Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Plant Biology

Stories that can't be told by SNPs; DNA methylation variation in plant populations

Jaclyn M Noshay and Nathan M Springer

Epigenetic variation has been observed in many plant populations. This variation can influence qualitative and quantitative traits. A key question is whether there is novel information in the epigenome that is not captured by SNP-based genetic markers. The answer likely varies depending on the sources and stability of epigenetic variation as well as the type of population being studied. We consider the epigenetic variation in several plant systems and how this relates to potential for hidden information that could increase our understanding of phenotypic variation.

Address

Department of Plant and Microbial Biology, University of Minnesota, Saint Paul, MN 55108, USA

Corresponding author: Springer, Nathan M (springer@umn.edu)

Current Opinion in Plant Biology 2021, **61**:101989

This review comes from a themed issue on **Epigenetics**

Edited by **Mary Gehring** and **François Roudier**

For a complete overview see the Issue and the Editorial

Available online 11th January 2021

https://doi.org/10.1016/j.pbi.2020.101989

The availability of low-cost genetic marker technologies enabled connections between genotype and phenotype in many plant populations. These resources allow for the detection and characterization of quantitative trait loci (QTL) as well as the development of genomic prediction approaches. However, in many cases the sum of QTL effects remains well below the heritability of the trait. This so-called 'missing heritability' can arise due to a variety of reasons including rare alleles, the inability to detect minor effect QTL and epistasis [1]. Epigenetic variation, heritable variation that is not solely explained by sequence differences, could also contribute to missing heritability. We will discuss the prevalence and potential impacts of epigenetic variation in plant populations.

The term epigenetics is used in different contexts to describe both biochemical and genetic phenomena. In this review we will focus on using the term epigenetic to refer to heritable variation that is not fully linked to genetic (sequence) differences. At the molecular level epigenetic variation can be associated with differences in chromatin modification (DNA methylation, chromatin accessibility, histone variants, histone modifications), small RNAs or even protein structure (prions). DNA methylation has received the most attention as a molecular marker for potential epigenetic variation due to it's relatively high heritability and the high-throughput methods for documenting genome-wide patterns [2]. As we consider the potential role of epigenetics we will largely focus on studies that have evaluated variation, heritability and impacts of DNA methylation but we expect that additional studies will highlight potential roles for other molecular mechanisms of epigenetic variation.

Our desire to understand the role of epigenetic variation in plant populations stems from the potential for its contribution to phenotypic variation. Identifying the causative basis for QTL, genetic or epigenetic, increases our understanding of how variation arises and how to create or alter alleles for crop improvement. Current SNP-based QTL or GWAS approaches may fail to identify key contributors to variation of a trait if the causative chromatin modification is not captured by genetic markers. The key question we explore here is the degree to which variation in a chromatin modification such as DNA methylation will be tagged by SNPs or other genetic markers. To address this question we must consider the sources and stability of chromatin variation and the structure of the population being considered.

Sources and stability of 'epigenetic' variation

As we seek to understand the importance of epigenetic variation in plant populations, it is necessary to consider the sources and stability of this variation. The inheritance of an epigenetic state could potentially range from high levels of stability, like a genetic variant, to complete instability from one generation to the next. Many studies that have sought to understand the sources and stability of epigenetic variation use DNA methylation in plant populations as a proxy for epigenetics. Genome-wide analyses of DNA methylation identify many differentially methylated regions (DMRs) among individuals [3–5]. While many quantitative differences in DNA methylation level for a locus have been identified, the regions changing between highly methylated and unmethylated states are most likely to represent heritable differences that may contribute to altered gene expression. It is tempting to consider all chromatin variation as true epigenetic variation, independent of sequence. However, there is

abundant evidence that a significant portion of this variation in DNA methylation might be explained by genetic changes [6]. Richards [7] created useful terminology for considering the interaction of genetic and epigenetic variation (Box 1). As we consider the potential for DNA methylation variation to provide additional information beyond SNPs, it might be helpful to separate the sources of epigenetic variation into local (cis-acting) variants and other factors. Examples of local genetic variants that are strongly associated with an altered chromatin state (i.e. obligatory epialleles) will likely be quite stable and offer limited potential for novel information based solely on the chromatin state. In contrast, other sources of epigenetic variation including trans-acting genetic variants that trigger stable chromatin variants, environmental factors or spontaneous epimutation have the potential to create variation that is not predicted based on SNPs or other genetic markers.

