The Plant Cell, Vol. 25: 2783–2797, August 2013, www.plantcell.org

LARGE-SCALE BIOLOGY ARTICLE

Epigenetic and Genetic Influences on DNA Methylation Variation in Maize Populations[C][W]

Steven R. Eichten,[a] Roman Briskine,[b] Jawon Song,[c] Qing Li,[a] Ruth Swanson-Wagner,[a] Peter J. Hermanson,[a] Amanda J. Waters,[a] Evan Starr,[a] Patrick T. West,[a] Peter Tiffin,[a] Chad L. Myers,[b] Matthew W. Vaughn,[c] and Nathan M. Springer[a,1]

[a] Microbial and Plant Genomics Institute, Department of Plant Biology, University of Minnesota, Saint Paul, Minnesota 55108
[b] Department of Computer Science and Engineering, University of Minnesota, Minneapolis, Minnesota 55455
[c] Texas Advanced Computing Center, University of Texas, Austin Texas 78758

ORCID IDs: 0000-0003-2268-395X (S.R.E.); 0000-0002-7301-4759 (N.M.S.).

DNA methylation is a chromatin modification that is frequently associated with epigenetic regulation in plants and mammals. However, genetic changes such as transposon insertions can also lead to changes in DNA methylation. Genome-wide profiles of DNA methylation for 20 maize (*Zea mays*) inbred lines were used to discover differentially methylated regions (DMRs). The methylation level for each of these DMRs was also assayed in 31 additional maize or teosinte genotypes, resulting in the discovery of 1966 common DMRs and 1754 rare DMRs. Analysis of recombinant inbred lines provides evidence that the majority of DMRs are heritable. A local association scan found that nearly half of the DMRs with common variation are significantly associated with single nucleotide polymorphisms found within or near the DMR. Many of the DMRs that are significantly associated with local genetic variation are found near transposable elements that may contribute to the variation in DNA methylation. Analysis of gene expression in the same samples used for DNA methylation profiling identified over 300 genes with expression patterns that are significantly associated with DNA methylation variation. Collectively, our results suggest that DNA methylation variation is influenced by genetic and epigenetic changes that are often stably inherited and can influence the expression of nearby genes.

INTRODUCTION

Heritable information is most commonly found as genetic variation between individuals and populations. However, recent studies have restored interest in epigenetic variation, heritable variation that is not directly connected to DNA sequence polymorphisms (Bird, 2007). Epigenetic variation is often associated with a variety of chromatin marks, such as histone tail modifications, small RNAs, and methylation of cytosine in genomic DNA (Jablonka and Raz, 2009). The methylation of cytosines in genomic DNA is one of the best-studied examples of chromatin variation with a large body of work investigating the mechanisms of heritability as well as the actions of chromatin-modifying enzymes on particular methylation states (Stroud et al., 2013).

DNA methylation often suppresses the expression of transposable elements, pseudogenes, repetitive sequences, and individual genes (Chan et al., 2005; Slotkin and Martienssen 2007). DNA methylation variation has been implicated in stable phenotypic variation in plants and animals (Rasmusson and Phillips, 1997; Richards, 2006, 2008; Feinberg, 2007; Jirtle and Skinner, 2007; Vaughn et al., 2007; Johannes et al., 2008). Well-characterized examples of natural variants displaying differing DNA methylation patterns, deemed epialleles, include *Lcyc*, a *cycloidea* homologue, (Cubas et al., 1999), *Colorless Nonripening* (Manning et al., 2006), *CmWIP1* (Martin et al., 2009), *FOLT1* (Durand et al., 2012), and *PAI* (Luff et al., 1999; Melquist et al., 1999). There are also examples of natural variation for epigenetic state associated with paramutation at several loci in maize (*Zea mays* ssp *mays*) (Chandler et al., 2000). Many examples of natural variation for DNA methylation exhibit occasional changes of methylation state, suggesting the possibility of semistable heritability for this epigenetic information (Vaughn et al., 2007; Gutzat and Mittelsten Scheid, 2012). There is increased interest in epigenetic phenomenon within plant systems as a possible source for previously unassessed heritable variation for improving phenotypic predictions and breeding methods. In order to fully evaluate the potential of epigenetics for understanding phenotypic variation, it is important both to characterize epigenetic variation and to understand the association between genetic and epigenetic changes at particular loci.

Genome-wide profiling of DNA methylation abundance in several plant species has documented variation among genotypes or mutant strains. In *Arabidopsis thaliana*, DNA methylation is frequently associated with repetitive DNA sequences, such as transposable elements and heterochromatic regions

[1] Address correspondence to springer@umn.edu.
The author responsible for distribution of materials integral to the findings presented in this article in accordance with the policy described in the Instructions for Authors (www.plantcell.org) is: Nathan M. Springer (springer@umn.edu).
[C] Some figures in this article are displayed in color online but in black and white in the print edition.
[W] Online version contains Web-only data.
www.plantcell.org/cgi/doi/10.1105/tpc.113.114793

(Lippman et al., 2004; Vaughn et al., 2007; Zilberman et al., 2007; Miura et al., 2009), although there is also evidence for intermediate levels of CG methylation within some gene bodies. Mutations in genes involved in chromatin modification or small interfering RNA generation and processing can have global or locus-specific effects on DNA methylation (Stroud et al., 2013). Several studies that generated epiRILs, genotypes that segregate for differences in DNA methylation with limited genetic changes (due to transposon mobilization), provide evidence that alterations in DNA methylation patterns have the potential to influence quantitative traits in *Arabidopsis* (Johannes et al., 2009; Reinders et al., 2009). The analysis of spontaneous mutation accumulation provided evidence that changes in DNA methylation at single nucleotides occurs at a much higher rate than mutations in nucleotide sequence (single nucleotide polymorphisms, or SNPs) but that changes in regional methylation levels (differentially methylated regions [DMRs]) occur at frequencies roughly similar to the single nucleotide mutation rate (Becker et al., 2011; Schmitz et al., 2011).

Comparison of DNA methylation levels in different ecotypes reveals relatively stable inheritance of DNA methylation that is associated with repetitive sequences but less stable inheritance of gene body methylation (Vaughn et al., 2007; Schmitz et al., 2013b). Detailed profiling of DNA methylation in 152 *Arabidopsis* ecotypes revealed many examples of single nucleotide differences in DNA methylation as well as examples of regions that are highly methylated in some ecotypes but lack methylation in other ecotypes (Schmitz et al., 2013b). Association mapping of these DMRs shows that many of these are locally (*cis*) controlled with some examples of *trans*-acting effects on DNA methylation (Schmitz et al., 2013b). Similar results were found by assessing DNA methylation variation in rice (*Oryza sativa*) and looking at allele-specific DNA methylation and expression in the F1 hybrid (Chodavarapu et al., 2012). The analysis of nearly 80 soybean (*Glycine max*) recombinant inbred line (RILs) found that the majority of variable DNA methylation was controlled by local methylQTL (quantitative trait loci) (Schmitz et al., 2013a). In maize, the comparison of DNA methylation in two genotypes revealed that some DNA methylation appears to be purely epigenetic (Eichten et al., 2011), whereas other regions of DNA methylation likely result from transposon insertion variation (Eichten et al., 2012). While many examples of DNA methylation variation are stably inherited in maize, there are examples of regions that have altered DNA methylation levels in offspring (Regulski et al., 2013).

DNA methylation is often considered to be an epigenetic mark. However, DNA methylation is specifically a chromatin modification that can be the result of genetic or epigenetic influences. Richards (2006) provided a classification scheme to consider the relative influence of genetics and epigenetics upon chromatin state at epialleles. Obligatory epialleles represent examples in which the chromatin state is directly correlated with a genetic change, such as a transposon insertion or a structural variant. Facilitated epialleles occur when a genetic change leads to a poised allelic state that could exist in either a silenced or active form. Pure epialleles describe instances in which there are no genetic changes that influence chromatin state. There is evidence that each of these types of epialleles exist in plant populations. Many of the known loci showing DNA methylation variation (Luff et al., 1999; Melquist et al., 1999; Martin et al., 2009; Durand et al., 2012) are located near sequence variation or transposable element insertions that could be causal factors determining the methylation state. Indeed, the hypermethylation of certain retrotransposon elements has been shown to spread DNA methylation to low copy regions flanking the element (Hollister et al., 2011; Eichten et al., 2012) and could lead to obligatory epiallele formation. Other methylation variants have been connected to structural variation for related sequences located in other areas of the genome in which small interfering RNAs may play a role in directing the methylation state at the observed loci (Luff et al., 1999; Melquist et al., 1999; Durand et al., 2012). By contrast, epialleles such as the *Lcyc*, a *cycloidea* homologue, locus, the *Colorless Nonripening* locus, and the *SUPERMAN*/*clark kent* alleles, do not display any sequence variation within 10 kb of the gene between lines displaying different methylation states (Jacobsen and Meyerowitz, 1997; Cubas et al., 1999; Manning et al., 2006). Similarly, some examples of differential DNA methylation in maize exist in genomic regions without genetic differences and are not influenced by other genomic regions (Eichten et al., 2011).

Understanding the relative contribution of these three classes of epialleles (obligatory, facilitated, and pure) is important for several reasons. First, if the majority of DMRs are obligatory, then it is possible that SNPs will be in linkage disequilibrium (LD) with the genetic change and could be used to predict or infer DNA methylation state. By contrast, facilitated and pure epialleles have information content that would not be captured by DNA genotyping approaches. Second, these different types of epialleles are expected to exhibit differences in stability and heritability. If the chromatin state is programmed by genetic features, then the chromatin state should be stable and reproducible across generations. By contrast, facilitated or pure epialleles might be much less stable and exhibit reversions in chromatin state at much higher frequencies.

In this study, we investigate the diversity of DNA methylation states and their association to genotype and gene expression in 51 diverse maize inbred lines. Maize is a highly diverse species (Buckler et al., 2006; Messing and Dooner, 2006; Chia et al., 2012) that provides a useful model to study the role of epigenetic variation due to the interspersed nature of genes and repetitive sequences that may create abundant opportunities for epigenetic variation (Rabinowicz and Bennetzen, 2006; Baucom et al., 2009; Schnable et al., 2009). Thousands of DMRs were found in these diverse maize lines, and many of these can be confirmed using MethylC-Seq (Lister et al., 2008). The comparison of DNA methylation levels and genotypes provides evidence that a substantial portion of DMRs are associated with local genotype. This heritability of the DNA methylation patterns for many of the DMRs was confirmed by analysis of methylation levels in RILs. The genes located near DMRs include a number of genes with very tissue-specific expression patterns, and there are over 300 genes located near DMRs that have expression patterns that are strongly correlated with the methylation state among genotypes.

