The role of histone modifications in gene regulation: from C4 genes to a genome-wide exploration along the maize leaf gradient

Von der Naturwissenschaftlichen Fakultät der
Gottfried Wilhelm Leibniz Universität Hannover

Zur Erlangung des Grades
Doktor der Naturwissenschaften (Dr. rer. nat.)

genehmigte Dissertation
von
M. Sc. Renke Perduns

geboren am 03.06.1984 in Wittmund

2016

Referent: Prof. Dr. Christoph Peterhänsel

Korreferent: Prof. Dr. Thomas Debener

Tag der Promotion: 15.12.2015

Abstract

Histone modifications contribute to gene regulation in eukaryotes. Previous studies investigated how histone modifications support the integration of multiple signals into one transcriptional output at the promoter of the phosphoenolpyruvate carboxylase (*Pepc*) gene of *Zea mays*. Two distinct sites were identified that were differentially modified in response to endogenous and exogenous stimuli: the transcription initiation site (TIS) and the upstream promoter region. Based on these studies, a histone code for *Pepc* was proposed.

The first aim of the present study was to investigate the conservation of this histone code at other genes. It was shown that the histone code of *Pepc* was also observed at other C4 genes in maize as well as at the promoters of *Pepc* and malic enzyme (*Me*) in the related monocotyledonous plants *Sorghum bicolor* and *Setaria italica*. It was concluded that the observed mechanism was potentially recruited from ancestral C3 plants to regulate the expression of C4 genes.

In the second part of this work, the genome-wide distribution of two histone modifications (histone 3 lysine 9 acetyl [H3K9ac] and histone 3 lysine 4 trimethyl [H3K4me3]) was compared between two developmental zones of the maize leaf (base and blade). This allowed for the analysis of dynamic changes in H3K9ac and H3K4me3 compared to static analyses at constant conditions in previous studies. Comparative analyses of the dynamics of both histone modifications and the transcription of associated genes revealed that the fold changes of both histone modifications correlated well with the fold changes of the transcription of associated genes. When focusing on regulated H3K9ac enriched regions (peaks), we identified regulated secondary upstream peaks (R-SUPs) as a new feature of H3K9ac. R-SUPs were enriched at genes that were induced in the blade compared to the base and especially on photosynthetic genes. Meta-analyses revealed that R-SUP associated genes were enriched within potential C4 genes that were identified by other groups. It was concluded that R-SUPs could be a consequence of the adaption of the regulation of ancestral C3 genes during the evolution of C4 photosynthesis.

In the last part of this thesis, the influence of light on gene regulation from the base to the blade of the maize leaf was analyzed. The regulation of genes along the leaf gradient was classified into light-dependent and light-independent regulation. It was shown that light is an important factor for the regulation of genes that were induced in the blade while the majority of repressed genes were regulated light-independently. A meta-analysis showed that genes from the defined classes were equally co-regulated with H3K9ac, underlining the general role of H3K9ac in gene regulation.

Keywords: Histone modification, gene regulation, C4 photosynthesis

Zusammenfassung

Histonmodifikationen sind an der Genregulation in Eukaryoten beteiligt. In vorangegangen Studien der Gruppe wurde untersucht, wie Histonmodifikationen an der Integration vielfältiger Reize in eine Transkriptionsantwort des Phosphoenolpyruvatcarboxylase (*Pepc*) Gens von *Zea mays* mitwirken. Es wurden zwei Promotorregionen identifiziert, die in Abhängigkeit von internen und externen Reizen differentiell modifiziert wurden: am Startpunkt der Transkription und an einer vorgeschalteten (distalen) Position des Promotors. Aus diesen Studien wurde ein Histonmodifikationsmuster für *Pepc* abgeleitet.

In der vorliegenden Arbeit wurde die Allgemeingültigkeit dieses Musters untersucht. Es konnte gezeigt werden, dass das Histonmodifikationsmuster von *Pepc* sowohl an anderen C4 Genen in Mais als auch an den Promotoren von *Pepc* und dem Malatenzym (*Me*) in den verwandten monokotylen Pflanzen *Sorghum bicolor* und *Setaria italica* zu finden war. Es wurde geschlussfolgert, dass dieses Muster von gemeinsamen Vorfahren der C4 Gräser übernommen wurde, um die Transkription von C4 Genen zu regulieren.

Im zweiten Teil der vorliegenden Arbeit wurde die genomweite Verteilung von zwei Histonmodifikationen (Histon 3 Lysin 9 Acetyl [H3K9ac] und Histon 3 Lysin 4 Trimethyl [H3K4me3]) zwischen zwei Entwicklungszonen des Maisblattes (Basis und Lamina) verglichen. Dies ermöglichte eine Analyse der dynamischen Veränderung beider Histonmodifikationen im Gegensatz zu vorangegangenen Analysen in statischen Situationen. Korrelationsanalysen zwischen den Histonmodifikationen und der Transkription assoziierter Gene zeigten, dass die Änderungen der Histonmodifikationen mit den Änderungen der Transkription assoziierter Gene korrelierten. Der Fokus auf H3K9ac-modifizierte Bereiche (Peaks), die differentiell zwischen Basis und Lamina reguliert waren, führte zur erstmaligen Identifikation von sekundären, regulierten distalen Peaks (R-SUPs) der Histonmodifikation H3K9ac. R-SUPs wurden vermehrt an Genen gefunden, die von der Basis zur Lamina induziert wurden; dabei insbesondere an photosynthetischen Genen. Durch eine Metaanalyse konnte zudem gezeigt werden, dass R-SUP assoziierte Gene unter C4 Kandidatengenen, die von anderen Gruppen zuvor identifiziert worden waren, angereichert waren. Basierend auf diesen Ergebnissen wurde diskutiert, ob R-SUPs während der Evolution von C4 Genen aus deren C3 Vorläufern entstanden sein könnten.

Im letzten Teil der Arbeit wurde der Einfluss des Lichts auf die Genregulation zwischen Basis und Lamina des Maisblattes untersucht. Diese Regulation wurde in lichtabhängig und lichtunabhängig eingeteilt. Es wurde gezeigt, dass Licht ein entscheidendes Signal für die Regulation von Genen war die in der Lamina induziert wurden, während der Großteil der reprimierten Gene lichtunabhängig reguliert wurde. Sowohl lichtabhängige als auch lichtunabhängige Gene zeigten Ko-Regulation von Transkription und H3K9ac. Dies bestätigt die grundlegende Rolle von H3K9ac in der Genregulation.

Schlagwörter: Histonmodifikationen, Genregulation, C4 Photosynthese

Table of contents

Appendix

Chapter 1 - General introduction

1.1 Chromatin

Eukaryotes have to compress their genome in some way to pack the large amount of DNA within their nuclei. For example, the diploid genome of the sequenced cultivar of maize (B73) consists of about 2,3 x 10^9 base pairs (Schnable et al., 2009). The whole genome would measure 1.7 m in an uncompressed form. To store this huge amount of information in the nucleus of every single cell the DNA is associated with certain proteins and tightly packed by physical interactions between these DNA-protein complexes. The resulting structure is named chromatin. The basic building blocks of the chromatin are nucleosomes (Kornberg, 1977). Each nucleosome consists of around 147 base pairs of DNA wrapped around an octamer of each two of the histone proteins H2A, H2B H3 and H4. The histone protein H1 acts as linker to stabilize the interaction of DNA with the octameric histone complex. The single nucleosomes are linked by a DNA stretch with varying length (Luger et al., 1997; Kouzarides, 2007). The nucleosomes are further compressed in a highly organized manner and build up the chromosomes in the end, which are visible under the light microscope. The structure of chromatin is highly dynamic and massively influences essential DNA-associated processes like replication, DNA repair, recombination and especially regulation of gene expression (Groth et al., 2007; Li et al., 2007). Chromatin associated with active or inducible genes is called euchromatin and associated with silenced genes heterochromatin. These states of chromatin are partially variable and the transition is fluid to adapt to the requirements of different developmental stages or changing conditions and needs of the individual cell (Shu et al., 2012).

1.1.1 Histone modifications and their role in the regulation of gene expression

Nucleosomes are mainly globular structures. However, the N-terminal tails of histone molecules are exposed and can be post-translationally modified to regulate the interaction between the histone proteins and the DNA. The alterations of the interaction are used in eukaryotes to regulate essential processes like replication, DNA-repair, condensation and transcription of associated genes. Examples for such histone modifications are methylation, acetylation, ubiquitination, ADP-ribosylation, and sumoylation of lysines, but also other amino acids are modified (Kouzarides, 2007; Li et al., 2007). One way of regulation is the recruitment of chromatin remodeling complexes that rearrange the nucleosomes in an ATP-

dependent manner and open the chromatin structure to facilitate DNA-binding of different enzymes (van Attikum et al., 2004; Wysocka et al., 2006).

In plants, especially the regulation of transcription by lysine (K) methylation and acetylation of histone 3 (H3) and 4 (H4) has been studied comprehensively. So far, 8 lysine residues of H3 and H4 are known that are targeted by methylation (me) and 10 that are targeted by acetylation (ac). 6 of these residues are targeted by both ac and me (Benhamed et al., 2006; Chen and Tian, 2007). The complexity even increases because lysine methylation can occur as mono-, di, or tri methylation (me1, me2 and me3). Acetylation is widely associated with active genes while methylation can also be found associated with inactive genes dependent on the modified K residue and the degree of methylation (Wang et al., 2009; Lauria and Rossi, 2011). Two models exist that explain the regulatory influence of histone modifications on the transcription of associated genes: the histone code model and the charge neutralization model. Histones have a positive net charge due to their high content of lysine and arginine. The incorporation of acetyl groups reduces the positive charge which leads to a diminished interaction with the negatively charged DNA. This can facilitate the binding of DNA binding proteins like transcription factors and RNA polymerases to activate transcription in an unspecific manner (Dion et al., 2005; Wang et al., 2009). However, K methylation does not change the net charge of histones, which shows that the charge neutralization model alone is not sufficient to explain the regulatory effects of histone modifications completely. The histone code model argues that histone modifications are directly recognized by certain factors to guide DNA binding proteins to their target sequences. For example, it was shown in yeast that bromodomains are able to recognize acetylated histones and might act as targeting modules for certain proteins (Hassan et al., 2007). Also a specific interaction of the TATA box binding-protein associated factor 3 with H3K4me3 to initiate transcription was observed in human cell lines (Vermeulen et al., 2007; Lauberth et al., 2013). Data for both models exist and potentially both models describe aspects of the integration of diverse signals into one transcriptional output.

1.1.2 Genome-wide analysis of histone modifications

Genome-wide analyses of histone modifications offer the opportunity to study the association of certain histone modifications with distinct regions of the genome on a global level. First attempts were carried out with tilling arrays. The tilling array is a subtype of microarrays using probes that cover specific regions of the genomic DNA. Labeled DNA fragments that are chromatin immunoprecipitated (ChIP) with antibodies against certain histone modifications are hybridized and quantified. In recent years, next generation sequencing

techniques have become the method of choice for genome-wide analysis of histone modifications (ChIP-Seq). The immunoprecipitated DNA is extensively sequenced and the sequencing reads are afterwards mapped to the reference genome. Complex algorithms are used to analyze genomic regions for the enrichment of analyzed histone modifications. ChIP-Seq enables genome-wide analyses of the association of histone modifications that is not biased by the nature of the spotted probes (Ho et al., 2011). Both methods have been used and correlated with genome-wide expression data to study the characteristics of various histone modifications and their impact on gene expression in plants (Ha, 2013).

Genome-wide profiles of histone modifications were comprehensively studied in *Arabidopsis* (Zhang et al., 2007; Zhang et al., 2009; Zhou et al., 2010; Roudier et al., 2011). The histone modifications H3K4me1, H3K4me2, H3K4me3, H3K9me3, H3K9ac, H3K27me3, H3K36me3 and H3K56ac were mostly found in euchromatic regions, indicating that they are associated with genes rather than transposable elements (TEs) and other repetitive and silenced regions that are located in heterochromatic regions. On the contrary, H4K20me1 and H3K9me2 were mainly found in heterochromatic regions and associated with silenced TEs. H3K27me1 and H3K27me2 were found in both hetero- and euchromatic regions. Most of the histone modifications that were associated with genes were mainly enriched at the 5´ end of gene bodies. Especially H3K4me3, H3K9ac and H3K56ac sharply peaked at the transcription initiation sites (TIS). H3K27me2 and H2K27me3 were evenly distributed and H3K27me2 was additionally located in the flanking regions of genes. The association of histone modifications with genes in euchromatic regions indicates that they might contribute to the regulation of gene expression. To test this hypothesis, the correlation between the enrichment of histone modifications and the transcription of associated genes was evaluated. The intensities of H3K9ac, H3K56ac, H3K4me3, H3K9me3 and H3K36me3 increased with higher transcription rates whereas the histone modifications H3K4me1, H3K4me2, H3K27me1, H3K27me2 and H3K27me3 showed decreasing intensities. Such correlations alone are not sufficient to define certain histone modifications as general “on” or “off” marker for associated genes. The mechanistic connection between each histone modification and the transcription of associated genes has to be evaluated in detail to understand the impact on the regulation of gene expression. Genome-wide analyses of histone modifications were also carried out in maize and rice. The described characteristics of H3K4me3, H3K9ac, H3K9me3, H3K27me3 and H3K36me3 were also observed in maize. H3K4me2, H3K4me3 and H3K9ac also showed equal characteristics in rice, indicating a deep evolutionary conservation of the functions (Wang et al., 2009; Du et al., 2013; West et al., 2014).

Most genome-wide analyses were used to study histone modifications in a static situation. The intensities of histone modifications were correlated with the gene expression of genes

with different transcriptional levels. Up to now, only two studies in the plant research field analyzed the dynamics of histone modifications from individual genes between two situations and correlated fold changes of the transcription with the fold changes of associated histone modifications. Brusslan et al. (2012) analyzed the dynamics of H3K4me3 and H3K27me3 between mature and senescent leaves of *Arabidopsis*. A positive correlation for H3K4me3 and a negative correlation for H3K27me3 with the transcription of associated genes were observed. He et al. (2013) studied the dynamics of H3K4me3 and H3K9ac between shoots and roots of maize more comprehensively. 70% to 90% of the genes that were associated with differential H3K4me3 and H3K9ac between shoots and roots were transcriptionally co-regulated. The fold changes of the histone modifications correlated well with the fold changes of the transcription of associated genes. Additional analyses and also alternative normalization approaches are needed to extend the knowledge about the dynamics of histone modifications in a genome-wide context.

1.2 C4 Photosynthesis

Plants are photoautotrophic organisms that fix atmospheric CO_2 into biomass using light as energy source in a process called photosynthesis. The photosynthetic pathway can be divided into two parts: light-dependent and light-independent reactions. Light-dependent reactions are characterized by the absorption of light energy to generate the energy and reducing equivalents adenosine triphosphate (ATP) and reduced nicotinamide adenine dinucleotide phosphate (NADPH). These equivalents are used in the light-independent reaction (Calvin cycle) to fix inorganic CO_2 into organic compounds like sugars. The enzyme that catalyzes the fixation of CO_2 is ribulose-1,5-bisphosphate carboxylase-oxygenase (RubisCO), which is the most abundant protein on our planet. RubisCO also has an oxygenase activity that results in the formation of toxic 2-phosphoglycolate that is detoxified in an energy consuming process called photorespiration (Peterhansel et al., 2010). In dry and warm climates, photorespiration can lower the photosynthetic efficiency up to 40% (Ehleringer, JR et al., 1997). Plants developed different strategies to reduce photorespiration. The most successful approach is C4 photosynthesis. In C4 photosynthesis, CO_2 is concentrated around RubisCO to lower the oxygenase activity. C4 photosynthesis has evolved independently in at least 66 lineages within the past 35 million years (Sage et al., 2012). Three independent sub types of C4 photosynthesis exist that were defined based on different pathways to release CO_2 in the bundle sheath cells. The economical most important C4 crops are monocotyledonous grasses like maize, *Sorghum bicolor* and *Setaria italica* that all perform NADP-malic enzyme (NADP-ME) C4 photosynthesis (Christin et al., 2009).

1.2.1 C4 photosynthesis in maize

C4 photosynthesis is characterized by the spatial separation of the initial CO_2 fixation, independent of RubisCO, and the Calvin cycle into mesophyll (M) and bundle sheath (BS) cells. A schematic overview of the NADP-ME C4 photosynthesis in maize is shown in Figure 1-1. CO_2 diffuses into M cells and is converted to bicarbonate which is fixed by phosphoenolpyruvate carboxylase (PEPC) into the C4 intermediate oxalacetate (OAA). OAA is reduced to malate in the chloroplasts by NADP-dependent malate dehydrogenase (MDH). Malate enters the BS cells through plasmodesmata and is decarboxylated to pyruvate by NADP-malic enzyme (ME) in the chloroplasts. The released CO_2 enters the Calvin cycle while pyruvate diffuses back into the chloroplasts of M cells where it is regenerated to the bicarbonate acceptor phosphoenolpyruvate by pyruvate-P_i-dikinase (PPDK). The cycle can be regarded as a pump that uses C4 acids as an intermediate to transport bound carbon into the BS cells where CO_2 is released around RubisCO. Thick cell walls around the BS cells and the enlarged distance of BS cells to the leaf surface compared to leaves from C3 plants prevent the loss of released CO_2 by diffusion. The increased CO_2 concentrations around RubisCO lead to a drastic decrease in photorespiration and therefore an enhanced photosynthetic efficiency, especially in dry and warm climates (Nelson, 2010; Langdale, 2011).

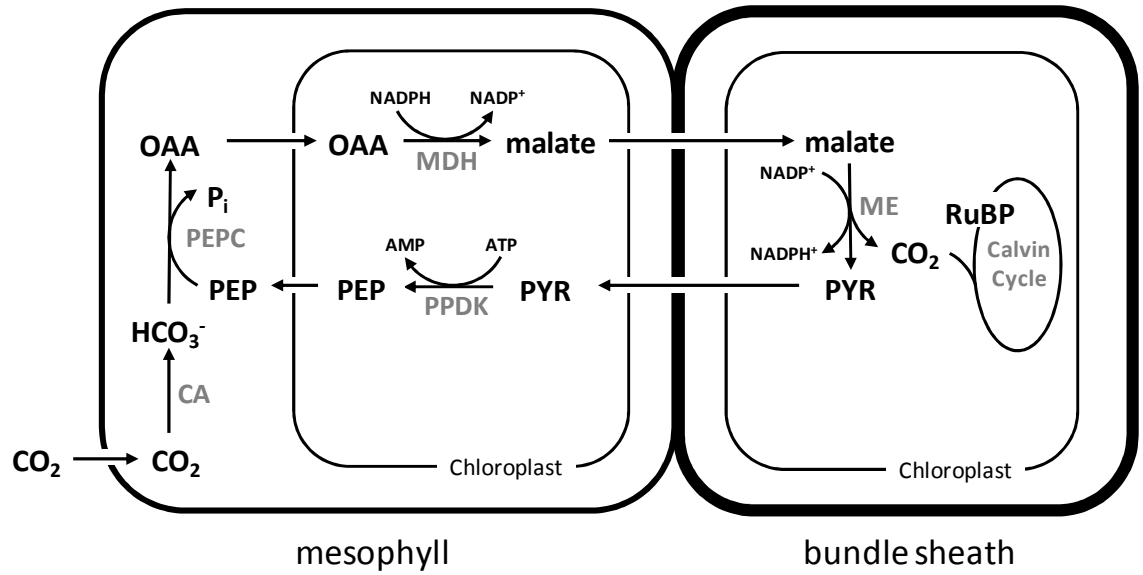


Figure 1-1: Schematic overview of the C4 photosynthesis of *Zea mays*. The bold line around the bundle sheath cell indicates a reinforced cell wall. CA, carbonic anhydrase; PEPC, phosphoenolpyruvate carboxylase; MDH, malate dehydrogenase; ME, malic enzyme; PPDK, pyruvate-P_i-dikinase; HCO_3-, bicarbonate; PEP, phosphoenolpyruvate; OAA, oxaloacetate; PYR, pyruvate; RuBP, ribulose-1,5-bisphosphate; NADPH, nicotinamide adenine dinucleotide phosphate (reduced); NAPD+, nicotinamide adenine dinucleotide phosphate (oxidized); AMP, adenosine monophosphate; ATP, adenosine triphosphate (Horst, 2009).

Several anatomical and physiological requirements have to be met for functional C4 photosynthesis: Kranz anatomy has to be established which is characterized by a high density of BS cells surrounded by M cells. The BS cells have to be enclosed by thick cell walls to reduce diffusion of released CO_2. Both cell types have to be equipped with differentiated chloroplasts specialized for the requirements within both cell types. Plasmodesmata have to be established to facilitate the excessive exchange of metabolites between M and BS cells. Additionally, transporters for the import and export of C4 specific metabolites have to be expressed and integrated into the chloroplast membranes (Gowik and Westhoff, 2011; Nelson, 2011). The development of these requirements is described in Chapter 1.3.1 of the introduction.

1.2.2 Regulation of C4 gene expression by histone modifications in maize

It is supposed that the genes needed to perform C4 photosynthesis derive from gene duplications in ancestral C3 plants. Gene duplications enable mutation of one of the duplicated genes without losing the primary function (Brautigam et al., 2011; van den Bergh, Erik et al., 2014). During C4 evolution, the expression of C4 genes was adapted to the requirements needed to perform C4 photosynthesis. C4 genes are appropriate models to study the regulation of gene expression because they are regulated by various stimuli and developmental signals. Relevant C4 genes are differentially expressed between M and BS cells, potentially, to keep the directional CO_2 pump running (Chang et al., 2012; Tausta et al., 2014). C4 genes, like most photosynthetic genes, are induced by light and show a diurnal expression pattern (Horst et al., 2009). Nitrogen availability regulates the expression of the C4 genes *Pepc*, *Ca* and *Ppdk*, potentially due to the fact that according proteins are highly abundant and their synthesis consumes a lot of nitrogen (Sugiharto et al., 1990; Sugiharto et al., 1992; Suzuki et al., 1994). Additionally, it has been shown that high sugar contents repress the expression of various C4 genes in a hexokinase dependent manner (Jang and Sheen, 1994). All these signals have to be integrated into one transcriptional output.

Comprehensive studies on the *Pepc* promoter in maize showed that the appearance of certain histone modifications at certain positions of the promoter can be associated with distinct signals that regulate *Pepc* transcription. Two sites at the *Pepc* promoter were identified that were associated with analyzed histone modifications: the core promoter which is localized adjacent to the TIS and the upstream promoter which is located further upstream of the TIS. It was shown that H3K9ac and H4K5ac were induced by light and highly co-regulated with transcription at the upstream promoter. These data are in line with the general association of histone acetylations with active genes, described in 1.1.2. However, these

histone modifications were differently regulated at the core promoter. H3K9ac and H4K5ac were only observed in the light, but not affected by the metabolic state and therefore not strictly co-regulated with the transcription (Offermann, 2006; Offermann et al., 2008). Interestingly, levels of H3K9ac and H4K5ac at the core promoter followed a diurnal rhythm and started to decrease in the second half of the light period when the light stimulus was still present (Horst et al., 2009).Thus, H3K9ac seemed to represent the light stimulus in a diurnal manner at the core promoter of *Pepc* that is independent of transcription. These data suggested different roles of H3K9ac and H4K5ac dependent on the modified region of the *Pepc* promoter. Additionally, H3K4me3, that has also been associated with active genes, was found to mark the *Pepc* core promoter in a tissue specific manner. *Pepc* is transcribed in M cells and H3K4me3 was found at the core promoter only in M cells, independent of the illumination state and the transcriptional activity. At the core promoter of the BS specific *Me* gene, H3K4me3 was only found in the BS cells (Danker et al., 2008). These data exemplify the complexity of the signal integration at C4 promoters and the demand to study histone modifications in different situations to understand their impact on gene expression. Further studies in other C4 species and genome-wide analyses could help understanding and generalizing the regulation of C4 gene expression by histone modifications.