Trans-acting genetic variation and genetic background can influence epigenetic variation and stability. There are well-characterized examples of genetic variants that create allelic interaction in trans (paramutation) or at unlinked genomic sites [8]. These may result in epigenetic variants that are initially predictable based on genetic variation such as SNPs. However, the behavior of this chromatin variation following segregation can create scenarios in which the epigenetic state is uncoupled with the original genetic trigger. For example, the altered epigenetic states following paramutation can be very stable or decay over the course of several generations [9]. Instances of genetic variants that trigger a stable trans-acting epigenetic change but are themselves lost due to segregation will create epigenetic variants that are not well-tagged by SNPs or other genetic markers.

Box 1 Spectrum of potential genetic influences on chromatin variation

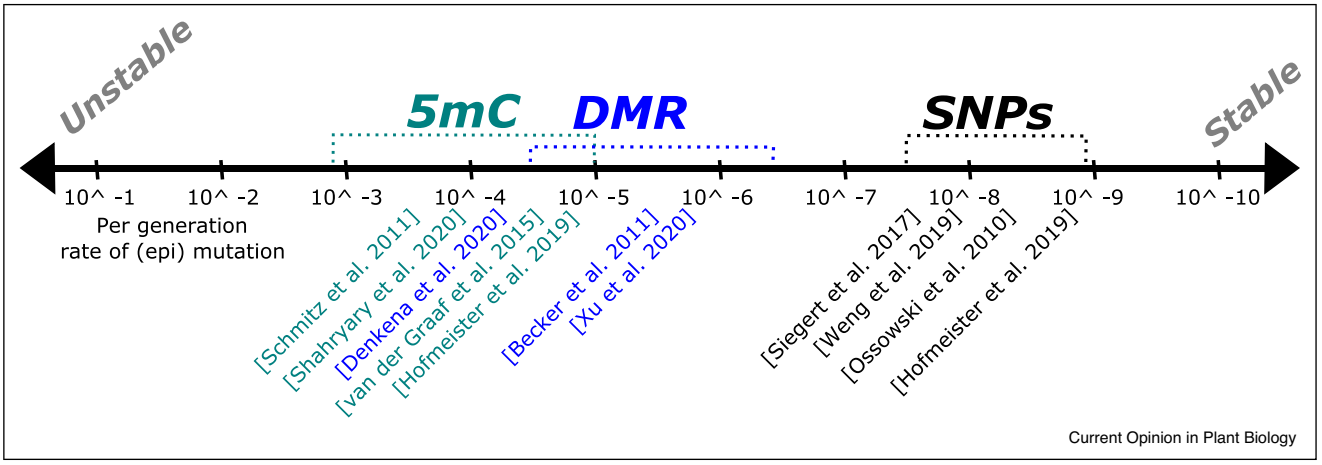
Variation for chromatin state in different haplotypes can range from completely dependent on genetic variation to being completely independent. According to Richards [7] an **obligatory epiallele** occurs when a genetic change (such as a transposon insertion or structural variant) triggers a chromatin change. For example, a transposon insertion may trigger high levels of DNA methylation for the transposon itself as well as the flanking sequences [65]. At the other extreme a **pure epiallele** would represent instances in which epigenetic variation arises with no genetic influence. Pure epialleles could arise through spontaneous epimutation or through variation triggered by environmental or developmental conditions. Between these two extremes there is the potential for several types of **facilitated epialleles**. One instance of facilitated epiallele would occur when the presence of a genetic variant, like a transposable element, predisposes a region to chromatin variation but is not completely penetrant. This leads to partial but not complete association of the genetic variant and the chromatin state. Another instance of a facilitated epiallele could occur with trans-acting effects. For example, a genetic variant that creates an inverted repeat (as seen at the PAI locus in Arabidopsis [66] could lead to small RNAs that could trigger high levels of DNA methylation at the other allele or at other genomic locations with high homology to the inverted repeat sequence. Importantly, if the hypermethylation state is heritable high levels of DNA methylation could be maintained even after segregation of the triggering locus. This would result in an apparent pure epiallele that originally was attributed to a genetic variant. These examples highlight the complexities in determining the linkage between chromatin variation and genetic changes.