RESULTS

Thousands of Differentially Methylated Regions Are Detected in Diverse Maize Inbred Lines

Genome-wide profiling of DNA methylation abundance using methylated DNA immunoprecipitation (meDIP)-chip profiling was performed on the B73 reference genome along with 19 diverse inbred genotypes of maize selected from the nested association mapping (NAM) population (process outlined in Figure 1A). The 19 genotypes include representatives of different heterotic groups as well as tropical germplasm and were selected to represent maize diversity (McMullen et al., 2009). Immunoprecipitation of methylated DNA was performed on DNA samples of each genotype from three biological replicates using a 5-methylcytosine antibody that allows for enrichment of DNA fragments that contain DNA methylation in any sequence context (CG, CHG, or CHH). The immunoprecipitated DNA and input total DNA were hybridized to a custom long-oligonucleotide microarray platform containing 2.1 M probes spaced every 200 bp across the low-copy portion of the maize genome sequence (Eichten et al., 2011). This meDIP-chip method allows for the analysis of DNA methylation abundance across the genome but does not provide information about the sequence context of DNA methylation.

The genomes of maize inbreds have high levels of structural variation, including copy number variation and presence-absence variation (Swanson-Wagner et al., 2010). A twofold approach was used to mitigate the influence of structural variation on the analysis of DNA methylation levels (see Methods for details). Any probes with more than twofold difference in hybridization intensity for genomic DNA of a genotype relative to the genomic DNA of the reference genotype (B73) were omitted from DNA methylation analysis for that genotype. For the remaining probes, the estimates of DNA methylation levels were normalized to account for slight differences in the hybridization intensities of genomic DNA between samples. The DNA methylation profile for each of the 19 genotypes was compared with the B73 DNA methylation profile in order to identify DMRs. The DNAcopy algorithm (Venkatraman and Olshen, 2007) was used to find regions consisting of multiple probes that exhibit at least twofold variation in the level of DNA methylation relative to B73 and resulted in 415 to 804 DMRs in each genotype relative to B73. The DMRs discovered for all 19 samples were condensed to a nonredundant list of 9899 DMRs that are detected in at least one genotype relative to B73 (see Supplemental Data Set 1 online).

In order to better survey the frequency of DNA methylation variation and to assess potential causes or consequences of this variation, the level of DNA methylation at these DMRs was assayed in a single biological replicate of 31 additional genotypes. These additional genotypes were part of a recent resequencing study (Chia et al., 2012) and include eight inbred lines derived from teosinte and 11 inbred lines developed from landraces (see Supplemental Table 1 online). The average per-probe DNA methylation level for each DMR was calculated for each of these genotypes. As only a single replicate of these 31 genotypes was profiled, we did not attempt to discover DMRs presented in these genotypes that were not also present in the original 20 genotypes. DMRs were validated in independent biological replicates by quantitative PCR (qPCR) (see Supplemental Figure 1 online), bisulfite sequencing (see below), and profiling in a recombinant inbred population (see below).

A series of filtering criteria were applied to the 9899 DMRs from the 19 original genotypes in order to identify a robust set of common and rare variants (Figure 1A; see Methods for details). The 1966 common variants include DMRs that have high/low levels of DNA methylation in at least three of the 51 genotypes surveyed (example in Figure 1B). The 1754 rare variants only exhibit altered DNA methylation levels in one or two of the genotypes (example in Figures 1C and 1D). The common and rare DMR variants are distributed throughout the maize genome with no obvious positional enrichments (Figure 1E). Most (2922 / 3720) of the DMRs are located in low-copy intergenic regions, and only 21% of the DMRs overlap with an annotated gene. Hierarchical clustering of both rare and common DMRs revealed that most genotypes have similar frequencies of DMRs (Figure 2A; see Supplemental Figure 2 online). There were not any particular genotypes that exhibited unique DNA methylation profiles relative to other genotypes. The relationships between genotypes based upon varying DNA methylation levels is generally similar to relationships of these genotypes observed from SNP data (Chia et al., 2012). The clustering of the rare genotypes revealed more examples of rare hypomethylation than rare hypermethylation (Figure 2A). Further analysis provided evidence for a significant ($P < 0.001$) enrichment for hypomethylation being more common for rare DMRs (see Supplemental Figure 2B online). This suggests that it is more common for methylated regions to lose DNA methylation in a small number of genotypes than for generally unmethylated regions to gain methylation in a small number of genotypes.

DNA Methylation Differences between Maize and Teosinte

Domestication may have resulted in altered DNA methylation profiles in maize. In order to assess the potential influence of domestication upon DNA methylation, average methylation levels in all maize genotypes (excluding landraces) were compared with the average levels in teosinte. There were 172 maize-teosinte DMRs that met the filtering criteria (see Supplemental Methods 1 and Supplemental Data Set 2 online). The clustering of all genotypes based on DNA methylation levels in these regions clearly separate maize and teosinte genotypes (Figure 2B). The 172 maize-teosinte DMRs are biased toward more examples of higher methylation levels in teosinte (81%) than in maize (19%; Figure 2C). Many (87%) of the maize-teosinte DMRs were not identified as either common or rare DMRs in the 19 replicated maize genotypes (Figure 2D). However, many of the maize-teosinte DMRs still have examples of both high and low DNA methylation levels within both maize and teosinte (Figure 2B). Seventeen (10%) of the maize-teosinte DMRs are found in genomic regions that were recently identified as candidates for selective sweeps during domestication (Hufford et al., 2012). This suggests that many maize-teosinte DMRs are not simply the result of selection on linked genetic variants during domestication. Seven maize-teosinte DMRs are located within 5 kb of genes that are differentially expressed between maize and

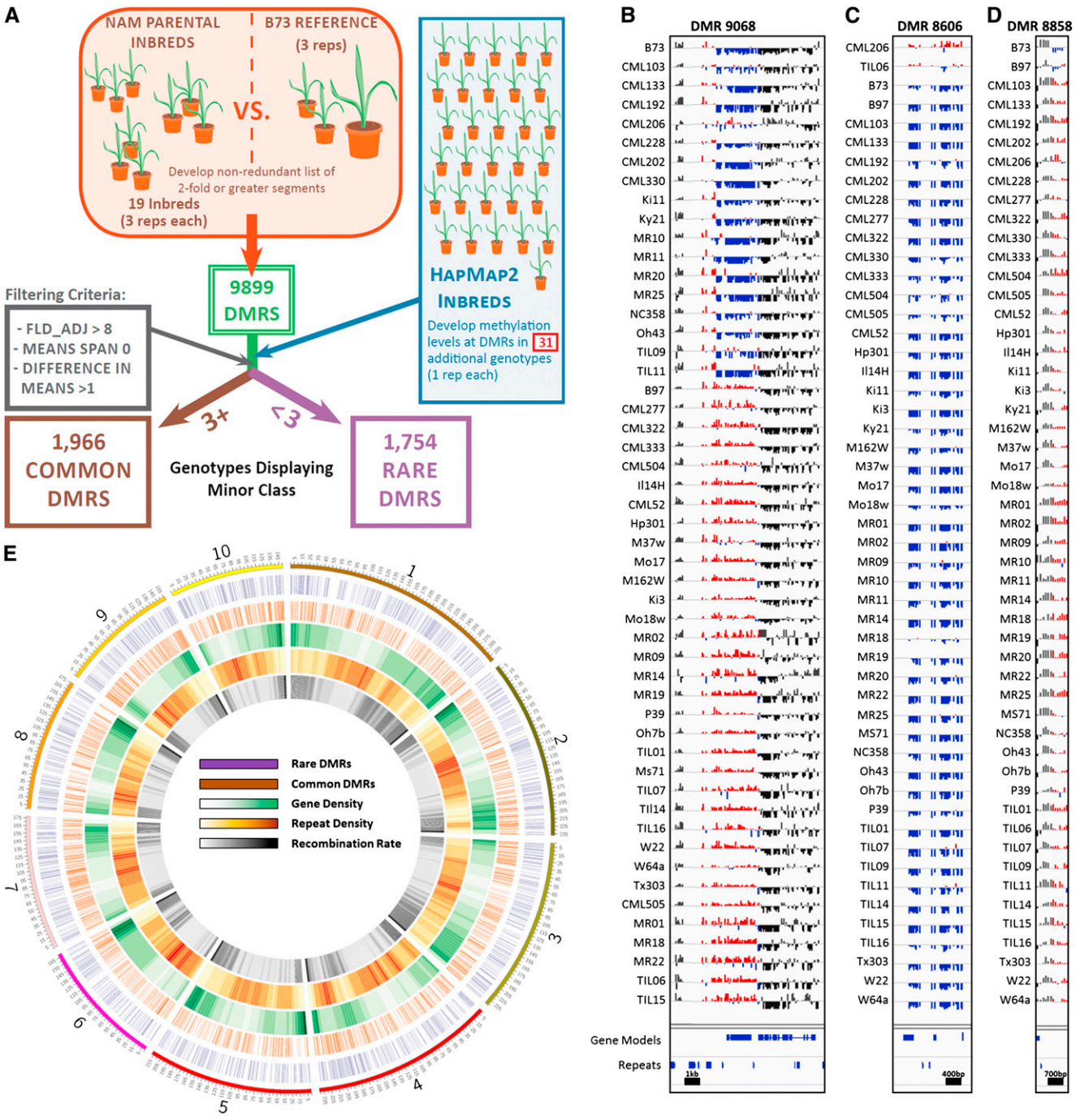

Figure 1. Identification of DMRs across Diverse Maize Lines.

(A) An outline of the methods to identify DMRs. DMR discovery was performed by contrasting replicated measurements of 19 NAM parental lines (McMullen et al., 2009) with B73. The methylation levels for these 9899 regions were also determined for a single replicate of 31 additional genotypes from the HapMap2 panel (Chia et al., 2012). A series of filtering criteria were applied to identify robust common (three or more genotypes with both high and low methylation) or rare (only one or two genotypes with alternate methylation state) DMRs.

(B) to **(D)** Examples of common **(B)**, rare methylated **(C)**, and rare unmethylated **(D)** DMRs are visualized in all 51 genotypes. In each case, the genotypes are grouped according to DNA methylation state. The red (high) and blue (low) indicate variable methylation with the DMR and probes with black coloring are outside of the DMR. The genes and repetitive elements annotated in each region are shown at the bottom.

(E) A genome-wide view of DMRs in relationship to other genomic features. Circos (Krzywinski et al., 2009) was used to show (outside to inside tracks) rare DMRs (purple ticks), common DMRs (brown ticks), gene density in 1-Mb windows based on annotation from Schnable et al. (2009) (green is high), repeat density in 1-Mb windows based on annotated from Schnable et al. (2009) (orange is high), and recombination rate as centimorgans/Mb from Liu et al. (2009) (black is high).

