DNA Methylation Changes during Maize Leaf Growth

Corresponding authors:
Hilde Nelissen
Department of Plant Systems Biology, VIB
Department of Biotechnology and Bioinformatics, Ghent University
Technologiepark 927, B-9052 Gent, Belgium
Tel.: +32 (0)9 331 35 27
Fax: +32 (0)9 331 38 09
Email: hinel@psb.vib-ugent.be

Research Area:
Genes, Development and Evolution
Differential Methylation during Maize Leaf Growth Targets Developmentally Regulated Genes

Author names:
Jasper Candaele, Kirin Demuynck, Douglas Mosoti, Gerrit T.S. Beemster, Dirk Inzé and Hilde Nelissen
Department of Plant Systems Biology, VIB, B-9052 Ghent, Belgium (J.C., K.D., D.M., D.I., H.N.); Department of Plant Biotechnology and Bioinformatics, Ghent University, B-9052 Ghent, Belgium (J.C., K.D., D.M., D.I., H.N.); Department of Biology, University of Antwerp, B-2000 Antwerp, Belgium (G.T.S.B.)

ONE-SENTENCE SUMMARY

The successive growth processes of cell division and cell elongation and the transition between them in the growing maize leaf are characterized by differential DNA methylation at the 5’ and 3’ regions of genes, of which only 5’ methylation changes are correlated with differential expression.
1 This work was supported by the Agency for Innovation by Science and Technology in Flanders ("Strategisch Basisonderzoek") (SB093354), by the Interuniversity Attraction Poles Programme (IUAP P7/29 "MARS"), initiated by the Belgian State, Science Policy Office, and by grants from Ghent University ("Bijzonder Onderzoeksfonds Methusalem project" no. BOF08/01M00408 and Multidisciplinary Research Partnership "Biotechnology for a Sustainable Economy" no. 01MRB510W).

2 Present address: Mount Kenya University, Kisii Campus, Kisii, Kenya.

3 These authors contributed equally to the article.

* Corresponding author; e-mail hinel@psb.vib-ugent.be.

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.plantphysiol.org) is: Hilde Nelissen (hinel@psb.vib-ugent.be).

[C] Some figures in this article are displayed in color online but in black and white in the print edition.

[W] The online version of this article contains Web-only data.

[OA] Open Access articles can be viewed online without a subscription.
DNA methylation is an important and widespread epigenetic modification in plant genomes, mediated by DNA methyltransferases (DMTs). DNA methylation is known to play a role in genome protection, regulation of gene expression, splicing, and was previously associated with major developmental reprogramming in plants, such as vernalization and transition to flowering. Here, we show that DNA methylation is also controlling the growth processes of cell division and cell expansion within a growing organ. The maize (*Zea mays*) leaf offers a great tool to study growth processes as the cells progressively move through the spatial gradient encompassing the division zone, transition zone, elongation zone, and mature zone. Opposite to *de novo* DMTs, the maintenance DMTs were transcriptionally regulated throughout the growth zone of the maize leaf, concomitant with differential CCGG methylation levels in the four zones. Surprisingly, the majority of differentially methylated sequences mapped on or close to gene bodies, and not to repeat-rich loci. Moreover, especially the 5’ and 3’ regions of genes, which show overall low methylation levels, underwent differential methylation in a developmental context. Genes involved in processes such as chromatin remodeling, cell cycle progression, and growth regulation were differentially methylated. The presence of differential methylation located upstream of the gene anti-correlated with transcript expression, while gene body differential methylation was unrelated to the expression level. These data indicate that DNA methylation is correlated with the decision to exit mitotic cell division and to enter cell expansion, which adds a new epigenetic level to the regulation of growth processes.
DNA methylation is the covalent modification of nucleotides in DNA by addition of a methyl group. In the nuclear genome of higher eukaryotes, 5-methylcytosine (5-mC) is the most important DNA modification (Goll and Bestor, 2005). It is a phenomenon of ancient origin predating plant-animal diversification. However, some differences exist between plant and animal methylome patterning and function, and DNA methylation has been found to be evolutionarily lost in a few species (Feng et al., 2010; Zemach et al., 2010). Eukaryotic DNA methylation is established by DNA methyltransferase enzymes that transfer a methyl group from S-adenosyl methionine to the fifth carbon of cytosine. These enzymes can largely be subdivided in maintenance and de novo DNA methyltransferases (DMTs), depending on whether the recognition site is already methylated or not. Maintenance DMTs conserve the methylation status of symmetrical (palindromic) sites after DNA replication, by recognizing the hemi-methylated locus and methylating the newly synthesized strand. In plants, there are two types of maintenance DMTs: DNA METHYLTRANSFERASE (MET) and CHROMOMETHYLASE (CMT). The former methylates CG sites during DNA replication, whereas the latter methylates CHG (H = A, C or T) sites located in chromatin in which histone 3 is dimethylated on lysine 9 (H3K9Me2) (Goll and Bestor, 2005). De novo DMTs methylate previously unmethylated DNA, and are DOMAINS REARRANGED METHYLTRANSFERASEs (DRMs) in plants, since the motifs in the catalytic methyltransferase domain are reshuffled through a circular permutation. DRMs are mainly known for methylating asymmetric CHH sites, but are able to de novo methylate cytosines in any sequence context, and are guided by 24-nt short interfering RNAs (siRNAs) in a process called RNA-directed DNA methylation (RdDM). These siRNAs are generated through the action of plant-specific RNA polymerases and the RNA interference (RNAi) machinery (Law and Jacobsen, 2010). Also DRM-like (DRML) proteins, catalytically mutated DRM paralogs, play an important role in de novo methylation (Saze et al., 2012). Lastly, DNMT2-type methyltransferases have been classified as DMTs because of the apparent presence of a DNA methylase domain, but are in fact most likely RNA methyltransferase enzymes (Goll et al., 2006). Plant DNA methylation is removed passively through DNA replication or actively by removal and replacement of the methylated C by a DNA glycosylase and the base excision repair mechanism (Law and Jacobsen, 2010).

Plant genomes bear methylation in CG, CHG, and CHH context, as opposed to only CG methylation in animals, with the notable exception of embryonic stem cells and neurons (Lister et al., 2009; Lister et al., 2013). Moreover, the highest levels of DNA methylation have been found in plants, with up to 50% in some species. Together with histone modifications and chromatin remodeling, DNA methylation determines the epigenetic state of the genome, not only at a global level, affecting large chromosomal domains or even entire chromosomes, but also at very specific sites, such as individual genes (Suzuki and Bird, 2008). Plant DNA methylation is known to play an important role in genomic imprinting and genomic protection from transposable elements and other repetitive DNA sequences. Furthermore, it regulates the expression of multiple genes (Jaenisch and Bird, 2003), and has been implied to play a role in gene splicing (Regulski et al., 2013). The influence of DNA methylation in development is especially evident in case of mutations in DNA methyltransferase genes, which are embryo-lethal in mice and lead to developmental abnormalities in plants (Goll and Bestor, 2005). The role of DNA methylation in mammalian development has been thoroughly studied in the context of embryonic development, differentiation of pluripotent stem cells, and aberrances associated with cancer progression (Smith and Meissner, 2013). In plants, DNA methylation plays a role in endosperm and embryo development (Choi et al., 2002), vernalization, hybrid vigor (Groszmann et al., 2011), and fruit ripening.
(Zhong et al., 2013). Tissue-specific differences in plant methylomes have been known for some time, such as the specific methylation changes during maize leaf development that help steer the tissue-specific expression of certain genes (Tolley et al., 2012). However, the identification of differentially methylated genes that are developmentally and tissue-specifically regulated in plants remains largely unexplored.

Maize (Zea mays) is a monocotyledonous grass with a bulky genome consisting largely of repetitive elements in which DNA methylation ranging from 86 to 65%, 74 to 50%, and 5.4 to 5% of CG, CHG, and CHH sequence contexts, respectively (Gent et al., 2013; Regulski et al., 2013). Much research on DNA methylation has been carried out in maize, but mostly in the framework of paramutation, imprinting, and transposon regulation (Regulski et al., 2013). Growing maize leaves are an interesting system to study plant organ growth as at a given time point during their development, all processes of growth are represented in a single leaf. At the basis of the leaf, cells are dividing (division zone or DZ) while at the more distally located transition zone (TZ), cells start to expand and differentiate (expansion zone or EZ). When cells stop expanding and reach their mature size, they become part of the mature zone (MZ) (Nelissen et al., 2013). Previous work has shown that the size of the DZ largely determines final leaf size (Rymen et al., 2007; Nelissen et al., 2012). Here, we use the growth zone of the maize leaf to study the dynamics of DNA methylation during leaf growth by methylation-sensitive amplified polymorphism (MSAP) analysis on dividing, transitioning, elongating or mature cells. An equal portion of hyper- and hypomethylation occurred throughout the growth zone, of which the majority of differential methylation was observed between the DZ and TZ. Surprisingly, only a minority of the differentially methylated sites was associated with transposable elements (TEs), and most differentially methylated sites mapped away from the centromeric regions. Almost 85% of these differentially methylated sites that changed over the developmental gradient, mapped in or close to coding sequences. Our data indicate that the majority of the genic differentially methylated sites in the growing maize leaf are at the beginning and end of the coding sequences, and immediately up- and downstream, opposite to what was seen for steady state methylated gene bodies (Lister et al., 2008). An inverse correlation between DNA methylation and expression level was observed in genes with differentially methylated sites in the promoter and 5′ region of the coding sequence. This correlation was not found for genes that are differentially methylated in the central and 3′ region of the gene body. Many genes that were found to be differentially methylated in the growing maize leaf, have functions in chromatin modification, gene regulation and development.

