect on expression levels, but high CHG methylation levels in the gene bodies of gymnosperm genes contradict this expectation [51]. The tendency for global methylation levels in each context to be differentially linked to gene expression holds across many species, but lineage-specific innovations seem to provide new twists on a common theme.

One impressive example that stands in stark contrast to the global trend of TE methylation reducing gene expression is the DNA methylation-mediated feed-forward regulation of *REPRESSOR OF SILENCING 1* (*ROS1*) [43••,52••]. *ROS1* is a DNA glycosylase that removes methylated cytosines from a number of genomic regions in somatic tissues [53]. Recent work by Williams *et al.* [43••] and by Lei *et al.* [52••] demonstrates that the expression level of this gene is finely tuned to ensure that genomes strike the proper balance between TE silencing and gene transcription (Figure 3a). Instead of reducing expression levels, RdDM-mediated methylation of an upstream TE is sufficient to increase the endogenous expression levels of *ROS1*. As expression levels increase, *ROS1* will then selectively demethylate this upstream sequence to achieve a balanced level of DNA methylation not only at this locus, but across the genome [43••]. Importantly, methylation sensitive expression of *ROS1* is conserved in *A. thaliana's* congener, *Arabidopsis lyrata*, due at least in part to maintenance of the upstream TE sequence over millions of years of evolution [43••]. The careful molecular dissection of methylation-mediated regulation of *ROS1* illustrates two critical points. First, this work clearly shows that TE methylation promotes expression in contrast to global patterns. Second, and more importantly, it shows that regulation of expression via DNA methylation is evolutionarily conserved. We conclude that despite the limited contribution of DNA methylation to the global regulation of gene expression, this regulatory mechanism is extremely important at a select number of loci.

Two recent examples also clearly demonstrate how variation in methylation at TEs can directly alter gene regulation. Probably the most remarkable example of epigenetically controlled transcription to date was recently reported in the clonally propagated oil palm (*Elaeis guineensis*). Ong-Abdullah *et al.* [54••] discovered that trees producing mantled fruits, which cannot be used for oil production, had reduced methylation at an intronic TE located in the homeotic gene *DEFICIENS* (Eg *DEF1*), which resulted in the aberrant splicing of this gene (Figure 3b). The second example illustrates how hyperosmotic stress in *A. thaliana* [26••] leads to selective demethylation of a TE downstream of the gene *CARBON/NITROGEN INSENSITIVE 1* (*CNI1*). Reduced methylation at this TE directly up-regulates the expression of an antisense long non-coding RNA (lncRNA), which in turn decreases the expression of *CNI1* itself (Figure 3c).

The examples of *ROS1* [43••,52••], *EgDEF1* [54••], *CNI1* [26••], and others nicely demonstrate that DNA methylation is sufficient to drive changes in expression, but

Figure 3



Recent examples illustrate the direct link between DNA methylation, gene expression, and phenotypic variation.
(a) Control of *REPRESSOR OF SILENCING 1* (*ROS1*) by DNA methylation in *A. thaliana.* Methylation of a TE upstream of *ROS1* leads to elevated gene expression levels [43••,52••]. TE methylation (black pins) at this locus is controlled by RNA-directed DNA methylation (RdDM). In the absence of RdDM, the TE becomes unmethylated (white pins), decreasing *ROS1* expression. In wild-type plants, *ROS1* selectively demethylates this TE to maintain the desired epigenetic status at this locus. **(b)** In oil palm, demethylation of an intronic TE named *KARMA* leads to alternative splicing of the *EgDEF1* transcript. The resulting aberrant transcript causes the formation of mantled flowers [54••]. **(c)** In *A. thaliana*, hyperosmotic stress induces loss of DNA methylation at a TE downstream of the *CNI1* locus. This activates antisense transcription of a long-noncoding RNA (lncRNA), resulting in down-regulation of *CNI1* [26••].

whether this holds true for most genes is still unclear. For the first time in plants, Meng *et al.* used statistically causal models to tease apart the direction of the relationships between genetic variation, methylation variation, and gene expression [28••]. Although experimental studies are needed to verify these findings, the authors show that the most likely model is one where genetic variation affects both DNA methylation and gene expression independently. In some cases, there was evidence both for DNA methylation affecting expression and for the inverse situation of expression affecting DNA methylation [28••]. In support of the latter, phosphate starvation of *Oryza sativa* caused stress-induced gene expression to precede changes in DNA methylation [55••]. Once the plants were resupplied with phosphate, gene expression returned to pre-stress levels while changes in DNA methylation were longer lasting [55••]. While a number of studies have examined the impact of stress of the levels of DNA methylation [26••,56–59], Secco *et al.* [55••] were the first to clearly detect this unconventional relationship, likely becuse of the extensive time-course that was generated. Although DNA methylation levels are predicted to control gene expression levels, more work is required to understand the function and mechanism of the reverse relationship.