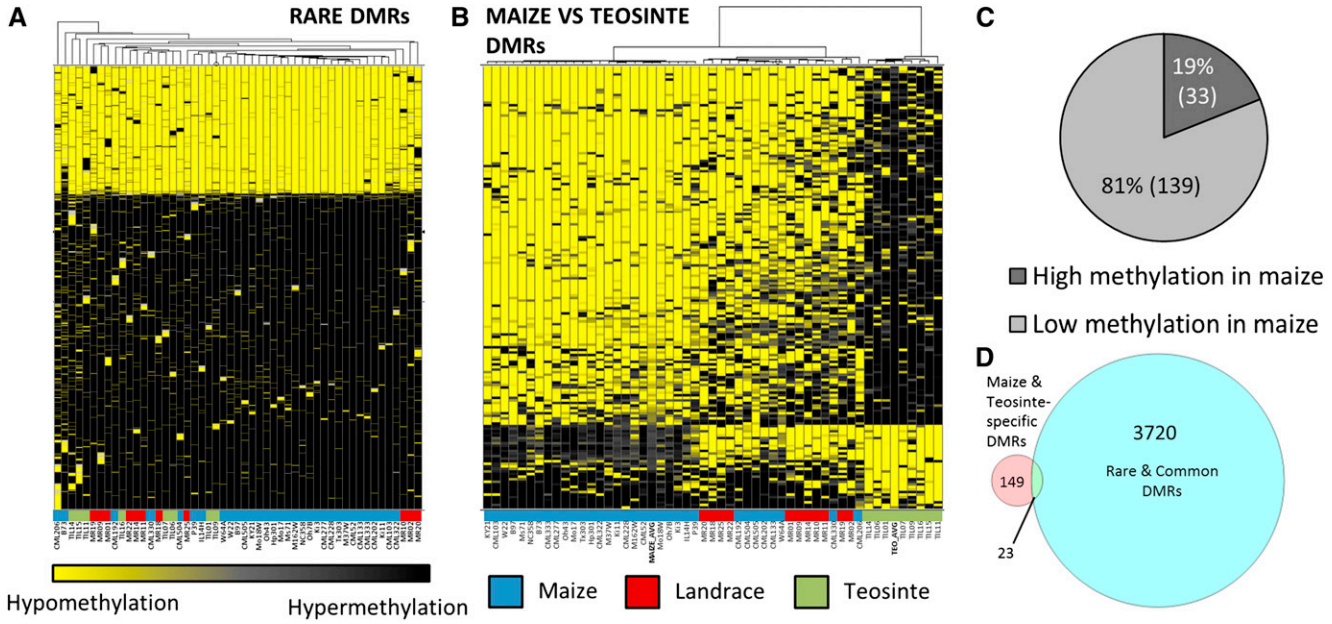

Figure 2. Hierarchical Clustering of DMRs in All Genotypes.

(A) The DMR state (yellow, hypomethylated; black, hypermethylated) was used to perform hierarchical clustering for the rare DMRs. There is not strong evidence for single genotypes exhibiting unique DNA methylation profiles relative to other genotypes for common or rare DMRs. The results illustrate the enrichment for the rare state to reflect hypomethylation in a few genotypes as opposed to hypermethylation in a few genotypes. The rare state can be observed in any of the genotypes.

(B) Hierarchical clustering of 172 maize and teosinte-specific DMRs across all maize, landrace, and teosinte samples studied. Clear separation of maize and teosinte lines is visible (coloring on bottom). Landrace samples often appear to have variable methylation state for maize and teosinte-specific DMRs (center cluster).

(C) Significant enrichment for hypomethylation in rare DMRs.

(D) The overlap between maize-teosinte DMRs and the DMRs discovered among the 19 NAM parents is shown.

teosinte (Swanson-Wagner et al., 2012). In each of these seven cases, higher expression is associated with lower DNA methylation levels. The comparison of DNA methylation levels in maize and teosinte reveals examples of differential methylation, and these may have been selected on during domestication. Further studies will be necessary to assess whether these DMRs are the result of selection during domestication.

DMRs Frequently Result from CG and CHG Methylation

MethylC-Seq (Lister et al., 2008) was performed on independently grown samples of B73 and Mo17 in order to validate the DNA methylation differences observed by array-based profiling and to characterize the sequence context of DNA methylation at DMRs (see Supplemental Data Set 1 online). There were 248 DMRs with differences between B73 and Mo17 methylation levels based on meDIP array estimates that were present in the common or rare set of DMRs and had at least 80% of their length covered by reads in both the B73 and Mo17 MethylC-Seq data. The majority (91%) of these B73-Mo17 DMRs had substantial (>50%) difference in the level of CG and/or CHG methylation within the DMR (Figure 3). Similar proportions of common (92%) and rare (89%) DMRs were confirmed by the bisulfite sequencing data. Most (84%) of the DMRs that were validated by bisulfite sequencing data had a substantial difference in DNA methylation in both the CG and CHG contexts. A small proportion of DMRs only exhibit differences in levels of either CG (9%) or CHG (7%) DNA methylation. None of the DMRs exhibit substantial variation in CHH DNA methylation (see Supplemental Figure 3 online) and they generally had very low (<10%) CHH methylation (see Supplemental Data Set 1 online).

DNA Methylation Variation Is Frequently Associated with Local Genetic Variation

Although DNA methylation is often considered an epigenetic mark, there is evidence that it can be highly influenced by genetic changes, such as transposon insertions (Law and Jacobsen, 2010; Hollister et al., 2011; Eichten et al., 2012). A local association scan was implemented to determine what proportion of DMRs might be associated with local genetic variation and represent information content that was already captured by SNP-based analyses. We did not attempt a genome-wide association scan due to the limited number of genotypes (51) that were sampled. All 51 genotypes with DNA methylation data have been resequenced and used to identify a set of 56 million SNPs (Chia et al., 2012). For each of the common DMRs, the SNPs

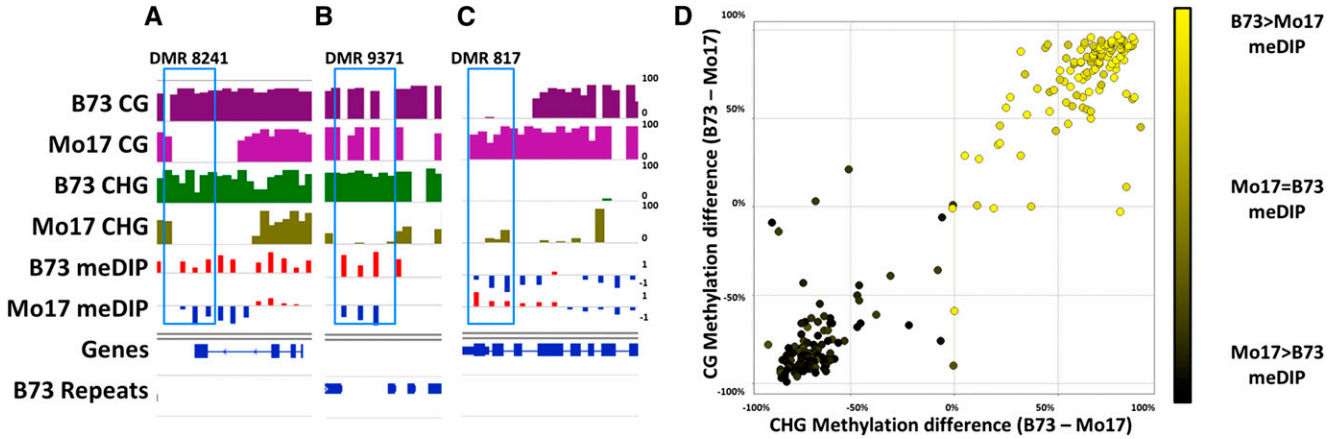

Figure 3. Validation of Differential Methylation Levels by MethylC-Seq Data on an Independent Sample of B73 and Mo17.

(A) to **(C)** Examples of DMRs that exhibit differences in CG and CHG **(A)**, CHG only **(B)**, or CG only **(C)**. Weighted DNA methylation levels (Schultz et al., 2012) were calculated for cytosines in CG, CHG, and CHH contexts and summarized for the DMR. The CG methylation levels are shown in different shades of purple, and CHG methylation levels are shown in different shades of green for B73 and Mo17. The meDIP array methylation estimates are shown for both B73 and Mo17 with high methylation indicated by red and low methylation indicated by blue. The annotation of genes and repetitive sequences for each region is shown at the bottom.

(D) For the 248 DMRs that exhibit significant meDIP array variation and have at least 80% coverage of MethylC-Seq reads in both B73 and Mo17, the relative levels of CG (*y* axis) and CHG (*x* axis) methylation are shown. The color coding of each DMR indicates the meDIP array difference in DNA methylation levels between B73 and Mo17 (yellow, higher in B73; black, higher in Mo17). The majority of DMRs show substantial differences in both CG and CHG methylation in the direction predicted by the meDIP array data. A small number of the DMRs only exhibit differences in CG or CHG methylation or do not show any difference in methylation levels in this independent sample of B73 and Mo17.

overlapping and within 1 kb of the DMR boundaries were used to perform an association analysis relative to the DNA methylation level in each genotype. Per-SNP significance was estimated using permutation to associate the same DNA methylation phenotype with random genomic regions, and the per-DMR significance was determined by comparing the frequency of significant SNPs in the DMRs compared with randomly selected genomic regions (Figures 4A and 4B; see Methods). This analysis was applied only to the common DMRs as the rare DMRs did not have statistical power for associating rare phenotypic variation in DNA methylation with SNPs. A substantial proportion (1003/1966) of the common DMRs was significantly associated with local SNPs (Figure 4C; see Supplemental Data Set 1 online). In some cases, the highly associated SNPs were located within the DMRs but SNPs were often outside the DMR itself (see Supplemental Figure 4 online). The remaining 963 common DMRs did not exhibit significant associations between the DNA methylation level and local SNPs. There are several potential reasons for lack of association including a lack of SNPs in LD with local causative genetic changes, potential *trans*-acting factors that influence DNA methylation level, or purely epigenetic influences on DNA methylation.

The significant associations of local SNPs with DNA methylation level for over half of the common DMRs suggests that the DNA methylation levels can be predicted based upon SNPs for these regions. The ability to accurately predict DNA methylation level based upon the associated SNPs was tested by determining DNA methylation levels in 12 additional genotypes (see Supplemental Table 1 online) that had SNP calls (Chia et al., 2012). These genotypes were hybridized to a separate microarray format that only contained probes to survey DMRs (methods). There were 535 DMRs assessed by this microarray platform that exhibit allelic variation for the most significantly associated SNP (see Supplemental Figure 5 online). In many cases, only one (205) or two (111) of the 12 inbreds contained a different SNP allele relative to the other inbreds (see Supplemental Figure 5B online). The remaining 219 DMRs had at least three inbreds that contain each of the genotypes at the most associated SNP. The genotype and methylation state were compared for each DMR (Figure 4D). For the majority (77%) of the DMRs, the average DNA methylation state is accurately predicted by the SNP genotype in the 12 genotypes (Figure 4D; see Supplemental Figure 5C online). These results indicate that DNA methylation patterns at DMRs associated to genetic variation can be predicted solely on genotype data at a high rate.