RESULTS

Maize Maintenance Methyltransferases Are Highly Expressed in Dividing Tissue

Since DNA methylation patterns are established by DMTs, we set out to profile the expression of all maize DMTs across the developing leaf. First, the encoding genes needed to be denominated, and the gene product correctly classified. As in other plant species, four types of DMTs are encoded by the maize genome: MET, CMT, DRM, and DNMT2. However, the nomenclature for the maize orthologs of every type has been assigned in a less straightforward manner. One representative of MET, CMT, and DRM was originally denominated ZMET1, ZMET2, and ZMET3, respectively (Cao et al., 2000). Furthermore, a total of seven genes encoding maize DNA methyltransferase domains were identified by the Chromatin Database consortium (ChromDB), and called DMT101 to DMT107 (Gendler et al., 2008). In order to eliminate further confusion concerning maize DNA methyltransferase nomenclature, we renamed them according to protein functionality.
and orthology (Fig. 1 and Supplemental Table S1). The maize B73 genome (release 5b.60) in fact encodes eight proteins with DNA methyltransferase domains. Two genes, GRMZM2G334041 and GRMZM2G333916, encode MET-type DMTs and can thus be denominated as ZmMET1 and ZmMET2, respectively. They are located inversely oriented on chromosome 7, 12.3 kb from each other, and are 99.5% identical in the coding sequence. The CMT-type maintenance DMTs are also represented by two genes in maize: GRMZM2G025592 and GRMZM2G005310, which we call ZmCMT1 and ZmCMT2, respectively. As is the case in Arabidopsis thaliana (Arabidopsis), there also are three DRM-type de novo DMT-encoding genes in the maize genome. ZmDRM1 (GRMZM2G092497) and ZmDRM2 (GRMZM2G137366) encode a functional DRM domain, and are just like the CMTs land plant-specific and quite different from the animal de novo methyltransferase enzymes, DNMT3A and DNMT3B (Cao et al., 2000). ZmDRML (GRMZM2G065599) encodes a DRM-like protein, similar to AtDRM3, in which the functional motifs of the methyltransferase domain have been mutated. Lastly, GRMZM2G157589 encodes ZmDNMT2, homologous to a protein originally called DNMT2 in animals. However, the animal protein was found to methylate a cytosine at position 38 in aspartic acid transfer RNA. Therefore, its name was changed to tRNA aspartic acid methyltransferase 1 (TRDMT1), the first RNA cytosine methyltransferase to be identified (Goll et al., 2006).

Samples for MSAP and expression analysis were prepared from growing maize leaves. Therefore, growth of the fourth leaf of maize B73 was measured over time. The first days after leaf emergence from the sheath are characterized by a steady-state growth, during which leaf length increases linearly, since the speed of leaf growth (leaf elongation rate, LER) is constant. After five days, the LER diminishes as the leaf approaches its final length. During the steady-state growth phase, the different zones of the fourth leaf are present and representative samples for the DZ, TZ, EZ and MZ were dissected according to the microscopic analysis that determined the position of the zones based on the average cell length profile and the position of the division zone (Fig. 2A). The most basal part (first cm) of the maize leaf represents proliferating tissue (the DZ), as cells are continuously dividing and maintain a constant average cell length (18.30 ± 0.15 µm). The last dividing cell of the B73 leaf lies in the second cm, at 1.23 ± 0.04 cm (n=4) from the basis of the leaf. At this stage, cells also start elongating, so the second centimeter of the maize leaf (the TZ) consists of both dividing and elongating tissue. In the following centimeters (the EZ), cell length increases rapidly. Further toward the tip of the leaf, cells stop elongating and attain their mature cell length (the MZ).

For expression analysis, a fine-sampling method was applied, harvesting consecutive samples across the first 10 cm of the maize leaf. Combined with a kinematic analysis, this allows us to correlate the expression levels obtained by qPCR to the cellular processes (Nelissen et al., 2012; Nelissen et al., 2013). The expression of maintenance DMT transcripts ZmMET (p = 1.3e-13), ZmCMT1 (p = 2.2e-16), and ZmCMT2 (p = 6.5e-14) follows more or less the same profile as that of ZmCDKB1;1 (p = 6.7e-8) (Fig. 2B), encoding a CYCLIN DEPENDENT KINASE and used here as a proxy for cell division (Rymen et al., 2007; Nelissen et al., 2012). The expression of ZmMET1 and ZmMET2 cannot be distinguished because of the high sequence similarity. This expression pattern is in agreement with their role in methylation maintenance as high levels of them are necessary to replicate methylation patterns on newly synthesized DNA strands (Goll and Bestor, 2005). On the other hand, expression of ZmDNMT2 (p = 0.251) and the de novo methyltransferases, ZmDRM1 (p = 0.216) and ZmDRM2 (p = 0.055), remains constant across the different zones (Fig. 2C). Fluctuations in transcript abundance are observed, especially in the younger tissue, but these expression changes are not significant at the 5% level.
Notably, the expression of \textit{ZmDRML} is broadly similar to that of the maintenance DMTs, with high expression in the dividing tissue and lower expression in the more mature zones (p = 6.7e-9; Fig. 2C). This reduction, however, occurs more gradually compared to the expression patterns of \textit{ZmMET}, \textit{ZmCMT1}, and \textit{ZmCMT2} that reach their minimal level immediately after the TZ.

**Most Differential DNA Methylation Takes Place in Young Developing Tissue**

Next we investigated DNA methylation within the developing maize leaf, and the differences between the four zones (DZ, TZ, EZ, and MZ). Therefore, genomic DNA of these maize leaf zones was extracted from maize seedlings during steady-state leaf growth. Methylation sensitive amplified polymorphism (MSAP) analysis was carried out, applying 64 primer combinations, which amplify approximately 9,168 fragments smaller than 800 bp (Rombauts et al., 2003), covering 1.56\% of all CCGG sites in the maize genome. The methylation state of these CCGG sites for every zone can be deduced after methylation-sensitive restriction cleavage with \textit{Hpa}II and \textit{Msp}I (Supplemental Table S2). A total of 28,714 patterns of restriction fragments on gel were scored, the majority of which (95.1\%) did not exhibit differences over the four zones. These monomorphic bands represent loci with a stable methylation state across the leaf, and are subdivided into three methylation states according to the sensitivity of the locus to \textit{Hpa}II/\textit{Msp}I restriction (McClelland et al., 1994). Almost half (48.1 ± 1.4\%) of the loci were found to be stably unmethylated, whereas regarding stable methylation in the maize leaf, full CG methylation was the dominant form, represented by 36.7 ± 1.19\% of all investigated CCGG sites. Hemi-CHG (where H = C) methylation, henceforth referred to as CHG methylation, represented 15.2 ± 0.3\%. Stable methylation of both cytosines in the CCGG tetramer inhibits restriction of both enzymes, and cannot be visualized on an MSAP gel, as it yields no bands. However, differential methylation of this locus can be identified, since changing its methylation state into one of the three other methylation states can be visualized (Fig. 3A).

For the remaining methylated bands, we did observe changes in the MSAP pattern, representing differential methylation, either between the two biological repeats or between the zones (1.9\% and 3.0\% of all observed bands, respectively). The latter represents reproducible changes in DNA methylation across the developing maize leaf, which is the interest of our research. A total of 217 reproducibly differentially methylated bands were identified and cut out from the gel for amplification. Of these, 81\% exhibited a change in methylation pattern between at least two zones. The remaining 19\% were marked by a gradual change in intensity of the MSAP pattern. This could be an increase (17 bands) or decrease (20 bands) in intensity, or a rather gradual transition from one MSAP pattern in the DZ to another in the MZ (4 bands). These gradual methylation changes are most likely caused by unequal changes in the methylation state of the different cell types that make up the tissues of the harvested samples (Xiong et al., 1999; Peraza-Echeverria et al., 2001; Cervera et al., 2002). For our objectives, the most interesting loci were the CCGG sites that abruptly changed their methylation from one zone to the next. Of these, 39\% exhibited a change in methylation between the DZ and TZ, 37\% between the TZ and EZ, and 24\% between the EZ and MZ. Most loci thus undergo a change in the young, growing tissue, either at the end of the division zone, or toward the beginning of the elongation zone.