1.3 The maize leaf gradient - a model to study development

Leaves of dicotyledonous plants develop in an asymmetric manner. The development is polarized from the base to the leaf edges in general but further divisions in all directions are inserted at later stages of leaf development, resulting in a mosaic-like spatial dispersion of cells with different developmental stages. This pattern impedes the analysis of developmental processes in leaves from dicotyledonous plants (Poethig, 1987; Nelson and Langdale, 1989; Efroni et al., 2010).

In contrast, the development of leaves of monocotyledonous grasses is polarized all over the leaf. In maize, cell division is restricted to the base of the leaf where the initial cells are recruited from around 40 meristematic cells (Poethig and Szymkowiak, 1995). The arrangement of newly generated cells remains aligned from the base to the tip of the maize leaf (Nelson and Langdale, 1989). The differentiation of cells is strictly organized and continuous from the tip to the base (basipetally); cells at the tip are thoroughly differentiated while they are undifferentiated at the base of the leaf (Sharman, 1942; Evert et al., 1996; Sylvester et al., 1990). Because of the structured and polarized development, the growing maize leaf is nowadays also referred to as the developmental gradient. The developmental gradient of maize was classified into three zones in dependence on the local activity of cells

(Figure 1-2). The base of the leaf is characterized by undifferentiated cells from the intercalary meristem and was defined as division zone (DZ). The size of the mature leaf is strongly controlled by the size of the DZ (Rymen et al., 2007; Nelissen et al., 2012). In the adjacent zone, cells expand and start to differentiate (expansion zone [EZ]). Cells become part of the mature zone when the expansion is completed (mature zone [MZ], Sylvester et al., 1990; Nelissen et al., 2013). Nelissen et al. (2012) showed that the developmental zones of the maize leaf were characterized by varying concentrations of endogenous plant hormones. Auxin and cytokinines peaked in the DZ while gibberellines peaked in the EZ. Environmental factors that might influence the development of the maize leaf have not been investigated so far.

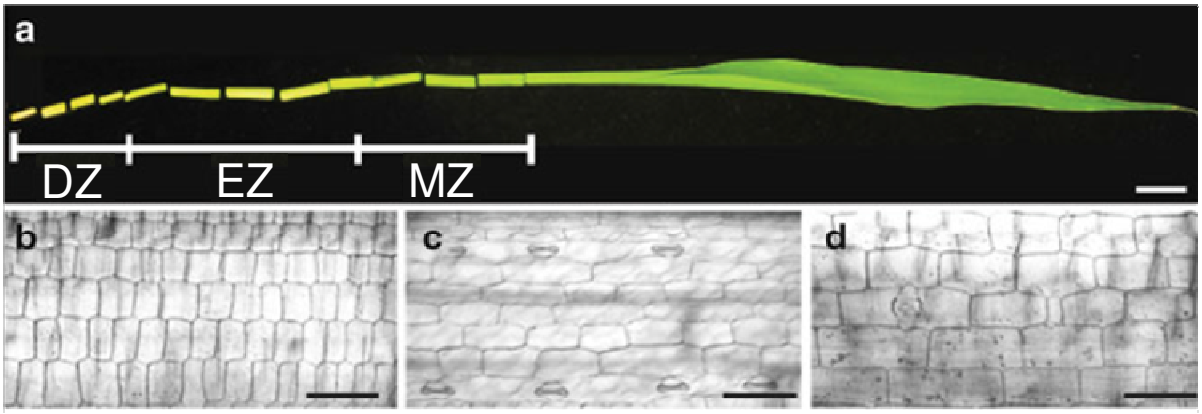


Figure 1-2: Classification of analyzed sections into defined zones (a) and corresponding microscopic images (b-d). DZ, division zone (b); EZ, expansion zone (c); MZ, mature zone (d). The nomenclature of the figure from Nelissen et al. (2013) was modified according to Candaele et al. (2014).

The strictly polarized and linear organization of the developmental gradient of the maize leaf offers the opportunity to comprehensively study the development of various biological processes within a single organ. Identifying factors required for the establishment of C4 photosynthesis at different developmental stages is of special interest. Additionally, the connection of different datasets can be used to analyze regulatory cascades from the regulation of gene expression to functional proteins and metabolites in general (Nelson, 2011; Candaele et al., 2014; Ponnala et al., 2014).

1.3.1 Omics along the developmental gradient of maize to characterize the establishment of C4 photosynthesis

In recent years, new technologies, like next generation sequencing technologies, have become available to a broad spectrum of researchers and large data sets on the developmental gradient of maize have been generated on a genome-wide level. Such datasets comprise transcriptomic, proteomic and metabolomic analyses (Majeran et al., 2010; Li et al., 2010; Pick et al., 2011; Aubry et al., 2014; Tausta et al., 2014; Wang et al., 2014). One aim was to understand the regulatory and anatomical factors that are responsible for the establishment of C4 photosynthesis along the developmental gradient. Technologies for genome-wide analyses to study the developmental gradient of maize were firstly used by Li et al. (2010) and Majeran et al. (2010). Li et al. (2010) applied RNA-Seq to analyze the transcriptomes of four representative zones of the maize leaf (base, expansion, maturing and mature). Majeran et al. (2010) used a shotgun proteomic approach to analyze the proteomes of identical four and two additional zones. The proteomic data was complemented with comprehensive microscopic analyses to analyze the establishment of anatomic features needed for C4 photosynthesis. Additionally, both studies also incorporated differential analyses of BS and M cells at different stages of development. Hierarchical clustering and analyses of known marker genes or proteins for certain processes were used to understand the establishment of C4 photosynthesis. The RNA and protein levels of both studies correlated well over the analyzed developmental zones (Ponnala et al., 2014).

Initially, the transition point between sink and source tissues was identified as a kind of reference point for developmental analyses of C4 photosynthesis. The base of the leaf is a heterotrophic sink tissue that depends on the supply with carbohydrates from the autotrophic source tissue with active photosynthesis (Evert et al., 1996). The sink-source transition point is defined as the zone of the leaf where the import and export of carbohydrates are level. The location was identified at the area of the leaf where cells are initially exposed to light (transition from EZ to MZ) by autoradiography of incorporated $^{14}CO_2$ (Majeran et al., 2010; Li et al., 2010).

The observations regarding the gradual establishment of C4 photosynthesis overlapped between both studies and are summarized in Figure 1-3 (Majeran et al., 2010). Plasmodesmata between future BS and M cells were already established within the first two centimeters of the division zone. Slightly afterwards, the maturation of the vascular elements occurred. The tissue up to that point was characterized as completely heterotrophic: the abundances of proteins involved in sucrose degradation, sink-specific transporters of metabolites and enzymes of the respiratory machinery peaked in that area. However, structural requirements for C4 photosynthesis already started to show up: the orientations of

cells in a Kranz anatomy specific manner and the beginning differentiation of proplastids were already recognized before the sink-source transition point. Also differential accumulation of BS and M specific proteins was observed for a small fraction of total BS and M specific proteins. In the following developmental phase, the leaf emerges into direct light and the sink-source transition point is included. In that area, the photosynthetic machinery rapidly evolved. Proteins associated with the establishment and homeostasis of the plastidal proteome and the synthesis of isoprenoids and tetrapyrroles peaked in that area. Also a rapid accumulation of BS and M specific proteins needed for C4 photosynthesis was observed. Cell elongation and cell wall thickening were still in progress in that area, indicated by a high accumulation of enzymes associated with cellulose and fatty acid synthesis. Thus, the photosynthetic machinery was basically prepared before the sink-source transition point and directly activated when light was available, although the development was still ongoing. The area after the sink-source transition point was characterized by continuous maturation of the photosynthetic machinery and required anatomical structures towards the tip of the leaf. The elongation of cells and the cell wall thickening of the BS cell wall were completed closely after the sink-source transition point. This was characterized by decreased abundances of cell wall and lignin synthesis associated proteins. Proteins associated with the biogenesis of chloroplasts also decreased while especially the expansion of M chloroplasts continued towards the tip of the leaf. The increasing accumulation of BS and M specific proteins and photosynthetic metabolite transporters towards the tip indicated increasing C4 metabolism. This was underpinned by strongly increasing abundances of proteins associated with starch metabolism, protection from reactive oxygen species and the visible accumulation of starch grains towards the tip of the leaf.

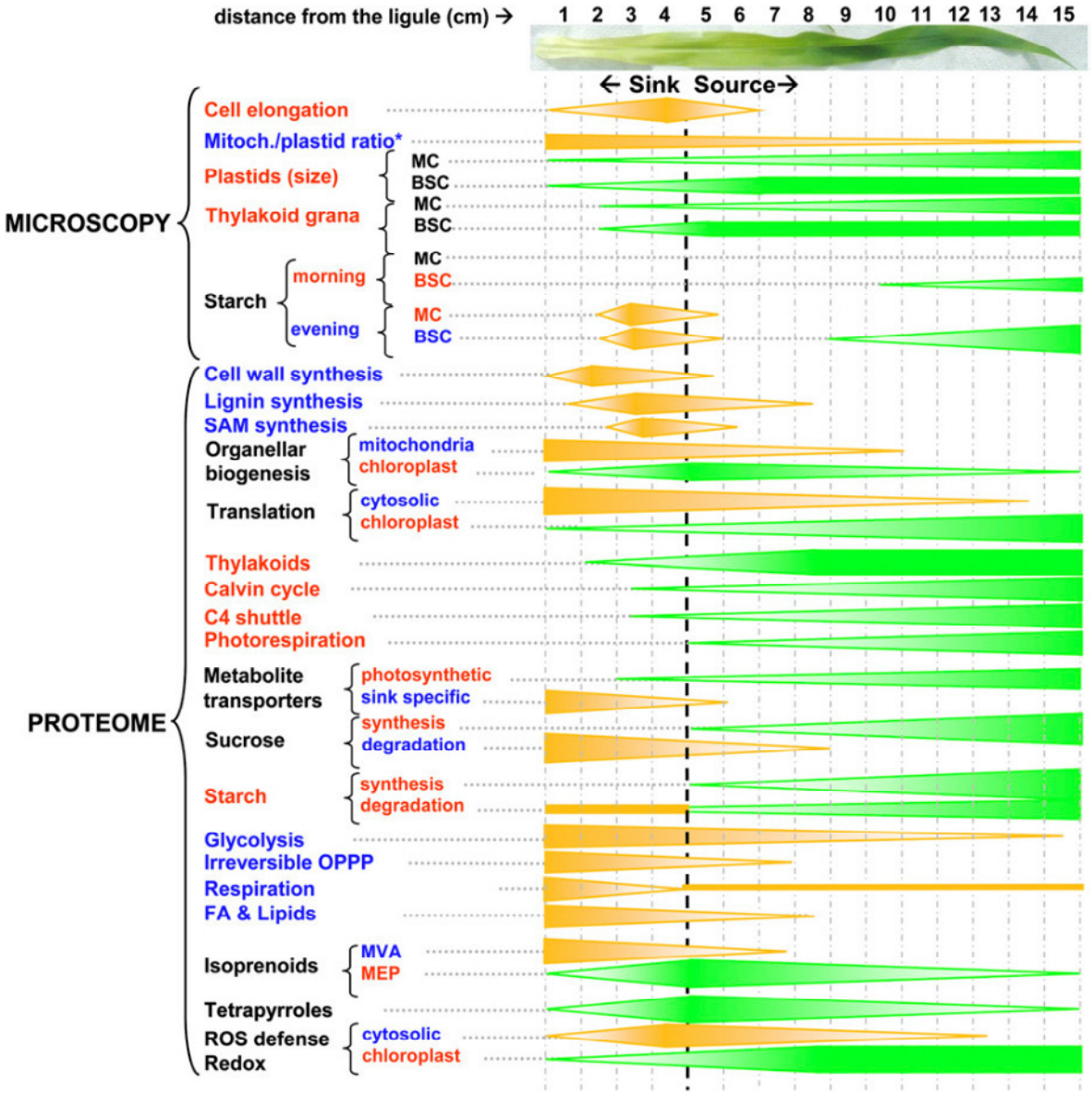


Figure 1-3: Overview of observed proteomic and anatomical changes along the developmental gradient of maize. The upper part summarizes anatomical features observed by light microscopy and the lower part changes within the indicated protein clusters measured by shot gun proteomics. The vertical dashed line indicates the transition point between sink and source tissues. The green lines and processes in red represent features that are preferably found in the source part while orange lines and blue processes are rather found in the sink part of the leaf. MC, mesophyll cells; BSC, bundle sheath cells; SAM, S-adenosylmethionine; OPPP, oxidative pentose phosphate pathway; FA, fatty acid; MVA, mevalonate pathway; MEP, methyl-erythritol-4-phosphate pathway; ROS, reactive oxygen species (Majeran et al., 2010).

The establishment of C4 photosynthesis along the development gradient of maize was also analyzed by Pick et al. (2011). The growing maize leaf was cut at the sink-source transition point and the upper part was subdivided into 10 continuous fragments that were mainly analyzed for their metabolites, enzyme activities and transcript levels. The results were in line with the observations described above. On the metabolite level, different clusters of co-regulated metabolites were identified. Clusters including metabolites linked to C4 photosynthesis and chlorophyll synthesis increased towards the tip of the leaf. Additionally,

enzyme activities of C4 associated proteins were highest at the tip of the leaf. On the other hand, one big cluster of metabolites containing proteinogenic amino acids, precursors of the cell wall, and minor sugars decreased. This cluster was defined as building block cluster, reflecting the needs of the included metabolites at the heterotrophic base of the leaf to establish cellular functions.

1.3.2 Identification of C4 photosynthesis related genes

Transcriptomic data can be used to identify unknown genes that might participate in certain biological processes by co-expression analyses of genes with known function. Pick et al. (2011) first used transcriptomic data of the developmental gradient from maize to identify transcription factors that might control the establishment and regulation of C4 photosynthesis. 32 transcription factors were found that were tightly co-expressed with the C4 genes *Pepc*, NADP-*Me* and *Ppdk*. Comparative analyses of transcriptomic data from different species have also been used to identify transcription factors that might be associated with C4 photosynthesis. Wang et al. (2014) compared the transcriptomic data of the developmental gradient of maize (Li et al., 2010) to transcriptomic data of the developmental gradient of the monocotyledonous C3 plant rice. The sink-source transition point and metabolomic data of rice and maize were used to calibrate the developmental gradients from both species on each other. Potential C4-related transcription factors were identified by four criteria: expression above background in both species, a clear correspondence of the orthologous genes, differential expression between BS and M cells in maize and different expression profiles along the developmental gradients from both species. 118 transcription factors were identified that fulfilled these criteria and that are candidates for genes that evolved from C3 plants to regulate C4 specific gene expression. Aubry et al. (2014) compared the transcriptomes of the developmental gradient of the dicotyledonous C4 plant *Gynandropsis* with the transcriptomic data of the developmental gradient from maize (Li et al., 2010). 18 transcription factors were identified that were co-expressed with known C4 photosynthesis associated genes along the developmental gradient in both species. Interestingly, nine of these transcription factors were equally distributed between BS and M cells in maize and *Gynandropsis*. However, co-expression and clustering analyses with known genes can only suggest potential association of genes with biological processes. The usage of multiple strategies can help to narrow candidate lists down, but in the end, functional association can only be indicated by functional assays.

Chapter 2 - Publication I

A Common Histone Modification Code on C4 Genes in Maize and Its Conservation in *Sorghum* and *Setaria italica*

Louisa Heimann, Ina Horst, Renke Perduns, Björn Dreesen, Sascha Offermann, and Christoph Peterhansel

Leibniz University Hannover, Institute of Botany, 30419 Hannover, Germany

Type of authorship:	co-author
Type of article:	research article
Share of the work:	20%
Contribution to the publication:	contributed to experiments, data analysis and writing of the paper
Journal:	Plant Physiology (www.plantphysiol.org)
Impact Factor:	7.394 (2013)
Date of Publication:	May, 2013
DOI:	10.1104/pp.113.216721
PubMed-ID:	23564230
Citation:	**Heimann L, Horst I, Perduns R, Dreesen B, Offermann S, Peterhansel C** (2013) A Common histone modification code on C4 genes in maize and its conservation in Sorghum and Setaria italica. Plant Physiol **162**: 456–469

A Common Histone Modification Code on C4 Genes in Maize and Its Conservation in Sorghum and *Setaria italica*[1[W][OA]]

Louisa Heimann[2], Ina Horst[2], Renke Perduns, Björn Dreesen[3], Sascha Offermann, and Christoph Peterhansel*

Leibniz University Hannover, Institute of Botany, 30419 Hannover, Germany

C4 photosynthesis evolved more than 60 times independently in different plant lineages. Each time, multiple genes were recruited into C4 metabolism. The corresponding promoters acquired new regulatory features such as high expression, light induction, or cell type-specific expression in mesophyll or bundle sheath cells. We have previously shown that histone modifications contribute to the regulation of the model C4 phospho*enol*pyruvate carboxylase (C4-*Pepc*) promoter in maize (*Zea mays*). We here tested the light- and cell type-specific responses of three selected histone acetylations and two histone methylations on five additional C4 genes (C4-*Ca*, C4-*Ppdk*, C4-*Me*, C4-*Pepck*, and C4-*RbcS*2) in maize. Histone acetylation and nucleosome occupancy assays indicated extended promoter regions with regulatory upstream regions more than 1,000 bp from the transcription initiation site for most of these genes. Despite any detectable homology of the promoters on the primary sequence level, histone modification patterns were highly coregulated. Specifically, H3K9ac was regulated by illumination, whereas H3K4me3 was regulated in a cell type-specific manner. We further compared histone modifications on the C4-*Pepc* and C4-*Me* genes from maize and the homologous genes from sorghum (*Sorghum bicolor*) and *Setaria italica*. Whereas sorghum and maize share a common C4 origin, C4 metabolism evolved independently in *S. italica*. The distribution of histone modifications over the promoters differed between the species, but differential regulation of light-induced histone acetylation and cell type-specific histone methylation were evident in all three species. We propose that a preexisting histone code was recruited into C4 promoter control during the evolution of C4 metabolism.

The current best estimate for the minimal number of independent evolutionary origins of C4 photosynthesis is 62. Thus, C4 photosynthesis belongs to the most prominent examples of parallel or convergent evolution in nature (Sage et al., 2012). C4 plants established a carbon pump that transports CO_2 from mesophyll (M) to bundle sheath (B) cells, where Rubisco and the Calvin cycle are active (von Caemmerer and Furbank, 2003). One major aspect of C4 evolution was the recruitment of preexisting genes to encode the enzymes of the C4 pathway (Hibberd and Covshoff, 2010). Key enzymes for C4 photosynthesis in M cells are carbonic anhydrase (C4-*Ca*), phospho*enol*pyruvate carboxylase (C4-*Pepc*), and pyruvate phosphate dikinase (C4-*Ppdk*). Conversely, in B cells, a decarboxylase such as NAD(P)-malic enzyme (C4-*Me*) and/or phospho*enol*pyruvate carboxykinase (C4-*Pepck*) and Rubisco (*RbcS* for the genes encoding the small subunit) are required at high levels (Langdale, 2011). Maize (*Zea mays*) makes use of both a chloroplastic NADP-*Me* and a cytosolic *Pepck* for C4 acid decarboxylation (Furumoto et al., 1999; Wingler et al., 1999). During the recruitment process, C4 genes acquired new regulatory features: the genes show much higher expression than their C3 counterparts (Ku et al., 1996), they are activated in response to light (Sheen and Bogorad, 1987; Sheen, 1999), and their activity is often modulated by additional metabolic stimuli such as nitrate availability (Sugiharto et al., 1992) or sugar accumulation (Sheen, 1990). Much of this regulation takes place on the level of promoter activity (Sheen, 1999; Hibberd and Covshoff, 2010). Furthermore, C4 proteins show selective accumulation in M or B cells. Information for cell type specificity is encoded in promoter sequences (Gowik et al., 2004; Akyildiz et al., 2007), in untranslated transcript regions (Patel and Berry, 2008; Kajala et al., 2012), and/or in the coding sequence (Brown et al., 2011), depending on the C4 gene and the species. There is evidence that some genes in C3 plants were predisposed to their recruitment into the C4 pathway, because their C3 orthologs show aspects of C4 regulation already in the C3 plant (Brown et al., 2010; Kajala et al., 2012). Also, C4 genes are often correctly regulated when overexpressed in a C3 host, indicating that the relevant trans-acting factors are available in C3 plants (Ku et al., 1996; Engelmann et al., 2003).

Research in recent years has highlighted the role of chromatin structure and modification in the control of transcription. The primary repeat unit of chromatin is the nucleosome particle that is formed by DNA wound

[1] This work was supported by the Deutsche Forschungsgemeinschaft (grant no. PE819/1–3 to C.P.).

[2] These authors contributed equally to the article.

[3] Present address: Projektträger Jülich, 52428 Julich, Germany.

* Corresponding author; e-mail cp@botanik.uni-hannover.de.

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: Christoph Peterhansel (cp@botanik.uni-hannover.de).

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

[OA] Open Access articles can be viewed online without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.113.216721

around a protein body. This protein body consists of two copies each of the histone proteins H2A, H2B, H3, and H4 (Kouzarides, 2007). Chromatin is a passive barrier for transcription and other DNA-associated biochemical processes (Kingston and Narlikar, 1999). Two major mechanisms that facilitate the transcription of chromatin have been identified. On the one hand, chromatin-remodeling complexes are recruited to specific chromatin domains and alter the mobility of nucleosomes in an ATP-dependent manner (Mellor, 2005). This results in differences in nucleosome occupancy (NO) at specific DNA positions. In both yeast and humans, active gene promoters are characterized by low NO (Lee et al., 2004; Nishida et al., 2006). On the other hand, histones can be posttranslationally modified in multiple ways (Bannister and Kouzarides, 2011; Tan et al., 2011). Using genome-wide correlation analyses, some of these modifications have been associated with transcriptionally active or inactive chromatin domains (Bernstein et al., 2007; Zhang, 2008). The best studied modification of histones is probably Lys acetylation. Multiple residues on the N-terminal tails of histones H3 and H4 can be acetylated. Histone acetylation is tightly correlated with gene transcription (Pokholok et al., 2005; Wang et al., 2009a). The activating properties of histone acetylation may be explained by two nonmutually exclusive hypotheses: either they are due to neutralization of the interaction of the positively charged Lys side chains with the negatively charged DNA, resulting in the mitigation of histone-DNA interaction (charge neutralization model; Dion et al., 2005; Henikoff and Shilatifard, 2011), or they are due to the provision of binding sites for transcription factors and other proteins that specifically bind to acetylated histones (histone code model; Berger, 2007; Hassan et al., 2007; Nelissen et al., 2007). The latter model also implies that histone modifications can be used for the storage of information about developmental and environmental cues on the promoters.

A second prominent modification of histones is Lys methylation. Again, multiple Lys residues on histones are prone to methylation (Sims et al., 2003). Depending on the Lys that is methylated, these modifications are either found in actively transcribed chromatin regions or in heterochromatic nontranscribed regions (Sims et al., 2003; Martin and Zhang, 2005). Furthermore, the terminal amino group of the Lys side chain can be monomethylated, dimethylated, or trimethylated, and this also affects the interpretation of the modification. Methylated Lys residues are recognized by specific protein domains (Kim et al., 2006) and, thus, can recruit proteins that, in turn, activate or repress transcription. The most prominent example of histone methylation associated with active genes is the trimethylation of Lys-4 on histone H3 (H3K4me3) at the start of the transcribed gene region (Santos-Rosa et al., 2002; Heintzman et al., 2007; Wang et al., 2009a).