While there were a number of common DMRs that exhibit associations with SNPs, it is not likely that SNPs themselves are causative for differential DNA methylation. Instead, SNPs are more likely in LD with nearby transposon insertions or structural variation that impact local DNA methylation profile. A subset of retrotransposable elements have been identified in maize that display a spreading of DNA methylation beyond the borders of the element into the low-copy sequence adjacent to the insertions (Eichten et al., 2012). This phenomenon may contribute to DNA methylation variation if the specific insertion is polymorphic among genotypes. To investigate the relationship between transposable elements and our identified DMRs, each DMR was mapped relative to its nearest transposon (Figure 4E). Each of the common or rare DMRs, as well as nearly 10,000

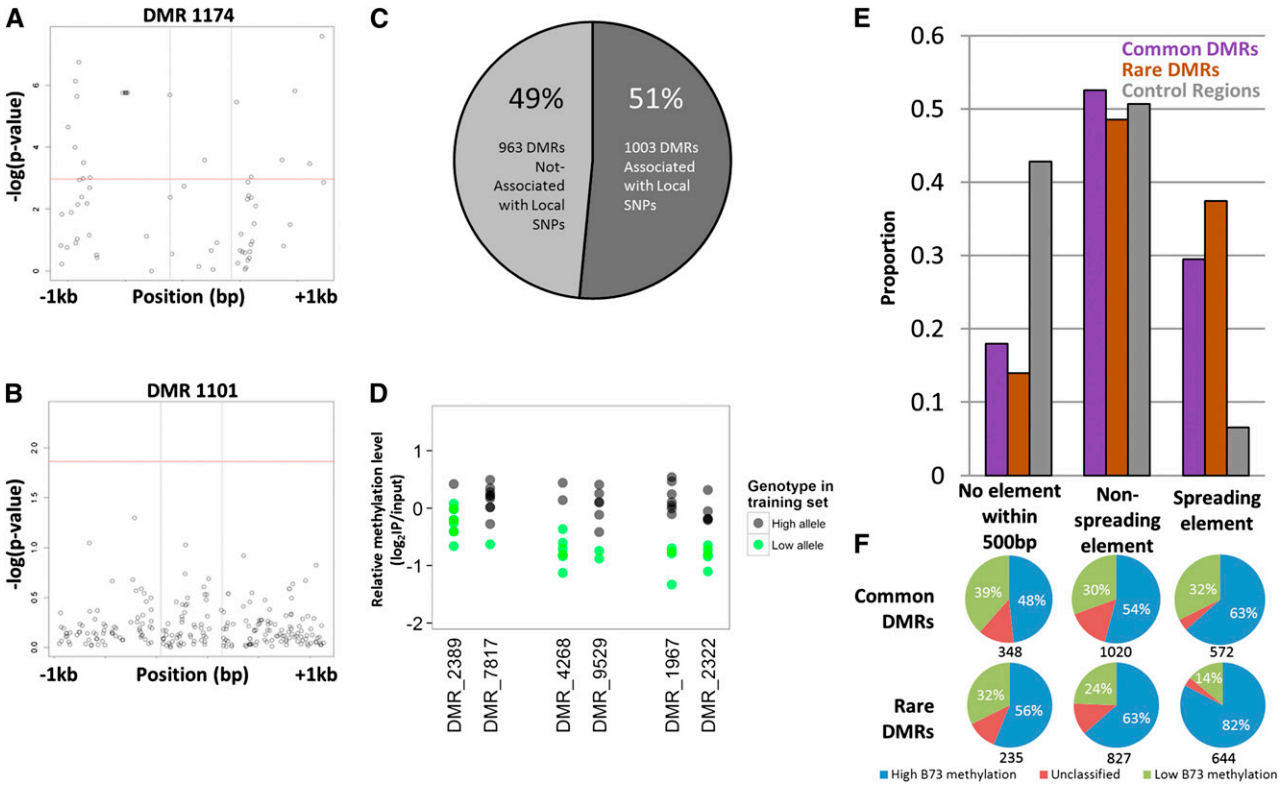

Figure 4. Many Common DMRs Are Associated with Local Genetic Variation.

(A) An example DMR displaying significant association to local SNPs. Vertical lines indicate DMR boundaries. Horizontal line (red) indicates 1% quantile P value cutoff based on permutation analysis of other SNPs with the methylation variation for this region. SNPs above this line display significant association to DMR methylation state.

(B) Example DMR displaying no significance for local SNPs with methylation state.

(C) The proportion of DMRs with and without significant association with local SNPs.

(D) Example relative methylation values [$\log_2$(IP/input)] for additional genotypes. Color indicates the predicted high methylation and low methylation allele based on their genotype calls.

(E) Enrichment for heterochromatin spreading transposable elements near both common and rare DMRs. Common and rare DMRs were mapped to nearby repetitive elements within 500 bp and classified as having no annotated repeat, nonspreading elements, or heterochromatin spreading elements (based on genome-wide B73 reference genome annotations from Schnable et al. [2009] and spreading assignments from Eichten et al. [2012]). The DMRs were compared with a set of 10,000 control regions selected to reproduce features of our experimental DMRs. The common and rare DMRs are enriched for having spreading transposable elements within 500 bp compared with the control regions. For each of the three groups, the proportion of common DMRs associated to local SNPs is presented (white text in purple bars).

(F) For each of the three repeat classes in Figure 3E, the proportion of high (>80%) and low (<20%) CG methylation in B73 is presented for both common and rare DMRs. The total number of DMRs in each class is presented below each chart. An increase in methylation level in B73 is observed for DMRs near spreading elements.

randomly selected control regions, were classified according to the presence and type of transposable element within 500 bp (Figure 4E). The common and rare DMRs are both more likely to be located near a spreading element relative to control regions (Figure 4E) and are less likely to be in a region in which there is no transposon within 500 bp.

The proximity to transposons was tested based on transposable element annotation of the B73 reference genome. However, it is likely that some of these insertions are polymorphic among the genotypes in this study. We hypothesized that DMRs found near spreading retrotransposable elements, which where annotated in the B73 reference background, should display a high level of methylation in the B73 sample. To test this, the B73 CG methylation level (from MethylC-Seq) across both common and rare DMRs was determined (Figure 4F). As expected, B73 has higher levels of DNA methylation for both common and rare DMRs that are near transposons compared with DMRs not located near transposons (Figure 4F). This increase is most pronounced when looking specifically at the spreading elements. These observations support the potential for obligatory epialleles to result from polymorphic transposon insertions.

A transposon insertion may be in LD with local SNPs and therefore may be captured by our local association tests. The proportion of the common DMRs associated with local SNPs

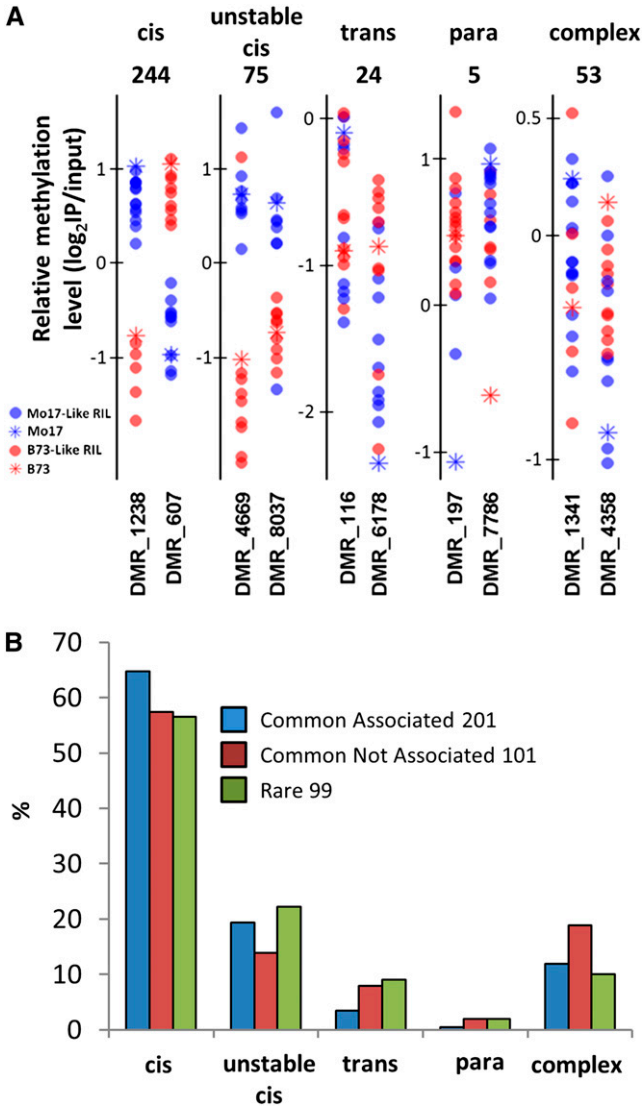

Figure 5. DMRs Appear Heritable across Diverse and Recombinant Inbred Lines.

(A) Examples of five classes of DMR stability across 17 B73-Mo17 RILs are shown. These are divided into locally inherited "*cis*" patterns, locally inherited "unstable *cis*" patterns with occasionally methylation state shift, remote "*trans*" inheritance of methylation state by nonlocal region of the genome, DMRs displaying paramutation-like states in which all lines regardless of local genetic content appear like one parent, and complex DMRs that display methylation state instability or multi-region control. The total number of DMRs for each category is displayed below.
(B) The proportion of each RIL inheritance class is presented based on DMR association class. Few changes in inheritance states were observed due to DMR class (common associated, common nonassociated, and rare).

was determined for the DMRs with no element, spreading elements, or nonspreading elements within 500 bp. Only 37% of the common DMRs that do not have transposon insertions within 500 bp are significantly associated with local SNPs. By contrast, 55% of the DMRs located near nonspreading and 52% of DMRs located near spreading elements have significant associations with local SNPs. Collectively, analysis of DMRs relative to nearby SNPs suggest that some DMRs are associated with genetic changes and that many of these may be the result of transposon insertions that influence the DNA methylation patterns of nearby low-copy sequences.