In contrast to epigenetic changes that are triggered by trans-acting variants there are also spontaneous epigenetic changes. The best knowledge of spontaneous epimutation frequencies in plants has come from analyses of mutation accumulation lines in Arabidopsis that have minimal genetic variation [10–13]. The rates of spontaneous epigenetic variation for single sites are several orders of magnitude higher than rates for SNPs [14,15] (Figure 1). Model based approaches designed to estimate epimutation rates suggest some variation for different plant species but place the estimates in a generally similar range [16••]. Similar estimates were obtained for maize based on population genetics based approaches [17•]. It is more difficult to estimate frequencies for differentially methylated regions but estimates suggest these occur at an overall frequency similar to site-specific methylation changes [18••]. These different approaches to estimate epimutation rates in various plant species suggest slightly different rates but still place epigenetic variation into a unique position of being quite heritable but far less stable than genetic changes. Several studies have investigated whether the frequency of spontaneous epimutations may be influenced by different environments or conditions [19–30]. While some treatments, such as tissue culture, have been associated with increased changes in DNA methylation, the influence of abiotic stresses on DNA methylation has varied from negligible to significant in different studies. The analysis of DNA methylation patterns in wild Arabidopsis populations suggested rates of accumulation of epimutations in natural environments that is similar to that observed in mutation accumulation lines, suggesting limited roles for environmental variation influencing epimutation rate [31]. It is likely that the specific rate of epimutations and their distribution in different genomic regions may be influenced by environmental factors as well as genetic background.

Different populations; different potential for epigenetic variation

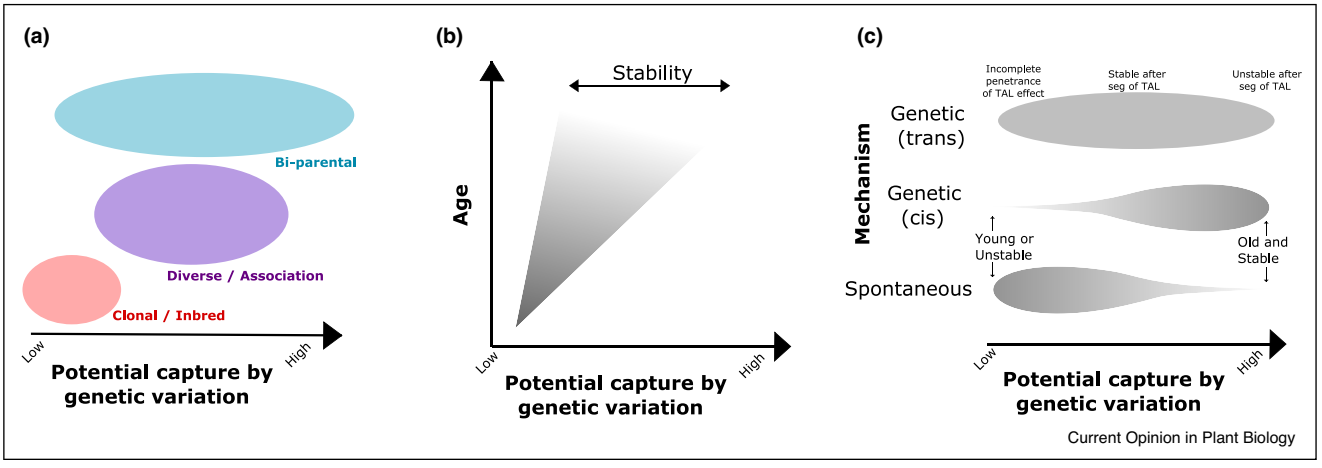
The stability of epigenetic variation and epimutation rates become critical factors as we consider the potential for untagged epigenetic variation in different types of plant populations and there are different factors that become important in disentangling genetic and epigenetic variation in these different populations. The potential to capture DNA methylation variation through the use of SNP-based profiling will vary in different populations as a consequence of the dynamics of the age, stability and mechanisms that generate epialleles (Figure 2). We will begin with plant

Figure 1



Relative stability of DNA methylation and SNPs. Several studies have monitored the epimutation rates for DNA methylation at specific sites (5mC - green) or for differentially methylated regions (DMRs-blue) in comparison to SNPs (black). The specific studies are referenced at the approximate position of the reported rates. Although DNA methylation levels are highly heritable they are orders of magnitude less stable than SNPs.