We have previously studied information storage by histone modifications on the promoter of C4-*Pepc* in maize. Our studies revealed that the acetylation of Lys-9 on histone H3 (H3K9ac) and Lys-5 on histone H4 (H4K5ac) on the promoter is controlled by illumination (Offermann et al., 2006, 2008). The modifications were even set on inactive gene copies as long as the light stimulus was received, but they were removed in the dark. Other Lys residues such as H3K18 were constitutively acetylated on C4-*Pepc* (Offermann et al., 2008). Conversely, H3K4me3 on C4-*Pepc* was not affected by the light stimulus but was controlled by a cell type-specific signal specifically in M cells. In B cells, H3K4me2 was found instead (Danker et al., 2008). Here, we tested whether this control of specific modifications is conserved on other C4 genes in maize and on the orthologous C4-*Pepc* and C4-*Me* genes of the C4 grasses sorghum (*Sorghum bicolor*) and *Setaria italica*. Whereas maize and sorghum share a common C4 origin, *S. italica* evolved C4 metabolism independently (Christin et al., 2009a; Sage et al., 2012) and separated from the maize/sorghum lineage approximately 25 million years ago (Vicentini et al., 2008). Our data reveal a close similarity of the histone code within maize C4 genes and, for at least two key genes, across the species. This suggests that this code was recruited into C4 in two independent C4 lineages.

RESULTS

Chromatin Modification Profiles of C4 Genes in Maize

We used chromatin immunoprecipitation (ChIP) from illuminated leaves harvested 4 h after the onset of illumination to identify regions on C4 gene promoters in maize that show high H3K9ac or H3K4me3. C4 genes were identified from the genome sequence based on their homology to complementary DNAs (cDNAs) that had been shown before to encode C4-specific transcripts (see "Materials and Methods"; Supplemental Fig. S1). All modification data were standardized for the number of nucleosomes in the tested chromatin region (NO). NO was determined by precipitating chromatin with an antibody directed to an invariant domain of histone H3 and is shown in Supplemental Table S1. Y-scales always show the relative enrichment of acetylation or methylation at the respective site compared with the *Actin1* housekeeping gene promoter (Haring et al., 2007; i.e. a value of 10 means that the modification is found 10 times more often on nucleosomes on the tested C4 gene promoter region than on nucleosomes on the *Actin1* promoter [see "Materials and Methods"]).

In the core promoter region (−200 relative to the transcription initiation site [TIS]) of the C4-*Pepc* promoter (Fig. 1A), we detected a 6-fold enrichment in acetylation compared with the *Actin1* promoter. This number slightly decreased at position −900 and then increased to 12-fold enrichment at position −1,300. Further upstream, acetylation constantly decreased. For C4-*Me* (Fig. 1D), the situation was similar, with a peak of acetylation on the core promoter (−240) and a second peak further upstream (−1,450). Also, relative

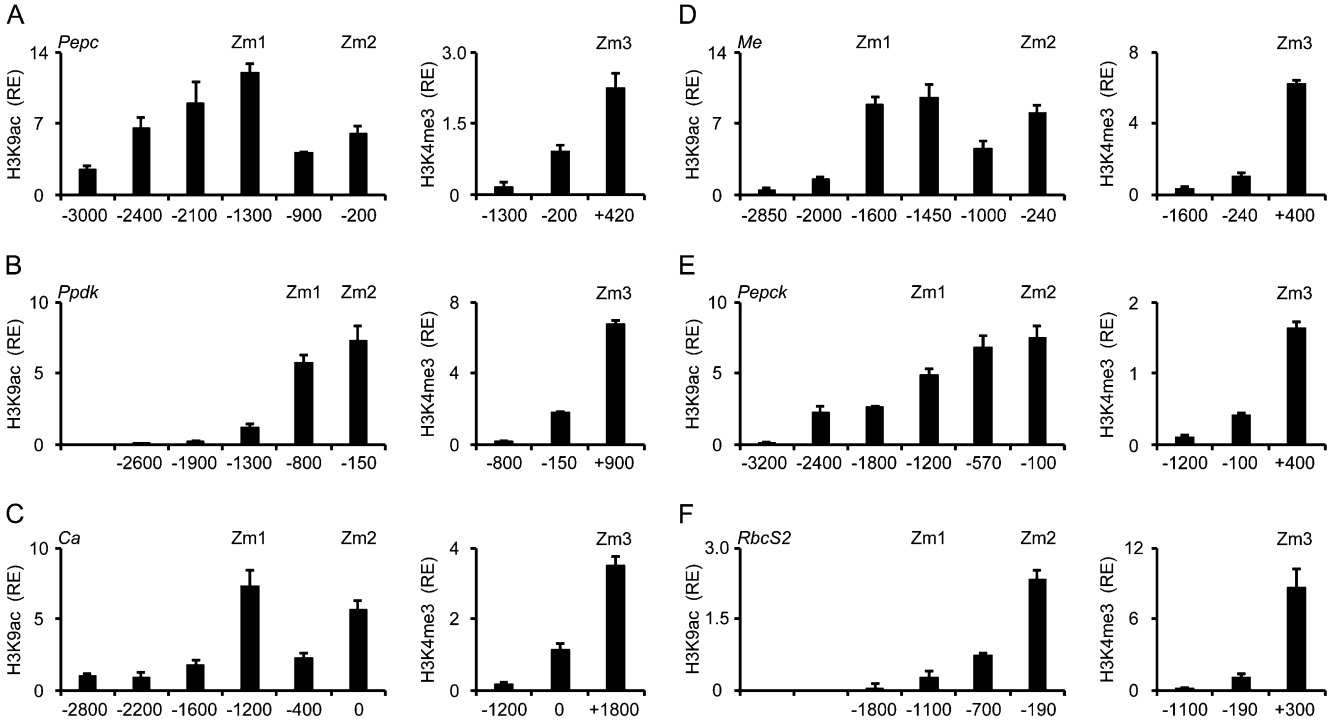


Figure 1. Histone modification profiles of six C4 genes from maize. Amounts of chromatin precipitated with an antibody specific for H3K9ac or H3K4me3 in illuminated leaves (4L) are shown. Numbers on the *x* axis indicate bp positions relative to TIS. Positions chosen for further analyses are designated Zm1, Zm2, and Zm3. Values are presented as the relative enrichment (RE) of modifications per nucleosome over modifications per nucleosome found on maize *Actin1*. All data points are based on at least three independent experiments. Vertical lines indicate SE.

enrichment compared with *Actin1* was similar on C4-*Pepc* and C4-*Me*. Therefore, for both genes, we selected a position on the upstream promoter (Zm1) and a position on the core promoter (Zm2) with high H3K9ac for the following analysis of light regulation (Fig. 2). These selected positions on the C4-*Pepc* and C4-*Me* promoters had already been shown to contain regulated chromatin modifications in previous studies (Offermann et al., 2006, 2008; Danker et al., 2008).

We further screened the promoters of the C4 genes C4-*Ca*, C4-*Ppdk*, C4-*Pepck*, and C4-*RbcS2* with the aim to define comparable chromatin positions (i.e. a core promoter position and an upstream promoter position with high H3K9ac). On C4-*Ppdk* (Fig. 1B), we observed high H3K9ac on the core promoter (position −150) and at position −800 but a rapid decline at more upstream positions. Accordingly, Zm1 was positioned at −800 and Zm2 at −150. On C4-*Ca* (Fig. 1C), we detected high H3K9ac signals at position −1,200 (Zm1) and directly at the TIS (Zm2). We were unable to establish a functioning PCR system directly in the core promoter region of C4-*Ca* because of the very high GC content (more than 75% between −300 and 0) and multiple repetitive sequence elements; therefore, we had to use a PCR system centered on the TIS instead. On C4-*Pepck* (Fig. 1E), H3K9ac was highest on the core promoter (−100) and constantly declined toward the upstream promoter. Zm1 was defined at the most upstream position with at least 5-fold higher acetylation compared with *Actin1* (−1,200), and Zm2 again was placed on the core promoter. On C4-*RbcS2* (Fig. 1F), H3K9ac levels were generally lower. Maximum levels on the core promoter were only 2.3-fold higher compared with the *Actin1* promoter. Moreover, H3K9ac rapidly declined toward more upstream promoter regions. In order to allow direct comparison with the other C4 genes, we decided to place positions Zm1 and Zm2 on C4-*RbcS2* at −1,100 (the most upstream position with detectable H3K9ac) and −190 (highest acetylation), respectively.

H3K4 trimethylation (H3K4me3) is another chromatin mark of active genes that usually peaks in the 5′ region of the transcribed sequence (Santos-Rosa et al., 2002; Heintzman et al., 2007; Wang et al., 2009a). Therefore, we tested an additional position in the 5′ region of the transcribed sequence. This position is designated as Zm3 henceforth (for exact positions, see Supplemental Fig. S1). On all genes, H3K4me3 levels were clearly increased at Zm3 compared with the promoter positions (Fig. 1, right graphs). On C4-*Pepc*, C4-*Ca*, and C4-*Pepck*, significant relative enrichments of H3K4me3 were also detected at Zm2, whereas H3K4me3 levels were near background at the Zm2 position on the other genes. At Zm1, H3K4me3 signals were undetectable (C4-*Ppdk*, C4-*Ca*, C4-*Me*, C4-*RbcS2*) or very low (C4-*Pepc*, C4-*Pepck*). We selected the Zm3

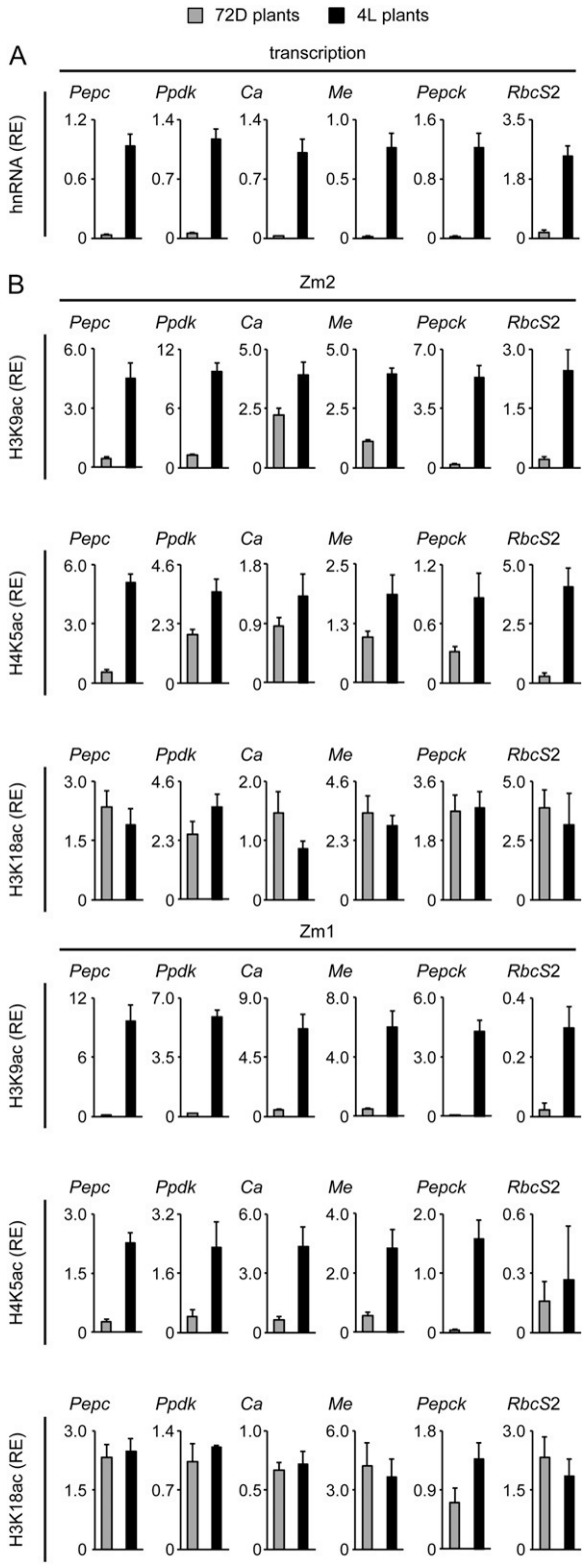


Figure 2. Light regulation of transcription and histone acetylation on six C4 genes in maize. A, Relative quantification of hnRNA expression levels of the six C4 genes in leaves from plants that were exposed to 72 h of darkness (72D; gray columns) and in plants that were illuminated for 4 h (4L; black columns). Transcription is standardized for maize *Actin1* expression (relative enrichment [RE]). hnRNA expression levels were determined by quantitative reverse transcription (RT)-PCR with a primer system specific for an intron (Supplemental Fig. S1). B, Light-dependent H3K9ac, H4K5ac, and H3K18ac on positions Zm1 and Zm2 of the six C4 genes. Values are presented as the relative enrichment of modifications per nucleosome over modifications per nucleosome found on the *Actin1* promoter. All data points are based on at least four independent experiments. Vertical lines indicate SE.

position for further studies of histone methylation on the maize C4 genes.

Environmental and Developmental Regulation of C4 Gene Chromatin in Maize

A key feature of C4 gene regulation is gene induction by light. We compared promoter activity and histone acetylation at the Zm1 and Zm2 positions on the six maize C4 genes in response to illumination (Fig. 2). Beside H3K9ac, H4K5ac and H3K18ac were included, because these modifications showed specific responses to illumination on the C4-*Pepc* promoter before (Offermann et al., 2008). The amounts of heterogeneous nuclear RNA (hnRNA) derived from the genes were used as an approximation for promoter activity, because these primary transcripts are rapidly spliced after synthesis and do not accumulate (Elferink and Reiners, 1996; Delany, 2001; Wu et al., 2009). For all six C4 genes, hnRNA levels were low in plants that were exposed to 72 h of darkness (72D plants) and clearly increased in plants that were illuminated for 4 h (4L plants), indicating light-induced promoter activity (Fig. 2A).

For chromatin analyses, we used the same standardization methods as described for Figure 1. On the Zm2 site proximal to the TIS, H3K9ac was strongly induced (3- to 5-fold) on all genes after illumination, with the exception of C4-*Ca* (only 1.5-fold induction; Fig. 2B). H4K5ac at position Zm2 also responded positively to illumination on all tested genes, albeit to different extents. Light induction was 10-fold or more for C4-*Pepc* and C4-*Rbcs2*, approximately 2-fold for C4-*Ppdk*, C4-*Me*, and C4-*Pepck*, and less than 2-fold for C4-*Ca*. In contrast, H3K18ac at Zm2 remained largely unchanged in 72D plants compared with 4L plants on all C4 genes. At the more upstream-located Zm1 position, H3K9ac was strongly induced by illumination on all C4 genes. Likewise, H4K5ac showed a clear induction at Zm1 on all genes with the exception of *RbcS2*, where H4K5ac levels remained low in 4L plants. H3K18ac at the Zm1 position was unaffected by the light treatment on five of the six tested genes. A slight increase was only observed on C4-*Pepck*. Thus, with few exceptions, H3K9ac and H4K5ac responded positively to illumination at both the Zm1 and Zm2 promoter positions of the C4 genes, whereas H3K18ac remained unaffected by light.

A second hallmark of C4 gene regulation is cell type-specific expression of C4 genes in M or B cells. We compared hnRNA levels of the tested C4 genes in leaves and isolated B strands (Fig. 3A). A rapid mechanical isolation protocol at low temperatures was used for B cell preparation in order to minimize hnRNA turnover during the preparation (Jenkins and Boag, 1985; Bassett et al., 1988). We did not isolate M cells for this assay, because the preparation of M cell protoplasts from leaves is a lengthy procedure during which C4 genes are often strongly suppressed. If promoter activity would contribute to gene regulation, a clear depletion of M-specific transcripts in isolated B cells compared with total leaves would be expected in our assay. In contrast, B-specific transcripts are expected to be more abundant in B preparations than in total leaves (assuming an even number of B cells and M cells in leaves, a 2-fold increase in B cells compared with total leaves would be expected). Accordingly, hnRNA levels were clearly lower in B cells compared with total leaves for C4-*Pepc*, C4-*Ppdk*, and C4-*Ca*, whereas C4-*Me*, C4-*Pepck*, and *RbcS2* transcripts were 1.8- to 2.6-fold enriched in B preparations.

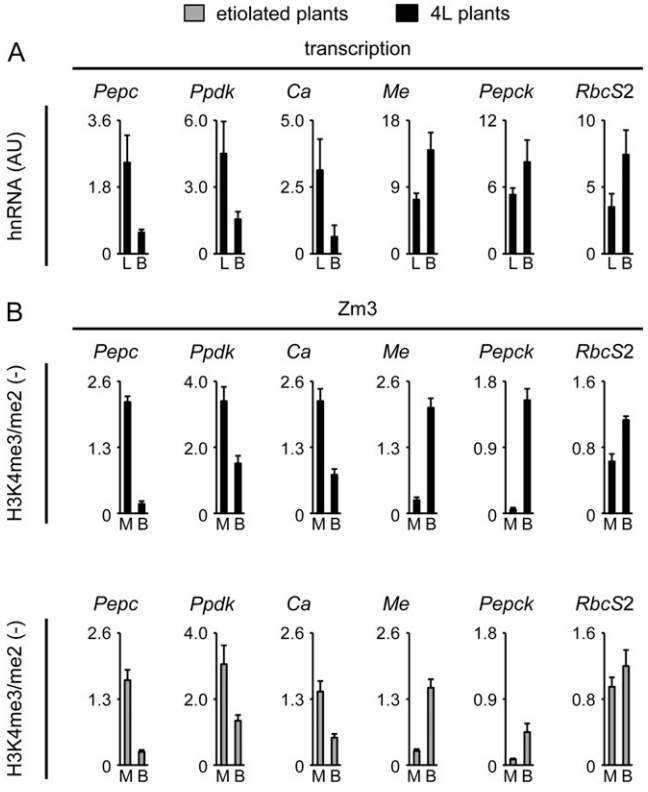


Figure 3. Cell type-specific transcription and histone methylation on six C4 genes in maize. A, Quantification of hnRNA expression levels of the six C4 genes in leaves (L) and B cells (B) isolated from plants that were illuminated for 4 h (4L plants). Values are arbitrary units (AU) derived from a cDNA standard dilution series. hnRNA expression levels were determined by quantitative RT-PCR with a primer system specific for an intron (Supplemental Fig. S1). B, Ratio of the amount of chromatin precipitated with an antibody specific for H3K4me3 and H3K4me2 in isolated M or B cells. Black bars indicate data from illuminated plants (4L), whereas gray bars indicate data from etiolated plants. Ratios are without dimension. All data points are based on at least four independent experiments. Vertical lines indicate SE.

We used chromatin isolated from M and B cells to test cell type-specific histone methylation on the C4 genes at position Zm3. Previous analyses indicated that methylation marks remain stable on M cell promoters during the preparation of protoplasts, even though C4 gene expression is suppressed (Danker et al., 2008). Figure 3B shows the ratio of H3K4me3 to H3K4me2 signals (H3K4me3/me2) at the Zm3 positions of the tested C4 genes in M and B cell chromatin. For C4-*Pepc* and C4-*Me*, we had shown before that high H3K4me3 signals and low H3K4me2 signals were established in the cell type where the particular gene can be activated by other stimuli (i.e. in M cells for C4-*Pepc* and in B cells for C4-*Me*). In each respective other cell type, we had observed the opposite pattern, with high H3K4me2 and low H3K4me3 instead. Consequently, an H3K4me3/me2 ratio of 1.9 was determined for C4-*Pepc* in chromatin from M cells, and this ratio dropped to 0.2 in chromatin from B cells. However, on C4-*Me*, the H3K4me3/me2 ratio was much higher in B cells than in M cells. For the other genes that showed preferential expression in either M cells (C4-*Ca*, C4-*Ppdk*) or B cells (C4-*Pepck*, C4-*RbcS2*), similar patterns were observed. C4-*Ca* and C4-*Ppdk* showed clearly higher H3K4me3/me2 ratios in chromatin from M cells than in chromatin from B cells, and C4-*Pepck* and C4-*RbcS2* showed an opposite distribution of these histone modifications.

In order to determine whether this H3K4 methylation pattern was established independently of the transcription of C4 genes, we also measured H3K4me3/me2 levels in M and B chromatin isolated from etiolated leaves. In etiolated leaves, C4 genes are inactive because they never received a light stimulus (compare with Fig. 2). As shown in Figure 3B, H3K4me3/me2 ratios in M and B cells derived from etiolated leaves were very near those obtained from illuminated leaves. For M cell-specific genes, signals were higher in M cells than in B cells. B cell-specific genes showed the opposite pattern, although absolute levels were somewhat reduced. Only on C4-*RbcS2* were ratios near 1 observed in chromatin from both M and B cells. Thus, cell type-specific histone methylation of H3K4 is found in maize on all tested C4 genes except C4-*RbcS2*. The specific role of the latter gene is revisited below in "Discussion."

Histone Code at the C4-*Pepc* and C4-*Me* Genes in Independent C4 Lineages

The data described so far were obtained to study whether different C4 genes in maize showed similar chromatin modification profiles. We further analyzed whether homologous C4 genes in different species also

showed similar chromatin modifications. To this end, we compared light-induced and cell type-specific chromatin marks on the C4-*Pepc* and C4-*Me* genes from the C4 model grasses sorghum and *S. italica* with those obtained from maize (Figs. 1–3). Identification of nearest gene homologs is described in "Materials and Methods" and Supplemental Figure S1. The three C4-*Pepc* genes and the three C4-*Me* genes showed very similar intron-exon organization in the coding regions (Supplemental Fig. S1) but no detectable sequence homology on the putative promoters. In order to delimit the maximum length of the promoters, we analyzed the distance to the next upstream gene. The next gene was predicted 30 kb upstream of the maize C4-*Pepc* gene, 100 kb upstream of the sorghum C4-*Pepc* gene, but only 7 kb upstream of the *S. italica* C4-*Pepc* gene. For C4-*Me*, distance to the next upstream gene was 13 kb (maize), 11 kb (sorghum), and 15 kb (*S. italica*).

In order to define suitable gene positions for chromatin analyses, we again profiled H3K9ac over the putative promoter regions. Furthermore, we tested H3K4me3 levels at selected positions. Figure 4A shows the profile obtained for sorghum C4-*Pepc*. H3K9ac peaked at an upstream promoter position (−1,400; Sb1). Unlike in maize, a second peak was not observed in the core promoter (−100) but rather at the start of the transcribed region (+300; Sb2). At this position, also high H3K4me3 levels were observed (Fig. 4B). Positions Sb1 and Sb2 were chosen for the analysis of light-dependent histone acetylation (Fig. 4, C–E); in addition, Sb2 was chosen for the analysis of cell type-specific histone methylation (Fig. 4, F and G).