Heritability of DMRs in RILs

DNA methylation profiling was performed in a biparental RIL population derived from B73 and Mo17 (Lee et al., 2002) to study the heritability of DMRs and the potential for paramutation or *trans*-acting influences on DNA methylation. A custom 12-plex microarray that includes a subset of the probes from the full array was used to profile DNA methylation in 17 RILs (see Supplemental Table 1 online) as well as independently grown replicates of B73 and Mo17. There were 401 DMRs that were discovered in the whole-genome profiling and also displayed significant ($P < 0.05$) differences in their methylation state in the profiling of these independently grown replicates. The estimated DNA methylation levels of these DMRs were analyzed relative to the genotype of the chromosomal region containing the DMR in each RIL. DMRs that are heritable and influenced by *cis*-acting genetic or epigenetic change are expected to have DNA methylation levels that are associated with the genotype (Li et al., 2013) of the region surrounding the DMR (Figure 5A, cis; see Supplemental Figure 6 online). If there is some instability in the inheritance of locally controlled DNA methylation, then it is possible that the methylation pattern in RILs will appear mostly associated with local genotype but some switches in epigenotype relative to genotype might be observed (Figure 5A, unstable cis). *Trans*-acting regulation of DNA methylation will result in two discrete levels of DNA methylation that are not correlated with local genotype but instead are correlated with the genotype of a locus elsewhere in the genome (Figure 5A, trans). Paramutation-like processes would be expected to result in all RILs exhibiting DNA methylation states that are similar to the levels observed in one parent (Figure 5A, para). The majority (79.6%) of DMRs assayed in the RILs exhibit local (*cis*) control of DNA methylation with 244 showing a complete association with local genotype and 75 showing some instability of the inheritance of DNA methylation (Figure 5A). There are relatively few examples of *trans*-acting control (24) or paramutation (5) among these 401 DMRs (Figure 5A). The remaining 53 DMRs exhibit complex patterns that resemble quantitative variation in DNA methylation levels that may be the result of multilocus control of DNA methylation. Alternatively, the DMRs classified as "complex" could be the result of hypervariability for DNA methylation levels in this region or the result of technical variation in the estimates of DNA methylation.

The DMR patterns in RILs were compared with the results from the local association scan for these 401 DMRs (Figure 5B). The 401 DMRs surveyed in RILs included 201 common DMRs with significant associations with local SNPs, 101 common DMRs without significant local association, and 99 rare DMRs (which were not tested for association). The common DMRs that exhibit significant association with local SNPs would be expected to show *cis*-acting inheritance of DNA methylation levels

if they are indeed under local genetic control. They also might be expected to be relatively stable as they were significantly associated with local SNPs in populations of diverse lines. Common DMRs without significant local SNP associations might be expected to include more examples of *trans*- or complex genetic control. Indeed, common DMRs that are significantly associated with nearby SNPs are enriched for *cis*-acting patterns in the RILs and have few examples of paramutation-like patterns or *trans*-acting control (Figure 5B). The DMRs that lacked significant associations had more examples of paramutation-like or *trans*-acting patterns of inheritance. However, DMRs that lack local SNP associations still include a large number of loci with relatively stable *cis*-acting inheritance. This may be due to assayed SNPs not in LD with causative genetic changes or that these are purely epigenetic variants. The rare DMRs exhibit frequencies of the different types of inheritance patterns in RILs that are similar to those observed for common DMRs (Figure 5B). These observations indicate the value of a biparental RIL population for studying the inheritance patterns of DMRs as the increased allele frequency for DMRs/SNPs provides an opportunity to study inheritance patterns for all DMRs.

Functional Consequences of DMRs

DNA methylation can impact gene expression and may lead to a change in phenotype. To investigate the functional consequences of DNA methylation variation, genes located adjacent to DMRs (the closest gene in each direction from the DMR) were identified (see Supplemental Data Set 3 online). The location of DMRs relative to genes was analyzed to assess whether DMRs are enriched near genes (see Supplemental Figure 7A online). Both common and rare DMRs are more likely to be located near (within 5 kb) genes than expected by chance, but DMRs are not often found overlapping the coding regions (see Supplemental Figure 7 online). DMRs were equally likely to be located 5′ or 3′ of the gene, reflecting the potential for low-copy sequences near genes to have variable methylation. Sequences that are located far away from genes tend to have uniformly high methylation, while sequences within genes have low methylation or have intermediate levels of DNA methylation that are not detectable by meDIP. Comparing the 2375 genes that are located within 10 kb of a DMR to all maize genes suggests that DMRs occur at equal frequency in both subgenomes of maize (the products of the recent whole-genome duplication event) and are not enriched for inserted or syntenic genes (see Supplemental Figures 7B and 7C online; based on annotations from Schnable et al., 2012). In addition, the DMRs occur at similar frequency in the high-recombination, gene-rich arms of chromosomes and in the central, gene-poor low recombination portions of chromosomes (see Supplemental Figure 7D online).

The DMRs may influence phenotypic variation by affecting expression of nearby genes. In order to determine possible consequences of variable methylation state, RNA-seq was performed on the same tissue samples that were used for DNA methylation profiling (see Supplemental Table 1 online). The transcript abundance for the gene nearest each DMR was compared with the DNA methylation level for all genotypes. The analysis of the 1966 common DMRs identified 277 genes with expression significantly correlated to DNA methylation levels (both rank-sum and Kendall's tau $q < 0.05$) (example shown in Figure 6A). For the 1754 rare DMRs, 111 significant correlations of transcript abundance and nearby DMR (significance calculated by z-score and Kendall's tau; example shown Figure 6B). The use of different statistical tests allowed for the identification

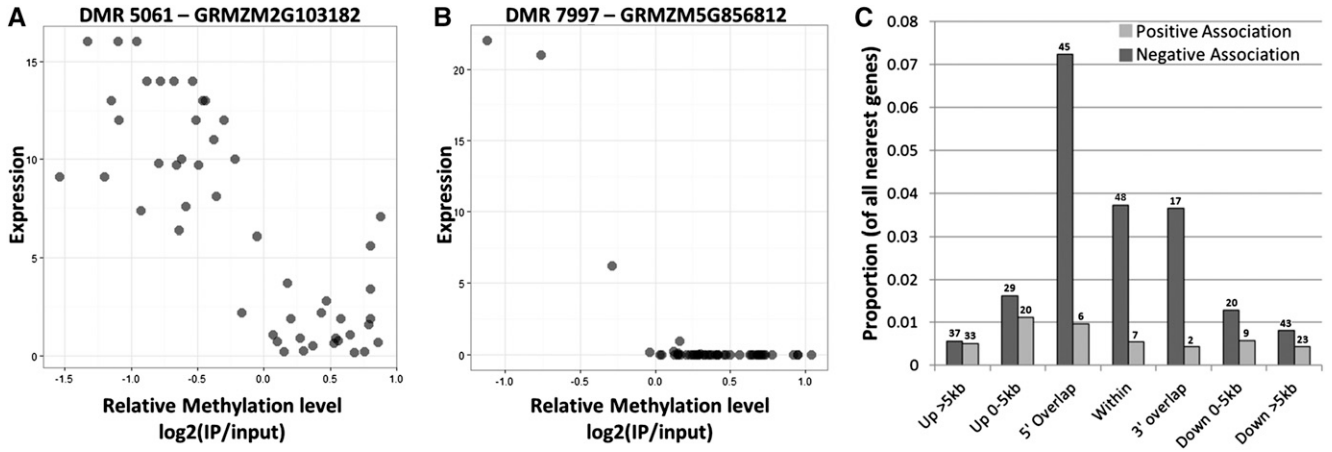

Figure 6. DMRs Associated with Gene Expression State.

(A) and **(B)** Examples of common **(A)** and rare **(B)** DMRs showing correlation to nearby gene (within 5 kb) expression state. The *y* axis displays log (reads per kilobase per million reads) values for the individual gene across 50 genotypes compared with array relative methylation value across 50 genotypes. **(C)** DMR-gene associations were grouped by the location of the DMR relative to the associated gene. DMRs were classified as being upstream (>5 kb or between 0 and 5 kb of a gene transcription start site), 5′ overlapping (DMR overlaps gene transcription start site), within (DMR falls completely within the borders of a gene), 3′ overlapping (DMR overlaps end of annotated gene), or downstream (0 to 5 kb or >5 kb from gene end). The position of DMRs in relationship to their associated gene is displayed for both positive (gray) and negative (black) correlations. An enrichment for negative methylation expression correlations for DMRs overlapping the 5′ end of genes is present.

of either qualitative or quantitative associations of expression and DNA methylation in both the common and rare DMRs. As expected, there was enrichment for significant negative (as opposed to positive) associations of DNA methylation with gene expression for both common (70%) and rare (73%) DMRs (Figure 6C). DNA methylation appears to be correlated with on/off states in gene expression for 26% of the DMRs that are negatively associated with DNA methylation levels (exemplified in Figure 6B). In other cases, the DNA methylation state is associated with quantitative differences in gene expression (Figure 6A). The DMRs for which methylation level is negatively correlated with gene expression are more likely to be located near or overlapping gene boundaries (Figure 6C) as opposed to DMRs that are located >5 kb from genes. In particular, the DMRs that overlap the transcription start site are most enriched for significant negative correlations with gene expression levels. The genes that have a negative association between gene expression levels and DNA methylation state are slightly depleted for syntenic genes and show some enrichment for inserted sequences or genes without homologs in other grass species (see Supplemental Figure 7C online). However, over 45% of the genes that have expression variation associated with nearby DMRs are conserved in other grasses and located in syntenic positions, which would suggest that at least a portion of the genes associated with variable DNA methylation and expression are not simply misannotated transposons.

DISCUSSION

Characterization of DMRs in Diverse Maize Lines

The profiling of DNA methylation in 51 diverse maize inbred lines identified 1966 common DMRs and 1754 rare DMRs. There are likely additional DMRs that did not meet our filtering criteria, were only present in one of the 30 lines not used for the DMR discovery, or that are present in regions of the genome not targeted by the microarray platform. However, the number of DMRs found among these maize genotypes are similar to the numbers identified in recent studies of *Arabidopsis* (Schmitz et al., 2013b) and rice (Chodavarapu et al., 2012). In *Arabidopsis*, it was noted that hypermethylation is the rare state for most DMRs that are within genes while for DMRs in transposons hypomethylation is the rare state (Schmitz et al., 2013b). In our sample of maize genotypes, hypomethylation is the rare state for most DMRs (see Supplemental Figure 2 online). Our profiling of maize focused on low-copy regions of the genome and therefore did not test most transposons. The majority (79%) of the maize DMRs identified in this study is located in intergenic regions and is often located near transposons, while the other 21% of the DMRs at least partially overlap an annotated gene. The DNA methylation levels for DMRs were highly consistent in different samples of the same genotypes. All three biological replicates of a genotype that were profiled for the 20 NAM parents exhibit very similar methylation profiles. In addition, the bisulfite sequencing of an independent replicate of B73 and Mo17 (seeds from a different plant) confirmed the majority of DMRs. It is worth noting that while many DMRs were identified in this study, the majority of the maize genome exhibits DNA methylation levels that are quite similar and reproducible even in diverse genotypes; in other words, the majority of methylated regions appear to be conserved within maize.