Conclusions

Since the discovery of epigenetic variation, there has been continued excitement and interest in its potential to induce phenotypic variation and to generate new opportunities for plants to respond to changing environmental conditions without the need for genetic innovation. Specific cases in multiple species have illustrated the potential of epivariants or epialleles to directly alter phenotypic traits of interest. Technological advances in sequencing and genome-wide analyses have allowed researchers to probe whether the direct link between epiallele and phenotype holds at the global level. Here, we have detailed recent work that has provided overwhelming evidence that genetic variation is a major source of epigenetic variation and that, outside of controlling TE proliferation, DNA methylation directly influences gene expression at only a small subset of specific loci.

It is important to note that the population epigenetics studies discussed here relied on tissue that was harvested from plants grown in a common greenhouse or growth chamber for at least one generation. Given that the vast majority of environmentally induced DNA methylation changes are reset during generational transitions [26••], the extent of GxE-driven epigenetic variation (Figure 1) in these populations may have been underestimated.

To fully comprehend the link between genotype, epigenotype, and phenotype, it will be the task of future research efforts to generate comprehensive high-resolution datasets from plants in the field, sampled under the prevailing environmental conditions.

Conflict of interest

The authors declare that no conflicts of interest exist.

Acknowledgements

This work was supported by the Max Planck Society, DFG (SFB 1101), and by the Austrian Academy of Sciences (CB). DKS is supported by an NSF Postdoctoral Research Fellowship in Biology (1609024).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- • of special interest
- •• of outstanding interest

1. Eichten SR, Schmitz RJ, Springer NM: **Epigenetics: beyond chromatin modifications and complex genetic regulation**. *Plant Physiol* 2014, **165**:933-947.

2. Richards EJ: **Opinion – inherited epigenetic variation – revisiting soft inheritance**. *Nat Rev Genet* 2006, **7** 395-U392.

3. Das OP, Messing J: **Variegated phenotype and developmental methylation changes of a maize allele originating from epimutation**. *Genetics* 1994, **136**:1121-1141.

4. Fujimoto R, Sasaki T, Kudoh H, Taylor JM, Kakutani T, Dennis ES: **Epigenetic variation in the FWA gene within the genus *Arabidopsis***. *Plant J* 2011, **66**:831-843.

5. Liu J, He YH, Amasino R, Chen XM: **siRNAs targeting an intronic transposon in the regulation of natural flowering behavior in *Arabidopsis***. *Genes Dev* 2004, **18**:2873-2878.

6. Martin A, Troadec C, Boualem A, Rajab M, Fernandez R, Morin H, Pitrat M, Dogimont C, Bendahmane A: **A transposon-induced epigenetic change leads to sex determination in melon**. *Nature* 2009, **461** 1135-U1237.

7. Miura K, Agetsuma M, Kitano H, Yoshimura A, Matsuoka M, Jacobsen SE, Ashikari M: **A metastable DWARF1 epigenetic mutant affecting plant stature in rice**. *Proc Natl Acad Sci U S A* 2009, **106**:11218-11223.

8. Cubas P, Vincent C, Coen E: **An epigenetic mutation responsible for natural variation in floral symmetry**. *Nature* 1999, **401**:157-161.

9. Manning K, Tor M, Poole M, Hong Y, Thompson AJ, King GJ, Giovannoni JJ, Seymour GB: **A naturally occurring epigenetic mutation in a gene encoding an SBP-box transcription factor inhibits tomato fruit ripening**. *Nat Genet* 2006, **38**:948-952.

10. Weigel D, Colot V: **Epialleles in plant evolution**. *Genome Biol* 2012, **13**.

11. Groszmann M, Greaves IK, Fujimoto R, Peacock WJ, Dennis ES: **The role of epigenetics in hybrid vigour**. *Trends Genet* 2013, **29**:684-690.

12. He GM, He H, Deng XW: **Epigenetic variations in plant hybrids and their potential roles in heterosis**. *J Genet Genom* 2013, **40**:205-210.

13. • Kawakatsu T, Huang SS, Jupe F, Sasaki E, Schmitz RJ, Urich MA, Castanon R, Nery JR, Barragan C, He Y *et al.*: **Epigenomic diversity in a global collection of *Arabidopsis thaliana* accessions**. *Cell* 2016, **166**:492-505.
This study, in combination with Ref. [14], constitutes of the one of largest resources of species-wide genome, DNA methylome and transcriptome data at single-nucleotide resolution, with ~1100 *A. thaliana* accessions.