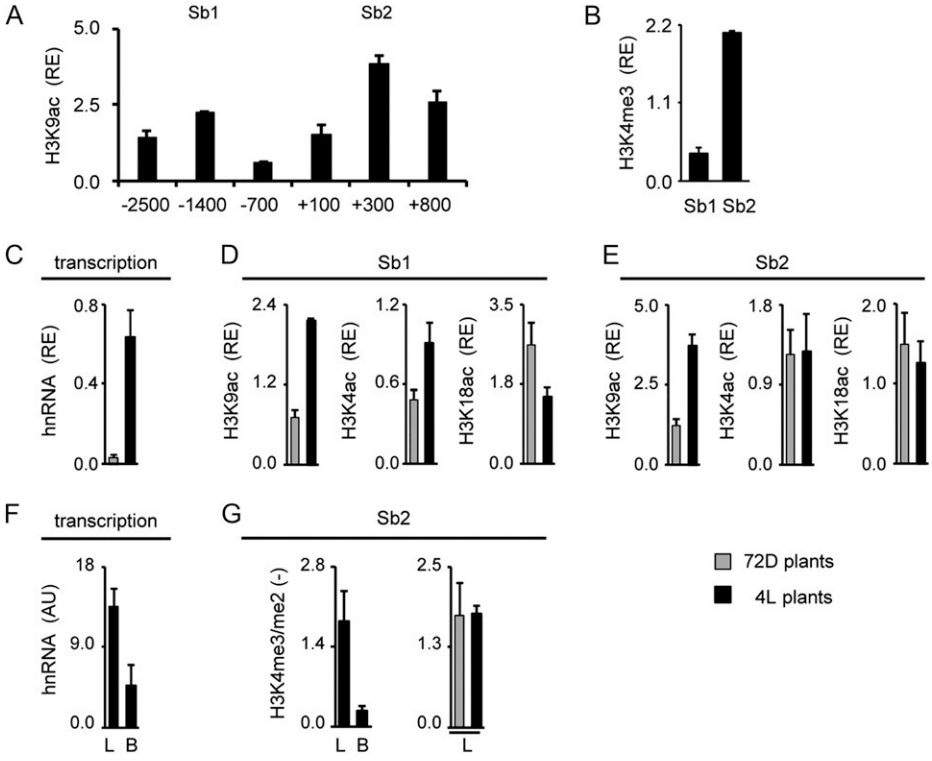


Figure 4. Histone modification profile of the C4-*Pepc* gene from sorghum. A and B, Amounts of chromatin precipitated with an antibody specific for H3K9ac and H3K4me3 in illuminated leaves (4L). Positions chosen for further analyses are designated Sb1 and Sb2. Values are presented as the relative enrichment (RE) of modifications per nucleosome over modifications per nucleosome found on sorghum *Actin1*. Numbers on the *x* axis indicate bp positions relative to TIS. C, Relative quantification of C4-*Pepc* hnRNA expression levels in sorghum leaves from plants that were exposed to 72 h of darkness (72D; gray columns) and in plants that were illuminated for 4 h (4L; black columns). Transcription is standardized for *Actin1* expression (RE). hnRNA expression levels were determined by quantitative RT-PCR with a primer system specific for an intron (Supplemental Fig. S1). D and E, Light-dependent H3K9ac, H4K5ac, and H3K18ac on positions Sb1 and Sb2. Values are presented as the relative enrichment of modifications per nucleosome over modifications per nucleosome found on sorghum *Actin1*. F, Quantification of C4-*Pepc* hnRNA expression levels in sorghum leaves (L) and B cells (B) isolated from plants that were illuminated for 4 h (4L plants). Values are arbitrary units (AU) derived from a cDNA standard dilution series. hnRNA expression levels were determined by quantitative RT-PCR with a primer system specific for an intron (Supplemental Fig. S1). G, Ratio of the amount of chromatin precipitated with an antibody specific for H3K4me3 and H3K4me2 in leaves and B cells. Black bars indicate data from illuminated plants (4L), whereas gray bars indicate data from plants that were exposed to 72 h of darkness (72D). Ratios are without dimension. All data points are based on at least three independent experiments. Vertical lines indicate SE.

As expected, sorghum C4-*Pepc* hnRNA levels were much higher in 4L than in 72D plants (Fig. 4C). On position Sb1, acetylation of H3K9 (3-fold) and H4K5 (2-fold) was enhanced by light, whereas H3K18ac was even slightly reduced in 4L compared with 72D plants (Fig. 4D). On Sb2, H3K9ac was 3-fold induced, but H4K5ac and H3K18ac remained unaffected by the light stimulus (Fig. 4E).

C4-*Pepc* hnRNA amounts in total leaves were about 3.5-fold higher than in isolated B strands (Fig. 4F). This is typical for genes that are preferentially transcribed in M cells (compare with Fig. 3). Unlike in maize, we analyzed cell type-specific histone methylation in sorghum by comparing methylation levels in total leaves and B cells, because we failed to prepare intact M cell protoplasts from sorghum leaves. H3K4me3/me2 ratios were more than 2-fold higher in total leaves compared with isolated B cells. In order to provide evidence that this H3K4 methylation pattern is established independent from the actual rate of transcription, we tested whether the H3K4me3/me2 ratio is different between chromatin from 4L and 72D leaves (Fig. 4H). No differences were detected, indicating that H3K4 methylations do not respond to the light stimulus or the rate of transcription.

An analogous data set for *S. italica* C4-*Pepc* is shown in Figure 5. On the *S. italica* C4-*Pepc* promoter, acetylation peaked at positions −600 (Sit1) and +300 (Sit2; Fig. 5A). Again, core promoter acetylation (−300) was relatively low. At Sit2, but not Sit1, also high H3K4me3 levels were detected. hnRNA levels of *S. italica* C4-*Pepc* were strongly induced by light (Fig. 5C). This correlated with light-induced H3K9ac and H4K5ac at both tested gene positions. Again, H3K18ac remained unaffected by the light stimulus (Fig. 5, D and E). When comparing hnRNA levels in total leaves and B cells, 4-fold higher levels were observed in leaves, indicating cell type-specific transcription (Fig. 5F). Consistently, H3K4me3/me2 ratio was much higher in total leaves compared with B cells but was unaffected by light and, therefore, transcription rates (Fig. 5G). Thus, the sorghum and *S. italica* C4-*Pepc* genes showed clearly different distributions of histone acetylation over the promoters compared with maize, but regulation of transcription, histone acetylation, and H3K4 methylation were highly similar to maize C4-*Pepc*.

In order to substantiate these observations, we also compared the transcription and chromatin regulation of the C4-*Me* genes from the three species. Data sets are organized identically to the C4-*Pepc* data in Figures 4 and 5. On sorghum C4-*Me*, H3K9ac peaked at positions −500 (Sb1) and +100 (Sb2; Fig. 6A). High H3K4me3 levels were again found at Sb2 (Fig. 6B). Promoter activity was induced 3-fold by light. On position Sb1, H3K9ac increased 6-fold after illumination, whereas H4K5ac did not change. On position Sb2, all tested histone acetylations were unaffected by illumination. Figure 6, F and G, shows the results regarding cell type specificity of this gene. The hnRNA for C4-*Me* accumulated to 1.8-fold higher levels in B cells compared with total leaves (Fig. 6F). A theoretical increase of 2-fold would be expected in these assays for genes that show B cell-specific transcription or modification, assuming even numbers of M and B cells in a leaf (compare with Fig. 3). The H3K4me3/me2 ratio was also 1.8-fold higher in B cells (Fig. 6G). Very similar H3K4me3/me2 ratios were recorded from leaves of 4L and 72D plants (Fig. 6G). Thus, H3K4 methylation levels were unaffected by light.

On *S. italica* C4-*Me*, the highest H3K9ac signals were observed at position +300 at the start of the transcribed sequence (Sit2). Acetylation declined toward the upstream promoter, but a second peak was detected at position −2,000 (Sit1; Fig. 7A). Comparable H3K9ac levels were also found at an even more upstream position (−2,500; data not shown). As for the other genes, high H3K4me3 signals were only detected at position Sit2 (Fig. 7B). A 2-fold increase in *S. italica* C4-*Me* gene transcription was induced by light. This was accompanied by a more than 2-fold increase in H3K9 acetylation at the more upstream Sit1 position. H4K5ac remained largely unaffected and H3K18ac was down-regulated by light at this position. At position Sit2 at the start of the transcribed sequence, H3K9ac and H4K5ac remained unaffected by the light stimulus, whereas H3K18ac was again down-regulated after illumination. When comparing B cells with total leaves, C4-*Me* hnRNA levels were 2.3-fold higher in B cells (Fig. 7F) and the H3H4me3/me2 ratio was 2.9-fold higher (Fig. 7G). Again, the H3K4me3/me2 ratio remained largely unaffected by light (Fig. 7G).

DISCUSSION

We wanted to analyze promoter histone modifications on C4 genes in maize. The definition of promoters in eukaryotic genomes is complicated, because of the large genome size and the resulting long distance to the next upstream gene that can be used to define maximal promoter size. Indeed, the next upstream gene was annotated between 5 kb and more than 100 kb distant from the predicted TIS of the maize genes analyzed (Goodstein et al., 2012; Supplemental Fig. S1). Promoter elements can act over a long distance, as described before for the element controlling M-specific expression of C4-*Pepc* in *Flaveria* spp. (Gowik et al., 2004) or the upstream promoter region of Flowering Locus T in Arabidopsis (Adrian et al., 2010). Promoter-deletion studies were instrumental in identifying such functional elements. However, these deletion studies require the use of transgenes that randomly integrate into plant genomes (Francis and Spiker, 2005; Kim et al., 2007). Epigenetic traits such as NO and histone modification can be strongly affected by the transgene integration site (Yan and Boyd, 2006; Yamasaki et al., 2011; Yin et al., 2012). Transgenic promoter studies, therefore, are of limited value for the analysis of epigenetic mechanisms controlling promoters. Chromatin signatures have been used instead

to identify functional promoter elements in humans (Heintzman et al., 2007; Müller-Tidow et al., 2010). Our analyses revealed that all maize C4 genes but *RbcS2* had extended promoter regions enriched in acetylation more than 1 kb upstream of the TIS (Fig. 1). This often coincided with reduced NO (Supplemental Table S1). This pattern is unexpected, as H3K9ac peaks around the TIS on the average maize gene and acetylation in upstream promoter regions is usually low (Wang et al., 2009a). However, the functional significance of these regions is suggested by the strong reaction of upstream promoter histone acetylation to illumination and, thus, gene transcription (Fig. 2). Moreover, the upstream promoter region of C4-*Pepc* also contains DNA methylation sites that are regulated by light, further supporting the involvement of these regions in gene regulation (Tolley et al., 2012). We cannot discriminate in this assay whether the increase in acetylation on the upstream promoter is necessary for transcriptional activation or just accompanies this process. However, we have shown before for C4-*Pepc* that upstream promoter acetylation can be induced even when gene activation is suppressed, supporting the autonomous regulation of histone acetylation in this promoter region (Offermann et al., 2006, 2008).

The distribution of acetylation over the C4-*Pepc* and C4-*Me* promoters was strikingly different in *S. italica* and sorghum compared with maize. The highest acetylation was detected in the 5′ part of the transcribed region in these species. Localization of the second acetylation

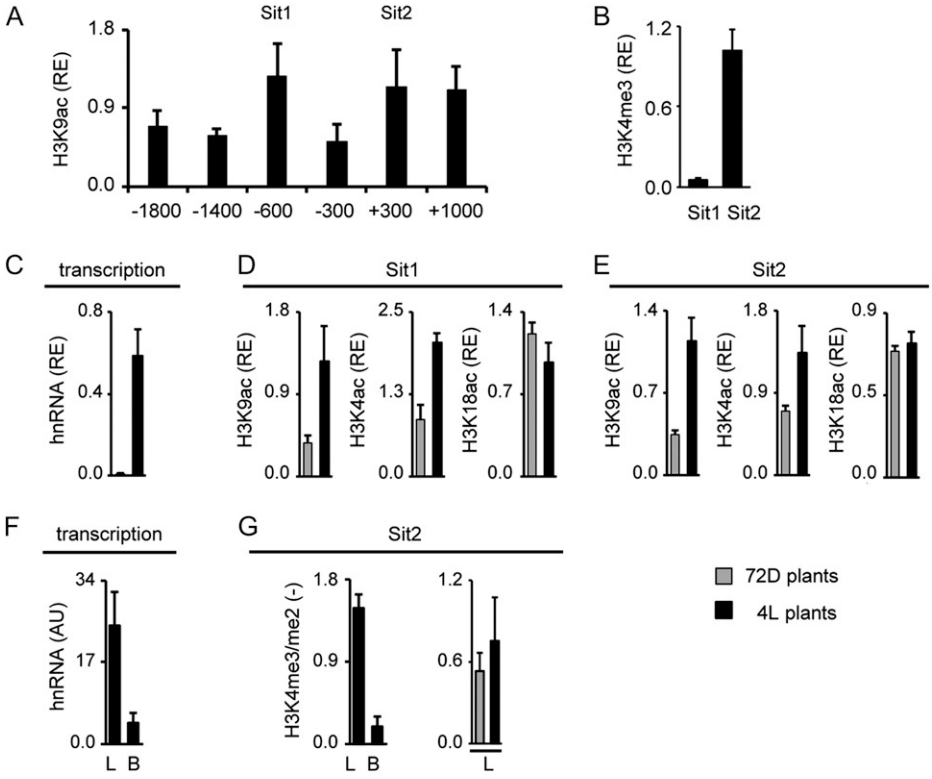


Figure 5. Histone modification profile of the C4-*Pepc* gene from *S. italica*. A and B, Amount of chromatin precipitated with an antibody specific for H3K9ac and H3K4me3 in illuminated leaves (4L). Positions chosen for further analyses are designated Sit1 and Sit2. Values are presented as the relative enrichment (RE) of modifications per nucleosome over modifications per nucleosome found on *S. italica Actin1*. Numbers on the *x* axis indicate bp positions relative to TIS. C, Relative quantification of C4-*Pepc* hnRNA expression levels in *S. italica* leaves from plants that were exposed to 72 h of darkness (72D; gray columns) and in plants that were illuminated for 4 h (4L; black columns). Transcription is standardized for *Actin1* expression (RE). hnRNA expression levels were determined by quantitative RT-PCR with a primer system specific for an intron (Supplemental Fig. S1). D and E, Light-dependent H3K9ac, H4K5ac, and H3K18ac on positions Sit1 and Sit2. Values are presented as the relative enrichment of modifications per nucleosome over modifications per nucleosome found on *S. italica Actin1*. F, Quantification of C4-*Pepc* hnRNA expression levels in *S. italica* leaves (L) and B cells (B) isolated from plants that were illuminated for 4 h (4L plants). Values are arbitrary units (AU) derived from a cDNA standard dilution series. hnRNA expression levels were determined by quantitative RT-PCR with a primer system specific for an intron (Supplemental Fig. S1). G, Ratio of the amount of chromatin precipitated with an antibody specific for H3K4me3 and H3K4me2 in leaves and B cells. Black bars indicate data from illuminated plants (4L), whereas gray bars indicate data from plants that were exposed to 72 h of darkness (72D). Ratios are without dimension. All data points are based on at least three independent experiments. Vertical lines indicate SE.

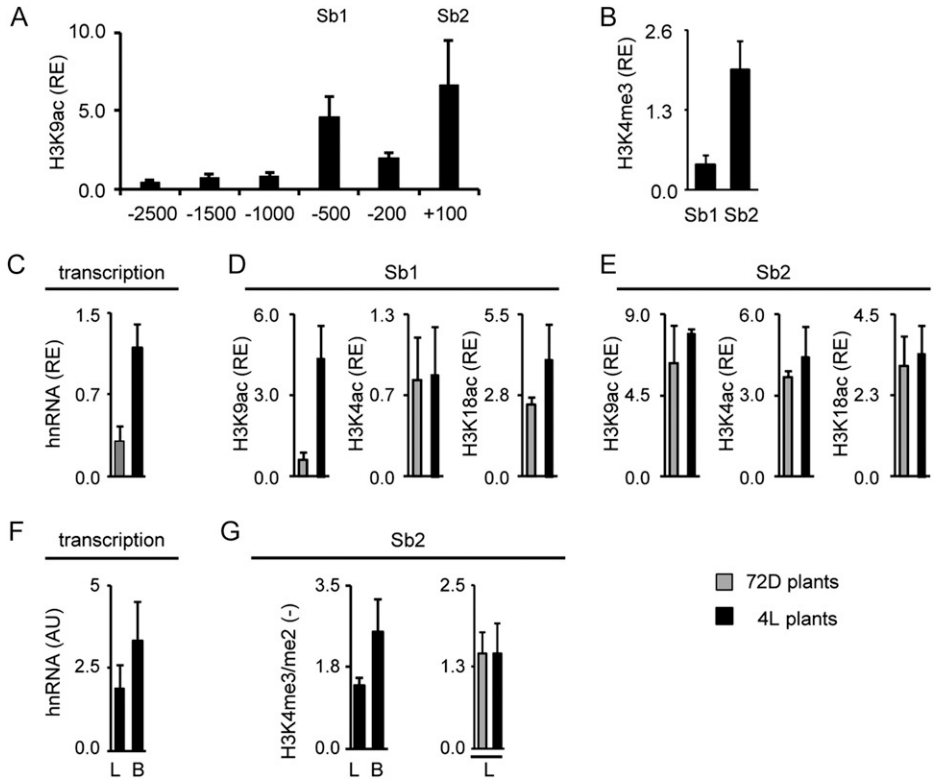


Figure 6. Histone modification profile of the C4-*Me* gene from sorghum. A and B, Amount of chromatin precipitated with an antibody specific for H3K9ac and H3K4me3 in illuminated leaves (4L). Positions chosen for further analyses are designated Sb1 and Sb2. Values are presented as the relative enrichment (RE) of modifications per nucleosome over modifications per nucleosome found on sorghum *Actin1*. Numbers on the *x* axis indicate bp positions relative to TIS. C, Relative quantification of C4-*Me* hnRNA expression levels in sorghum leaves from plants that were exposed to 72 h of darkness (72D; gray columns) and in plants that were illuminated for 4 h (4L; black columns). Transcription is standardized for *Actin1* expression (RE). hnRNA expression levels were determined by quantitative RT-PCR with a primer system specific for an intron (Supplemental Fig. S1). D and E, Light-dependent acetylation of H3K9ac, H4K5ac, and H3K18ac on positions Sb1 and Sb2. Values are presented as the relative enrichment of modifications per nucleosome over modifications per nucleosome found on sorghum *Actin1*. F, Quantification of C4-*Me* hnRNA expression levels in sorghum leaves (L) and B cells (B) isolated from plants that were illuminated for 4 h (4L plants). Values are arbitrary units (AU) derived from a cDNA standard dilution series. hnRNA expression levels were determined by quantitative RT-PCR with a primer system specific for an intron (Supplemental Fig. S1). G, Ratio of the amount of chromatin precipitated with an antibody specific for H3K4me3 and H3K4me2 in leaves and B cells. Black bars indicate data from illuminated plants (4L), whereas gray bars indicate data from plants that were exposed to 72 h darkness (72D). Ratios are without dimension. All data points are based on at least three independent experiments. Vertical lines indicate SE.

peak was highly variable, with positions between −500 (sorghum C4-*Me*) and −2,000 (*S. italica* C4-*Me*) relative to the TIS (Figs. 4A–7A). One obvious explanation for the lack of core promoter acetylation on some of the genes is the clearly smaller genome size of *S. italica* (490 Mb; Doust et al., 2009) and sorghum (730 Mb; Paterson et al., 2009) compared with maize (2,300 Mb; Schnable et al., 2009), which might limit promoter sizes. Indeed, on *S. italica* C4-*Pepc*, the gene with the shortest distance to the next upstream gene, the upstream acetylation peak was already found at position −600.

Independent of the variable positions of acetylation peaks, light regulation of acetylation at these peak positions was highly similar when comparing the three C4 grasses. All C4-*Pepc* and all C4-*Me* genes showed clear light regulation of H3K9ac at least at one of the tested positions, mostly at the more upstream position. Light regulation of H4K5ac was evident on the C4-*Pepc* genes but not on C4-*Me* genes in sorghum and *S. italica*. The lack of H4K5ac regulation correlates with a rather weak light induction of C4-*Me* transcription in these species (only 2- to 3-fold light induction for C4-*Me* compared with 20- to 60-fold for C4-*Pepc*), suggesting that transcription levels might contribute to the degree of modification of this site. Remarkably, H3K18ac was never induced by light on any of the C4 genes but remained unchanged or even declined after illumination. Thus, the positive regulation of selected acetylation sites, most notably H3K9ac, by light is a

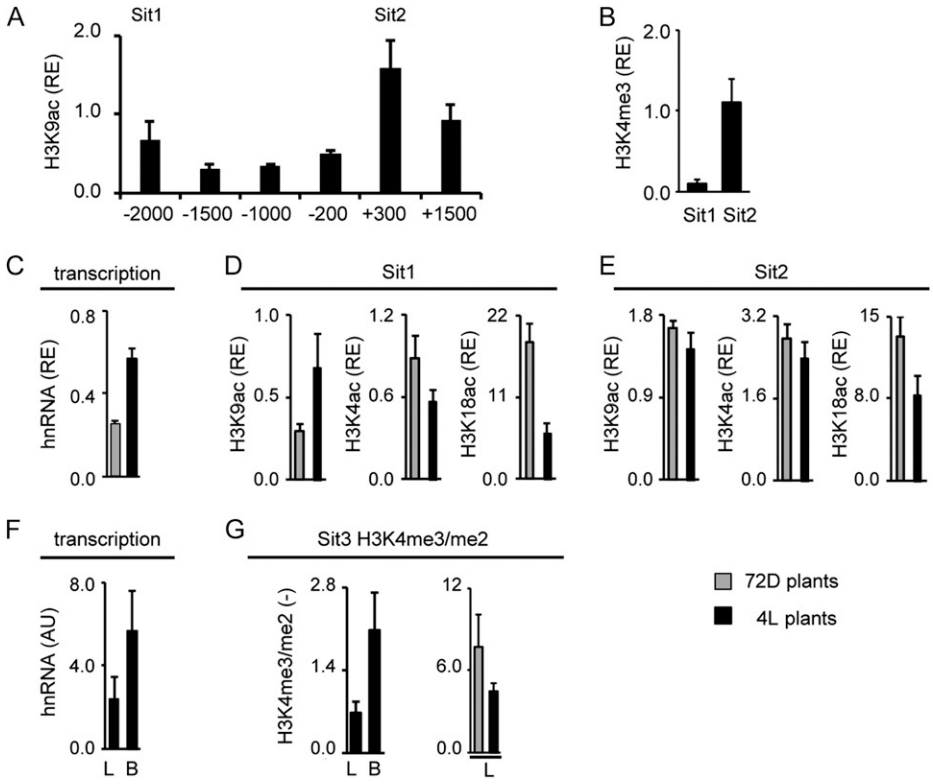


Figure 7. Histone modification profile of the C4-*Me* gene from *S. italica*. A and B, Amount of chromatin precipitated with an antibody specific for H3K9ac and H3K4me3 in illuminated leaves (4L). Positions chosen for further analyses are designated Sit1 and Sit2. Values are presented as the relative enrichment (RE) of modifications per nucleosome over modifications per nucleosome found on *S. italica Actin1*. Numbers on the *x* axis indicate bp positions relative to TIS. C, Relative quantification of C4-*Me* hnRNA expression levels in *S. italica* leaves from plants that were exposed to 72 h of darkness (72D; gray columns) and in plants that were illuminated for 4 h (4L; black columns). Transcription is standardized for *Actin1* expression (RE). hnRNA expression levels were determined by quantitative RT-PCR with a primer system specific for an intron (Supplemental Fig. S1). D and E, Light-dependent acetylation of H3K9ac, H4K5ac, and H3K18ac on positions Sit1 and Sit2. Values are presented as the relative enrichment of modifications per nucleosome over modifications per nucleosome found on *S. italica Actin1*. F, Quantification of C4-*Me* hnRNA expression levels in *S. italica* leaves (L) and B cells (B) isolated from plants that were illuminated for 4 h (4L plants). Values are arbitrary units (AU) derived from a cDNA standard dilution series. hnRNA expression levels were determined by quantitative RT-PCR with a primer system specific for an intron (Supplemental Fig. S1). G, Ratio of the amount of chromatin precipitated with an antibody specific for H3K4me3 and H3K4me2 in leaves and B cells. Black bars indicate data from illuminated plants (4L), whereas gray bars indicate data from plants that were exposed to 72 h of darkness (72D). Ratios are without dimension. All data points are based on at least three independent experiments. Vertical lines indicate SE.

common feature of the histone code on C4 genes in maize, sorghum, and *S. italica*. These results are in line with other observations in plants. Earley et al. (2007) reported that histone H4 is progressively acetylated from the inner modification Lys-16 to the N-terminal modification Lys-5. When yeast Lys residues on the N-terminal tail of H4 were replaced by Arg residues, mimicking unacetylated Lys residues, the inner Lys→Arg-16 mutation affected gene expression differently from all the other mutations, which rather showed additive effects on gene expression (Dion et al., 2005). Thus, acetylations on more C-terminal residues of the histone tails might play different roles than on N-terminal residues. However, in a genome-wide analysis of changes in histone modifications during deetiolation in Arabidopsis, the outer H3K9ac and the inner H3K27ac showed a high degree of coregulation (Charron et al., 2009). More chromatin analyses in dynamic, and not static, configurations (Roudier et al., 2009) are required to analyze whether the light-induced acetylation pattern observed on C4 genes can be generalized for all genes that respond to illumination.