Figure 2



The potential to capture or tag epialleles using SNPs or other genetic variants is influenced by the population type, age, stability and mechanistic basis of the epiallele. **(a)** The potential to capture epiallelic variation using genetic variation is relatively low in clonal or inbred populations. Bi-parental population and diverse association panels will exhibit a range in which some epialleles are well tagged while others are not captured. **(b)** In all types of populations the age and stability of inheritance for epialleles will influence the ability to capture this information using SNPs. Relatively young variants will also be difficult to tag using genetic variation scans. Epialleles that are older will exhibit a range potential capture by genetic variation resulting from different stabilities of inheritance. **(c)** The ability to tag epialleles with genetic variation will also vary based on the mechanistic basis of the epiallele origin. Both spontaneous epialleles and cis-linked genetic causes of epialleles will have a range of potential capture depending on the age and stability. It is likely that there are greater frequencies of young, unstable spontaneous epialleles and older stable genetic (cis) epialleles. Genetic epialleles induced by a trans-acting locus (TAL) trigger will exhibit a range of potential capture by genetic variation. If the TAL trigger has incomplete penetrance for inducing a chromatin change there will be low capture by genetic variation. Loci with high penetrance of the TAL effect but unstable inheritance of the induced epigenetic state will be well tagged by genetic variation in bi-parental or association populations. However, in cases in which the TAL induces a stable effect that remains after the TAL is segregated away there will only be partial association between the TAL and the chromatin change.

populations with minimal genetic variation and progress to populations with higher levels of genetic variation.

Clonal populations and inbred lines

The simplest populations from a genetic perspective will be clonal populations or inbred lines. Many crop species are clonally propagated which result in a population with relatively little genetic variation and this means that new variants (genetic or epigenetic) will not be tagged by SNPs (Figure 2a). In these species, characterizing spontaneous epigenetic variation will likely be very important for understanding sports or somaclonal variants [32]. One prominent recent example was the discovery of the *Bad karma* locus in somaclonal variants of oil palms [33,34]. This epigenetic variant arises at moderate (5–20% of individuals) frequency in clonally propagated oil palms and epigenetic assays have been developed to use for culling of affected individuals. Several recent studies have also provided insights into the epigenetic variability in some fruit species such as apple [35–37], grape [38,39] or poplar [40–42]. While these studies suggest promise for linking chromatin variants to novel phenotypes that have arisen in specific sports, it is worth noting that in general there is quite limited breeding progress using selection on inbred materials, suggesting limited potential for spontaneous epigenetic variants that allow for rapid shifts in quantitative traits. However, in some special cases there can be major epigenetic variation within inbred populations. The epiRILs were intentionally generated through crossing plants that are homozygous mutant for factors critical for DNA methylation with wild-type plants [43,44]. The off-springthat are homozygous wild type (lacking mutant alleles for DNA methylation pathways) segregate for genomic regions that have experienced loss of methylation. These populations give insight into the stability of epigenetic variants and show highly variable behavior for different loci. Some loci quickly regain wild-type methylation levels while others show stochastic rare recovery or stable unmethylated states [45]. These populations also show quantitative trait variation suggesting that segregation for varying chromatin states can influence many traits [46•,47–50]. EpiRILs provide examples of how epigenetic variants that are not tagged by SNPs could result in quantitative trait variation.

Bi-parental populations

As we shift to consider populations with segregating genetic variation it becomes more challenging to disentangle genetic and epigenetic sources of phenotypic variation [6]. These populations often have much higher rates of chromatin variation but this is present in haplotypes that also have genetic variants. Since genetic variants can have local effects on chromatin or trans effects at allelic positions (i.e. paramutation) or elsewhere in the genome, it becomes important to be able to resolve whether chromatin state differences are controlled by other genetic variants. One major challenge is that the lower stability for epigenetic variants relative to genetic variants changes the potential to use imputation approaches. While high-quality information on genetic variation from parents of a population can be accurately imputed to off-spring based on a smaller number of markers that define recombinations this approach should not be applied to high resolution maps of chromatin variants from parental genomes due to the lower stability of these variants.

Several types of genetically variable plant populations offer distinct potential for considering the role of epigenetic variation (Figure 2a). Bi-parental populations (including F2, recombinant inbred lines and near isogenic lines) provide opportunities to monitor the stability of chromatin variants and the potential for trans-acting control of DNA methylation. In general, studies in these populations have revealed widespread evidence for relatively stable inheritance of DNA methylation levels based on the stable inheritance of epialleles with some examples of unstable inheritance [5,51–54]. While these bi-parental populations can be quite useful for insights into inheritance patterns of chromatin variation, they offer very limited ability to separate genetic and epigenetic variation. Since these populations often have limited genetic resolution each chromatin variant is often in linkage disequilibrium with many nearby SNPs or other genetic changes. Each QTL will potentially contain multiple genetic and epigenetic variants and isolating the causative variant can be challenging. In addition, the potential to tag epialleles using SNP variation will be influenced by the age of the epiallele, the stability of the epigenetic state and the mechanistic basis of the epiallele (Figure 2b,c).