In *Arabidopsis* (Schmitz et al., 2013b) and rice (Chodavarapu et al., 2012), DMRs were classified based on the sequence context of DNA methylation (CG, CHG, or CHH). The DMRs that only exhibit differences in CG DNA methylation in *Arabidopsis* often occurred in genes, showed high levels of variation, and rarely were associated with low levels of gene expression (Vaughn et al., 2007, Schmitz et al., 2013b). By contrast, DMRs associated with methylation in contexts beyond just CG (termed C-DMRs) were more stable and were often associated with lower levels of expression. The protocol that we used for profiling should have the potential to discover either CG or C-DMRs. However, the bisulfite sequence analysis of B73 and Mo17 (Figure 3) reveals that the majority of maize DMRs exhibit both CG and CHG differences. Previous studies (Gent et al., 2013) found relatively few sites with high levels of CHH methylation in maize. Recent studies in *Arabidopsis* suggest two independent pathways for de novo methylation (Zemach et al., 2013), but maize appears to lack orthologs of the CMT2 gene that is implicated in the sRNA-independent CHH methylation pathway. There were few examples of CG- or CHG-specific differences in DNA methylation. It is worth noting that some types of DMRs that might exist in maize are not likely to be detected with the assay we employed. For example, the partial differences in DNA methylation (i.e., 20% versus 60% DNA methylation) often found at CG-DMRs are likely not strong enough to be detected by our approach. In addition, CHH-specific DMRs that have DNA methylation in all three contexts (CG, CHG, and CHH) in one genotype but lose CHH methylation due to loss of active targeting of de novo methylation in the other genotype likely would not be detected as both alleles would still have substantial DNA methylation. It is likely that deep bisulfite sequencing coverage may allow differences in CHH methylation among maize genotypes to be properly identified.

The most likely mechanism for differential DNA methylation levels to alter phenotype is via effects on gene expression. The general view is that DNA methylation within gene promoters or regulatory elements can result in repression of gene expression. However, there is evidence that gene body methylation is associated with moderately expressed genes (Zilberman et al., 2007; Law and Jacobsen, 2010). We found that DMRs were associated with variable gene expression in matched seedling tissue for 11.5% of the common DMRs and 6.3% of the rare DMRs. The majority (70%) of associated genes exhibit expression levels that are negatively associated with the DNA methylation levels, suggesting that a portion of gene expression variation among maize genotypes is attributable to differences in DNA methylation. It is possible that the expression of additional genes is affected by variable DNA methylation and might be detected by analysis of gene expression in other tissues.

Potential Causes of DNA Methylation

DNA methylation is often considered to be an epigenetic mark, and there is certainly evidence that DNA methylation can be a major contributor to epigenetic regulation at some loci (Cubas

et al., 1999; Manning et al., 2006; Jacobsen and Meyerowitz, 1997). However, there is also strong evidence that certain genetic features can lead to DNA methylation. We are interested in understanding the causes of variable DNA methylation in maize in order to understand how to capture this information in breeding programs. Richards (2006) classified three major types of epialleles based on the interaction of genetic and epigenetic influences. We propose further nuances to these classifications and define six potential subtypes of DMRs: epigenetic-stable, epigenetic-unstable, paramutation, genetic-local, genetic-remote, and genetic-polygenic (Figure 7).

The first two types of DMRs are both classified as purely epigenetic and are distinguished based on whether they are highly stable (epigenetic-stable) or include some instability (epigenetic-unstable). Paramutation would define another type of epigenetic DMR in which one allele could influence the methylation state of the other when present together in a heterozygote (Erhard and Hollick, 2011). This type of DNA methylation pattern is likely a subtype of one of the other classes but is expected to show unique behavior in populations and is therefore separately classified. The remaining three groups are classified as genetic as they are all affected by the genetic content of the individual. Genetic-local DMRs are influenced by *cis*-acting genetic changes, such as structural rearrangements or transposon insertions, and could result in stochastic DNA methylation variation in the presence of the genetic change or could show complete agreement between DNA methylation and genetic state. Genetic-remote DMRs exhibit changes in DNA methylation that are influenced by a *trans*-acting region of the genome. It is most likely that this *trans*-acting signal is based on small RNAs that are generated by related sequences elsewhere in the genome, such as in the well characterized example of PAI silencing (Bender and Fink, 1995). The third type of genetically influenced DMR has polygenic influences on the DNA methylation state. This could be due to influences from multiple *trans*-acting sites or a combination of a *cis*-acting and *trans*-acting feature. It is useful to separate these potential types of DMRs in order to address which types are most prevalent and to consider how to most efficiently capture this methylation variation when predicting phenotype in large populations.

We evaluated the heritability of DNA methylation variation both in an association panel of 51 genotypes as well as in 17 RILs from a biparental population. A major goal was to identify the proportion of DMRs that are associated to the local genetic content (i.e., SNPs). In order to detect a significant association in the panel of 51 genotypes, it would be necessary to have relatively stable inheritance of DNA methylation levels and to have a SNP in LD with either the causative genetic change or the altered DNA methylation level. The finding that approximately half of the common DMRs had significant associations with local SNPs (Figure 4) suggests either stable epigenetic DMRs or genetic-local DMRs are relatively common. The analysis of DNA methylation levels in the RILs provides further support for these two classes. The genetic-remote and paramutation types of DMR patterns are expected to have unique patterns in the RILs compared with genetic-local or epigenetic DMRs. However, there were few examples of DMRs that exhibit patterns that would reflect paramutation or simple *trans*-acting control (Figure 5). There is evidence for paramutation-like patterns or *trans*-chromosomal DNA methylation in rice (Chodavarapu et al., 2012) and *Arabidopsis* (Greaves et al., 2012), but these seem to be rare. A recent study in maize found that as many of 10% of DMRs had switches in DNA methylation state and suggested that many of these may be the result of paramutation (Regulski et al., 2013), although some of these may also be the result of

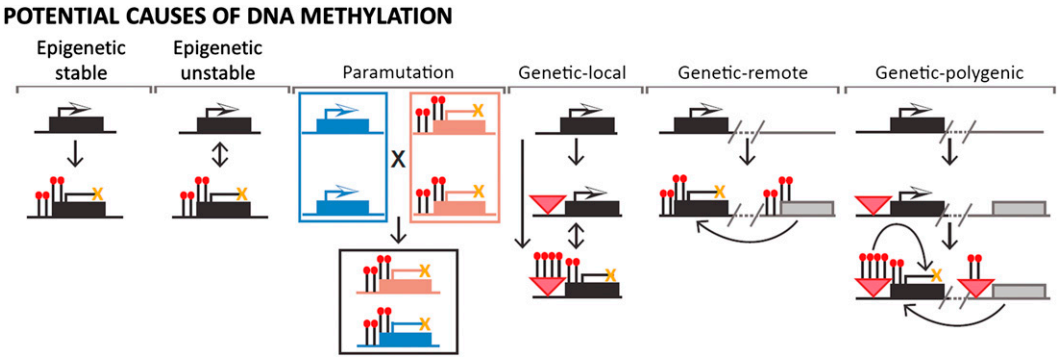

Figure 7. Diagram of Potential Causes of DNA Methylation.

DNA methylation variation may act as either an epigenetic mark independent from genetic variation (left) or as a heterochromatin mark linked to genetic variation (right). Epigenetic-stable variation acts independently of genetic context and is stable through generations. Epigenetic-unstable acts in a similar fashion; however, the instability of methylation states allows for reversion of methylation state at some frequency. Paramutation acts through a unique mechanism due to the activity of differently methylated alleles in the same nucleus of a hybrid. DNA methylation linked to genetic variation can be separated into three distinct categories. Genetic-local is where DNA methylation may be either controlled or facilitated by a local genetic variation such as a transposon insertion. Genetic-remote is where a genetic variant in *trans* acts to change the methylation state at the observed loci. The final class, genetic-polygenic, involves the actions of both local and *trans*-acting factors to initiate a DNA methylation change at the observed loci. This mechanism is expected to be complex in nature, requiring the knowledge of all controlling site genotypes in order to predict DNA methylation state. [See online article for color version of this figure.]

trans-control of DNA methylation. Our somewhat more restrictive classification of paramutation finds fewer examples of paramutation. Of the five DMRs identified in this study showing a paramutation inheritance pattern, all display differences in the CG and CHG contexts but not in CHH methylation levels. Many of the examples that we classified as "*trans*" or "complex" may have been classified as "paramutation-like" in Regulski et al. (2013). In both studies, the term "paramutation" is used to describe a certain type of pattern, but we have not actually tested for specific allelic interactions or compared other properties that would be necessary for actual evidence of true paramutation. A common theme to emerge from analysis of DNA methylation inheritance in maize (Regulski et al., 2013; this study) and soybean (Schmitz et al., 2013a) RILs is that the majority of DNA methylation variation in two parents is faithfully inherited in offspring. A small portion of the DNA methylation variation is not under stable, local control and may exhibit other patterns indicative of remote control or paramutation-like processes.

The finding that some RILs have altered DNA methylation levels relative to their genotype (unstable *cis*) suggests some stochastic behavior for DNA methylation levels relative to genotype and may highlight the potential for either epigenetic-unstable behavior or the genetic local influences that result in a stochastic change. It is important to note that highly unstable behavior for a DMR may result in patterns that we have classified as "complex" or even "*trans*" and understanding the instability of inheritance for DNA methylation levels is quite important. Schmitz et al. (2013b) performed a genome-wide association study to map the factors that influence DNA methylation for DMRs in *Arabidopsis*. They found evidence for local effects on some DMRs and found that a number of these had nearby insertion/deletions. However, they also found a number of cases with complex control by more than one genomic region. Examples of DMRs that are influenced by multiple genomic regions may be reflected by the DMRs that exhibit "complex" patterns in the RILs.

In contrast with both *Arabidopsis* and rice, the maize genome has a much higher number of transposable elements and many maize genes are located near transposons (Schnable et al., 2009). It has been hypothesized that a significant function of epigenetic regulation is the control of transposons (Hollister et al., 2011). Studies in maize have found that some retrotransposons exhibit spreading of heterochromatin to neighboring sequences, while other retrotransposon are marked by DNA methylation that does not spread beyond their borders (Eichten et al., 2012). We found that DMRs are more likely to be located near retrotransposons that exhibit spreading of heterochromatin than is expected by chance (Figure 4D). This suggests that a number of the DMRs are influenced by local variation for the presence of these elements. As these retrotransposons have been annotated only in the B73 reference genome, it is possible that many DMRs that do not appear near an annotated retrotransposon may have retrotransposon insertions across in other genotypes. It is unclear whether the insertion of one of these spreading retrotransposons results in an obligatory change in DNA methylation levels or simply facilitates the potential to acquire higher levels of DNA methylation.