14. Consortium G: **1,135 Genomes reveal the global pattern of polymorphism in *Arabidopsis thaliana***. *Cell* 2016, **166**:481-491.

15. Schmitz RJ, Schultz MD, Urich MA, Nery JR, Pelizzola M, Libiger O, Alix A, McCosh RB, Chen H, Schork NJ *et al.*: **Patterns of population epigenomic diversity**. *Nature* 2013, **495**:193-198.

16. Ellis J, Dodds P, Pryor T: **Structure, function and evolution of plant disease resistance genes**. *Curr Opin Plant Biol* 2000, **3**:278-284.

17. Michelmore RW, Meyers BC: **Clusters of resistance genes in plants evolve by divergent selection and a birth-and-death process**. *Genome Res* 1998, **8**:1113-1130.

18. Cao J, Schneeberger K, Ossowski S, Gunther T, Bender S, Fitz J, Koenig D, Lanz C, Stegle O, Lippert C *et al.*: **Whole-genome sequencing of multiple *Arabidopsis thaliana* populations**. *Nat Genet* 2011, **43** 956-U960.

19. •• Dubin MJ, Zhang P, Meng D, Remigereau MS, Osborne EJ, Paolo Casale F, Drewe P, Kahles A, Jean G, Vilhjalmsson B *et al.*: **DNA methylation in *Arabidopsis* has a genetic basis and shows evidence of local adaptation**. *Elife* 2015, **4**:e05255.
In accessions from north and south Sweden, allelic variation in the *A. thaliana* methyltransferase CMT2 was found to be associated with differential CHH methylation in transposable elements. This finding illustrates how genetic regulation of an epigenetic pathway might contribute to adaptation to changing environments.

20. • Hagmann J, Becker C, Muller J, Stegle O, Meyer RC, Wang G, Schneeberger K, Fitz J, Altmann T, Bergelson J *et al.*: **Century-scale methylome stability in a recently diverged *Arabidopsis thaliana* lineage**. *PLoS Genet* 2015, **11**:e1004920.
This study uses a unique collection of genetically near-identical *A. thaliana* accessions, grown for more than 100 generations in different geographic locations, to show that there is no detectable signal of epigenetic adaptation to the environment.

21. Schmitz RJ, He Y, Valdes-Lopez O, Khan SM, Joshi T, Urich MA, Nery JR, Diers B, Xu D, Stacey G *et al.*: **Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population**. *Genome Res* 2013, **23**:1663-1674.

22. Eichten SR, Briskine R, Song J, Li Q, Swanson-Wagner R, Hermanson PJ, Waters AJ, Starr E, West PT, Tiffin P *et al.*: **Epigenetic and genetic influences on DNA methylation variation in maize populations**. *Plant Cell* 2013, **25**:2783-2797.

23. Eichten SR, Stuart T, Srivastava A, Lister R, Borevitz JO: **DNA methylation profiles of diverse *Brachypodium distachyon* aligns with underlying genetic diversity**. *Genome Res* 2016, **26**:1520-1531.

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

25. Schmitz RJ, Schultz MD, Lewsey MG, O'Malley RC, Urich MA, Libiger O, Schork NJ, Ecker JR: **Transgenerational epigenetic instability is a source of novel methylation variants**. *Science* 2011, **334**:369-373.

26. •• Wibowo A, Becker C, Marconi G, Durr J, Price J, Hagmann J, Papareddy R, Putra H, Kageyama J, Becker J *et al.*: **Hyperosmotic stress memory in *Arabidopsis* is mediated by distinct epigenetically labile sites in the genome and is restricted in the male germline by DNA glycosylase activity**. *Elife* 2016, **5**.
This study presents a mechanism of stress-induced epigenetic gene regulation: hyperosmosis-induced demethylation of a TE leads to increased expression of a long non-coding RNA, which in turn down-regulates transcription of a protein-coding gene.

27. Treangen TJ, Salzberg SL: **Repetitive DNA and next-generation sequencing: computational challenges and solutions**. *Nat Rev Genet* 2012, **13**:36-46.