Diverse association panels

Moving to diverse association panels with GWAS can provide increased genetic resolution. In these populations there are both increased numbers of crossovers as well as additional generations that provide additional opportunity for the accumulation of spontaneous epimutations. This increases the opportunity to identify chromatin variants that are not well tagged by genetic variants. To date there have been relatively few scans of chromatin profiles in very large diverse plant populations. The only comprehensive profiling of DNA methylation at true population scales has been performed in Arabidopsis [3,4,55]. While there are many single-base methylation polymorphisms that are detected in this population, the phylogeny based on methylation polymorphisms is highly similar to SNP-based phylogeny suggesting overall stable inheritance of DNA methylation patterns [3]. Many of the differences in DNA methylation in Arabidopsis populations appear to have a genetic basis with examples of local, cis-acting variants as well as trans-acting variants that frequently map to genomic locations of genes known to play a role in

the regulation of DNA methylation [4,55]. More limited scans have been performed in brachypodium [56], rice [57•], soybean [58•] and maize [5,17•,59••]. The analysis of 45 soybean methylomes from wild accessions and domesticated lines reveals many changes in DNA methylation that are often associated with higher levels of nearby genetic variation [58•]. Similarly, the analysis of nearly 100 maize and teosinte methylomes identifies many differences in DNA methylation that include many examples that are associated with genomic regions that have undergone selection during domestication [17•]. These DNA methylation differences may have functional consequences that were the basis of selection or may simply reflect variants that 'hitch-hiked' with selection for nearby genetic variants. A capture-based profiling of DNA methylation at selected regions of the maize genome in combination with high-depth SNP panels on the same population reveals that only about half of the differentially methylated regions were effectively captured by SNPs [59••]. Importantly, using the DNA methylation variants could effectively predict variation for some gene expression or metabolite traits that were not strongly associated with SNPs providing evidence for potential value of DNA methylation profiles for predicting traits.

Conclusions

The analyses of diverse populations in several plant species highlight the potential for novel epigenetic variants that are not well captured in SNP-based scans to influence plant traits. Further studies will be necessary to document the full role of epigenetics in quantitative trait variation in plant populations. It will be important to design these studies in a fashion that can disentangle the effects of chromatin variation as opposed to hitch-hiking of a stable chromatin variant with nearby genetic changes. Many of the current population genetics based analyses of loci involved in domestication or adaptation will not have sufficient power to fully resolve the genetic and epigenetic variation at selected loci. However, in some cases these studies have pointed to intriguing potential for epigenetic variation at these loci. Several recent studies have reported potential technologies for targeted addition or removal of DNA methylation at specific loci [60••,61–64]. These approaches provide new opportunities for disentangling the role of DNA methylation and genetic variation by providing ways to trigger a methylation change with no genetic variation. A more complete understanding of the sources and stability of epigenetic variation in different plant populations will be critical as we seek to determine the importance, and potential value, of chromatin profiles for understanding phenotypic variation in plants.

Funding

This work is supported by the National Science Foundation (grants IOS-1934384 and IOS-1802848.

Conflict of interest statement

Nothing declared.

Acknowledgements

The authors gratefully acknowledge Robert Schmitz and Peter Crisp for helpful discussions and editing suggestions. We are grateful to our colleagues for the inspiring work on this field and apologize for not being able to reference all works relevant to this topic due to space limitations.