For crop improvement, a major question is how to fully capture the heritable information in populations in order to improve predictions of phenotypes. It is relatively inexpensive to collect dense SNP genotypes, but collecting data on DNA methylation profiles is currently more expensive and complex. Therefore, there is interest in understanding what proportion of DNA methylation variation might be captured by genotype information. Our results suggest that a substantial portion of the DMRs might be significantly associated with local SNPs. Our results also suggest that we could improve our ability to predict DNA methylation state if genotyping information provided the ability to survey the presence-absence of specific transposon insertions throughout the genome. The potential for certain retrotransposons to spread heterochromatin to flanking sequences suggests that being able to identify the genomic positions of these elements might improve our predictions of DNA methylation variation at low-copy sequences among individuals.

METHODS

Biological Materials

Three replications of the 20 NAM parental genotypes and single replicates of the additional 31 maize (*Zea mays*) and teosinte HapMap2 genotypes (see Supplemental Table 1 online) were grown in a randomized block design with a single replicate in each block. Similar experimental designs were used for the growth of 17 IBM RILs (Lee et al., 2002) and for 12 additional HapMap2 genotypes that were grown to assess heritability of DMRs (Supplemental Table 1 online). For each replication, four seedlings of each line were grown in a single pot that was randomly assigned to a location within the block. Seedlings were grown under controlled conditions with 15 h light:9 h dark at the University of Minnesota Agricultural Research station, Saint Paul, MN. Seedlings were watered daily. After 18 d of growth, the 3rd leaf (L3) of each plant was harvested and pooled with other plants from the same pot/replication or harvested independently and immediately frozen in liquid nitrogen. DNAs were isolated using the CTAB method (Doyle 1987) from frozen leaf tissue as described by Eichten et al. (2011). RNAs were isolated using Trizol (Invitrogen) per the manufacturer's protocol.

meDIP-Chip Epigenomic Profiling

Methods were adapted from Eichten et al. (2011). Briefly, methylated DNA was immunoprecipitated with an anti-5-methylcytosine monoclonal antibody from 400 ng sonicated DNA using the Methylated DNA IP Kit (Zymo Research). To serve as a negative control, water was substituted for the monoclonal antibody on B73 sonicated DNA. For each replication and genotype, whole-genome amplification was conducted on 50 to 100 ng immunoprecipitated DNA and 50 to 100 ng of sonicated DNA (input control) using the Whole Genome Amplification kit (Sigma-Aldrich). For each amplified immunoprecipitation input sample, 3 μg amplified DNA was labeled using the Dual-Color Labeling Kit (Roche NimbleGen) according to the array manufacturer's protocol (Roche NimbleGen Methylation User Guide v7.0). Each immunoprecipitation sample was labeled with Cy5, and each input/control sonicated DNA was labeled with Cy3. Samples were hybridized to the custom 2.1 M, 1.4 M, or 270k probe array, depending on sample set, for 16 to 72 h at 42°. Slides were washed and scanned according to NimbleGen's protocol for the GenePix4000B (2.1 M platform) and NimbleGen MS200 (1.4 M and 270k platform) scanner. Images were aligned and quantified using NimbleScan software (Roche NimbleGen), which produced raw pair reports containing fluorescence intensity readings for each probe on the array.

For details on normalization and DMR discovery, see Supplemental Methods1 online. The methods for the comparison of DMRs with other genomic features, such as genes and repetitive elements, are available in the supporting online text.

Bisulfite Sequencing and qPCR DMR Validation

Whole-genome bisulfite sequencing was performed adapting methods from Schmitz et al. (2011). Briefly, 14-d-old whole-seedling genomic DNA was isolated from B73 and Mo17 inbred lines planted independently from samples used in meDIP analysis. Samples were fragmented, and TruSeq-methylated adapters were ligated to DNA fragments. Five hundred nanograms of adapter-ligated DNA underwent bisulfite conversion using the MethylCode bisulfite conversion kit (Life Technologies) per the manufacturer's protocol. Converted DNA was split into four reactions and amplified using Pfu Turbo Cx DNA polymerase (Agilent) for four cycles and subsequently pooled. Libraries were sequenced on the HiSequation 2000 (Illumina) for 100 cycles, paired end. Sequencing reads were processed to identify and filter poor 3′ quality and incomplete conversion. Sequenced were aligned using the Bismark aligner (v0.7.2; Krueger and Andrews, 2011) against the B73 RefGen v2 genome under the following parameters (-n 2, -l 50). Methylated cytosines were extracted from aligned reads using the Bismark methylation extractor under standard parameters. The proportion of CG, CHG, and CHH methylation was determined as weighted methylation levels (Schultz et al., 2012) in 100-bp windows across the genome. The methylation levels across DMRs were created by averaging methylation levels from intersecting 100-bp windows using BEDTools (Quinlan and Hall, 2010). qPCR validation of DMRs was adapted from Eichten et al. (2011). Details on the 10 DMRs selected and methods are available in the supporting online text (see Supplemental Table 2 and Supplemental Figure 1 online).

DMR/SNP Association Analysis

In order to identify DMRs that display significant association to local genetic variation, a custom-designed pipeline was developed. SNPs from the maize HapMapv2 project (Chia et al., 2012) were extracted for all DMRs and flanking 1 kb of genomic sequence. The 36 DMRs with no SNPs among the 51 genotypes were not tested, and SNPs present in less than three genotypes were not tested. For each DMR and regional SNP pairing, genotypes with ambiguous SNP calls were omitted and a two-tailed *t* test was performed for each SNP call and methylation values [$\log_2$(IP/input)] for the DMR region. Given the number of tests performed, it is important to define a threshold of significance and control for false associations. A set of 100 regions of 1000 SNPs were randomly selected throughout the maize genome and were used to test for random associations between these control SNPs and the methylation values for each DMR. For each test of control SNPs and genotype methylation states, a significance value was determined. The threshold for significance was set at the 1% of all control *t* test P values, and a DMR was identified as putatively associated with local genetic state if at least three SNPs in the DMR region displayed significance based on the 1% cutoff. In order to further define a false discovery rate for the classification of DMRs as a whole, the proportion of SNPs passing the 1% threshold for all 100 control regions was calculated. The rank for the experimental DMR relative to all control regions was evaluated, and only experimental DMRs in the top 5% of all regions were classified as associated with local genetic state.

DMR Profiling in RILs and Additional HapMap2 Genotypes

DNA isolated from 17 IBM RILs and 12 additional HapMap2 genotypes (see Supplemental Table 1 online) were hybridized to a custom 12-plex microarray format (GPL17181) that contains 270,000 of the same probes present in the array used above that detect DMRs and can be used to differentiate B73 and Mo17 genotypes. meDIP, hybridization, and normalization protocols were the same as described above. Prior high-density genotypes for each RIL (Li et al., 2013) were used to impute the genotype for each DMR in each genotype, and RIL genotypes were validated with a subset of array probes selected to show differential hybridization between B73 and Mo17 (Springer et al., 2009). The details on the classification of inheritance patterns for DMRs in the RILs are available in Supplemental Method s 1 online. For the analysis of methylation state of DMRs in the 12 additional HapMap2 genotypes, only DMRs that contain allelic variation (Chia et al., 2012) for the SNPs associated with DMRs were assessed. Inbreds were separated into two groups based on the genotype of the SNP. Average methylation levels between those two groups were contrasted. If the same allele showed association with high-methylation state as in the panel used to identify the association, the association was called "confirmed" in the new panel; otherwise, "not confirmed" was assigned.

RNA-Seq and Expression Analysis

RNA isolated from 50 genotype seedling L3 leaf samples used for meDIP profiling was prepared for sequencing at the University of Minnesota Genomics Center in accordance with the TruSeq library creation protocol (Illumina). Samples were sequenced on the HiSequation 2000 developing 8 to 24 million reads per replicate. Raw reads were filtered to eliminate poor-quality reads using CASAVA (Illumina). Transcript abundance was calculated by mapping reads to the maize reference genome (AGPv2) using TopHat (Trapnell et al., 2009) under standard parameters. Reads per kilobase per million reads' values were developed using "BAM to Counts" across the exon space of the maize genome reference working gene set (ZmB73_5a) within the iPlant Discovery Environment (www.iplantcollaborative.org). Common and rare DMRs were profiled separately for significant association to their nearest gene expression state. The significant DMR gene pairs from the common data set were chosen based on the union between rank sum ($q < 0.05$) and Kendall's tau ($q < 0.05$) significance tests corrected for multiple testing using Storey's false discover rate approach (Storey, 2002). The pairs from the rare data set were selected based on the union between z-score ($q < 0.05$) and Kendall's tau ($q < 0.05$) after false discover rate control.

Accession Numbers

All processed data files formatted for the Integrative Genomics Viewer are available for download from http://genomics.tacc.utexas.edu/data/dmr_genetic_influences/. The meDIP microarray pair files (platforms GPL13499, GPL5621, and GPL17181) are deposited with the National Center for Biotechnology Information (NCBI) Gene Expression Omnibus under accession number GSE46949. RNA-seq reads are deposited with the NCBI Short Read Archive under accession number SRP018088. Whole-genome bisulfite reads are deposited with the NCBI Short Read Archive under accession number SRP022569.

Supplemental Data

The following materials are available in the online version of this article.

Supplemental Figure 1. Validation of DMRs by Methylation-Sensitive qPCR.

Supplemental Figure 2. Hierarchical Clustering of DMRs in All Genotypes.

Supplemental Figure 3. MethylC-Seq Levels for CHH Methylation across DMRs.

Supplemental Figure 4. Assessment of associated SNP position relative to DMRs.

Supplemental Figure 5. Prediction of Methylation State Based on Genotype.

Supplemental Figure 6. IGV View of RIL Lines Displayed in Figure 5C.

Supplemental Figure 7. Analysis of Genes Located Near DMRs.

Supplemental Table 1. Summary of Lines and Data Collected for Analysis.

Supplemental Table 2. Primers for qPCR Validation.

Supplemental Data Set 1. List of High-Confidence DMRs and Their Association Test Results.

Supplemental Data Set 2. List of Maize and Teosinte-Specific DMRs.

Supplemental Data Set 3. DMRs Associated to Nearest Maize Gene.

Supplemental Methods 1. Normalization, Validation, and Analysis of DMRs.