Above data support that many sequences were differentially methylated at the transition from cell division to cell expansion, a transition that was previously shown to play a crucial role in determining leaf size (Nelissen et al., 2012). To identify more sequences that are differentially methylated between the DZ and TZ, we
decided to zoom in on the transition between division and elongation, using additional primer combinations on pools of DZ and TZ samples, now covering a total of 6.25% of all CCGG sites (± 62,778 restriction fragments smaller than 800 bp (Rombauts et al., 2003)) (Fig. 3A). A total of 208 additional bands representing differentially methylated loci between the DZ and TZ were cut for sequencing. These compiled data over the two screens indicate that simple methylation changes are most common (95.6%) between two zones (Fig. 3B, methylation changes occurring at the edges). Within these, methylation changes affecting CG sites (54.2%) are more abundant than those occurring at CHG sites (41.4%). Changes between mCG and mCHG, on the other hand, are extremely rare (blue arrows). More specifically, a replacement of CmCG by mCCG occurred only once in our dataset, and the inverse was not found. These represent the occurrence of both a hypo-and hypermethylation event at the same CCGG site. Also changes between heavy and low methylation are quite rare (4.1%). This means that hypo- (1%) or hypermethylation (3.1%) of both Cs at the same CCGG site between two zones is not a common form of differential methylation in the maize leaf. The relative abundance between the different zones, of all twelve possible methylation changes that can be identified using MSAP, is summarized in Supplemental Table S3.

Most Differentially Methylated Single-Locus Sequences Map In or Close to Genes and Away from the Centromeres

Successfully amplified bands were sequenced and BLAST-searched against the maize B73 genome (version 5b.60). Sequences that map to multiple locations, have a high probability of being a transposable element (TE) and thus were BLASTed to the maize TE database (www.maizetedb.org) (Fig. 4). The latter multi-locus sites represent 37.7% of the sequences in our dataset, and since these sequences mapped to two or more loci in the genome, the exact site of differential methylation could not be identified. A minority of the multi-locus sequences mapped to a limited number of sites in the maize genome, some of which had also a copy on an organellar genome. The majority of the multi-locus sequences were highly repetitive TEs for which the location could not be identified. These sequences could be classified according to the transposon type (Supplemental Table S4). Type II (DNA) transposons were mostly found in the hypermethylated fraction (9/26), and less in the more abundant hypomethylated class (3/35). Overall, type I (retro)transposons are the most abundant in our data, as they are also most abundant in the maize genome (Baucom et al., 2009). However, the distribution of retroelement families within the dataset does differ significantly (p < 0.001) from the natural distribution. This is mainly because the most prevalent maize TE family, Huck, is underrepresented (Supplemental Table S5, adapted from Baucom et al. (2009)).

Of all sequences obtained from differentially methylated bands, more than 60% could be mapped to a single location in the maize genome (Fig. 4). Most of these mapped on or in proximity (up to 5 kb) of coding sequences. If the differentially methylated CCGG sequence lies within the transcript-coding region, this is referred to as gene body methylation, and was almost always found to encode a protein. In eight instances, a differentially methylated gene body was identified by more than one sequence. Also for the sites that mapped to a locus in proximity of one or more genes, in most cases at least one of these genes was protein-coding. All unique loci were mapped on the ten maize chromosomes showing an equal distribution over the chromosomes and the chromosome arms (Fig. 5; Supplemental Fig. S1A). However, on average, the sites of differentially methylated sequences were located away from the centromeres and the pericentromeric regions (Supplemental Fig. S1B), which are generally gene-poor.
Next, we studied the distribution of differential methylation that mapped to a single location with respect to the gene body (Fig. 4). Of the genic hits, two thirds map in the gene body and one third up- or downstream (up to 5 kb up- or downstream). The majority of the differentially methylated CCGG sites that map to a gene lie within an exon. However, when examining the site of differential methylation within exons and introns, the distribution of methylation is not significantly different from a random distribution (p > 0.25). When mapping the sites of differential methylation in and close to coding sequences (Fig. 6), we found that the sites of differential DNA methylation are not distributed equally throughout the gene body (p < 0.001). The number of differentially methylated sites is the highest around the gene extremities. More specifically, the number of differentially methylated sites is higher than expected in the first 10% (p < 0.02) and the last 20% (p < 0.001) of the gene body. Also, the highest amount of differentially methylated sites in the up- and downstream sequences was found in the first 0.5 kb before and 1.0 kb after the coding sequences. Similar results are obtained when mapping the hits with respect to the start and stop codon (Supplemental Fig. S2A), or when mapping hits without scaling the gene body around the gene start (Supplemental Fig. S2B) and stop sites. Several genes were identified by multiple differentially methylated sequences, and in two cases two different CCGG sites within the same gene were found to be differentially methylated. The first, a CTC-interacting domain-encoding gene (GRMZM5G829738) orthologous to human ATAXIN2, is hypermethylated in two exons between the DZ and TZ. The second gene (GRMZM2G139157) encodes a protein kinase with ubiquitin conjugating domain, which is differentially methylated at two consecutive CCGG sites in an exon and an intron between the EZ and MZ.

Many Differentially Methylated Genes Are Involved in Gene Regulation, Transcription, and Development

The MSAP dataset was mined for functional enrichment using Plaza (http://bioinformatics.psb.ugent.be/plaza) (Van Bel et al., 2012). Electron transport was enriched when considering all differentially methylated genes (p = 0.01), whereas zinc ion binding (p = 0.021) was enriched in the TZ-EZ transition, and ATP-dependent helicase (p = 0.025), chlorophyll binding (p = 0.042); and light-harvesting complex (p = 0.027) activities were enriched in the EZ-MZ transition. The DZ-TZ dataset is enriched for binding (7.8E-4) and catalytic activity (0.0017), and the genes that were hypomethylated between the DZ and TZ, were enriched for ATP-binding (p = 0.007) and helicase activity (0.046). To determine the function of these genes we identified the function of each of the gene products through the presence of protein domains and orthology with Arabidopsis genes. Information about the 95 sequences found to be hyper- and hypomethylated between the DZ and TZ is summarized in Supplemental Tables S6, A and B, respectively. The 54 remaining sequences were found to be differentially methylated between the TZ and EZ, the EZ and MZ, and between multiple zones, and are represented in Supplemental Tables S7, A-C, respectively.

At least seven genes with various functions in chromatin remodeling were found to be differentially methylated, or in the vicinity of a differentially methylated locus. A histone acetyltransferase (GRMZM2G371912) orthologous to Arabidopsis IDM1 (INCREASED DNA METHYLATION1; Qian et al. (2012)), and three plant homeodomain (PHD)-encoding genes were undergoing exonic differential methylation. PHD domains are found in chromatin modifiers and transcriptional regulators, where they are often responsible for binding of methylated histones (Sanchez and Zhou, 2011). The telomere-binding SINGLE MYB HISTONE3 (SMH3 - GRMZM2G023667; Marian et al. (2003)), a gene highly orthologous to the ARGONAUTE-encoding gene (AC189879.3_FGT003) involved in RNA-induced silencing (Hutvagner and Simard, 2008), and a SNF2-encoding gene (GRMZM2G313553) were hypomethylated. The SNF2 domain-encoding proteins are helicase-
related ATPases that drive chromatin remodeling complexes (Ryan and Owen-Hughes, 2011). Other DNA-interacting domains are found in two zinc finger proteins, a DNA ligase and a DNA helicase. Also, several genic hits are involved in transcription and transcriptional regulation: an RNA polymerase clamp, a TATA-binding interacting protein, four transcription factors, two RNA helicases, RNaseH and two splicing factors. Several cytoskeleton-related proteins were identified, such as a tubulin (GRMZM2G407869), kinesin (GRMZM2G338928), exosin (GRMZM2G172602), augmin-like (GRMZM2G041878), and two actin organizing proteins (GRMZM2G142779 and GRMZM2G113174). Moreover, alpha-tubulin-encoding genes, of which one was present in this dataset, have been shown to be differentially methylated in different maize tissues (Lund et al., 1995). In addition, several classes of growth related genes were represented in the dataset: a beta-expansin (EXPB3), GALACTURONOSYLTRANSFERASE4 (GAUT4 - cell wall synthesis), PHOSPHATIDYLINOSITOL SYNTHASE2 (PIS2 - cell membrane synthesis), two genes involved in vesicle-associated transport (Ankyrin repeat protein50 and syntaxin132), a CYCLIN D4 (CYCD4) and an ANAPHASE PROMOTING COMPLEX10 (APC10)-like gene (Eloy et al., 2011), indicate regulation of transcripts involved in cell division and expansion processes.