H3K4me3 was always weak on the upstream promoter, but strong signals were obtained at the start of the transcribed region. Surprisingly, H3K4me3 did not respond to the light stimulus either on the six maize C4 genes or on C4-*Pepc* or C4-*Me* in sorghum and *S. italica* (Figs. 3–7). The observed pattern is unexpected, because C4 gene activity is very low in etiolated plants and plants exposed to prolonged darkness (Fig. 2) and H3K4me3 is frequently used as the key epigenetic

indicator of active genes (Santos-Rosa et al., 2002; Heintzman et al., 2007; Wang et al., 2009a). Our previous analyses had shown that high H3K4me3 (and low H3K4me2) were found on C4-*Pepc* in M cells and on C4-*Me* in B cells, suggesting a function in the establishment of cell type specificity (Danker et al., 2008). The data presented here strongly support this hypothesis, because all the C4 genes in maize, sorghum, and *S. italica* showed cell type-specific but light-independent regulation of H3K4 methylation (Figs. 3–7). Developmentally regulated H3K4 trimethylation in one of the two photosynthetic cell types, therefore, constitutes a second element of the common histone code in leaves of C4 grasses.

Promoter histone modifications rather contribute to gene regulation on the transcriptional level than on the posttranscriptional level. In accordance with the histone methylation data described above, hnRNA accumulation patterns from leaves and isolated B cells recorded in this study suggested that the cell type specificity of C4 gene expression was controlled on the transcriptional level. However, different from our observation, it had been repeatedly shown for B cell-specific genes in maize such as *RbcS* and C4-*Me* that posttranscriptional mechanisms control the cell type specificity of gene expression (Viret et al., 1994; Sheen, 1999; Brown et al., 2011). On the other hand, a reporter construct containing the promoter and the 5′ untranslated region of maize C4-*Me* was exclusively expressed in B cells of transgenic maize plants, indicating an important role of the promoter in B cell specificity (Nomura et al., 2005). In this line, transient promoter-reporter assays with maize leaves suggested that an *RbcS* promoter element, together with sequence elements in the transcribed region, contributed to the repression of gene expression in M cells (Xu et al., 2001). The gene analyzed in the study by Xu et al. (2001) is identical to *RbcS1* as defined by Ewing et al. (1998), whereas we studied *RbcS2* here. However, both *RbcS1* and *RbcS2* showed B cell-specific expression in the latter study. Together, B-specific gene expression seems to be regulated simultaneously on multiple levels. H3K4 methylation might constitute a first level of this regulation that primes genes for possible activation by other stimuli.

Chromatin patterns on the maize *RbcS2* gene investigated here differed in several respects from the other maize C4 genes analyzed. The acetylated promoter region was shorter than the acetylated region on the other promoters (Fig. 1), and the light response of acetylation was stronger at the core promoter position Zm2 than at Zm1 (Fig. 2). In addition, different from the other C4 genes, illumination contributed to the establishment of the histone methylation pattern on *RbcS2* (Fig. 3). The latter observation is in accordance with in situ hybridization studies on *RbcS* expression in etiolated maize leaves that showed basal *RbcS* expression in both M and B cells. After illumination, M cell expression was suppressed and B cell expression was enhanced. Both processes together established a cell type-specific expression pattern (Langdale et al., 1988). *RbcS* expression in both M and B cells was also observed in very young *Amaranthus* spp. leaves (Wang et al., 1992). This type of *RbcS* gene regulation would explain the relatively high H3K4me3 levels on *RbcS2* in etiolated M cells.

It is an unsolved question how the promoters of C4 genes acquired the regulatory elements necessary for efficient functioning of the C4 pathway. Within a single species, several genes must have evolved C4 expression patterns in parallel, and the demands for regulation of the new promoters such as high expression, light inducibility, and B or M specificity were highly overlapping. Thus, it is tempting to speculate that common regulatory elements were recruited by the different C4 genes. Such corecruitment was not detectable by an analysis of primary DNA sequences, although it might exist, taking the low conservation and short sequence lengths of transcription factor-binding sites (Sandelin et al., 2004) into consideration. Instead, we observed a high degree of similarity on the level of regulated histone modifications on the different C4 genes in maize. Thus, we propose that C4 promoters rather jointly acquired a histone code than a DNA code. This hypothesis is supported by the extensive conservation of this code on orthologous C4-*Pepc* and C4-*Me* genes from two separate C4 lineages, the maize/sorghum lineage and the *S. italica* lineage (Figs. 4–7). Because these two lineages evolved C4 metabolism independently (Brutnell et al., 2010), a preexisting epigenetic mechanism for promoter control was probably recruited into C4. This hypothesis is analogous to what has been proposed for regulatory DNA sequences in the transcribed region of C4 genes (Brown et al., 2010, 2011; Kajala et al., 2012). In this respect, it will be interesting to see whether light-induced and tissue-specific genes in C3 plants share the described histone code.

CONCLUSION

Analysis of histone modification profiles on C4 genes in maize revealed a common histone modification code associated with light induction and cell type-specific gene expression. Comparative modification profiling on two selected C4 genes in sorghum and *S. italica* suggested that this code is used in independent C4 lineages and, thus, was probably recruited into C4 from an ancient mechanism already existing in C3 plants.

MATERIALS AND METHODS

Plant Material and Growth Conditions

Maize (*Zea mays* 'Montello'), sorghum (*Sorghum bicolor* 'BTx623'), and *Setaria italica* 'Set20' were cultivated in growth chambers with a 16-h photoperiod and a day/night temperature regime of 25°C/20°C. Seedlings were grown in soil (VM; Einheitserde) with a photon flux density of 120 to 180 μmol m^{-2} s^{-1} or in complete darkness (etiolated plants) until the third leaf was fully

expanded. 72D plants were grown in the normal light rhythm but darkened for 3 d before harvest.

Sequence Identification

Coding sequences for maize C4 genes were derived from the literature. References for individual genes are given in Supplemental Figure S1. Nearest C4 gene homologs were identified for sorghum by Wang et al. (2009b) and for *S. italica* by Christin et al. (2007, 2009b) and Besnard et al. (2003). The corresponding genes including exon-intron predictions and surrounding genome sequences were derived from www.phytozome.net (Goodstein et al., 2012). Coordinates of the respective loci are also listed in Supplemental Figure S1.

Isolation of B Cells

For gene expression analyses in maize, B strands were isolated mechanically as described before by Hahnen et al. (2003), but without diethylether treatment. For the isolation of B strands from sorghum and *S. italica*, leaves were washed extensively in ice-cold water and homogenized in a Waring Blendor three times for 3 s each time. The mixture was sieved through a household sieve, and the homogenization step was repeated with the filter residue. The suspension was then filtered through Miracloth (VWR), and the residue was washed extensively with ice-cold water. The isolated B strands were shortly dried with paper and frozen in liquid nitrogen.

The material used for ChIP from isolated maize M and B cells was already described by Danker et al. (2008). Leaves were treated with formaldehyde as described below and afterward incubated in SMC buffer (0.5 M sorbitol, 5 mM MES, and 10 mM $CaCl_2$, pH 5.8) containing 15% (w/v) Rohament CL (AB Enzymes), 10% (w/v) Rohament PL (AB Enzymes), and 0.6% (w/v) Macerozyme R-10 (Serva) for 2.5 h at 25°C.

For ChIP analysis from isolated sorghum and *S. italica* B strands, 4 g of leaves was cross linked as described below and afterward incubated in SMC buffer containing 3% (w/v) Cellulase Onozuka R-10 (Serva) and 0.6% (w/v) Macerozyme R-10 (Serva) for 20 h at 25°C under constant agitation. M protoplasts and remaining epidermal strips were separated manually in ice-cold water. The quality of each preparation was evaluated microscopically.

ChIP

As described previously by Horst et al. (2009), 6 g of leaves from 10- to 12-d-old maize seedlings was harvested and cross linked. For sorghum and *S. italica*, 4 g of leaves from 14- to 16-d-old seedlings was harvested and vacuum infiltrated with 1% (v/v) formaldehyde instead of 3% (v/v) for maize. ChIP was performed as described by Haring et al. (2007).

The material was ground, resuspended in extraction buffer (10 mM sodium-butyrate, 400 mM Suc, 10 mM Tris-HCl, pH 8.0, 5 mM β-mercaptoethanol, 0.1 mM phenylmethylsulfonyl fluoride [PMSF], and 1× Complete [Roche Applied Science]), and incubated for 15 min at 4°C. Afterward, the solution was filtered through four layers of Miracloth (VWR), and the residue was washed with purification buffer 1 (10 mM sodium-butyrate, 250 mM Suc, 10 mM Tris-HCl, pH 8.0, 5 mM β-mercaptoethanol, 0.1 mM PMSF, 10 mM $MgCl_2$, 1% [w/v] Triton X-100, and 1× Complete) and afterward with purification buffer 2 (10 mM sodium-butyrate, 1.64 M Suc, 10 mM Tris-HCl, pH 8.0, 5 mM β-mercaptoethanol, 0.1 mM PMSF, 2 mM $MgCl_2$, 0.15% [w/v] Triton X-100, and 1× Complete). After purification, nuclei were resuspended in nuclei lysis buffer (25 mM Tris-HCl, pH 8.0, 5 mM EDTA, 0.5% [w/v] SDS, 0.1 mM PMSF, and 1× Complete).

Chromatin was sheared with a Bioruptor (Diagenode) for 10 min (setting, high; interval, 30/30 s) under constant cooling. The sheared chromatin solution was diluted 2-fold with ChIP buffer (50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 150 mM NaCl, and 0.1% [w/v] Triton X-100) and precleared with 40 μL of protein A agarose (Roche Applied Science). Precleared chromatin was split into aliquots of 400 μL for immunoprecipitation and one aliquot of 40 μL for determination of the amount of input. The chromatin aliquots were added to 30 μL of protein A agarose, and modified histones were detected with 5 μL of anti-acetyl H4K5 (07-327; Millipore), 5 μL of anti-acetyl H3K9 (07-352; Millipore), 1 μL of anti-acetyl H3K18 (07-354; Millipore), 5 μL of anti-dimethyl H3K4 (07-030; Millipore), 2.5 μL of anti-trimethyl H3K4 (04-745 [Millipore] and ab8580 [Abcam]), and 1 μL of anti-H3 C-term (ab1791; Abcam). The control serum for the determination of background precipitation was derived from rabbits immunized with an unrelated protein from potato (*Solanum tuberosum*).

After washing, the antibody-bound complexes were released and decross linked by incubation in elution buffer (62.5 mM Tris-HCl, pH 6.8, 200 mM NaCl, 2% [w/v] SDS, and 10 mM dithiothreitol) at 65°C overnight. The coprecipitated DNA was purified using the MSB Spin PCRapace kit (Invitek). Typically, 2 μL of eluted DNA was used as a template for quantitative PCR analysis.

Data Normalization

Real-time PCR signals obtained from an immunoprecipitate with an antibody directed against a specific histone acetylation or methylation were first corrected for the real-time PCR signals precipitated using a negative control serum (see above). The negative control serum signal was never more than 10% of the signal obtained with a specific antibody. The signal obtained with the antibody against an invariant domain of histone H3 (anti-H3 C-term; see above) was defined as NO. The acetylation or methylation signal at a gene position was divided by NO at the same position to obtain the modification signal per nucleosome (MN). MN is always shown as a relative enrichment compared with the MN on the promoter of the *Actin1* housekeeping gene. For H3K4me3/me2 ratios, the MN obtained with the antibody directed to H3K4me3 was divided by the MN obtained with the antibody directed to H3K4me2. The resulting ratio is dimensionless.

RNA Isolation and Reverse Transcription

Total RNA isolation was performed by phenol-chloroform extraction as described by Haring et al. (2007). About 25 to 30 mg of ground plant material was dissolved in 1 mL of Trizol and agitated for 15 min. After the addition of 0.2 volume of chloroform and agitation for 10 min, phases were separated by centrifugation (13,000 rpm, 4°C, 15 min). The aqueous phase was transferred to a new reaction tube and washed twice with 1 volume of chloroform. RNA was precipitated with 2 volumes of ice-cold ethanol (96%) for 20 min at −20°C and following centrifugation (13,000 rpm, 4°C, 15 min). After washing with 70% ethanol, the RNA was dissolved in 30 μL of water. The quality of the isolated RNA was controlled by electrophoresis, and the concentration was determined photometrically.

One unit of DNaseI (Fermentas) per microgram of RNA and $MgCl_2$ to a final concentration of 2 mM were added, and reactions were incubated for 30 min at 37°C, followed by a denaturation step of 15 min at 70°C to remove traces of contaminating DNA. cDNA synthesis was performed with approximately 1 μg of total RNA and 50 pmol of random nonamer primer. Reactions were incubated for 5 min at 70°C and cooled down on ice before adding 200 units of Moloney murine leukemia virus reverse transcriptase (Promega) and 1 mM deoxyribonucleotide triphosphates in reaction buffer as specified by the manufacturer. hnRNAs were amplified from cDNA using primer systems specific for introns (Supplemental Fig. S1). A dilution series of cDNA from illuminated leaves was used as a standard.

Quantitative PCR

Quantitative PCR was performed on an ABI PRISM 7300 sequence detection system (Life Technologies) using SYBR Green fluorescence (Platinum SYBR Green QPCR Mix; Life Technologies) for detection. Oligonucleotides were purchased from Metabion. Oligonucleotide sequences are given in Supplemental Figure S1. Amplification conditions were 2 min of initial denaturation at 95°C, followed by 40 cycles of 15 s at 95°C and 1 min at 60°C. Afterward, a melting curve was recorded. General reaction conditions were 3 mM $MgCl_2$ and 200 nM of each oligonucleotide. Sizes of the amplified molecules were confirmed by gel electrophoresis.

Sequence data from this article can be found in the GenBank/EMBL data libraries under accession numbers NM_001111948, NM_001112268, U08401.1, NM_001111843, AB018744, Y092214.1, J01238, XM_002438476, XM_002454985, XM_002456645, AF495586, FN397881, and AF288226.

Supplemental Data

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

Supplemental Figure S1. Nucleosome occupancy on the gene promoters investigated in this study.

Supplemental Table S1. Gene information and oligonucleotide sequences.

Received February 22, 2013; accepted April 4, 2013; published April 5, 2013.

LITERATURE CITED

Adrian J, Farrona S, Reimer JJ, Albani MC, Coupland G, Turck F (2010) cis-Regulatory elements and chromatin state coordinately control temporal and spatial expression of FLOWERING LOCUS T in *Arabidopsis*. Plant Cell **22:** 1425–1440

Akyildiz M, Gowik U, Engelmann S, Koczor M, Streubel M, Westhoff P (2007) Evolution and function of a cis-regulatory module for mesophyll-specific gene expression in the C4 dicot *Flaveria trinervia*. Plant Cell **19:** 3391–3402

Bannister AJ, Kouzarides T (2011) Regulation of chromatin by histone modifications. Cell Res **21:** 381–395

Bassett CL, Rawson JR, Jernstedt JA (1988) DNA and RNA levels in bundle sheath and mesophyll cells of pearl millet (*Pennisetum americanum*). Plant Physiol **87:** 307–310

Berger SL (2007) The complex language of chromatin regulation during transcription. Nature **447:** 407–412

Bernstein BE, Meissner A, Lander ES (2007) The mammalian epigenome. Cell **128:** 669–681

Besnard G, Pinçon G, D'Hont A, Hoarau JY, Cadet F, Offmann B (2003) Characterisation of the phosphoenolpyruvate carboxylase gene family in sugarcane (Saccharum spp.). Theor Appl Genet **107:** 470–478

Brown NJ, Newell CA, Stanley S, Chen JE, Perrin AJ, Kajala K, Hibberd JM (2011) Independent and parallel recruitment of preexisting mechanisms underlying C_4 photosynthesis. Science **331:** 1436–1439

Brown NJ, Palmer BG, Stanley S, Hajaji H, Janacek SH, Astley HM, Parsley K, Kajala K, Quick WP, Trenkamp S, et al (2010) C acid decarboxylases required for C photosynthesis are active in the mid-vein of the C species Arabidopsis thaliana, and are important in sugar and amino acid metabolism. Plant J **61:** 122–133

Brutnell TP, Wang L, Swartwood K, Goldschmidt A, Jackson D, Zhu XG, Kellogg E, Van Eck J (2010) *Setaria viridis*: a model for C4 photosynthesis. Plant Cell **22:** 2537–2544

Charron J-BF, He H, Elling AA, Deng XW (2009) Dynamic landscapes of four histone modifications during deetiolation in *Arabidopsis*. Plant Cell **21:** 3732–3748

Christin P-A, Salamin N, Kellogg EA, Vicentini A, Besnard G (2009a) Integrating phylogeny into studies of C4 variation in the grasses. Plant Physiol **149:** 82–87

Christin P-A, Salamin N, Savolainen V, Duvall MR, Besnard G (2007) C4 photosynthesis evolved in grasses via parallel adaptive genetic changes. Curr Biol **17:** 1241–1247

Christin P-A, Samaritani E, Petitpierre B, Salamin N, Besnard G (2009b) Evolutionary insights on C4 photosynthetic subtypes in grasses from genomics and phylogenetics. Genome Biol Evol **1:** 221–230

Danker T, Dreesen B, Offermann S, Horst I, Peterhänsel C (2008) Developmental information but not promoter activity controls the methylation state of histone H3 lysine 4 on two photosynthetic genes in maize. Plant J **53:** 465–474

Delany AM (2001) Measuring transcription of metalloproteinase genes: nuclear run-off assay vs analysis of hnRNA. Methods Mol Biol **151:** 321–333

Dion MF, Altschuler SJ, Wu LF, Rando OJ (2005) Genomic characterization reveals a simple histone H4 acetylation code. Proc Natl Acad Sci USA **102:** 5501–5506

Doust AN, Kellogg EA, Devos KM, Bennetzen JL (2009) Foxtail millet: a sequence-driven grass model system. Plant Physiol **149:** 137–141

Earley KW, Shook MS, Brower-Toland B, Hicks L, Pikaard CS (2007) In vitro specificities of Arabidopsis co-activator histone acetyltransferases: implications for histone hyperacetylation in gene activation. Plant J **52:** 615–626

Elferink CJ, Reiners JJ Jr (1996) Quantitative RT-PCR on CYP1A1 heterogeneous nuclear RNA: a surrogate for the in vitro transcription run-on assay. Biotechniques **20:** 470–477

Engelmann S, Bläsing OE, Gowik U, Svensson P, Westhoff P (2003) Molecular evolution of C4 phosphoenolpyruvate carboxylase in the genus Flaveria: a gradual increase from C3 to C4 characteristics. Planta **217:** 717–725

Ewing RM, Jenkins GI, Langdale JA (1998) Transcripts of maize RbcS genes accumulate differentially in C3 and C4 tissues. Plant Mol Biol **36:** 593–599

Francis KE, Spiker S (2005) Identification of Arabidopsis thaliana transformants without selection reveals a high occurrence of silenced T-DNA integrations. Plant J **41:** 464–477

Furumoto T, Hata S, Izui K (1999) cDNA cloning and characterization of maize phosphoenolpyruvate carboxykinase, a bundle sheath cell-specific enzyme. Plant Mol Biol **41:** 301–311

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, et al (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res **40:** D1178–D1186

Gowik U, Burscheidt J, Akyildiz M, Schlue U, Koczor M, Streubel M, Westhoff P (2004) cis-Regulatory elements for mesophyll-specific gene expression in the C4 plant *Flaveria trinervia*, the promoter of the C4 phosphoenolpyruvate carboxylase gene. Plant Cell **16:** 1077–1090

Hahnen S, Joeris T, Kreuzaler F, Peterhänsel C (2003) Quantification of photosynthetic gene expression in maize C(3) and C(4) tissues by real-time PCR. Photosynth Res **75:** 183–192

Haring M, Offermann S, Danker T, Horst I, Peterhänsel C, Stam M (2007) Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization. Plant Methods **3:** 11

Hassan AH, Awad S, Al-Natour Z, Othman S, Mustafa F, Rizvi TA (2007) Selective recognition of acetylated histones by bromodomains in transcriptional co-activators. Biochem J **402:** 125–133

Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, Van Calcar S, Qu C, Ching KA, et al (2007) Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet **39:** 311–318

Henikoff S, Shilatifard A (2011) Histone modification: cause or cog? Trends Genet **27:** 389–396

Hibberd JM, Covshoff S (2010) The regulation of gene expression required for C4 photosynthesis. Annu Rev Plant Biol **61:** 181–207

Horst I, Offermann S, Dreesen B, Niessen M, Peterhansel C (2009) Core promoter acetylation is not required for high transcription from the phospho*enol*pyruvate carboxylase promoter in maize. Epigenetics Chromatin **2:** 17

Jenkins CL, Boag S (1985) Isolation of bundle sheath cell chloroplasts from the NADP-ME type C4 plant *Zea mays*: capacities for CO_2 assimilation and malate decarboxylation. Plant Physiol **79:** 84–89

Kajala K, Brown NJ, Williams BP, Borrill P, Taylor LE, Hibberd JM (2012) Multiple Arabidopsis genes primed for recruitment into C_4 photosynthesis. Plant J **69:** 47–56

Kim J, Daniel J, Espejo A, Lake A, Krishna M, Xia L, Zhang Y, Bedford MT (2006) Tudor, MBT and chromo domains gauge the degree of lysine methylation. EMBO Rep **7:** 397–403

Kim S-I, Veena , Gelvin SB (2007) Genome-wide analysis of Agrobacterium T-DNA integration sites in the Arabidopsis genome generated under non-selective conditions. Plant J **51:** 779–791

Kingston RE, Narlikar GJ (1999) ATP-dependent remodeling and acetylation as regulators of chromatin fluidity. Genes Dev **13:** 2339–2352

Kouzarides T (2007) Chromatin modifications and their function. Cell **128:** 693–705

Ku MSB, Kano-Murakami Y, Matsuoka M (1996) Evolution and expression of C4 photosynthesis genes. Plant Physiol **111:** 949–957

Langdale JA (2011) C4 cycles: past, present, and future research on C4 photosynthesis. Plant Cell **23:** 3879–3892

Langdale JA, Zelitch I, Miller E, Nelson T (1988) Cell position and light influence C4 versus C3 patterns of photosynthetic gene expression in maize. EMBO J **7:** 3643–3651

Lee CK, Shibata Y, Rao B, Strahl BD, Lieb JD (2004) Evidence for nucleosome depletion at active regulatory regions genome-wide. Nat Genet **36:** 900–905