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. Maher B: **Personal genomes: the case of the missing heritability**. *Nature* 2008, **456**:18-21.
2. Plongthongkum N, Diep DH, Zhang K: **Advances in the profiling of DNA modifications: cytosine methylation and beyond**. *Nat Rev Genet* 2014, **15**:647-661.
3. 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.
4. 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.
5. 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.
6. Taudt A, Colome-Tatche M, Johannes F: **Genetic sources of population epigenomic variation**. *Nat Rev Genet* 2016, **17**:319-332.
7. Richards EJ: **Inherited epigenetic variation—revisiting soft inheritance**. *Nat Rev Genet* 2006, **7**:395-401.
8. Melquist S, Luff B, Bender J, Arabidopsis PAI: **Gene arrangements, cytosine methylation and expression**. *Genetics* 1999, **153**:401-413.
9. Hollick JB: **Paramutation and related phenomena in diverse species**. *Nat Rev Genet* 2017, **18**:5-23.
10. van der Graaf A, Wardenaar R, Neumann DA, Taudt A, Shaw RG, Jansen RC, Schmitz RJ, Colome-Tatche M, Johannes F: **Rate, spectrum, and evolutionary dynamics of spontaneous epimutations**. *Proc Natl Acad Sci U S A* 2015, **112**:6676-6681.
11. 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.
12. 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.
13. Hofmeister BT, Lee K, Rohr NA, Hall DW, Schmitz RJ: **Stable inheritance of DNA methylation allows creation of epigenotype maps and the study of epiallele inheritance patterns in the absence of genetic variation**. *Genome Biol* 2017, **18**:155.
14. Ossowski S, Schneeberger K, Lucas-Lledo JI, Warthmann N, Clark RM, Shaw RG, Weigel D, Lynch M: **The rate and molecular spectrum of spontaneous mutations in *Arabidopsis thaliana***. *Science* 2010, **327**:92-94.
15. Schmid-Siegert E, Sarkar N, Iseli C, Calderon S, Gouhier-Darimont C, Chrast J, Cattaneo P, Schütz F, Farinelli L, Pagni M *et al.*: **Low number of fixed somatic mutations in a long-lived oak tree**. *Nat Plants* 2017, **3**:926-929.

16. •• Shahryary Y, Symeonidi A, Hazarika RR, Denkena J, Mubeen T, Hofmeister B, van Gurp T, Colomé-Tatché M, Verhoeven KJF, Tuskan G, Schmitz RJ, Johannes F: **AlphaBeta: computational inference of epimutation rates and spectra from high-throughput DNA methylation data in plants**. *Genome Biol* 2020, **21**:260

This work developed a framework for estimating spontaneous epimutation rates and applies this approach to several plant species.

17. • Xu G, Lyu J, Li Q, Liu H, Wang D, Zhang M, Springer NM, Ross-Ibarra J, Yang J: **Evolutionary and functional genomics of DNA methylation in maize domestication and improvement**. *Nat Commun* 2020, **11**:5539

A population level survey of variable methylation levels in teosinte and maize to identify potential roles of epigenetic change during domestication and selection.

18. •• Denkena J, Johannes F, Colomé-Tatché M: *Region-level Epimutation Rates in **Arabidopsis thaliana***. 2020 http://dx.doi.org/10.1101/2020.08.18.255919

The first estimation of spontaneous epimutation rates for regional changes in DNA methylation (DMRs).

19. 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.

20. Stroud H, Ding B, Simon SA, Feng S, Bellizzi M, Pellegrini M, Wang GL, Meyers BC, Jacobsen SE: **Plants regenerated from tissue culture contain stable epigenome changes in rice**. *eLife* 2013, **2**:e00354.

21. Han Z, Crisp PA, Stelpflug S, Kaeppler SM, Li Q, Springer NM: **Heritable epigenomic changes to the maize methylome resulting from tissue culture**. *Genetics* 2018, **209**:983-995.

22. Nguyen HM, Kim M, Ralph PJ, Marín-Guirao L, Pernice M, Procaccini G: **Stress memory in seagrasses: first insight into the effects of thermal priming and the role of epigenetic modifications**. *Front Plant Sci* 2020, **11**.

23. Ganguly DR, Crisp PA, Eichten SR, Pogson BJ: **The Arabidopsis DNA methylome is stable under transgenerational drought stress**. *Plant Physiol* 2017, **175**:1893-1912.

24. Zheng X, Chen L, Xia H, Wei H, Lou Q, Li M, Li T, Luo L: **Transgenerational epimutations induced by multi-generation drought imposition mediate rice plant's adaptation to drought condition**. *Sci Rep* 2017, **7**:39843.

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

26. Lämke J, Bäurle I: **Epigenetic and chromatin-based mechanisms in environmental stress adaptation and stress memory in plants**. *Genome Biol* 2017, **18**:124.

27. 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**.

28. Secco D, Wang C, Shou H, 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** http://dx.doi.org/10.7554/eLife.09343.

29. Ji L, Mathioni SM, Johnson S, Tucker D, Bewick AJ, Do Kim K, Daron J, Keith Slotkin R, Jackson SA, Parrott WA *et al.*: **Genome-wide reinforcement of DNA methylation occurs during somatic embryogenesis in soybean**. *Plant Cell* 2019, **31**:2315-2331.

30. Colicchio JM, Kelly JK, Hileman LC: **Parental experience modifies the Mimulus methylome**. *BMC Genomics* 2018, **19**:746.