Interestingly, at least five genes were found that are known to be involved in development. The MYB-like transcription factor DIVARICATA (GRMZM2G079458) determines dorsoventral asymmetry, promoting ventral identity (Galego and Almeida, 2002). This gene got promoter hypermethylated between the DZ and TZ. GRMZM2G076257 is an ortholog of Arabidopsis DOT4 (DEFECTIVELY ORGANIZED TRIBUTARIES4), a pentatricopeptide repeat protein that regulates vasculature development (Petricka et al., 2008). Also this gene was hypermethylated between the DZ and TZ, in its gene body. Trehalose-6-phosphate (T6P) synthase (TPS, GRMZM2G099860) is hypermethylated between DZ and TZ. TPS is involved in the biosynthesis of T6P and trehalose, both having important functions in plant growth and development. Trehalose is an osmoprotectant that influences maize inflorescence architecture, whereas T6P is an important signaling molecule involved in embryo development, vegetative growth, and leaf senescence (O’Hara et al., 2013). Also the presence of a GIBBERELLIN 20 OXIDASE2 ortholog (GRMZM2G099467), which is hypermethylated in the TZ, was notable, as our previous experiments have revealed a pivotal role for gibberellins in maize leaf development, specifically in the TZ (Nelissen et al., 2012). A hypomethylated CCGG site between the EZ and MZ maps onto the ZmMiR396a stem-loop precursor. In Arabidopsis, MiR396 regulates leaf growth through regulation of GROWTH REGULATING FACTORs (GRFs) (Zhang et al., 2009). More specifically, seven out of nine AtGRFs are targeted by MiR396 (Debernardi et al., 2012). Similarly, all but two out of 18 maize GRFs carry the MiR396 target motif. These two genes, ZmGRF4 and ZmGRF10 (Zhang et al., 2008), are not downregulated in the elongating tissue (Supplemental Fig. S3). Other abundant protein domains are associated with protein turnover, regulation of protein function, sugar metabolism, and transport.

**Only Differential Methylation in Upstream Genic Regions is Correlated with Gene Expression**

We then addressed the question whether the expression of differentially methylated genes was correlated to the methylation state. A total of 43 genes was selected for expression analysis using quantitative reverse transcriptase PCR (q-RT-PCR). The expression for genes that undergo differential methylation between the DZ and TZ was investigated over the first four cm, whereas expression of genes that undergo methylation changes in the TZ-EZ or the EZ-MZ was analyzed over the full first 10 centimeters of the developing maize leaf.
For 10 out of 11 genes undergoing differential methylation in the promoter and 5’ region of the gene body (up to the start codon), an inverse correlation between DNA methylation and expression was found, eight of which were significant at the 5% level (Supplemental Fig. S4). All four genes undergoing hypermethylation upstream of the gene between DZ and TZ (Supplemental Fig. S4A), were upregulated in the following centimeters. EXPB3 (GRMZM2G169967), for example, was 20-fold upregulated (p = 6.5e-05) in the second centimeter, and downregulated afterwards (Fig. 7A). Similarly, two genes undergoing promoter hypomethylation between TZ and EZ were upregulated in the more mature zones (Supplemental Fig. S4C). The most extreme case (GRMZM2G125934) encodes a bZIP protein orthologous to Arabidopsis bZIP65/TGA10 (TGACG MOTIF-BINDING PROTEIN10). This gene was virtually not expressed in dividing and early elongating tissue, but was highly upregulated in late elongating tissue, after hypomethylation, and subsequently downregulated in mature tissue (p = 8.7e-08). Four out of five genes undergoing 5’ hypermethylation between DZ and TZ were subsequently downregulated (Supplemental Fig. S4, B and D). For three of these genes, this change in expression was significant. One gene (GRMZM2G123585), encoding a leucine zipper and a domain of unknown function (DUF547), underwent differential methylation at a site 5 bp downstream of the start codon. This site was not only hypermethylated between DZ and TZ, but also hypomethylated between EZ and MZ. The expression was drastically reduced between DZ and MZ: almost sevenfold between DZ and EZ, and drops toward 1% of its original expression level in the MZ (p = 1.3e-08). DIVARICATA (GRMZM2G031441) was downregulated after having its promoter hypermethylated between DZ and TZ (p = 0.072) (Supplemental Fig. S4B).

Thirteen genes for which differential methylation was found in the remainder of the gene body were profiled (Supplemental Fig. S5), and only for three of these, the expression was anti-correlated with their methylation status. Expression of five of these genes actually showed a significant positive correlation with DNA methylation. For example, an ATP binding microtubule motor family protein-encoding gene (GRMZM2G338928) was 26-fold downregulated when a site in its first exon became hypomethylated (Fig. 7B). Oppositely, hypermethylation was sometimes also associated with an increase in gene expression. Hypermethylation of a site in the second intron of a GATA zinc finger transcription factor (GRMZM2G052616) between TZ and EZ was associated with upregulation of the gene in the more mature zones. For several of the genes with a positive correlation between methylation and expression, multiple CCGG sites were identified in the MSAP analysis as differentially methylated. An ARGONAUTE-like PAZ/PIWI domain encoding gene (AC189879.3_FGT003), was downregulated (3.4-fold, p = 0.0016) after hypomethylation in its next-to-last exon, but showed a second CCGG site in its last exon. Lastly, for one of the few genes for which methylation negatively correlated with expression, we found two hypomethylated sites. The gene encoding both a protein kinase and ubiquitin conjugating domain (GRMZM2G139157), had both its second exon and adjacent intron hypomethylated between EZ and MZ. Expression of the gene was upregulated 4.4-fold in the mature tissue (p = 0.02). This might be evidence of a large scale hypomethylation event across the gene body, causing upregulation. Lastly, we analyzed the expression of genes for which differential methylation was found downstream of the stop codon (Supplemental Fig. S6A), or for which a methylation change was found to be gradual (Supplemental Fig. S6B). Also in these cases, no correlation between methylation and expression was found.
By compiling all expression data (Fig. 8), it becomes evident that only 5’ methylation correlates negatively with expression. Of the 28 transcriptionally profiled genes, expression of 14 transcripts anti-correlated with expression, ten of which were found in the 5’ region of a gene. Eight genes showed a positive correlation between gene expression and methylation, all but one found in the gene body. Six genes were not found to be differentially expressed, only one of which showed differential methylation in the 5’ region.

DISCUSSION

The Methylation State of the Developing Maize Leaf

Here we present the analysis of DNA methylation over the different growth zones of the developing maize leaf. More than 50% of the investigated CCGG sites were found to be stably methylated. Because the use of restriction enzymes does not allow us to distinguish some forms of methylation, MSAP results tend to yield an underestimation of genomic DNA methylation (Xiong et al., 1999). Indeed, recent whole genome bisulfite sequencing of the maize genome in unfertilized ears revealed that of all C’s in CG and CHG context, 86% and 74%, respectively, are methylated (Gent et al., 2013). A second whole genome methylome profiling effort of coleoptile tissue found lower, albeit still substantial, values: 65%, 50%, and 5% (Regulski et al., 2013). Generally, we found that full CG methylation was more represented than hemi-CHG methylation in the leaf, which is in agreement with previous MSAP experiments carried out on leaves of other grass species, such as sorghum (Zhang et al., 2011) and rice (Wang et al., 2011). Moreover, all plant species for which whole-methylome profiles have been established, predominantly have methylation in a CG context (Lister et al., 2008; Zemach et al., 2010; Gent et al., 2013). Furthermore, we showed that, although the majority of genomic sequences remained stably (un)methylated in the leaf cells, a small number (3%) of the loci underwent methylation changes as the cells change from a dividing into elongating state or as they differentiate into mature leaf cells.