Martin C, Zhang Y (2005) The diverse functions of histone lysine methylation. Nat Rev Mol Cell Biol **6:** 838–849

Mellor J (2005) The dynamics of chromatin remodeling at promoters. Mol Cell **19:** 147–157

Müller-Tidow C, Klein HU, Hascher A, Isken F, Tickenbrock L, Thoennissen N, Agrawal-Singh S, Tschanter P, Disselhoff C, Wang Y, et al (2010) Profiling of histone H3 lysine 9 trimethylation levels predicts transcription factor activity and survival in acute myeloid leukemia. Blood **116:** 3564–3571

Nelissen H, Boccardi TM, Himanen K, Van Lijsebettens M (2007) Impact of core histone modifications on transcriptional regulation and plant growth. Crit Rev Plant Sci **26:** 243–263

Nishida H, Suzuki T, Kondo S, Miura H, Fujimura Y, Hayashizaki Y (2006) Histone H3 acetylated at lysine 9 in promoter is associated with

low nucleosome density in the vicinity of transcription start site in human cell. Chromosome Res **14:** 203–211

Nomura M, Higuchi T, Ishida Y, Ohta S, Komari T, Imaizumi N, Miyao-Tokutomi M, Matsuoka M, Tajima S (2005) Differential expression pattern of C4 bundle sheath expression genes in rice, a C3 plant. Plant Cell Physiol **46:** 754–761

Offermann S, Danker T, Dreymüller D, Kalamajka R, Töpsch S, Weyand K, Peterhänsel C (2006) Illumination is necessary and sufficient to induce histone acetylation independent of transcriptional activity at the C_4-specific phosphoenolpyruvate carboxylase promoter in maize. Plant Physiol **141:** 1078–1088

Offermann S, Dreesen B, Horst I, Danker T, Jaskiewicz M, Peterhansel C (2008) Developmental and environmental signals induce distinct histone acetylation profiles on distal and proximal promoter elements of the C4-Pepc gene in maize. Genetics **179:** 1891–1901

Patel M, Berry JO (2008) Rubisco gene expression in C4 plants. J Exp Bot **59:** 1625–1634

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, et al (2009) The Sorghum bicolor genome and the diversification of grasses. Nature **457:** 551–556

Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolsheimer E, et al (2005) Genome-wide map of nucleosome acetylation and methylation in yeast. Cell **122:** 517–527

Roudier F, Teixeira FK, Colot V (2009) Chromatin indexing in Arabidopsis: an epigenomic tale of tails and more. Trends Genet **25:** 511–517

Sage RF, Christin P-A, Edwards EJ (2012) The C4 plant lineages of planet Earth. J Exp Bot **62:** 3155–3169

Sandelin A, Alkema W, Engström P, Wasserman WW, Lenhard B (2004) JASPAR: an open-access database for eukaryotic transcription factor binding profiles. Nucleic Acids Res **32:** D91–D94

Santos-Rosa H, Schneider R, Bannister AJ, Sherriff J, Bernstein BE, Emre NC, Schreiber SL, Mellor J, Kouzarides T (2002) Active genes are trimethylated at K4 of histone H3. Nature **419:** 407–411

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, et al (2009) The B73 maize genome: complexity, diversity, and dynamics. Science **326:** 1112–1115

Sheen J (1990) Metabolic repression of transcription in higher plants. Plant Cell **2:** 1027–1038

Sheen J (1999) C4 gene expression. Annu Rev Plant Physiol Plant Mol Biol **50:** 187–217

Sheen JY, Bogorad L (1987) Differential expression of C4 pathway genes in mesophyll and bundle sheath cells of greening maize leaves. J Biol Chem **262:** 11726–11730

Sims RJ III, Nishioka K, Reinberg D (2003) Histone lysine methylation: a signature for chromatin function. Trends Genet **19:** 629–639

Sugiharto B, Burnell JN, Sugiyama T (1992) Cytokinin is required to induce the nitrogen-dependent accumulation of mRNAs for phosphoenolpyruvate carboxylase and carbonic anhydrase in detached maize leaves. Plant Physiol **100:** 153–156

Tan M, Luo H, Lee S, Jin F, Yang JS, Montellier E, Buchou T, Cheng Z, Rousseaux S, Rajagopal N, et al (2011) Identification of 67 histone marks and histone lysine crotonylation as a new type of histone modification. Cell **146:** 1016–1028

Tolley BJ, Woodfield H, Wanchana S, Bruskiewich R, Hibberd JM (2012) Light-regulated and cell-specific methylation of the maize PEPC promoter. J Exp Bot **63:** 1381–1390

Vicentini A, Barber J, Aliscioni S, Ciussani L, Kellogg E (2008) The age of the grasses and clusters of origins of C4 photosynthesis. Glob Change Biol **14:** 2963–2977

Viret J-F, Mabrouk Y, Bogorad L (1994) Transcriptional photoregulation of cell-type-preferred expression of maize rbcS-m3: 3′ and 5′ sequences are involved. Proc Natl Acad Sci USA **91:** 8577–8581

von Caemmerer S, Furbank RT (2003) The C(4) pathway: an efficient $CO_{(2)}$ pump. Photosynth Res **77:** 191–207

Wang JL, Klessig DF, Berry JO (1992) Regulation of C4 gene expression in developing Amaranth leaves. Plant Cell **4:** 173–184

Wang X, Elling AA, Li X, Li N, Peng Z, He G, Sun H, Qi Y, Liu XS, Deng XW (2009a) Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and small RNA transcriptomes in maize. Plant Cell **21:** 1053–1069

Wang X, Gowik U, Tang H, Bowers JE, Westhoff P, Paterson AH (2009b) Comparative genomic analysis of C4 photosynthetic pathway evolution in grasses. Genome Biol **10:** R68

Wingler A, Walker RP, Chen ZH, Leegood RC (1999) Phosphoenolpyruvate carboxykinase is involved in the decarboxylation of aspartate in the bundle sheath of maize. Plant Physiol **120:** 539–546

Wu H, Zhang X, Ling G, D'Agostino J, Ding X (2009) Mechanisms of differential expression of the CYP2A13 7520C and 7520G alleles in human lung: allelic expression analysis for CYP2A13 heterogeneous nuclear RNA, and evidence for the involvement of multiple cis-regulatory single nucleotide polymorphisms. Pharmacogenet Genomics **19:** 852–863

Xu T, Purcell M, Zucchi P, Helentjaris T, Bogorad L (2001) TRM1, a YY1-like suppressor of rbcS-m3 expression in maize mesophyll cells. Proc Natl Acad Sci USA **98:** 2295–2300

Yamasaki S, Oda M, Daimon H, Mitsukuri K, Johkan M, Nakatsuka T, Nishihara M, Mishiba K (2011) Epigenetic modifications of the 35S promoter in cultured gentian cells. Plant Sci **180:** 612–619

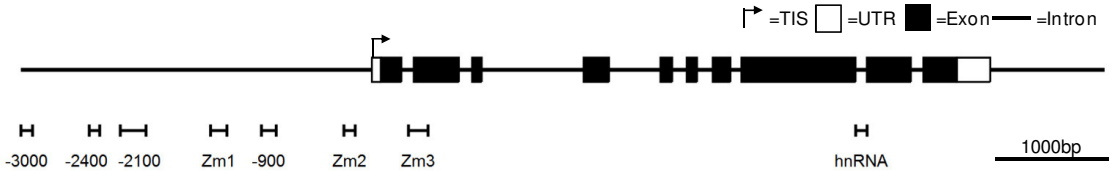
Yan C, Boyd DD (2006) Histone H3 acetylation and H3 K4 methylation define distinct chromatin regions permissive for transgene expression. Mol Cell Biol **26:** 6357–6371

Yin Z, Kong QR, Zhao ZP, Wu ML, Mu YS, Hu K, Liu ZH (2012) Position effect variegation and epigenetic modification of a transgene in a pig model. Genet Mol Res **11:** 355–369

Zhang X (2008) The epigenetic landscape of plants. Science **320:** 489–492

Supplemental data 1: Gene information and oligonucleotide sequences.

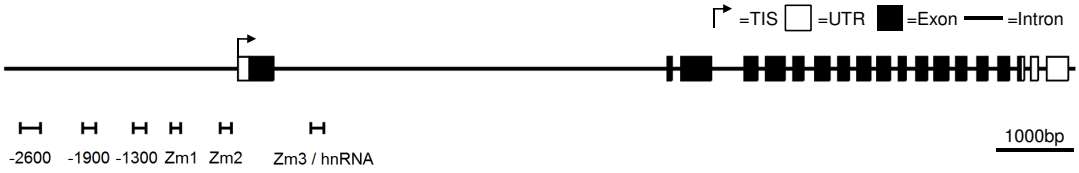
***Zea mays* – Phosphoenolpyruvate carboxylase (C4-*Pepc*)**	
Locus	GRMZM2G083841
Transcript variant	T01
Chromosome	9
Next upstream gene	~ 30 kb
Reference/Identification	genomic locus homologous to mRNA NM_001111948 (Wang et al., 2009) was identified by whole genome BLAST on www.phytozome.net (Goodstein et al., 2012)



Name	bp relative to TIS	Sequence (5'→3')
	-3000	GTATTGTGATGGTGACCCTAGGAAC
		TGTTTATTTGGGATGTACTTCCTTTG
	-2400	TATCCTTCTGCCTAGGTTGAGTAGCT
		TGTTGACACCAAATCCTAACCAAA
	-2100	GTCACAATTGAAGATTCGTGCAAGG
		CAGTTTGAACTAAACGACTTCCAAC
Zm1	-1300	GTACAAATGAGGTGCCGGATTGATG
		CGGCCATGGCATGATACAATTCTCA
	-900	CAAGTGCCAACAACACATCGC
		GAAGGGCACCATACATATAGGG
Zm2	-200	CGATTGCCGCCAGCAGT
		GAACCGGCTGTGGCTGAG *
Zm3	+420	GCTCGTGTCGTGTGCTCGCT *
		ATGGAGCTCGCCACGAGGATGG
hnRNA	+4300	GTATGCTGCCATTGCCCATTGC
		TAGCCTGATAGTGAGTGACGCACA

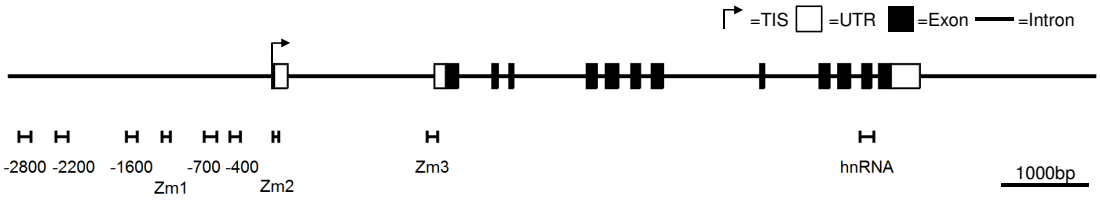
* oligonucleotide shows mismatches to database genome sequence, but efficiently amplifies DNA from the genotype used in this work

***Zea mays* – Pyruvate phosphate dikinase (C4-*Ppdk*)**	
Locus	GRMZM2G306345
Transcript variant	as described by (Sheen, 1991)
Chromosome	6
Next upstream gene	~ 34 kb
Reference/Identification	genomic locus homologous to mRNA NM_001112268 (Wang et al., 2009) was identified by whole genome BLAST on www.phytozome.net (Goodstein et al., 2012)



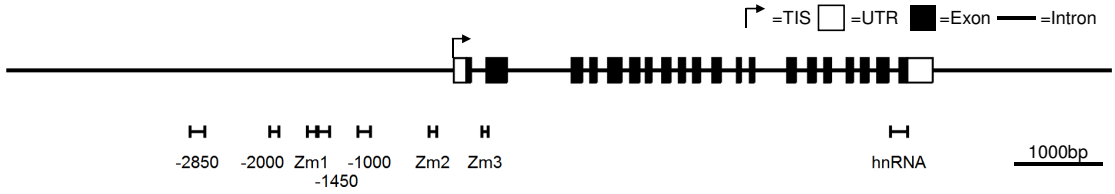
Name	bp relative to TIS	Sequence (5'→3')
	-2600	CGATCATCTCTCCAGTCAACTG
		CAAGCTCAGGGTGCTAAAATCAC
	-1900	GATTCATCAGTAGTTAGACTTAGTC
		CCTGGTAAGTCTTCATTCATAACC
	-1300	AGGGGTATTGTGAACAAGAGGATG
		CCAATTCCTCGCAAAGACACTTCAC
Zm1	-800	TGGAGGCGTTGGCTAAAGTAC
		AGAGGTAAATCAGATGACTACAAAAGAAAG
Zm2	-150	CACTATAGCCACTCGCCGCAAG
		CTGCTCACCTTATCCCGGACGT
Zm3	+900	CGTGTCAAGGTGTCCTCGCAAG
		CACAGGTGTTGTAACGCAAACGTTG
hnRNA	+900	CGTGTCAAGGTGTCCTCGCAAG
		CACAGGTGTTGTAACGCAAACGTTG

***Zea mays* – Carbonic anhydrase (C4-*Ca*)**	
Locus	GRMZM2G121878
Transcript variant	T02
Chromosome	3
Next upstream gene	~ 4 kb
Reference/Identification	genomic locus homologous to mRNA U08401.1 (Wang et al., 2009) was identified by whole genome BLAST on www.phytozome.net (Goodstein et al., 2012)



Name	bp relative to TIS	Sequence (5'→3')
	-2800	GAGGCGGCGGGAACTCC
		CGACGTGAGGTGTTCGGTG
	-2200	CATGCACGACAAAGGGAAAACG
		GGCGACATCATAAGCACATGAG
	-1600	CTATACCACCCCTCACTTGTTCTG
		GATTGGCTGGCTGCTCATGTAC
Zm1	-1200	GATCTGACAGCACCACCGAAC
		GTTCTAGGCATCATTCATCATCACG
	-700	CAGGGTCAGGGAGACCGC
		CTACGAGAGAGACGTGCTTAC
	-400	GCAGTAGCATCGCGTCCAC
		GGGAAATGATGAAACGCGCGG
Zm2	0	CGGCACTCGCACGATCAATG
		GCGAGGCTGGCGACGATG
Zm3	+1800	TGGGCGCGCGCG CGTG
		CTCTCGTTTGACTCCTCAGCTGC
hnRNA	+6700	GGTTTGGTGTGGTGTACGTACG
		GTAACTGCTCGACGAATGTACAAC

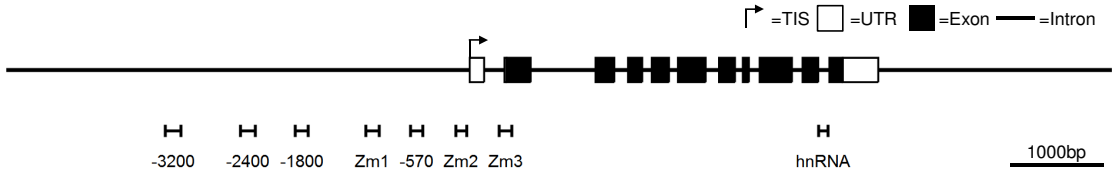
***Zea mays* – Malic Enzyme (C4-*Me*)**	
Locus	GRMZM2G085019
Transcript variant	T01
Chromosome	3
Next upstream gene	~ 13 kb
Reference/Identification	genomic locus homologous to mRNA NM_001111843 (Wang et al., 2009) was identified by whole genome BLAST on www.phytozome.net (Goodstein et al., 2012)



Name	bp relative to TIS	Sequence (5'→3')
	-2850	CGTTCAGGTAAAGACACGCAAACTC
		GGTATATGTTCTTGATCTGGTGTG
	-2000	TGATGGCTACAGTTTGCCGCTAG
		TAG CAGCCGCGCCCTTCATC
Zm1	-1600	TGCACCGTTTCGTCCTGAGC
		TGCGTCCAATTGATCGGCAC
	-1450	GACTGGTGAAAAGATTCAACTTCGC
		CATATCGTACCACCACTGCCTC
	-1000	CGTCTTCTTCCAGAGGCGG
		CTGATCACCGATAGAAAAGCGG
Zm2	-240	GGATATGATCGTCCCGCCCAACG
		GCTGCCGACCACGGGTATTGAC
Zm3	+400	CAGGTTGCCACCGCCTCATC
		CGCTTCGTCCTCCCTGCTT *
hnRNA	+4900	GCAGCACTACCGGTAGTTGCGG
		GTTTGGCTTTGCTTTGCTTTGC *

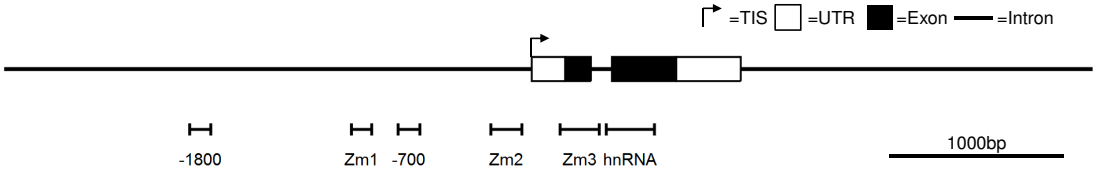
* oligonucleotide shows mismatches to database genome sequence, but efficiently amplifies DNA from the genotype used in this work

***Zea mays* – Phosphoenolpyruvate carboxykinase (C4-*Pepck*)**	
Locus	GRMZM2G001696
Transcript variant	T01
Chromosome	1
Next upstream gene	~ 18 kb
Reference/Identification	genomic locus homologous to mRNA AB018744 (Furumoto et al., 1999) was identified by whole genome BLAST on www.phytozome.net (Goodstein et al., 2012)



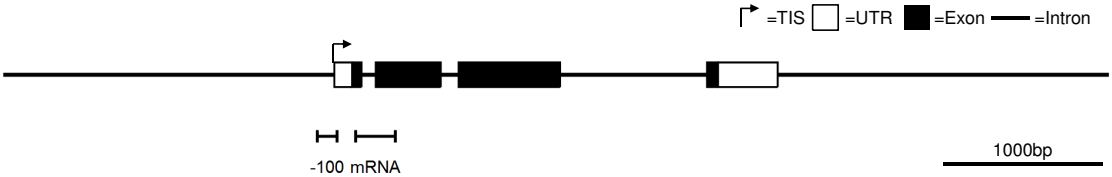
Name	bp relative to TIS	Sequence (5'→3')
	-3200	GACACAAAGAGGATCAATTACAGAG
		CTTTGTAAGCCGCAGGAGATAAC
	-2400	CCGGTATGGTCACATTATCTGTG
		CACGACCGTAAAAACTTATATAGTAG
	-1800	GCTGTCATATGCGTGGTACCC
		GAAATCCAACCACAACCAAGAGG
Zm1	-1200	GATTTACTCTCATGAGCGCCATATGG
		CGCTCTAAACCTGATGCTCCTAG
	-570	GTGGTGTCGGCGCAGTCTG
		GAAATAGAACACGCAACCTACAGATTC
Zm2	-100	GAGTATTAGCAAGCATACAGGAGT
		CACTCTGCAGGAGCAGCAG
Zm3	+400	TGGTCCCATCCCAGCAGGG
		GTCCTGCTTCTTCTTCCCGG
hnRNA	+3800	GCTCTGAGTCTCACCTCACG
		CATTTGTCAAGTTCGAGATTGGTC

***Zea mays* – Ribulose-1,5-bisphosphate carboxylase oxygenase (C4-*RbcS*2)**	
Locus	GRMZM2G113033
Transcript variant	T01
Chromosome	2
Next upstream gene	>100 kb
Reference/Identification	genomic locus homologous to mRNA Y092214.1 (Ewing et al., 1998) was identified by whole genome BLAST on www.phytozome.net (Goodstein et al., 2012)



Name	bp relative to TIS	Sequence (5'→3')
	-1800	CTACACCACACCCTCCAAATAG
		CATGGTCGCGCGGTTGGG
Zm1	-1100	GGCTGATGTTAGCGCTATACTCTG
		CGAAGTGTGCCTTTTAGATTACATGC
	-700	CTTTTTCATAATAATGGCTGAGGCG
		CGGACCCGCCGAAATTTCAG
Zm2	-190	CCTAGTTTCCATTGTCGTACGTTC
		GCCACCACTTGTCGCCTTATCG
Zm3	+300	CATACTAGCCAGCCTGCCAGC
		GCTGCAGCCTGCAAAGAAGATG
hnRNA	+570	CGATGATGTACCATGTGTGTGCG
		GTACACCTGCGTGGCGTCGG

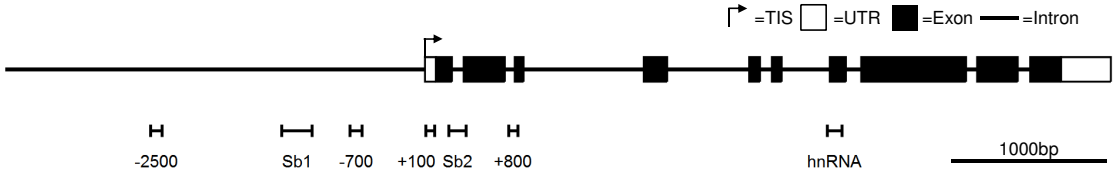
***Zea mays* – *Actin*1**	
Locus	GRMZM2G126010
Transcript variant	T03
Chromosome	8
Next upstream gene	~ 16 kb
Reference/Identification	genomic locus homologous to mRNA J01238 (Haring et al., 2007) was identified by whole genome BLAST on www.phytozome.net (Goodstein et al., 2012)



Name	bp relative to TIS	Sequence (5'→3')
	-100	TTTAAGGCTGCTGTACTGCTGTAGA
		CACTTTCTGCTCATGGTTTAAGG
mRNA	+170 **	CCTATCGTATGTGACAATGGCACT
		GCCTCATCACCTACGTAGGCAT

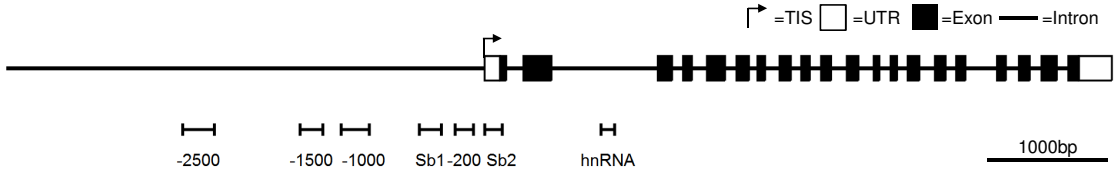
** intron not included

***Sorghum bicolor* – Phosphoenolpyruvate carboxylase (C4-*Pepc*)**	
Locus	Sb10g021330
Transcript variant	Sb10g021330.1
Chromosome	10
Next upstream gene	>100 kb
Reference/Identification	genomic locus homologous to mRNA Sb10g021330 (Wang et al., 2009) was identified by whole genome BLAST on www.phytozome.net (Goodstein et al., 2012)