28. •• Meng D, Dubin M, Zhang P, Osborne EJ, Stegle O, Clark RM, Nordborg M: **Limited contribution of DNA methylation variation to expression regulation in *Arabidopsis thaliana***. *PLoS Genet* 2016, **12**:e1006141.
Using a panel of natural *A. thaliana* accessions, this paper clearly outlines the role of DNA methylation in modulating gene expression variation. The two major conclusions from this paper are (1) that DNA methylation is tightly linked to genetic variation and (2) that not only can DNA

methylation direct changes in gene expression but expression can also cause changes in methylation.

29. Stroud H, Greenberg MVC, Feng SH, Bernatavichute YV, Jacobsen SE: **Comprehensive analysis of silencing mutants reveals complex regulation of the *Arabidopsis* methylome**. *Cell* 2013, **152**:352-364.

30. Stroud H, Do T, Du JM, Zhong XH, Feng SH, Johnson L, Patel DJ, Jacobsen SE: **Non-CG methylation patterns shape the epigenetic landscape in *Arabidopsis***. *Nat Struct Mol Biol* 2014, **21** 64-+.

31. Zemach A, Kim MY, Hsieh PH, Coleman-Derr D, Eshed-Williams L, Thao K, Harmer SL, Zilberman D: **The *Arabidopsis* nucleosome remodeler DDM1 allows DNA methyltransferases to access H1-containing heterochromatin**. *Cell* 2013, **153**:193-205.

32. Quadrana L, Silveira AB, Mayhew GF, LeBlanc C, Martienssen RA, Jeddeloh JA, Colot V: **The *Arabidopsis thaliana* mobilome and its impact at the species level**. *Elife* 2016, **5**.

33. Willing EM, Rawat V, Mandakova T, Maumus F, James GV, Nordstrom KJ, Becker C, Warthmann N, Chica C, Szarzynska B *et al.*: **Genome expansion of *Arabis alpina* linked with retrotransposition and reduced symmetric DNA methylation**. *Nat Plants* 2015, **1**:14023.

34. Bewick AJ, Ji LX, Niederhuth CE, Willing EM, Hofmeister BT, Shi XL, Wang L, Lu ZF, Rohr NA, Hartwig B *et al.*: **On the origin and evolutionary consequences of gene body DNA methylation**. *Proc Natl Acad Sci U S A* 2016, **113**:9111-9116.

35. Lister R, O'Malley RC, Tonti-Filippini J, Gregory BD, Berry CC, Millar AH, Ecker JR: **Highly integrated single-base resolution maps of the epigenome in *Arabidopsis***. *Cell* 2008, **133**:523-536.

36. 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**:61-69.

37. Cheng CY, Tarutani Y, Miyao A, Ito T, Yamazaki M, Sakai H, Fukai E, Hirochika H: **Loss of function mutations in the rice chromomethylase OsCMT3a cause a burst of transposition**. *Plant J* 2015, **83**:1069-1081.

38. Hu LJ, Li N, Xu CM, Zhong SL, Lin XY, Yang JJ, Zhou TQ, Yuliang AZ, Wu Y, Chen YR *et al.*: **Mutation of a major CG methylase in rice causes genome-wide hypomethylation, dysregulated genome expression, and seedling lethality**. *Proc Natl Acad Sci U S A* 2014, **111**:10642-10647.

39. Blevins T, Pontvianne F, Cocklin R, Podicheti R, Chandrasekhara C, Yerneni S, Braun C, Lee B, Rusch D, Mockaitis K *et al.*: **A two-step process for epigenetic inheritance in *Arabidopsis***. *Mol Cell* 2014, **54**:30-42.

40. Li Q, Song J, West PT, Zynda G, Eichten SR, Vaughn MW, Springer NM: **Examining the causes and consequences of context-specific differential DNA methylation in maize**. *Plant Physiol* 2015, **168** 1262-U1260.

41. Roessler K, Takuno S, Gaut BS: **CG methylation covaries with differential gene expression between leaf and floral bud tissues of *Brachypodium distachyon***. *PLoS One* 2016, **11**.

42. Shibuya K, Fukushima S, Takatsuji H: **RNA-directed DNA methylation induces transcriptional activation in plants**. *Proc Natl Acad Sci U S A* 2009, **106**:1660-1665.

43. •• Williams PB, Pignatta D, Henikoff S, Gehring M: **Methylation-sensitive expression of a DNA demethylase gene serves as an epigenetic rheostat**. *PLoS Genet* 2015, **11**:e1005142.
This paper, together with Ref. [52••], highlights a specific case in which DNA methylation at a TE upstream of *ROS1* enhances gene expression levels. The authors provide extensive molecular support for this observation and show that this regulatory mechanism is conserved in *A. thaliana*'s close relative *A. lyrata*.

44. 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**:1419-1428.