Moreover, almost 40% of the methylation changes were found between cells that are fully dividing and those that are transitioning into elongating cells. Both the cell cycle machinery and cytosine methylation are strongly conserved systems in eukaryotes. However, DNA methylation has been evolutionary lost in cell cycle model species, such as *Saccharomyces cerevisiae* and *Caenorhabditis elegans* (Goll and Bestor, 2005). This suggests that DNA methylation itself does not play a role in cell cycle regulation in these lower eukaryotes. In higher eukaryotes, on the other hand, cytosine methylation and DNA methylation are strongly interconnected, and DNA methylation might impose an additional level of regulation. The progress from dividing to elongating tissue also implies that gradually fewer cells are dividing and undergoing maintenance methylation. This means that the shift between dividing and elongating cells is also a shift between a continuous interchange in hemimethylation and full methylation in the DZ and a more stable methylation state in the EZ (Goll and Bestor, 2005). However, this shift is also associated with important developmental changes determining growth. The cell cycle is characterized by a tightly orchestrated regulation of specific genes, and an exit from the cell cycle leads to the expression of different classes of genes altogether (Inzé and De Veylder, 2006). Moreover, previous experiments from our lab have indicated that the TZ, which contains both dividing and elongating tissue, largely determines the speed of leaf growth and final leaf length in maize (Nelissen et al., 2012). Also Arabidopsis leaf growth is characterized by a robust spatiotemporal regulation of cell division and cell elongation (Andriankaja et al., 2012). Therefore, we focused further screening for methylation changes on the DZ and TZ.
It is also between these two zones that maize maintenance methyltransferase expression was drastically reduced. Yet, although hypomethylation seemed to be slightly more abundant than hypermethylation between DZ and TZ, overall hypo- or hypermethylation were found to occur in more or less equal amounts. Most likely, the maintenance-type enzymes are required in high number during cell division, during which a great number of cytosine methylation needs to occur after every round of DNA replication (Goll and Bestor, 2005). At later stages of cellular development, only basal levels of these enzymes are necessary to reinforce cytosine methylation by de novo DMTs, which are stably expressed, and other chromatin modifications established by histone modifying enzymes (Vaillant and Paszkowski, 2007). Interestingly, the functional DMT genes (MET, CMT and DRM) are all present in duplicate in the maize genome. Moreover, they all share high homology. This is especially true for the MET genes, which are 99.4% identical at the protein level. Most likely, these genes were quite recently duplicated to compensate for the expansion of the maize genome, and its large proportion of repetitive (transposable) elements (Baucom et al., 2009).

Not all methylation changes identifiable with the MSAP technique occurred with the same frequency. Most strikingly, changes between hemi 6mCCG and full C6CG, indicative of a demethylation and a remethylation event at the same CCGG site, were found to be almost non-existent. Similar results were found in MSAP experiments conducted in Arabidopsis (Cervera et al., 2002), pepper (Portis et al., 2004), and maize (Lu et al., 2008). Also, the occurrence of two methylation changes between two zones was quite uncommon. Going from one zone to the next, almost all CCGG sites would thus undergo simple one-step hypo- or hypermethylation. Most of these differential methylation events occurred at CG sites, in which most DNA methylation takes place (Law and Jacobsen, 2010). It is important to note that not all MSAP pattern changes can be deduced in a straightforward manner. If one or more CCGG sites are present within the amplified MSAP band, the differential methylation pattern might not be caused by a change in the methylation state of the outer CCGG, but by an opposing change in (one of) the inner CCGG site(s). Indeed, for 42 of the 149 single-locus differentially methylated sites, one or more CCGG sites were found inside the amplified MSAP sequence (Supplemental Tables S6 and S7).

**Differential Methylation Targets 5’ and 3’ Edges of Genic Regions, but Only 5’ Correlates with Expression**

Genic loci that underwent differential methylation showed enrichment at the 5’ and 3’ edges of the gene body. This pattern of differential methylation is the exact opposite of the stable methylation patterns that have been identified in several tissues of higher plant species such as Arabidopsis, rice, and poplar (Feng et al., 2010). In these species, the steady-state gene body methylation is generally high in the central part of genes, with 5’ and 3’ terminal regions being devoid of DNA methylation, which implies an important function in gene regulation. This pattern has been proposed to play a role in transcriptional regulation, with moderately expressed genes being highly methylated in the gene body. An alternative role might lie in exon definition, since the majority of gene body methylation is found in exons (Saze and Kakutani, 2011), or the inhibition of transcriptional initiation from spurious promoters (Lauria and Rossi, 2011). Indeed, almost three quarters of the differentially methylated sites identified in this experiment, were found in an exon. However, in our study a high prevalence of methylation changes was found at the start and stop of the gene. This might be an indication of versatile gene regulation, whereas the central body methylation is steadily maintained depending on the overall expression...
level of the gene. There are different ways in which gene body DNA methylation can influence RNA polymerase action and general transcript formation. Firstly, DNA methylation could shield spurious transcript start sites within the gene, in order to avoid formation of truncated gene products, which could be deleterious to cellular function (Lauria and Rossi, 2011). Secondly, DNA methylation might simply identify which sequences within the transcript are exons, and which introns, as splicing generally coincides with transcription. Indeed, more and more evidence is accumulating that epigenetic modifications, such as DNA methylation, are influencing gene splicing, also in maize (Regulski et al., 2013). In this case, differential methylation between tissues could coincide with differential splicing, giving rise to different transcripts of the same gene. However, we did not find compelling evidence that suggests that differential methylation in the gene bodies influences differential splicing of the transcripts. Further research will be needed to investigate this phenomenon.

The upstream, or 5’, sequences encompassing the promoter as well as the downstream, or 3’, sequences can contain gene regulatory elements and are subjected to DNA methylation (Suzuki and Bird, 2008; Yu et al., 2013). In the current study, we found a correlation between gene expression and DNA methylation of the 5’ portion of the gene. Differential methylation of both the promoter and the portion before and around the start codon has an adverse effect on gene expression. This is reminiscent of CpG island methylation in animals, which is associated with gene promoters, but often extends into the 5’ untranslated region and even the first intron (Suzuki and Bird, 2008). No correlation could be found between differential methylation of the central and 3’ part of the gene body, or sequences downstream of the gene. Although, in animals, several studies of the methylome during development of both somatic and cancer cells has indicated a correlation between intragenic methylation and gene expression (Kulis et al., 2013). No correlation could be observed between gene expression and differential methylation of the 3’ part of the gene body, or sequences downstream of the gene. However, previous steady-state methylome data indicated that gene body methylation might not have a direct influence on the transcription of the gene per se, but is rather correlated with the polymerase function in a parabolic fashion (Zemach et al., 2010). This means that moderately expressed genes are most likely to be heavily methylated, whereas highly and lowly expressed genes tend to bear little gene body methylation.

As the MSAP technique only yields information about the methylation state of one CCGG site, it does not give information about the methylation status of surrounding cytosines. Therefore we cannot show the causality of the changes in methylation status at the 5’ part of the gene to the difference in expression level, but merely show an increased correlation. In addition, there might also be other factors, such as histone modifications and binding of transcription factors which are influenced by changes in methylation, and hence have an effect on their expression (Jaenisch and Bird, 2003). Nevertheless, the current study demonstrates a strong correlation between methylation and developmental transitions, that provides a basis for further more detailed studies. For example, bisulfite sequencing would yield valuable information on the neighboring sequences and the overall methylation status of genes.

The Majority of Multi-Locus Sequences Are (Retro)Transposons

Roughly 25% of the maize genome consists of non-TE DNA, and only 6% of protein-coding genes (Schnable et al., 2009). Maize TEs can be categorized in two classes: RNA (or retro- or type I) transposons and DNA (type II) transposons. The former class is by far the largest, occupying 76% of the genome, whereas DNA transposons make up 8.6% of the genome (Schnable et al., 2009). However, more than 70% of the obtained
sequences did not exhibit TE-like features. It is unlikely that this is the consequence of a bias caused by the used primers, since we used all 64 possible HpaII/MspI-specific primers in combination with five EcoRI-specific primers, covering a wide range of restriction sites in the genome. A problematic amplification of repetitive sequences could in part explain the bias. However, even if all failed amplifications were TEs, non-TE sequences would still make up more than 40% of the dataset. Moreover, at least 34.7% of the hits in our dataset were mapped to protein-coding genes, and only 9.6% were not in the vicinity (5 kb in either direction) of a gene. Judging from these data, we could see that genic sequences were more prone to differential methylation than repetitive DNA. Similar results were obtained from MSAP experiments in sorghum (Zhang et al., 2011), rice (Wang et al., 2011), and maize (Lauria et al., 2004).

Less than 38% of the acquired sequences could not be mapped to a single locus in the maize genome, and most of those turned out to be TEs. When combining transposons from both the single-locus and multi-locus datasets, only 24.1% of the MSAP hits were found to be TEs. From these, 66.3% are type I TEs, 23.7% type II, and 10% were TEs of unknown origin (i.e. annotated as being a TE in maize_sequence, but not when blasted to mizatedb). There seems to be more type II TEs than expected in the maize genome. This can be explained by the fact that these TEs are generally of lower copy number and silenced less robustly than the very abundant RNA retrotransposons. Several members of the type II TE class are known for their mutagenic abilities. For example, CACTA TEs (11/19 type II TE in our dataset) are known to be mobilized in mutants with a compromised DNA methylation machinery in Arabidopsis (Kato et al., 2004). In maize, a process called developmental relaxation of TE silencing causes a burst of TE transcriptional activity in the shoot apical meristem, after which silencing needs to be re-established (Martínez and Slotkin, 2012). Indeed, 75% of type II TEs are found in the hypermethylated fraction of our dataset. This could be evidence of a progressive silencing of DNA TEs in older zones, in order to prevent mutation due to transposition.