31. 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.

32. Latutrie M, Gourcilleau D, Pujol B: **Epigenetic variation for agronomic improvement: an opportunity for vegetatively propagated crops**. *Am J Bot* 2019, **106**:1281-1284.

33. 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.

34. Sarpan N, Taranenko E, Ooi S-E, Low E-TL, Espinoza A, Tatarinova TV, Ong-Abdullah M: **DNA methylation changes in clonally propagated oil palm**. *Plant Cell Rep* 2020, **39**:1219-1233.

35. Jiang S-H, Sun Q-G, Chen M, Wang N, Xu H-F, Fang H-C, Wang Y-C, Zhang Z-Y, Chen X-S: **Methylome and transcriptome analyses of apple fruit somatic mutations reveal the difference of red phenotype**. *BMC Genomics* 2019, **20**:117.

36. Daccord N, Celton J-M, Linsmith G, Becker C, Choisne N, Schijlen E, van de Geest H, Bianco L, Micheletti D, Velasco R *et al.*: **High-quality de novo assembly of the apple genome and methylome dynamics of early fruit development**. *Nat Genet* 2017, **49**:1099-1106.

37. Li W-F, Ning G-X, Mao J, Guo Z-G, Zhou Q, Chen B-H: **Whole-genome DNA methylation patterns and complex associations with gene expression associated with anthocyanin biosynthesis in apple fruit skin**. *Planta* 2019, **250**:1833-1847.

38. Ocaña J, Walter B, Schellenbaum P: **Stable MSAP markers for the distinction of *Vitis vinifera* cv Pinot noir clones**. *Mol Biotechnol* 2013, **55**:236-248.

39. Xie H, Konate M, Sai N, Tesfamicael KG, Cavagnaro T, Gilliham M, Breen J, Metcalfe A, Stephen JR, De Bei R *et al.*: **Global DNA methylation patterns can play a role in defining terroir in grapevine (*Vitis vinifera* cv. Shiraz)**. *Front Plant Sci* 2017, **8**:1860.

40. Lu W, Xiao L, Quan M, Wang Q, El-Kassaby YA, Du Q, Zhang D: **Linkage-linkage disequilibrium dissection of the epigenetic quantitative trait loci (epiQTLs) underlying growth and wood properties in *Populus***. *New Phytol* 2020, **225**:1218-1233.

41. Schönberger B, Chen X, Mager S, Ludewig U: **Site-dependent differences in DNA methylation and their impact on plant establishment and phosphorus nutrition in *Populus trichocarpa***. *PLoS One* 2016, **11**:e0168623.

42. Hofmeister BT, Denkena J, Colomé-Tatché M, Shahryary Y, Hazarika R, Grimwood J, Mamidi S, Jenkins J, Grabowski PP, Sreedasyam A *et al.*: **A genome assembly and the somatic genetic and epigenetic mutation rate in a wild long-lived perennial *Populus trichocarpa***. *Genome Biol* 2020, **21**:259.

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

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

45. Catoni M, Griffiths J, Becker C, Zabet NR, Bayon C, Dapp M, Lieberman-Lazarovich M, Weigel D, Paszkowski J: **DNA sequence properties that predict susceptibility to epiallelic switching**. *EMBO J* 2017, **36**:617-628.

46. • Zhang Y-Y, Latzel V, Fischer M, Bossdorf O: **Understanding the evolutionary potential of epigenetic variation: a comparison of heritable phenotypic variation in epiRILs, RILs, and natural ecotypes of *Arabidopsis thaliana***. *Heredity* 2018, **121**:257-265

A detailed comparison of potential methylation changes associated with phenotype in distinct types of Arabidopsis populations.

47. Zhang YY, Fischer M, Colot V, Bossdorf O: **Epigenetic variation creates potential for evolution of plant phenotypic plasticity**. *New Phytol* 2013, **197**:314-322.

48. Kooke R, Johannes F, Wardenaar R, Becker F, Etcheverry M, Colot V, Vreugdenhil D, Keurentjes JJ: **Epigenetic basis of morphological variation and phenotypic plasticity in *Arabidopsis thaliana***. *Plant Cell* 2015, **27**:337-348.

49. Cortijo S, Wardenaar R, Colome-Tatche M, Gilly A, Etcheverry M, Labadie K, Caillieux E, Hospital F, Aury JM, Wincker P *et al.*: **Mapping the epigenetic basis of complex traits**. *Science* 2014, **343**:1145-1148.