45. Eichten SR, Ellis NA, Makarevitch I, Yeh CT, Gent JI, Guo L, McGinnis KM, Zhang XY, Schnable PS, Vaughn MW *et al.*: **Spreading of heterochromatin is limited to specific families of maize retrotransposons**. *PLoS Genet* 2012, **8**.

46. Cokus SJ, Feng SH, Zhang XY, Chen ZG, Merriman B, Haudenschild CD, Pradhan S, Nelson SF, Pellegrini M, Jacobsen SE: **Shotgun bisulphite sequencing of the *Arabidopsis* genome reveals DNA methylation patterning**. *Nature* 2008, **452**:215-219.

47. Zhang XY, Yazaki J, Sundaresan A, Cokus S, Chan SWL, Chen HM, Henderson IR, Shinn P, Pellegrini M, Jacobsen SE *et al.*: **Genome-wide high-resolution mapping and functional analysis of DNA methylation in *Arabidopsis***. *Cell* 2006, **126**:1189-1201.

48. Takuno S, Gaut BS: **Body-methylated genes in *Arabidopsis thaliana* are functionally important and evolve slowly**. *Mol Biol Evol* 2012, **29**:219-227.

49. Gent JI, Ellis NA, Guo L, Harkess AE, Yao YY, Zhang XY, Dawe RK: **CHH islands: de novo DNA methylation in near-gene chromatin regulation in maize**. *Genome Res* 2013, **23**:628-637.

50. West PT, Li Q, Ji LX, Eichten SR, Song JW, Vaughn MW, Schmitz RJ, Springer NM: **Genomic distribution of H3K9me2 and DNA methylation in a maize genome**. *PLoS One* 2014, **9**.

51. Takuno S, Ran JH, Gaut BS: **Evolutionary patterns of genic DNA methylation vary across land plants**. *Nat Plants* 2016, **2**.

52. •• Lei MG, Zhang HM, Julian R, Tang K, Xie SJ, Zhu JK: **Regulatory link between DNA methylation and active demethylation in *Arabidopsis***. *Proc Natl Acad Sci U S A* 2015, **112**:3553-3557.
Together with Ref. [43••], this paper provides molecular support for the alternative regulatory mechanism of *ROS1*.

53. Gong ZH, Morales-Ruiz T, Ariza RR, Roldan-Arjona T, David L, Zhu JK: **ROS1, a repressor of transcriptional gene silencing in *Arabidopsis*, encodes a DNA glycosylase/lyase**. *Cell* 2002, **111**:803-814.

54. •• Ong-Abdullah M, Ordway JM, Jiang N, Ooi SE, Kok SY, Sarpan N, Azimi N, Hashim AT, Ishak Z, Rosli SK *et al.*: **Loss of Karma transposon methylation underlies the mantled somaclonal variant of oil palm**. *Nature* 2015, **525**:533-537.
Together with Refs. [26••] and [55••], this study shows how genome-wide profiling of DNA methylation can identify specific epigenetically regulated loci. The authors identify the *KARMA* locus, which, when demethylated, causes aberrant splicing and the formation of mantled flowers in oil palm trees.

55. •• Secco D, Wang C, Shou HX, Schultz MD, Chiarenza S, Nussaume L, Ecker JR, Whelan J, Lister R: **Stress induced gene expression drives transient DNA methylation changes at adjacent repetitive elements**. *Elife* 2015, **4**.
Together with Refs. [26••] and [54••], this study shows how genome-wide profiling of DNA methylation can identify specific epigenetically regulated loci. In addition, the authors provide time series data that indicate that stress-induced transcriptional changes precede methylation changes.

56. Eichten SR, Springer NM: **Minimal evidence for consistent changes in maize DNA methylation patterns following environmental stress**. *Front Plant Sci* 2015, **6**:308.

57. Jiang C, Mithani A, Belfield EJ, Mott R, Hurst LD, Harberd NP: **Environmentally responsive genome-wide accumulation of de novo *Arabidopsis thaliana* mutations and epimutations**. *Genome Res* 2014, **24**:1821-1829.

58. Seymour DK, Koenig D, Hagmann J, Becker C, Weigel D: **Evolution of DNA methylation patterns in the Brassicaceae is driven by differences in genome organization**. *PLoS Genet* 2014, **10**:e1004785.

59. Dowen RH, Pelizzola M, Schmitz RJ, Lister R, Dowen JM, Nery JR, Dixon JE, Ecker JR: **Widespread dynamic DNA methylation in response to biotic stress**. *Proc Natl Acad Sci U S A* 2012, **109**: E2183-E2191.