The LTR retrotransposon order occupies by far the most space in the maize genome (74.6%). This is almost entirely caused by the presence of two superfamilies, Gypsy (46.4%) and Copia (23.7%) (Baucom et al., 2009). Also in our dataset, all but one of the retrotransposons found are members of the order of the Long Terminal Repeat (LTR) retrotransposons. However, within this order, the prevalence of members of the different families differs from the abundance in the maize genome (Supplemental Table S5). Most LTR-type retroelements identified in this study fall within the 20 most prevalent TE families of the maize genome. However, the most prevalent maize TE, Huck, is found only once in our dataset. Moreover, the average GC content of Huck elements is 60%, which is much higher than the average GC content of the maize repetitive fraction (48%), and than the average genomic GC content (47%) (Meyers et al., 2001). The Giepum TE, on the other hand, is more than eight times less present in the maize genome, but eight times more present in our dataset. The reason for this bias remains to be investigated.

The remaining multi-hit sequences mapped to two or more loci of the genome encoding the same sequence. Some of these sequences have at least one copy on a plastid genome. Combined with the four single-locus genes of organellar origin (all found in the EZ-MZ dataset), a total of 17 organelle-related genes were found using the MSAP approach. Since organellar genomes are of bacterial origin, they exhibit mainly methylation of adenines, and not of cytosines (Vanyushin and Ashapkin, 2011). Integration of organellar sequences in the nuclear genome is very common in many plant species, especially those who are outbreeding and have large genomes (Ayliffe et al., 1998). Seeing that over 99% of plastid and 95% of mitochondrial
genomic sequences have at least one copy in the maize nuclear genome (Kumar and Bendich, 2011), we most likely are visualizing DNA methylation in the nuclear copies.

**Differential Methylation Along the Developing Maize Leaf Potentially Affects Many Processes**

An interesting finding is the fact that some of the genes that are differentially methylated, are themselves involved in epigenetic processes, either directly (IDM1, SMH3, AGO, SNF2, and PHD-domain genes) or indirectly. For example, GRMZM2G311883 is a gene homologous to Arabidopsis MOS14, which is known to form a link between splicing and RNA-dependent DNA methylation (Zhang et al., 2013). Also genes involved in cell cycle regulation, such as CYCD4, APC10, TIP120-encoding (Wang et al., 2011), FORMIN 8-like (Xue et al., 2011), and AUGMIN 6-like (Hotta et al., 2012), are represented in the dataset. As are developmental processes affecting the vasculature (Petricka et al., 2008), meristem maintenance (Liu et al., 2009), cell shape and plant architecture (O'Hara et al., 2013). Several genes that are differentially methylated in the older zones play a known role in development. For example, the family of corn cystatins is known to play a role in development, and more specifically, CC5, which was found to be hypermethylated, is induced under drought stress (Massonneau et al., 2005). Also MiR396a influences development, by targeting the destruction of the GRF genes (Debernardi et al., 2012). Another interesting gene is PHOSPHATIDYLOSIOTOL SYNTHASE 2, which underwent two hypomethylation events in the older zones. Recently, it was found that overexpression of ZmPIS in maize enhanced ABA-mediated drought stress tolerance (Liu et al., 2013).

It is therefore tempting to speculate that DNA methylation regulates aforementioned processes in the developmental framework of the growing maize leaf. However, our data show that especially differential methylation affecting the 5' region of the gene is correlated with gene expression, whereas this is not the case for the remainder of the gene. This is especially striking in the case of the MiR396a-encoding gene, which is specifically upregulated in the TZ, and again downregulated in the EZ. Hypomethylation of the CCGG site within the gene body, however, occurs between EZ and MZ. Exactly how these methylations affect gene function remains unknown. They might be involved in processes other than gene expression as such, or counteracted by other epigenetic marks, such as histone modification.

**CONCLUSION**

It is already known that maize leaf growth is strongly regulated by fine-tuning the development of the different zones in the maize leaf. Especially the size of the DZ is an important factor affecting growth and this is regulated through differential gene expression and precise hormone deposition (Rymen et al., 2007; Nelissen et al., 2012). Here, we present that the consecutive developmental zones of the growing maize leaf are correlated to differential DNA methylation. Especially when occurring upstream of the gene start and the 5’ end of the gene, differential DNA methylation is associated with expression changes of genes which need to be expressed in a specific developmental context.

**MATERIALS AND METHODS**

**Plant Material and Growth Conditions**

Wild-type maize (Zea mays L.) cv. B73 plants were grown in soil (Saniflor - NV Van Israel) in a growth room under controlled conditions: photoperiod (16/8-h light/dark), temperature (25.5°C), relative
humidity (50%), and light intensity 170 mmol m$^{-2}$ s$^{-1}$ photosynthetically active radiation at plant level. The growth room was illuminated by a combination of high-pressure sodium vapor lamps (RNP-T/LR/400W/S/230/E40; Radium) and metal halide lamps with quartz burners (HRI-BT/400W/D230/E40; Radium).

**Growth, Cellular Analysis and Sampling**

Plants were measured daily to determine leaf elongation rates of leaf 4 (n=20). For experimental manipulations, leaf 4 was harvested 2 d after appearance (12 days after sowing, V-stage 2), at which point the ligule is only a few mm from the base, during steady-state leaf growth. For cellular analysis, the first 10 cm of the leaf was cut into 1 cm pieces. These samples were cleared overnight with ethanol:acetic acid (3:1). Epidermal cell length profiles of cell files along the proximal-distal axis were established with a DIC microscope (AxioImager; Zeiss) and image analysis software (AxioVision; Zeiss) on samples fixed using lactic acid (n=3). The size of the DZ was determined as the distance between the leaf base and the most distally located mitotic figure in DAPI-stained leaves with a fluorescence microscope (n=4). For MSAP analysis, the first, second, fourth and tenth cm were harvested as representatives of the DZ, TZ, EZ, and MZ. For expression analysis, the most basal 4 cm of the maize leaf was harvested in 5 mm pieces, whereas the following 6 cm were harvested in 1 cm pieces. Samples for MSAP and expression analysis were harvested and flash-frozen in liquid N$_2$.

**Nucleotide Isolation**

Total DNA for MSAP analysis was isolated form individual samples with the CTAB method. This DNA was quantified using the quantIT™ dsDNA High-sensitivity Assay Kit (Invitrogen) and LUMIstar Galaxy (BMG Labtech) according to the manufacturer’s instructions. Total RNA for expression analysis was isolated from pooled samples (5 plants) with the guanidinium thiocyanate-phenol-chloroform extraction method using TRI-reagent (Sigma-Aldrich). First-strand cDNA was synthesized from 1 µg of total RNA with the iScript kit (Bio-Rad Laboratories) according to the manufacturer’s instructions.

**MSAP Analysis of DNA Methylation**

The maize methylome was screened using the MSAP protocol (Xiong et al., 1999). $Msp$I and $Hpa$II are used as frequent isoschizomeric restriction enzymes with differential sensitivity towards CCGG methylation, each in combination with $Eco$RI as a cytosine methylation-insensitive restriction enzyme (Supplemental Table S8). Briefly, 390 ng of genomic DNA underwent restriction-ligation (R-L) using $Eco$RI-$Hpa$II/$Msp$I and $Eco$RI (E) - and $Hpa$II/$Msp$I (HM)-specific adapters (Supplemental Table S9) at 37°C for 4 h, and diluted 10-fold. The R-L mix was template for pre-amplification PCR, using primers with 1 selective nucleotide, and diluted 20-fold. This dilution was template for selective amplification, using primers with 3 selective nucleotides, in which the $Eco$RI primers were labeled with [$\gamma$-3$^3$P] ATP. For the first screen, 65 selective primers were applied in all 64 possible combinations (E+AAC x HM+NNN) to two biological replicates of DZ, TZ, EZ, and MZ. For the second screen, 68 selective primers (E+AAA/G/T and E+ACA x HM+NNN) were used in all 256 possible combinations on pools of DZ and TZ samples. Hence, for DZ-TZ, 320 primer combinations were tested. Samples (3 µl) were loaded onto a 4.5% denaturing polyacrylamide gel, and electrophoresis was performed at
constant power of 100 W. Labeled bands were blotted onto Whatman™ 3MM blotting paper and visualized on Amersham Hyperfilm ECM (GE Healthcare).