50. Furci L, Jain R, Stassen J, Berkowitz O, Whelan J, Roquis D, Baillet V, Colot V, Johannes F, Ton J: **Identification and characterisation of hypomethylated DNA loci controlling quantitative resistance in Arabidopsis**. *eLife* 2019, **8**.

51. 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.

52. Li Q, Eichten SR, Hermanson PJ, Springer NM: **Inheritance patterns and stability of DNA methylation variation in maize near-isogenic lines**. *Genetics* 2014, **196**:667-676.

53. Regulski M, Lu Z, Kendall J, Donoghue MT, Reinders J, Llaca V, Deschamps S, Smith A, Levy D, McCombie WR *et al.*: **The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA**. *Genome Res* 2013, **23**:1651-1662.

54. Vaughn MW, Tanurd Ic M, Lippman Z, Jiang H, Carrasquillo R, Rabinowicz PD, Dedhia N, McCombie WR, Agier N, Bulski A *et al.*: **Epigenetic natural variation in *Arabidopsis thaliana***. *PLoS Biol* 2007, **5**:e174.

55. Dubin MJ, Zhang P, Meng D, Remigereau MS, Osborne EJ, Casale FP, 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.

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

57. • Zhao L, Xie L, Zhang Q, Ouyang W, Deng L, Guan P, Ma M, Li Y, Zhang Y, Xiao Q *et al.*: **Integrative analysis of reference epigenomes in 20 rice varieties**. *Nat Commun* 2020, **11**:2658

In depth analysis of epigenome variation among rice varieties.

58. • Shen Y, Zhang J, Liu Y, Liu S, Liu Z, Duan Z, Wang Z, Zhu B, Guo Y-L, Tian Z: **DNA methylation footprints during soybean domestication and improvement**. *Genome Biol* 2018, **19**:128

Identification of methylation changes that are potentially associated with domestication or crop improvement in soybean.

59. •• Xu J, Chen G, Hermanson PJ, Xu Q, Sun C, Chen W, Kan Q, Li M, Crisp PA, Yan J *et al.*: **Population-level analysis reveals the widespread occurrence and phenotypic consequence of DNA methylation variation not tagged by genetic variation in maize**. *Genome Biol* 2019, **20**:243

A population level survey of variable DNA methylation at a subset of loci in the maize genome identifies variation in methylation that is not captured by SNPs that is associated with trait variation.

60. •• Gallego-Bartolomé J, Gardiner J, Liu W, Papikian A, Ghoshal B, Kuo HY, Zhao JM-C, Segal DJ, Jacobsen SE: **Targeted DNA demethylation of the Arabidopsis genome using the human TET1 catalytic domain**. *Proc Natl Acad Sci U S A* 2018, **115**: E2125-E2134

This work describes the development of a system for targeted demethylation of specific loci in the Arabidopsis genome.

61. Ji L, Jordan WT, Shi X, Hu L, He C, Schmitz RJ: **TET-mediated epimutagenesis of the *Arabidopsis thaliana* methylome**. *Nat Commun* 2018, **9**:895.

62. Johnson LM, Du J, Hale CJ, Bischof S, Feng S, Chodavarapu RK, Zhong X, Marson G, Pellegrini M, Segal DJ *et al.*: **SRA- and SET-domain-containing proteins link RNA polymerase V occupancy to DNA methylation**. *Nature* 2014, **507**:124-128.

63. Papikian A, Liu W, Gallego-Bartolomé J, Jacobsen SE: **Site-specific manipulation of Arabidopsis loci using CRISPR-Cas9 SunTag systems**. *Nat Commun* 2019, **10**.

64. Ghoshal B, Vong B, Picard CL, Suhua F, Tam JM, Jacobsen SE: **A viral guide RNA delivery system for CRISPR-based transcriptional activation and heritable targeted DNA demethylation in *Arabidopsis thaliana***. *PLoS Genet* 2020, **16**: e1008983.

65. Noshay JM, Anderson SN, Zhou P, Ji L, Ricci W, Lu Z, Stitzer MC, Crisp PA, Hirsch CN, Zhang X *et al.*: **Monitoring the interplay between transposable element families and DNA methylation in maize**. *PLoS Genet* 2019, **15**:e1008291.

66. Melquist S, Bender J: **An internal rearrangement in an Arabidopsis inverted repeat locus impairs DNA methylation triggered by the locus**. *Genetics* 2004, **166**:437-448.