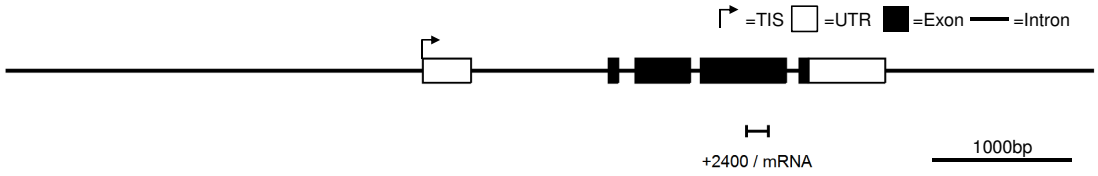
Name	bp relative to TIS	Sequence (5'→3')
	-2500	AGTTACTTCATTATCATAAATTTCTTGGCC
		GGTTTAGAAGATTTTGGCCATGAAGAC
Sb1	-1400	CGGGACATGTAATAAGGAGTTAGG
		GGTGGTGGTGAAGATATGCGG
	-700	GCCTTCCTCCAGCGCCATGCATCCTC
		CTGTTTGCAGTCAAGGCCGGATTCTGGGG
	-100	GCATGCCTTTCCAATCCCGCG
		CAATGCAGGGCGCCGGCC
Sb2	+300	GGACCTCCATGGCCCCAGCCTTCGCG
		CGACACCTCGTAGCACTCCTGGACC
	+600	CCAACCTGGCGGAGGAAGTGG
		GGACTCGGTGGTGGCGGAG
hnRNA	+3900	AAGAGTATTTGATGCTGGCGCAG
		CAGAAAATCCAGTTGCCAGCAG

***Sorghum bicolor* – Malic Enzyme (C4-*Me*)**	
Locus	Sb03g003230
Transcript variant	Sb03g003230.1
Chromosome	3
Next upstream gene	~ 11 kb
Reference/Identification	genomic locus homologous to mRNA Sb03g003230 (Wang *et al.*, 2009) was identified by whole genome BLAST on www.phytozome.net (Goodstein *et al.*, 2012)



Name	bp relative to TIS	Sequence (5'→3')
	-2500	CGCCACCTTGCGCCACCTCT
		GGACTCGATAGGGCATGGTATGC
	-1500	GTTAAGGACATGTTCAACAAATGCAA
		AGGCAGCAAGAGCTAGCCATGC
	-1000	CCACCTAAAACCTTGAGTCCTCACAAA
		GCACCATGGAATGAAAAGCTACTATTTT
Sb1	-500	GTCGTTGTTAGTGTACGTGGCACAAG
		GCTTACACTTCCAAAAAACAAGCGCC
	-200	GCGGCGTTCTAGTTTTCCGCGT
		GGTCAGTCCCAAGGTTCAGCAAAC
Sb2	+100	TCCACACTACTACTGCCCCTG
		GGCGCGAGCGGAGATCATGGT
hnRNA		GCGTTGTTGTAGAATTCTGAATCGAGT
		AGGACAAATCTAAAGCAAAGCAGACAA

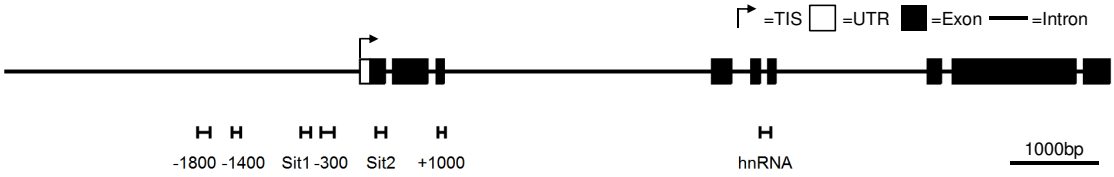
Sorghum bicolor* – *Actin	
Locus	Sb03g040880
Transcript variant	Sb03g040880.1
Chromosome	3
Next upstream gene	~ 4 kb
Reference/Identification	nearest homologue to *Actin*1 from *Zea mays* on www.phytozome.net (Goodstein et al., 2012)



Name	bp relative to TIS	Sequence (5'→3')
	+2400	GTGCTATTCCAGCCATCCTTCATTGG
		GCGGTCAGCAATACCAGGGAAC *
mRNA	+2400	GTGCTATTCCAGCCATCCTTCATTGG
		GCGGTCAGCAATACCAGGGAAC *

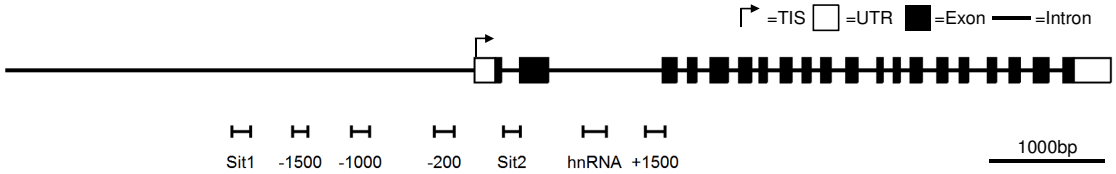
* oligonucleotide shows mismatches to database genome sequence, but efficiently amplifies DNA from the genotype used in this work

***Setaria italica* – Phosphoenolpyruvate carboxylase (C4-*Pepc*)**	
Locus	Si005789m.g
Transcript variant	Si005789m
Chromosome	scaffold 4
Next upstream gene	~ 7 kb
Reference/Identification	genomic locus homologous to mRNA AF495586 (Besnard et al., 2003; Christin et al., 2007, suppl. table1) was identified by whole genome BLAST on www.phytozome.net (Goodstein et al., 2012)



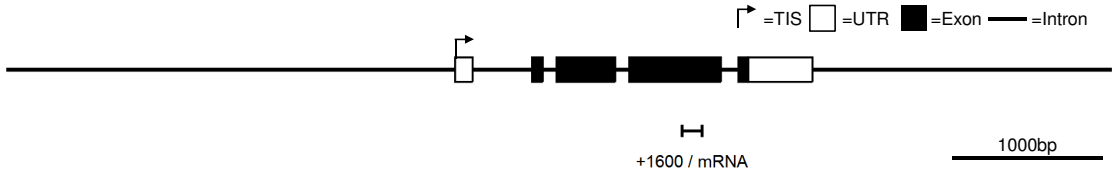
Name	bp relative to TIS	Sequence (5'→3')
	-1800	CGGAAGACAACATCATGCATGTGC
		ATGGATGCTCAGCCAGCTACC
	-1400	ATCAATGCCCTCGGCGCCAATC
		GATTATATATCGGCTCCATGTGTG
Sit1	-600	GAAAAAAACCTGGGAACAAGCC
		GTCTTCCTCTTCCCCTACCCG
	-300	GGGTATGTGGCTGGCTGTGCA
		GGACTCAAGCATGTGCTATATAGGAC
Sit2	+300	GGCAAGGTCTCCGAGGACGAC
		AGGACGATTACGAATTCACGGATG
	+1000	GCAGCTGTATGCCAAGGACATCAC
		GTGTTGAAATGTAAGTACCTCCCTC
hnRNA	+4500	CGAGGTATTGTAGAAATGTGTTTGA
		CACGGAGTTCGTCATTGCAGCG

***Setaria italica* – Malic Enzyme (C4-*Me*)**	
Locus	Si000645m.g
Transcript variant	Si000645m
Chromosome	scaffold 5
Next upstream gene	~ 15 kb
Reference/Identification	nearest homologue to *Malic Enzyme* from *Setaria viridis* FN397881(Christin et al., 2009, suppl. table1) was identified by whole genome BLAST on www.phytozome.net (Goodstein *et al.*, 2012)



Name	bp relative to TIS	Sequence (5'→3')
Sit1	-2000	TTTGACAATGTGGTGCTACATATTTAG
		CAGTCTCTTGAATGTGTCGTAAAC
	-1500	ATAAGGCACAAACCTCCTCAAAACC
		CCGCTGTCGAGCACATGTCG
	-1000	GCCGCTACAACAACGTTGTAC
		CATAGCTGAAATCACACTATGTGG
	-200	TGATCAAATGGTTGGTCAGGACCG
		GTTAGTCGGCTGGAGATGGAAT
Sit2	+300	GCCTATACCCCCTTACCGTTTCC
		GCCTTCCACTGCGGAGACAAAAAAA
	+1500	CAATATGACCATGCCACCCAAAG
		GATCCCTCAAAAGGGTGTAACCAC
hnRNA	+1100	GGATTCTCCCTTCACCTACGTTTAC
		GACCTACGCAAATCTGATTCCTAAACT

Setaria italica* – *Actin	
Locus	Si010361m.g
Transcript variant	Si010361m
Chromosome	scaffold 7
Next upstream gene	~ 3 kb
Reference/Identification	nearest homologue to *Actin*1 from *Zea mays* on www.phytozome.net (Goodstein et al., 2012)



Name	bp relative to TIS	Sequence (5'→3')
	+1600	GGTATGGAGTCGCCTGGAATCC
		GCGGTCAGCAATACCAGGGAAC
mRNA	+1600	GGTATGGAGTCGCCTGGAATCC
		GCGGTCAGCAATACCAGGGAAC

Literature cited

Besnard G, Pinçon G, D'Hont A, Hoarau JY, Cadet F, Offmann B (2003) Characterisation of the phosphoenolpyruvate carboxylase gene family in sugarcane (Saccharum spp.). TAG **107:** 470-478

Christin P-A, Salamin N, Savolainen V, Duvall M, Besnard G (2007) C4 Photosynthesis evolved in grasses via parallel adaptive genetic changes. Curr Biol **17:** 1241-1247

Christin P-A, Samaritani E, Petitpierre B, Salamin N, Besnard G (2009) Evolutionary Insights on C4 Photosynthetic Subtypes in Grasses from Genomics and Phylogenetics. Genome Biol Evol **2009:** 221-230

Ewing RM, Jenkins GI, Langdale JA (1998) Transcripts of maize RbcS genes accumulate differentially in C3 and C4 tissues. Plant Mol Biol **36:** 593-599

Furumoto T, Hata S, Izui K (1999) cDNA cloning and characterization of maize phosphoenolpyruvate carboxykinase, a bundle sheath cell-specific enzyme. Plant Mol Biol **41:** 301-311.

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N, Rokhsar DS (2012) Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res **40:** D1178-1186

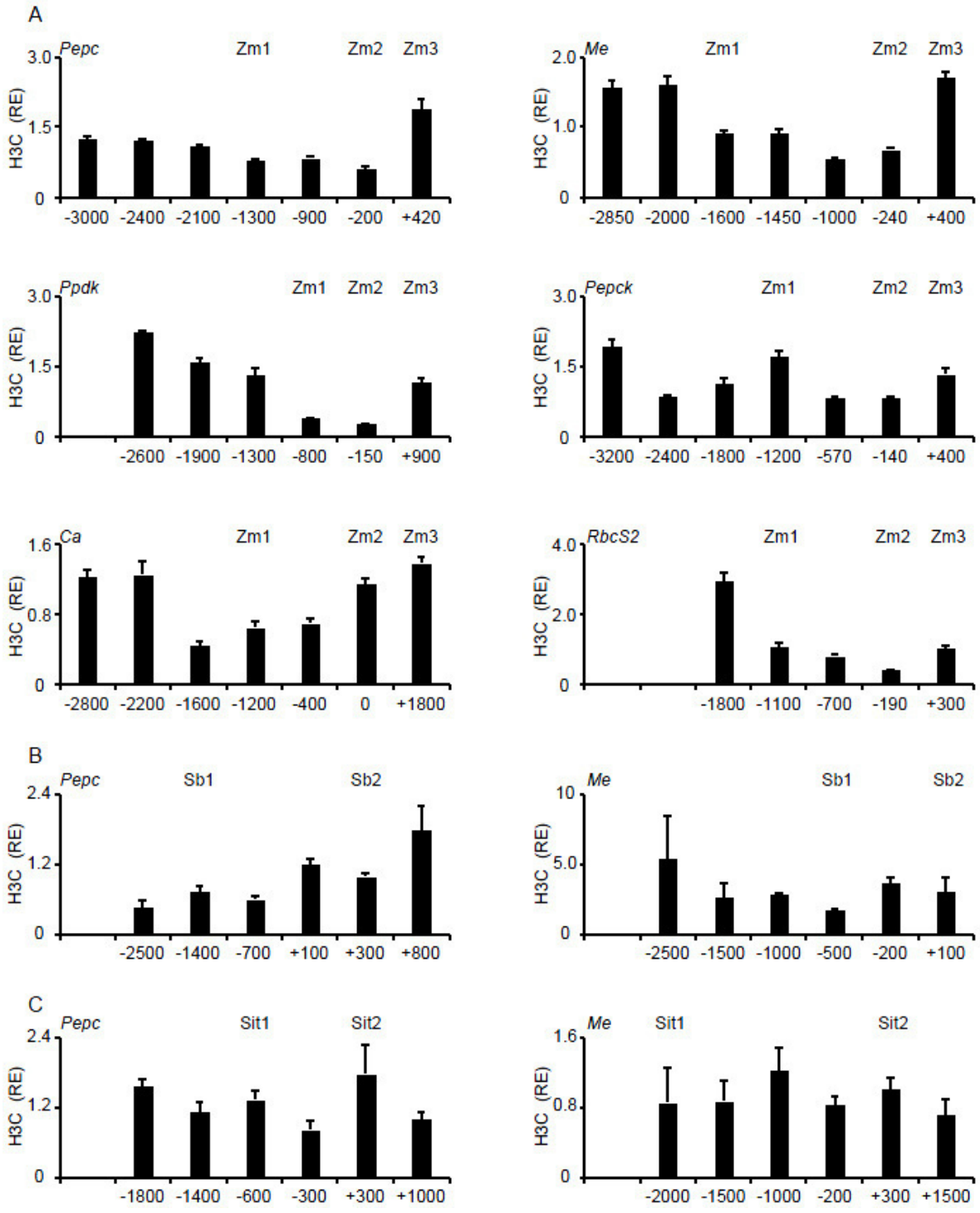
Haring M, Offermann S, Danker T, Horst I, Peterhänsel C, Stam M (2007) Chromatin immunoprecipitation: quantitative analysis and data normalization. Plant Methods **2:** 11

Sheen J (1991) Molecular mechanisms underlying the differential expression of maize pyruvate, orthophosphate dikinase genes. Plant Cell **3:** 225-245

Wang X, Gowik U, Tang H, Bowers J, Westhoff P, Paterson A (2009) Comparative genomic analysis of C4 photosynthetic pathway evolution in grasses. Genome Biol **10:** R68

Supplemental data 2: Nucleosome occupancy on the gene promoters investigated in this study.

A, *Zea mays*, B, *Sorghum bicolor and* C, *Setaria italica*. NO is defined as the amount of chromatin precipitated with an antibody specific for an invariant C-terminal epitope on histone H3 (H3C) divided by the amount of chromatin subjected to immunoprecipitation (Input). NO is shown as relative enrichment (RE) compared to the NO on the promoter of the *Actin*1 gene. All data points are based on at least three independent experiments. Vertical lines indicate standard errors.



Chapter 3 - Publication II

Photosynthetic Genes and Genes Associated with the C4 Trait in Maize Are Characterized by a Unique Class of Highly Regulated Histone Acetylation Peaks on Upstream Promoters

Renke Perduns, Ina Horst-Niessen, and Christoph Peterhansel

Leibniz University Hannover, Institute of Botany, 30419 Hannover, Germany

Type of authorship:	first author
Type of article:	research article
Share of the work:	80%
Contribution to the publication:	performed all experiments, analyzed the data, prepared all figures and contributed to writing of the paper
Journal:	Plant Physiology (www.plantphysiol.org)
Impact Factor:	7.394 (2013)
Date of Publication:	August, 2015
DOI:	10.1104/pp.15.00934
PubMed-ID:	26111542
Citation:	**Perduns R, Horst-Niessen I, Peterhansel C** (2015) Photosynthetic Genes and Genes Associated with the C4 Trait in Maize Are Characterized by a Unique Class of Highly Regulated Histone Acetylation Peaks on Upstream Promoters. Plant Physiol **168**: 1378–1388

Photosynthetic Genes and Genes Associated with the C4 Trait in Maize Are Characterized by a Unique Class of Highly Regulated Histone Acetylation Peaks on Upstream Promoters[1][OPEN]

Renke Perduns, Ina Horst-Niessen, and Christoph Peterhansel*

Institute of Botany, Leibniz University Hannover, 30419 Hannover, Germany

Histone modifications contribute to gene regulation in eukaryotes. We analyzed genome-wide histone H3 Lysine (Lys) 4 trimethylation and histone H3 Lys 9 acetylation (two modifications typically associated with active genes) in meristematic cells at the base and expanded cells in the blade of the maize (*Zea mays*) leaf. These data were compared with transcript levels of associated genes. For individual genes, regulations (fold changes) of histone modifications and transcript levels were much better correlated than absolute intensities. When focusing on regulated histone modification sites, we identified highly regulated secondary H3 Lys 9 acetylation peaks on upstream promoters (regulated secondary upstream peaks [R-SUPs]) on 10% of all genes. R-SUPs were more often found on genes that were up-regulated toward the blade than on down-regulated genes and specifically, photosynthetic genes. Among those genes, we identified six genes encoding enzymes of the C4 cycle and a significant enrichment of genes associated with the C4 trait derived from transcriptomic studies. On the DNA level, R-SUPs are frequently associated with ethylene-responsive elements. Based on these data, we suggest coevolution of epigenetic promoter elements during the establishment of C4 photosynthesis.

The transcription rate of genes is regulated by their promoters. In eukaryotes, genes are associated with proteins that condense the DNA molecule. The association of DNA and proteins is called chromatin. The basic repeat structure of chromatin is an octamer of two of the histone proteins H2A, H2B, H3, and H4 plus approximately 160 bp of DNA (Kouzarides, 2007). This structure not only forms a passive barrier to binding of other proteins to DNA but also, actively participates in the regulation of transcription. Posttranslational modifications of histones are important elements of this regulatory role (Lauria and Rossi, 2011). A wealth of different modifications was identified. Studies in different model organisms revealed that promoters and transcribed regions of active genes fundamentally differ in their histone modification profile compared with inactive genes (Berger, 2007; Zhang, 2008). Two major theories have been suggested of how histone modifications can regulate gene transcription. Acetylation of histones might neutralize the positive charge of lysines on the flexible histone N-terminal tails and by this, reduce the interaction with negatively charged DNA. This would enhance DNA accessibility for transcription factors and the RNA polymerase. Therefore, histone acetylation was mostly observed on actively transcribed genes (Dion et al., 2005; Wang et al., 2009). Other histone modifications do not neutralize the positive charge of Lys. The best studied of these modifications is Lys methylation. The three hydrogen atoms on the ε-amino group of lysines can be partially or fully replaced by methyl groups; thus, mono-, di-, and trimethylated forms of lysines exist. Different forms of Lys methylation have been associated with active or inactive genes (Wang et al., 2009; Roudier et al., 2011). The most obvious role for histone modifications in this scenario is the provision of binding sites for transcription factors. This has been exemplified by the binding of TATA box binding protein-associated factor3, a subunit of the general transcription initiator factor TFIID, to trimethylated Lys 4 on histone H3 (H3K4me3) to initiate transcription (Vermeulen et al., 2007; Lauberth et al., 2013) or the recognition of acetylated histones by bromodomain proteins (Hassan et al., 2007). The combination of all different histone modifications on a gene might form a histone code or epigenetic memory that collects and integrates information about developmental and environmental signals perceived by the organism (Turner, 2002; Song et al., 2012; To and Kim, 2014).

The leaf gradient of monocotyledonous plants is a useful tool to study developmental dynamics of

[1] This work was supported by the Deutsche Forschungsgemeinschaft (grant no. PE819/5–1) and the European Union FP7 Program (Project 3to4; contract no. 289582).

* Address correspondence to cp@botanik.uni-hannover.de.

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: Christoph Peterhansel (cp@botanik.uni-hannover.de).

R.P. performed most of the experiments and analyzed the data; I.H.-N. supported the experimental work and helped with data analysis; C.P. conceived the project and wrote the article with contributions of all the authors.

[OPEN] Articles can be viewed without a subscription.

www.plantphysiol.org/cgi/doi/10.1104/pp.15.00934

biological processes. Monocot leaves develop from undifferentiated meristematic cells at the base to mature photosynthetic cells at the tip. Thus, different developmental stages can be isolated from the same leaf (Nelissen et al., 2012; Wang et al., 2014). Development of the maize (*Zea mays*) leaf is of special interest, because the photosynthetic tissue functionally differentiates during maturation into mesophyll (M) and bundle sheath (BS) cells: the two cell types required to perform C4 photosynthesis (Langdale, 1998; Majeran et al., 2005). Thus, the dynamics of transcriptomes, proteomes, metabolite levels, and photosynthesis on the maize leaf gradient have been studied (Li et al., 2010; Majeran et al., 2010; Pick et al., 2011; Tausta et al., 2014; Wang et al., 2014). Moreover, this type of analysis has also been transferred to dicotyledonous C4 plants, where leaves of different growth stages (Külahoglu et al., 2014) or different zones of an individual developed leaf were compared (Aubry et al., 2014). All of these studies aimed at identifying genes that were coregulated with photosynthetic genes and specifically, C4 genes.

We previously analyzed the function of histone modifications on C4 genes in maize. We defined specific modifications that were associated with light regulation (Offermann et al., 2006), differentiation of M and BS cells (Danker et al., 2008), circadian regulation (Horst et al., 2009), or metabolic regulation of C4 genes (Offermann et al., 2008). This code was, at least in part, conserved in other C4 grasses (Heimann et al., 2013; Horst et al., 2013). Most significantly, we found that acetylation of H3K9 (H3K9ac) on core promoters only occurred in the light but was independent of other stimuli. In contrast, H3K4me3 was associated with M and BS cell differentiation but independent of light (Offermann et al., 2008).

In this study, we compared transcript levels and histone modification levels between base and blade of maize leaves on a genome-wide level. Our data reveal a quantitative association between changes in transcription and changes in histone modifications of individual genes. The focus on regulated histone modifications in this study allowed the identification of regulated secondary upstream peaks (R-SUPs) on many developmentally regulated genes as a unique feature of H3K9ac. R-SUPs showed distinct properties compared with transcription initiation site (TIS) region peaks and were associated with ethylene-responsive elements (EREs) on the DNA level. Moreover, R-SUPs were significantly enriched on photosynthetic genes and found on some key genes encoding enzymes of the C4 cycle. Meta-analyses of R-SUP genes and lists of potential C4-associated genes derived from different transcriptomic studies revealed a substantial overlap.

RESULTS AND DISCUSSION

Assessment of Data Quality

We isolated RNA and chromatin from the base and the blade of the maize leaf (corresponding to segments −4 and +4 in Li et al., 2010). The abundance of mRNAs was measured by RNA sequencing (RNA-Seq), and the abundance of histone modifications was measured by chromatin immunoprecipitation sequencing (ChIP-Seq) after precipitation of chromatin with antibodies directed to H3K9ac and H3K4me3. Comparison of transcript abundance at the base and the blade revealed 7,634 transcripts that were significantly up- or down-regulated (false discovery rate [FDR] ≤ 0.05; fold change $\geq |2|$) in our data set.

Quality of RNA data was analyzed by comparison with previously published data on transcription on the leaf gradient (Li et al., 2010). Pearson correlation coefficients of the two replicates generated here and the previous data set are given in Supplemental Figure S1; 88% of the regulated genes from our data set showed identical regulation in the data set in Li et al., 2010. To concentrate on robustly regulated genes, we further only considered genes that were identically regulated in both studies (2,865 up-regulated genes and 3,823 down-regulated genes when comparing base and blade).