Identification of Differentially Methylated Loci

Polymorphic banding patterns, representing differential methylation, were identified by the eye. One representative band was cut from the blotting paper using a razor blade, resuspended in ddH2O, re-amplified with Pwo SuperYield DNA Polymerase (Roche) and sequenced using the E-specific primer. Successfully sequenced loci were BLASTed to the maize genome (www.maizesequence.com, currently replaced by ensembl.gramene.org/Zea_mays), and the differentially methylated CCGG site identified as the CCGG site at the 3’ end of the sequence or the closest CCGG site. When more than one genomic site was identified, the sequence was BLASTed to the maize transposable element database (maizetedb.org) in order to determine if the underlying sequence was transposon-derived, and which type it was. If only one locus was identified, and not found to overlap with a gene, we scanned 5 kb up- and downstream for the presence of one. The largest transcript was chosen to represent each gene, and the position of each single locus CCGG site was calculated with respect to both the transcription and translational start and stop sites of this transcript.

Transcriptional Analysis

The maize CDKB1;1 has been identified earlier as a marker for mitotic activity (Rymen et al., 2007). All remaining primers (Supplemental Table S10) were designed with the Beacon Designer 4.0 software and the default settings (Premier Biosoft international). The transcripts were quantified using quantitative reverse transcription PCR with a Lightcycler 480 (Roche Applied Science) and SYBR Green I Master Kit (Roche Applied Science). PCR reactions were done in triple technical replicates, according to the manufacturer’s guidelines. Melting curves were generated to check primer specificity, and relative quantification was carried out using the ΔCt method, with 18S rRNA (5’-ACCTTACCAGCCCTTGACATATG-3’ and 5’-GACTTGACCAAACATCTCACGAC-3’) as housekeeping gene. Three biological replicates were averaged for analysis.

Statistical Analysis

Distribution of CCGG sites within the gene body and occurrence of transposons in the dataset were statistically analyzed using a Pearson’s χ² test. Transcript profiles were statistically evaluated based on an ANOVA for the factor position in R (www.r-project.org). P-values for distances between DML on the maize chromosomes were calculated using a two-sample T-test assuming unequal variance in Excel. All error bars depict standard error data.

Supplemental Data

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

Supplemental Table S1. The DMT-encoding genes, and alternative names.

Supplemental Table S2. MSAP patterns and possible DNA methylation states.

Supplemental Table S3. Relative abundance of differential methylation patterns.

Supplemental Table S4. Differentially methylated TEs in the multi-locus dataset.
Supplemental Table S5. Abundance of 20 most prevalent maize TEs.
Supplemental Figure S1. Distribution of differentially methylated single-locus sequences across the chromosomal arms and between different chromosomal sites.
Supplemental Figure S2. Location of differentially methylated sites in or close to coding sequences.
Supplemental Table S6. Single locus differentially methylated sequences, identified between DZ and TZ.
Supplemental Table S7. Single-locus differentially methylated sequences, identified between TZ and EZ, EZ and MZ, and between multiple zones.
Supplemental Figure S3. Expression of the maize MiR396a and GROWTH REGULATORY FACTOR (GRF) encoding genes.
Supplemental Figure S4. Expression of genes undergoing differential methylation in the promoter and 5’ region of gene body.
Supplemental Figure S5. Expression of genes undergoing differential methylation in the central part of the gene body.
Supplemental Figure S6. Relative expression of genes undergoing differential methylation downstream of the coding sequence and gradual methylation change in the gene body.
Supplemental Table S8. DNA methylation sensitivity of MSAP restriction enzymes (REs).
Supplemental Table S9. Primers used for MSAP analysis.
Supplemental Table S10. Primers used for q-RT-PCR analysis.

ACKNOWLEDGMENTS
The authors thank the Systems Biology of Yield group for the stimulating scientific environment and fruitful discussions, Dr. Annick Bleys for help in preparing the manuscript, and Wilson Ardiles Diaz for help in sequencing the MSAP bands.

LITERATURE CITED


FIGURE LEGENDS

Figure 1. Phylogeny of the DNA methyltransferase domain enzymes of maize (Zm), Arabidopsis (At) and humans (Hs). Three types of methyltransferase enzymes can be distinguished: maintenance (green), de novo (blue) and DNMT2-like (red). The maize proteins are indicated in bold. Maintenance methyltransferase enzymes fall into two categories: the DMTs that methylate the CG motif (dark green) and are conserved between animals (DNMT1) and plants (MET), and the plant-specific DMTs that methylate the CHG motif (CMTs, light green). ZmMET1 and ZmMET2 are 99.4% identical and are therefore indicated as one branch. Animal and plant de novo methyltransferase differ in the arrangement of the methylase domains, causing a difference in target motifs: CG in animals (DNMT3, light blue), CHH in plants (DRM, dark blue). The DNMT2 lineage encodes a DNA methyltransferase domain, but are most likely RNA methyltransferase enzymes, as animal DNMT2 was found to methylate aspartic acid tRNA, and was renamed TRDMT1. See online article for color version of this figure.

Figure 2. Cellular profile and DNA methyltransferase expression along the maize leaf.
(A) Cellular behavior throughout the first ten centimeters of maize leaf 4, two days after its emergence, with the first centimeter consisting of small, dividing cells (DZ), followed by a transition (TZ) toward elongating tissue (EZ), and finally matured cells (MZ). The samples for MSAP representing aforementioned tissue are indicated on top (black braces), as is the end of the DZ (arrowhead). DNA methyltransferase expression was checked across the whole maize leaf.
(B and C) Maize maintenance methyltransferase (ZmMET, ZmCMT1 and ZmCMT2) expression (B) is highest in the dividing cells at the base of the maize leaf. The expression sharply decreases until reaching minimal levels 3 to 4 cm from the leaf base. A very analogous expression pattern is found for CDKB1;1, the expression of which is highly correlated with cell division activity (Rymen et al., 2007; Nelissen et al., 2012). The de novo methyltransferases (C) ZmDRM1 and ZmDRM2, are expressed steadily across the zones, as is ZmDNMT2. ZmDRML, on the other hand, has a similar expression profile as the maintenance DMTs.

Figure 3. Example and explanation of MSAP gel banding pattern.
(A) MSAP gel comparing pools of DZ and TZ samples. Each sample is restricted by EcoRI+MspI (EM) and EcoRI+HpaII (EH). Both stably (I, II and III) and differentially (1: I→III, 2: 0→III) methylated sites are represented, of which the explanation is provided in (B).
(B) I, non-methylated CCGG; II, CHG methylation, where H = C; III, CG methylation; 0, heavy CCGG methylation. See Supp Table S2 for additional information. Briefly, CCGG means either methylation of both cytosines (CCGG) or only the outer cytosine (CCGG). Differential methylation is indicated by arrows: hypomethylation (red), hypermethylation (green), and both hypo- and hypermethylation (blue). A change from 0 to any other MSAP pattern from one zone to the next means hypomethylation. Similarly, a change from I to any other MSAP pattern signifies hypermethylation. Only a change from II to III and III to II can be interpreted as both hypo- and hypermethylation of the locus. The overall occurrence of each transition (sum of both hypo- and hypermethylation percentages) is represented. The sequence context affected, being either CG or CHG, is indicated at the edges. See online article for color version of this figure.

Figure 4. Identification of differentially methylated sequences. Differentially methylated bands were amplified and sequenced. Successful sequences were BLASTed against the maize B73 genome (version 5b.60). Sequences mapping to a single location (single site) do either map in or in close proximity of a gene (genic) or not (intergenic). Genic sequences map either to a transcript-coding region (gene body) or up/downstream from it (5'/3'). Most gene body methylation was found in exons of protein-coding genes. Sequences mapping to multiple sites in the maize genome, were either transposable elements (TEs), mapping to numerous sites in the maize genome, or mapped to a limited number of sites in the genome (non-TE). Most TEs are type I (retro)transposons, and about half of the oligo-mapping sequences had at least one copy on the plastid of mitochondrial genome (organellar).

Figure 5. Mapping of differentially methylated single locus sequences. All single-locus differentially methylated sequences that mapped to the ten maize chromosomes are represented. Hypermethylated sequences are indicated in green, hypomethylated ones in red and sequences that undergo both hypo- and hypermethylation are indicated in blue. Hits that are not in the vicinity of coding regions are indicated as non-coding (nc) in grey. Pseudogenes and TEs are also indicated in grey. Genes that have a known function, homology or discernible domain, are indicated as such, the other genes are indicated by their gene code. If the differentially methylated locus lies in the vicinity of two protein-coding genes, both are mentioned. Centromeres and pericentromeric regions are indicated in black and dark grey, respectively.
Figure 6. Location of genic hits in respect to the gene body. The position of all genic hits was plotted up to 5 kb up- or downstream of the gene. To adjust for different gene length and presence of introns, the genic position is expressed as a percentage of the gene body length, with 0% and 100% representing the start and end of the transcript, respectively. A canonical gene model is represented above the graph, with exons, introns and untranslated regions indicated in black, white and dark grey, respectively.