ACKNOWLEDGMENTS

Sequencing support and library preparation were performed at the University of Minnesota Genomics Center. Computational support was provided by the Minnesota Supercomputing Institute and the Texas Advanced Computing Center. We thank Kit Leffler for assistance with figure development. This work was created using resources or cyberinfrastructure provided by iPlant Collaborative. The iPlant Collaborative is funded by a grant from the National Science Foundation (DBI-0735191; www.iplantcollaborative.org). We thank the editors and anonymous reviewers for the helpful suggestions in improving the article.

AUTHOR CONTRIBUTIONS

S.R.E., M.W.V., and N.M.S. designed research. S.R.E., Q.L., R.S.-W., P.J.H., A.J.W., E.S., and P.T.W. performed research. S.R.E., R.B., J.S., Q.L., P.T., C.L.M., M.W.V., and N.M.S. analyzed the data. S.R.E. and N.M.S. wrote the article.

Received June 12, 2013; revised July 12, 2013; accepted July 15, 2013; published August 6, 2013.

REFERENCES

Baucom, R.S., Estill, J.C., Chaparro, C., Upshaw, N., Jogi, A., Deragon, J.M., Westerman, R.P., Sanmiguel, P.J., and Bennetzen, J.L. (2009). Exceptional diversity, non-random distribution, and rapid evolution of retroelements in the B73 maize genome. PLoS Genet. **5:** e1000732.

Becker, C., Hagmann, J., Müller, 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.

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

Buckler, E.S., Gaut, B.S., and McMullen, M.D. (2006). Molecular and functional diversity of maize. Curr. Opin. Plant Biol. **9:** 172–176.

Chan, S.W.-L., Henderson, I.R., and Jacobsen, S.E. (2005). Gardening the genome: DNA methylation in *Arabidopsis thaliana*. Nat. Rev. Genet. **6:** 351–360.

Chandler, V.L., Eggleston, W.B., and Dorweiler, J.E. (2000). Paramutation in maize. Plant Mol. Biol. **43:** 121–145.

Chia, J.-M., et al. (2012). Maize HapMap2 identifies extant variation from a genome in flux. Nat. Genet. **44:** 803–807.

Chodavarapu, R.K., Feng, S., Ding, B., Simon, S.A., Lopez, D., Jia, Y., Wang, G.-L., Meyers, B.C., Jacobsen, S.E., and Pellegrini, M. (2012). Transcriptome and methylome interactions in rice hybrids. Proc. Natl. Acad. Sci. USA **109:** 12040–12045.

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

Doyle, J.J. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. Phytochemical Bulletin **19:** 11–15.

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

Eichten, S.R., et al. (2011). Heritable epigenetic variation among maize inbreds. PLoS Genet. **7:** e1002372.

Erhard, K.F., Jr., and Hollick, J.B. (2011). Paramutation: A process for acquiring trans-generational regulatory states. Curr. Opin. Plant Biol. **14:** 210–216.

Feinberg, A.P. (2007). Phenotypic plasticity and the epigenetics of human disease. Nature **447:** 433–440.

Gent, J.I., Ellis, N.A., Guo, L., Harkess, A.E., Yao, Y., Zhang, X., and Dawe, R.K. (2013). CHH islands: de novo DNA methylation in near-gene chromatin regulation in maize. Genome Res. **23:** 628–637.

Greaves, I.K., Groszmann, M., Ying, H., Taylor, J.M., Peacock, W.J., and Dennis, E.S. (2012). Trans chromosomal methylation in *Arabidopsis* hybrids. Proc. Natl. Acad. Sci. USA **109:** 3570–3575.

Gutzat, R., and Mittelsten Scheid, O. (2012). Epigenetic responses to stress: Triple defense? Curr. Opin. Plant Biol. **15:** 568–573.

Hollister, J.D., Smith, L.M., Guo, Y.-L., Ott, F., Weigel, D., and Gaut, B.S. (2011). Transposable elements and small RNAs contribute to gene expression divergence between *Arabidopsis thaliana* and *Arabidopsis lyrata*. Proc. Natl. Acad. Sci. USA **108:** 2322–2327.

Hufford, M.B., et al. (2012). Comparative population genomics of maize domestication and improvement. Nat. Genet. **44:** 808–811.

Jablonka, E., and Raz, G. (2009). Transgenerational epigenetic inheritance: Prevalence, mechanisms, and implications for the study of heredity and evolution. Q. Rev. Biol. **84:** 131–176.

Jacobsen, S.E., and Meyerowitz, E.M. (1997). Hypermethylated SUPERMAN epigenetic alleles in *Arabidopsis*. Science **277:** 1100–1103.

Jirtle, R.L., and Skinner, M.K. (2007). Environmental epigenomics and disease susceptibility. Nat. Rev. Genet. **8:** 253–262.

Johannes, F., et al (2009). Assessing the impact of transgenerational epigenetic variation on complex traits. PLoS Genet. **5:** e1000530.

Johannes, F., Colot, V., and Jansen, R.C. (2008). Epigenome dynamics: A quantitative genetics perspective. Nat. Rev. Genet. **9:** 883–890.

Krueger, F., and Andrews, S.R. (2011). Bismark: A flexible aligner and methylation caller for Bisulfite-Seq applications. Bioinformatics **27:** 1571–1572.

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., Jones, S.J., and Marra, M.A. (2009). Circos: An information aesthetic for comparative genomics. Genome Res. **19:** 1639–1645.

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.

Lee, M., Sharopova, N., Beavis, W.D., Grant, D., Katt, M., Blair, D., and Hallauer, A. (2002). Expanding the genetic map of maize with the intermated B73 x Mo17 (IBM) population. Plant Mol. Biol. **48:** 453–461.

Li, L., Petsch, K., Shimizu, R., Liu, S., Xu, W.W., Ying, K., Yu, J., Scanlon, M.J., Schnable, P.S., Timmermans, M.C.P., Springer, N.M., and Muehlbauer, G.J. (2013). Mendelian and non-Mendelian regulation of gene expression in maize. PLoS Genet. **9:** e1003202.

Lippman, Z., et al. (2004). Role of transposable elements in heterochromatin and epigenetic control. Nature **430:** 471–476.

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.

Liu, S., Yeh, C.T., Ji, T., Ying, K., Wu, H., Tang, H.M., Fu, Y., Nettleton, D., and Schnable, P.S. (2009). Mu transposon insertion sites and meiotic recombination events co-localize with epigenetic marks for open chromatin across the maize genome. PLoS Genet. **5:** e1000733.

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, K., Tör, 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.

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.

McMullen, M.D., et al. (2009). Genetic properties of the maize nested association mapping population. Science **325:** 737–740.

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

Messing, J., and Dooner, H.K. (2006). Organization and variability of the maize genome. Curr. Opin. Plant Biol. **9:** 157–163.

Miura, A., Nakamura, M., Inagaki, S., Kobayashi, A., Saze, H., and Kakutani, T. (2009). An *Arabidopsis* jmjC domain protein protects transcribed genes from DNA methylation at CHG sites. EMBO J. **28:** 1078–1086.

Quinlan, A.R., and Hall, I.M. (2010). BEDTools: A flexible suite of utilities for comparing genomic features. Bioinformatics **26:** 841–842.

Rabinowicz, P.D., and Bennetzen, J.L. (2006). The maize genome as a model for efficient sequence analysis of large plant genomes. Curr. Opin. Plant Biol. **9:** 149–156.

Rasmusson, D.C., and Phillips, R.L. (1997). Plant breeding progress and genetic diversity from de novo variation and elevated epistasis. Crop Sci. **37:** 303–310.

Regulski, M., et al. (June 5, 2013). The maize methylome influences mRNA splice sites and reveals widespread paramutation-like switches guided by small RNA. Genome Res. doi/10.1101/gr.153510.112.

Reinders, J., Wulff, B.B.H., Mirouze, M., Marí-Ordóñez, 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. (2008). Population epigenetics. Curr. Opin. Genet. Dev. **18:** 221–226.

Schmitz, R.J., He, Y., Valdés-López, O., Khan, S.M., Joshi, T., Urich, M.A., Nery, J.R., Diers, B., Xu, D., Stacey, G., and Ecker, J.R. (June 5, 2013a). Epigenome-wide inheritance of cytosine methylation variants in a recombinant inbred population. Genome Res. doi/10.1101/gr.152538.112.

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., and Ecker, J.R. (2013b). Patterns of population epigenomic diversity. Nature **495:** 193–198.

Schnable, J.C., Freeling, M., and Lyons, E. (2012). Genome-wide analysis of syntenic gene deletion in the grasses. Genome Biol. Evol. **4:** 265–277.

Schnable, P.S., et al. (2009). The B73 maize genome: Complexity, diversity, and dynamics. Science **326:** 1112–1115.

Schultz, M.D., Schmitz, R.J., and Ecker, J.R. (2012). 'Leveling' the playing field for analyses of single-base resolution DNA methylomes. Trends Genet. **28:** 583–585.

Slotkin, R.K., and Martienssen, R. (2007). Transposable elements and the epigenetic regulation of the genome. Nat. Rev. Genet. **8:** 272–285.

Springer, N.M., et al. (2009). Maize inbreds exhibit high levels of copy number variation (CNV) and presence/absence variation (PAV) in genome content. PLoS Genet. **5:** e1000734.

Storey, J.D. (2002). A direct approach to false discovery rates. J. R. Stat. Soc. Series B Stat. Methodol. **64:** 479–498.

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

Swanson-Wagner, R., Briskine, R., Schaefer, R., Hufford, M.B., Ross-Ibarra, J., Myers, C.L., Tiffin, P., and Springer, N.M. (2012). Reshaping of the maize transcriptome by domestication. Proc. Natl. Acad. Sci. USA **109:** 11878–11883.

Swanson-Wagner, R.A., Eichten, S.R., Kumari, S., Tiffin, P., Stein, J.C., Ware, D., and Springer, N.M. (2010). Pervasive gene content variation and copy number variation in maize and its undomesticated progenitor. Genome Res. **20:** 1689–1699.

Trapnell, C., Pachter, L., and Salzberg, S.L. (2009). TopHat: Discovering splice junctions with RNA-Seq. Bioinformatics **25:** 1105–1111.

Vaughn, M.W., et al. (2007). Epigenetic natural variation in *Arabidopsis thaliana*. PLoS Biol. **5:** e174.

Venkatraman, E.S., and Olshen, A.B. (2007). A faster circular binary segmentation algorithm for the analysis of array CGH data. Bioinformatics **23:** 657–663.

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.

Zilberman, D., Gehring, M., Tran, R.K., Ballinger, T., and Henikoff, S. (2007). Genome-wide analysis of *Arabidopsis thaliana* DNA methylation uncovers an interdependence between methylation and transcription. Nat. Genet. **39:** 61–69.