Figure 7. Expression of genes undergoing differential methylation in the 5’ region (A) and in the remainder of the gene body (B).

(A) Relative expression of genes undergoing differential methylation in the promoter or 5’ region of gene body. The CCGG sites of EXPB3, TGA10 and CYCD4 are found in the promoter, whereas for DUF547 it is located immediately downstream of the start codon (ATGOCGG). (B) Relative expression of genes undergoing differential methylation in the remainder of the gene body. The CCGG sites of MTM and AGO are found in an exon, and of GATA in an intron. The PK/UbiC-encoding gene has two differentially methylated CCGG sites, found in an exon and the following intron. Zones between which methylation changes take place (DZ-TZ, TZ-EZ or EZ-MZ) are indicated by a black arrow. The location of differential methylation with respect to the gene is represented by a black arrowhead on the gene model. If other possible differential methylation CCGG sites are present, they are indicated by a grey arrowhead. *, p = 0.01 – 0.05; **, p = 0.001 – 0.01; ***, p = <0.001.

Figure 8. Summary of qPCR expression data in relation to differential methylation. Correlation between differential methylation and expression is indicated as colored arrows. A green arrow indicates that a methylation change at this side led to up- or downregulation of a profiled gene, after it was hypo- or hypermethylated, respectively. A light green color means the differential expression was not found to be significant at the 5%-level. A red arrow indicates an opposing effect: hypomethylation leading to downregulation or hypermethylation leading to upregulation. A blue arrow means the gene was not found to be differentially expressed and the grey arrows indicate genes for which no expression in the sampled zones of the maize leaf could be found. Arrows that are linked represent methylation changes within the same gene. The gene is represented, from start- to stop codon, as a rectangle with exons in black and introns in white. The side of the start codon is indicated as an arrow.
Figure 1. Phylogeny of the DNA methyltransferase domain enzymes of maize (Zm), Arabidopsis (At) and humans (Hs). Three types of methyltransferase enzymes can be distinguished: maintenance (green), de novo (blue) and DNMT2-like (red). The maize proteins are indicated in bold. Maintenance methyltransferase enzymes fall into two categories: the DMTs that methylate the CG motif (dark green) and are conserved between animals (DNMT1) and plants (MET), and the plant-specific DMTs that methylate the CHG motif (CMTs, light green). ZmMET1 and ZmMET2 are 99.4% identical and are therefore indicated as one branch. Animal and plant de novo methyltransferase differ in the arrangement of the methylase domains, causing a difference in target motifs: CG in animals (DNMT3, light blue), CHH in plants (DRM, dark blue). The DNMT2 lineage encodes a DNA methyltransferase domain, but are most likely RNA methyltransferase enzymes, as animal DNMT2 was found to methylate aspartic acid tRNA, and was renamed TRDMT1. See online article for color version of this figure.
Figure 2. Cellular profile and DNA methyltransferase expression along the maize leaf.

(A) Cellular behavior throughout the first ten centimeters of maize leaf 4, two days after its emergence, with the first centimeter consisting of small, dividing cells (DZ), followed by a transition (TZ) toward elongating tissue (EZ), and finally matured cells (MZ). The samples for MSAP representing aforementioned tissue are indicated on top (black braces), as is the end of the DZ (arrowhead). DNA methyltransferase expression was checked across the whole maize leaf.

(B and C) Maize maintenance methyltransferase (ZmMET, ZmCMT1 and ZmCMT2) expression (B) is highest in the dividing cells at the base of the maize leaf. The expression sharply decreases until reaching minimal levels 3 to 4 cm from the leaf base. A very analogous expression pattern is found for CDKB1;1, the expression of which is highly correlated with cell division activity (Rymen et al., 2007; Nelissen et al., 2012). The de novo methyltransferases (C) ZmDRM1 and ZmDRM2, are expressed steadily across the zones, as is ZmDNMT2. ZmDRML, on the other hand, has a similar expression profile as the maintenance DMTs.
Figure 3. Example and explanation of MSAP gel banding pattern.

(A) MSAP gel comparing pools of DZ and TZ samples. Each sample is restricted by EcoRI+MspI (EM) and EcoRI+HpaII (EH). Both stably (I, II and III) and differentially (1→III, 2→III) methylated sites are represented, of which the explanation is provided in (B).

(B) I, non-methylated CCGG; II, CHG methylation, where H = C; III, CG methylation; 0, heavy CCGG methylation. See Supplemental Table S2 for additional information. Briefly, CCGG means either methylation of both cytosines (CCGG) or only the outer cytosine (CCGG). Differential methylation is indicated by arrows: hypomethylation (red), hypermethylation (green), and both hypo- and hypermethylation (blue). A change from 0 to any other MSAP pattern from one zone to the next means hypomethylation. Similarly, a change from I to any other MSAP pattern signifies hypermethylation. Only a change from II to III and III to II can be interpreted as both hypo- and hypermethylation of the locus. The overall occurrence of each transition (sum of both hypo- and hypermethylation percentages) is represented. The sequence context affected, being either CG or CHG, is indicated at the edges. See online article for color version of this figure.
Figure 4. Identification of differentially methylated sequences. Differentially methylated bands were amplified and sequenced. Successful sequences were BLASTed against the maize B73 genome (version 5b.60). Sequences mapping to a single location (single site) do either map in or in close proximity of a gene (genic) or not (intergenic). Genic sequences map either to a transcript-coding region (gene body) or up/downstream from it (5'/3’). Most gene body methylation was found in exons of protein-coding genes. Sequences mapping to multiple sites in the maize genome, were either transposable elements (TEs), mapping to numerous sites in the maize genome, or mapped to a limited number of sites in the genome (non-TE). Most TEs are type I (retro)transposons, and about half of the oligo-mapping sequences had at least one copy on the plastid of mitochondrial genome (organellar).
**Figure 5.** Mapping of differentially methylated single locus sequences. All single-locus differentially methylated sequences that mapped to the ten maize chromosomes are represented. Hypermethylated sequences are indicated in green, hypomethylated ones in red and sequences that undergo both hypo- and hypermethylation are indicated in blue. Hits that are not in the vicinity of coding regions are indicated as non-coding (nc) in grey. Pseudogenes and TEs are also indicated in grey. Genes that have a known function, homology or discernible domain, are indicated as such, the other genes are indicated by their gene code. If the differentially methylated locus lies in the vicinity of two protein-coding genes, both are mentioned. Centromeres and pericentromeric regions are indicated in black and dark grey, respectively.
Figure 6. Location of genic hits in respect to the gene body. The position of all genic hits was plotted up to 5 kb up- or downstream of the gene. To adjust for different gene length and presence of introns, the genic position is expressed as a percentage of the gene body length, with 0% and 100% representing the start and end of the transcript, respectively. A canonical gene model is represented above the graph, with exons, introns and untranslated regions indicated in black, white and dark grey, respectively.
**Figure 7.** Expression of genes undergoing differential methylation in the 5' region (A) and in the remainder of the gene body (B).

(A) Relative expression of genes undergoing differential methylation in the promoter or 5' region of gene body. The CCGG sites of *EXPB3*, *TGA10*, and *CYCD4* are found in the promoter, whereas for *DUF547* it is located immediately downstream of the start codon (ATGCCGG).

(B) Relative expression of genes undergoing differential methylation in the remainder of the gene body. The CCGG sites of *MTM* and *AGO* are found in an exon, and of *GATA* in an intron. The *PK/UbiC*-encoding gene has two differentially methylated CCGG sites, found in an exon and the following intron. Zones between which methylation changes take place (DZ-TZ, TZ-EZ or EZ-MZ) are indicated by a black arrow. The location of differential methylation with respect to the gene is represented by a black arrowhead on the gene model. If other possible differential methylation CCGG sites are present, they are indicated by a grey arrowhead.

* *, p = 0.01 – 0.05; ***, p = 0.001 – 0.01; ***, p = <0.001.
Figure 8. Summary of qPCR expression data in relation to differential methylation. Correlation between differential methylation and expression is indicated as colored arrows. A green arrow indicates that a methylation change at this side led to up- or downregulation of a profiled gene, after it was hypo- or hypermethylated, respectively. A light green color means the differential expression was not found to be significant at the 5%-level. A red arrow indicates an opposing effect: hypomethylation leading to downregulation or hypermethylation leading to upregulation. A blue arrow means the gene was not found to be differentially expressed and the grey arrows indicate genes for which no expression in the sampled zones of the maize leaf could be found. Arrows that are linked represent methylation changes within the same gene. The gene is represented, from start- to stop codon, as a rectangle with exons in black and introns in white. The side of the start codon is indicated as an arrow.