For the comparative analyses and the statistical evaluation of quantitative ChIP-Seq data, we used DiffBind (Stark and Brown, 2011). This program was initially developed for the identification of differential transcription factor binding to DNA (Ross-Innes et al., 2012) but recently, also used in a study on genome-wide induced alterations in histone modifications in mice (Papait et al., 2013). DiffBind normalizes ChIP-Seq samples for differences in peak calling and the amount of mapped reads enabling a direct quantitative comparison of histone modifications from different biological situations ("Materials and Methods"). As a quality test, we determined the percentage of consensus peaks derived from expected or unexpected sample combinations (Table I). Expected combinations were merged from either all four samples (two replicates from the base and two replicates from the blade) or at least two associated samples (either the two replicates from the base or the two replicates from the blade). Unexpected combinations included detection of peaks in one of the base samples and one of the blade samples. The percentages of consensus peaks that were built from expected combinations were more than 95% for both tested modifications. For H3K9ac, 61% of all peaks were detected in both the blade and the base, whereas 10% were only detected at the leaf base and 8% were only in the blade. For H3K4me3, similar numbers were observed, but a higher number of peaks (75%) was detected both at the base and in the blade, indicating less regulation of H3K4me3 compared with H3K9ac. The percentages of consensus peaks that were derived from unexpected combinations were always lower than 5%. This indicated that peak calling was consistent in all samples and allowed for a comparative analysis of histone modifications at the leaf base and the blade.

It has previously been described that H3K9ac and H3K4me3 were associated with actively transcribed

Table I. *Sources of merged consensus peaks of H3K9ac and H3K4me3 from DiffBind*

Histone modification peaks called in replicates (A and B) from the leaf base (Base) or the leaf blade (Blade) were analyzed with DiffBind. Number and percentage of consensus peaks merged from the indicated peak combinations of samples are shown.

Combinations	Base A	Base B	Blade A	Blade B	H3K9ac Peaks		H3K4me3 Peaks	
					No.	*%*	*No.*	*%*
	X	X	X	X	26,742	61.1	24,661	74.9
	X	X			4,329	9.9	1,655	5.0
	X	X	X		1,362	3.1	1,980	6.0
	X	X		X	317	0.7	226	0.7
			X	X	6,258	14.3	1,197	3.6
	X		X	X	3,507	8.0	962	2.9
		X	X	X	27	0.06	639	1.9
Expected					42,542	97.3	31,320	95.2
	X		X		940	2.1	846	2.6
	X			X	236	0.5	91	0.3
		X	X		10	0.02	608	1.8
		X		X	6	0.01	43	0.1
Unexpected					1,192	2.7	1,588	4.8

genes. Modifications were mainly enriched around the TIS and with lower abundance in the transcribed region of maize genes (Wang et al., 2009; He et al., 2013). To investigate if our ChIP-Seq data from two different zones from the maize leaf confirm these findings, we tested whether histones on transcribed genes were methylated and/or acetylated (Fig. 1A). From the 26,055 genes that were expressed with reads per kilobase per million mapped reads (RPKM) > 0.1 in at least one of the investigated tissues, 20,841 (80%) showed H3K9ac modification, and 20,456 (79%) showed H3K4me3 modification. Both modifications were found on 19,560 (75%) transcribed genes. Thus, both modifications were highly associated with transcribed genes. When analyzing the distribution of H3K4me3 and H3K9ac signals, both modifications peaked between the TIS and position +1,000 bp of the transcribed region (Fig. 1B). This region was defined as the TIS region. The detected distributions were in accordance with other data sets on positioning of the investigated histone modifications on active genes in maize (Wang et al., 2009; He et al., 2013).

A Dynamic Analysis of Histone Modifications Revealed a Unique Class of Upstream Acetylation Peaks

Whereas a static view on histone modifications on genes with low and high transcript levels provided important insights into their function, much less is known about the dynamics of histone modifications during gene activation or inactivation of individual genes. Our quantitative determination of transcript abundances and histone modifications in two different developmental zones allowed for the comparison of both parameters on the level of individual genes. Figure 2 shows a correlation analysis of transcript levels and histone modifications. Correlation between the absolute intensity of the tested histone modifications and the absolute transcript level of the corresponding gene (Fig. 2, Intensities) was low and only showed a slight tendency for more intense histone modification signals on higher transcribed genes. However, when plotting the fold change in histone modifications between leaf base and blade against the fold change in transcript levels from the same samples (Fig. 2, Fold changes), we observed significant correlation for both H3K9ac and H3K4me3. This was also exemplified by the much better Pearson correlation coefficients (P values) for fold changes compared with intensities (Fig. 2). Thus, although a rough correlation between the intensity of a histone modification peak and the abundance of the associated transcript in a given biological situation exists, changes in histone modification between two biological situations or developmental stages provide much better information about the activity of the gene. Our observations are supported by another study in maize, where transcript abundance and histone modification levels were compared between root and shoot tissues of maize. Best correlations were found for relative changes and not absolute values (He et al., 2013). In a recent study on histone modifications and leaf senescence in Arabidopsis (*Arabidopsis thaliana*), histone modifications increased during gene activation for a subset of genes, but other genes were already premarked with histone modifications before activation and did not show such increases (Brusslan et al., 2015). The very good correlation between fold changes in the tested histone modifications and fold changes in transcript levels in our study argues against widespread premarking of genes at the leaf base for later activation in the blade of the maize leaf.

Our data analysis up to this point suggested that regulation of histone modifications between different biological situations was highly associated with regulation of transcript levels. We, therefore, focused

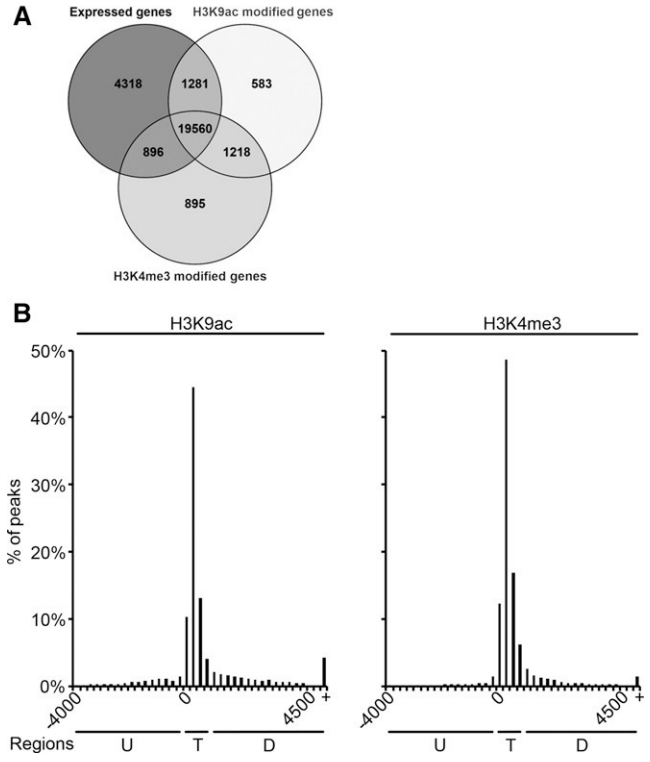


Figure 1. Overlap between expressed and modified genes and distribution of H3K9ac and H3K4me3 peaks. A, Overlap between expressed genes and genes modified by H3K9ac and H3K4me3. Expressed genes showed a mean RPKM $>$ 0.1 in one of the analyzed tissues. B, Distribution relative to the TIS. H3K9ac and H3K4me3 peaks were grouped into 250-bp bins according to their position. Percentages of all peaks are plotted on the *y* axes. Three distinct regions are indicated: upstream (U), TIS (T), and downstream (D). +, More than 4,500 bp downstream of the TIS.

our analyses on H3K9ac and H3K4me3 peaks that were regulated between leaf base and blade (FDR $\leq$ 0.05, peak fold change $\geq$ |2|; Fig. 3A, regulated peaks). A distribution of these peaks is shown in Figure 3A. For H3K4me3 (Fig. 3A, right), we observed a distribution that was very similar to the distribution observed for all peaks (Fig. 1B; i.e. most regulated peaks were found in the TIS region, the frequency slowly declined toward the downstream region, and only few regulated peaks were found in the upstream promoter region). However, for H3K9ac, we observed beside the TIS region and the downstream region a separate region of regulated acetylation peaks in the upstream promoter region. This region was distinct from the TIS region, because the number of regulated peaks declined between TIS and −500 bp and then increased again toward more upstream-located promoter regions. The highest number of these upstream acetylation peaks was found around position −1,000 bp relative to the TIS.

Whereas H3K9ac and H3K4me3 peaks in the TIS region and downstream region were already described in different species (Wang et al., 2009; Roudier et al., 2011; Du et al., 2013; He et al., 2013; Hussey et al., 2015), our observation of regulated upstream H3K9ac peaks on an genome-wide basis was, to our knowledge, unique. We first wanted to test whether regulated upstream acetylation was simply caused by false annotations of TIS. To exclude this possibility, we evaluated whether the 1,084 genes with regulated upstream H3K9ac peaks had additional H3K9ac peaks in the annotated TIS region or the downstream region (Fig. 3B). For comparison, data for the 251 upstream H3K4me3 peaks were included; 69% (53% + 16%) of all genes with a regulated upstream H3K9ac peak also had an H3K9ac peak in the annotated TIS region. On those genes, upstream H3K9ac peaks were distinct secondary peaks that were separate from the primary TIS region H3K9ac peak. For 31% (24% + 7%) of all tested genes, no TIS region acetylation peaks were detected, and it was unclear whether the TIS was correctly annotated. For H3K4me3, the distribution was different, and most genes with a regulated upstream H3K4me3 did not show a TIS region peak. We, therefore, concentrated on H3K9ac and analyzed R-SUPs in more detail. A list of the 748 (576 + 172) genes with R-SUPs is provided in Supplemental Table S1.

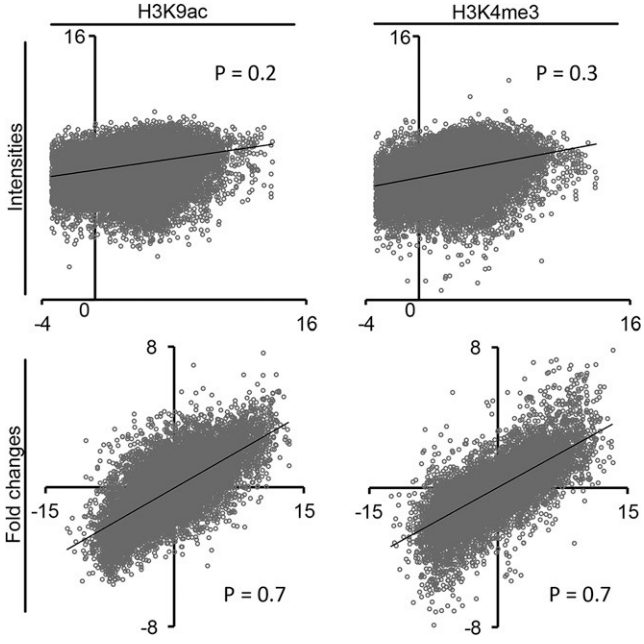


Figure 2. Correlation of H3K9ac or H3K4me3 signals and transcript levels for individual genes. Log_2-transformed values of absolute signal intensities or fold changes between base and blade were plotted against each other. The values for H3K9ac and H3K4me3 are plotted on the *y* axes, and the values for transcription are plotted on the *x* axes. For absolute signal intensities, only the values from the tissue with the higher transcription are shown. *P*, Pearson correlation coefficient.

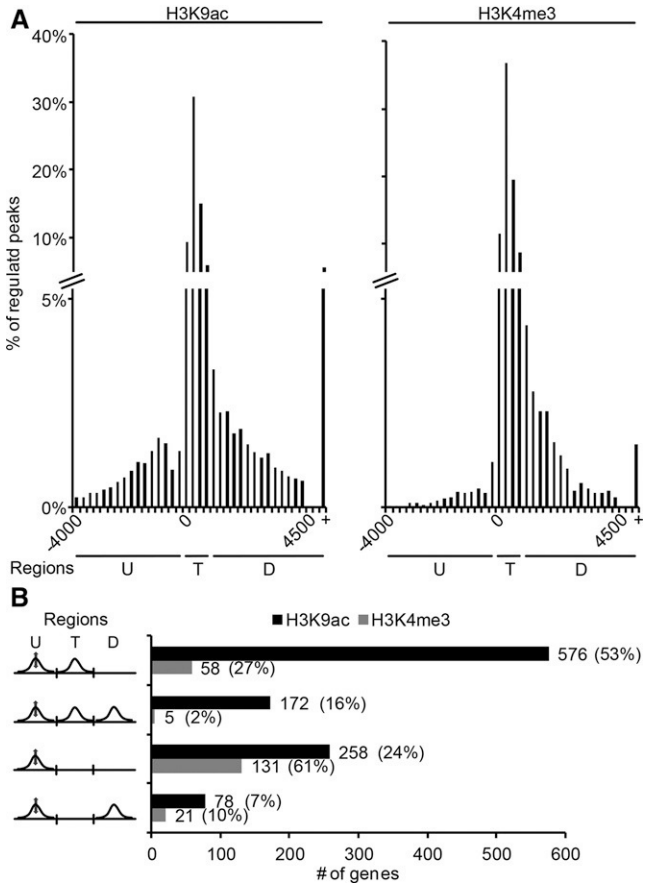


Figure 3. Distribution of regulated H3K9ac and H3K4me3 peaks relative to the TIS and peak composition of genes with a regulated upstream peak. A, Regulated H3K9ac and H3K4me3 peaks (fold change $\geq |2|$, FDR ≤ 0.05) were grouped into 250-bp bins according to their position. Percentages of all peaks were plotted on the *y* axes. +, More than 4,500 bp downstream of the TIS. B, Possible H3K9ac peak compositions of genes with a regulated H3K9ac upstream peak are indicated in left. The bars indicate absolute values. The fractions of the indicated composition from all genes with regulated upstream peaks are given as percentages. Three distinct regions are indicated: upstream (U), TIS (T), and downstream (D).

Distinct Properties of R-SUPs of H3K9ac and Associated Genes

We first tested whether R-SUPs differed in shape or the degree of regulation from the associated TIS region peaks on the same genes (Fig. 4A). Whereas both peak classes showed an average width of 600 to 700 bp, the peak intensity of R-SUPs was 3 times lower than the intensity of associated TIS region peaks. However, R-SUPs were clearly more regulated between leaf base and blade (factor 5.5× versus 3.1× for the TIS region peaks), suggesting that they might also be important for gene regulation.

We furthermore tested if R-SUPs and associated genes showed the same ratio of induction and repression in the blade compared with the majority of H3K9ac peaks and genes (Fig. 4B). Interestingly, the ratio of induced and repressed H3K9ac peaks differed notably between R-SUPs and all H3K9ac peaks or all TIS region peaks. The ratio of induced to repressed H3K9ac peaks in the blade was only 0.7 for all H3K9ac peaks and 0.5 for all TIS region peaks, indicating that more H3K9ac peaks were repressed than induced when comparing leaf base and blade. However, the ratio for R-SUPs was 2.1, indicating that they were more often induced than repressed. Comparable ratios were observed for the transcriptional regulation of associated genes. More genes were down-regulated between leaf blade and base, but for R-SUP genes, we observed more up-regulation.

Histone modifications might facilitate access of transcription factors to DNA on promoters (Hassan

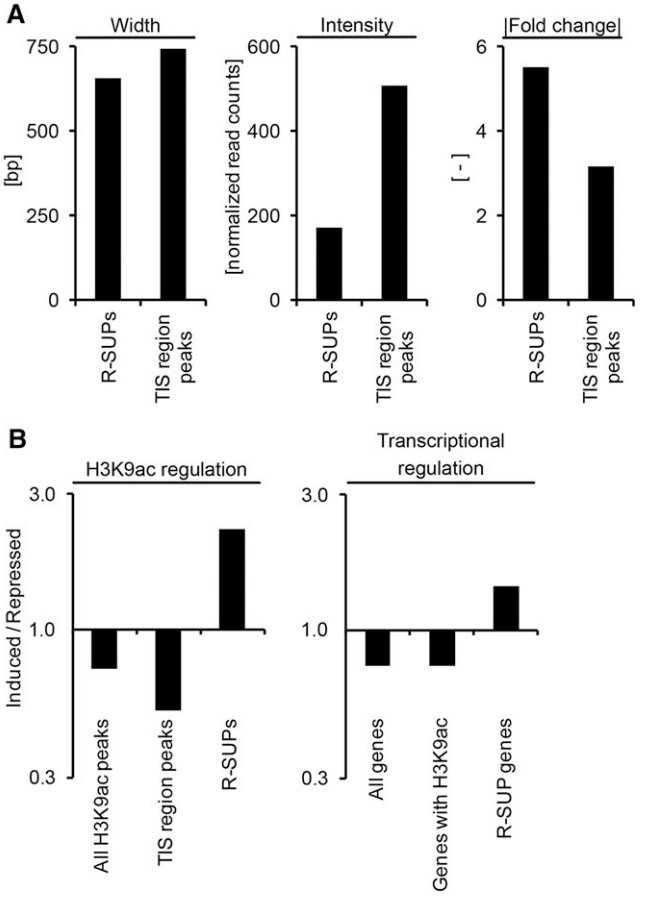


Figure 4. Characterization of R-SUPs and TIS region peaks from R-SUP-associated genes and comparison of ratios from induced to repressed H3K9ac peaks and genes from different groups. A, Mean widths, intensities and fold changes from R-SUPs and TIS region H3K9ac peaks of R-SUP-associated genes were calculated and plotted as indicated. The width is given as base pairs, and the intensities are given as normalized read counts. B, Ratios from induced to repressed H3K9ac peaks (left) and genes (right) from the indicated groups (*x* axes) were calculated and are plotted on the $\log_{10}$-transformed *y* axes. Genes and H3K9ac peaks are regulated with a fold change $\geq |2|$ and an FDR ≤ 0.05.

et al., 2007; Vermeulen et al., 2007; Lauberth et al., 2013). We, therefore, analyzed whether specific DNA sequences were enriched under R-SUPs. We used the MEME and DREME algorithms of the MEME-ChIP Suite (Machanick and Bailey, 2011) for de novo identification of enriched sequences. These algorithms are insensitive for biases (e.g. caused by the C/G content of a genome, because the nucleotide compositions of the input sequences are used to generate a randomized reference sequence for the statistical enrichment analyses; Bailey and Elkan, 1994; Bailey, 2011). Both algorithms identified highly related sequences as the most overrepresented motifs (Table II). These motifs are most similar to EREs with a (A/C)GCCGCC core sequence (Hart et al., 1993). EREs are binding sites for ethylene response element binding factors (ERFs; also ERE binding proteins) and have so far been mainly associated with biological processes, such as pathogen response (Zhou et al., 1997) and abiotic stress (Gutterson and Reuber, 2004).

Functional Classification of R-SUP-Associated Genes

To analyze if the occurrence of R-SUPs can be associated with specific processes, we also tested biological processes as defined by MapMan bins (Thimm et al., 2004) for enrichment within R-SUP-associated genes (Fig. 5). To be more specific, we focused on those R-SUP genes that showed the typical expression and acetylation pattern defined in Figure 4B: induction of both transcription and R-SUP acetylation in the leaf blade compared with the base (201 genes; Supplemental Table S2). Six biological processes contained sufficient genes to reliably test for enrichment by χ^2 test (expectation ≥ 5). Only the photosynthesis bin showed significant enrichment, with almost 2 times as many R-SUP genes as expected (Fig. 5). This was not caused by a general overrepresentation of photosynthesis genes among genes that were up-regulated in the leaf blade, because enrichment was calculated relative to all genes that were up-regulated in the leaf blade and not relative to all genes independent of their regulation.

R-SUP-associated genes in the photosynthesis bin are listed in Supplemental Table S2. We found genes from all different subprocesses of photosynthesis, such as light reaction and the Calvin cycle. However, the most striking observation was that five annotated genes encoding for enzymes of the C4 cycle were found in this group: two genes encoding carbonic anhydrases (CAs) and one gene each for phosphoenolpyruvate carboxylase (PEPC), malic enzyme (ME), and pyruvate-inorganic phosphate-dikinase. In addition, a gene encoding phosphoenolpyruvate carboxykinase was on the list of R-SUP-associated genes. This gene was assigned to the MapMan bin gluconeogenesis but has been shown to participate in C4 metabolism in maize (Pick et al., 2011). Thus, in total, six genes for C4 enzymes contained R-SUPs.

We verified this observation for five of the six genes by comparing ChIP-Seq data from this study with previously published chromatin immunoprecipitation-quantitative polymerase chain reaction (ChIP-qPCR) data from our laboratory (Offermann et al., 2008; Heimann et al., 2013). An exemplary profile for ME is shown in Figure 6A, and data for four other genes are shown in Supplemental Figure S2. Red and green dots in Figure 6A represent ChIP-Seq reads, whereas the black line in Figure 6A represents the ChIP-qPCR data. In the blade, beside the TIS region peak, an R-SUP was evident that peaked at around −1,500 relative to the TIS in both ChIP-Seq and ChIP-qPCR data. In this zone, ME transcript levels were highest. At the leaf base, where ME was weakly transcribed, the TIS region peak only slightly changed, whereas the R-SUP completely disappeared. To quantify this effect, we determined H3K9ac fold changes between leaf base and blade for R-SUP and TIS peaks on these C4 genes (Fig. 6B; "Materials and Methods" gives a definition of the R-SUP on one of the CA genes). Whereas fold changes for R-SUPs varied between 9- and 200-fold dependent on the gene, fold changes for TIS region peaks were only between 2- and 10-fold. For each individual C4 gene, regulation of the R-SUP was at least 6-fold stronger than regulation of the TIS region peak. This indicated a strong association of R-SUP regulation and transcriptional regulation for these C4 genes. Another link between R-SUPs and C4 metabolism was provided by our finding that EREs were enriched within the DNA sequences associated with R-SUPs (Table II). Pick et al. (2011) identified transcription factors that were coregulated with prominent C4

Table II. *Enrichment of motifs within sequences from R-SUPs*

Genomic sequences covered by R-SUPs were subjected to MEME ChIP for de novo identification of motifs and subsequent alignment of discovered motifs to known binding sites from the JASPAR CORE Plantae database.

Most Overrepresented Motif	Algorithm	E Value	Transcription Factor Family with the Most Similar Motif
	MEME	1.9e-13	ERF1
	DREME	3.3e-4	ERF1

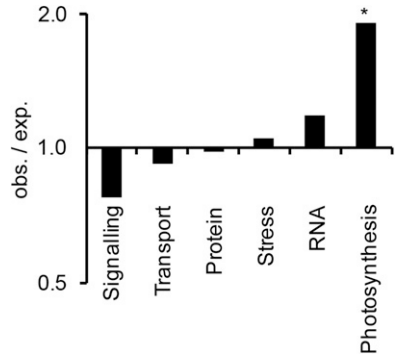


Figure 5. Enrichment analysis of R-SUP-associated genes with induction of both transcription and R-SUP acetylation in the leaf blade compared with the base within processes defined by MapMan bins. All genes that are transcriptionally induced and linked to an induced H3K9ac peak were chosen as reference for the analysis. The degree of enrichment is given as a ratio of expected (exp.) and observed (obs.) counts (log_{10}-transformed y axis). The significances of overrepresentations were evaluated by χ^2 test, and derived P values were adjusted using the Benjamini-Hochberg correction. Processes were analyzed only if at least five counts were expected in the tested group. *, FDR ≤ 0.05.

genes, and 9 of 18 functionally annotated transcription factors were ERF proteins.

Meta-Analysis of the Connection between Up-Regulated R-SUP Genes and Potential C4 Genes

We hypothesized that R-SUPs were common elements of genes related to C4 photosynthesis and might be involved in coregulation of these genes on the leaf gradient. However, a manual inspection revealed that other C4 genes, such as malate dehydrogenase and small subunit of Rubisco, did not show R-SUPs. We, therefore, screened previously generated lists of candidate genes associated with C4 metabolism based on their expression profile on the leaf developmental gradient or their specificity for M or BS cells. The analysis by Aubry et al. (2014) identified genes that were M or BS specific in both dicotyledoneous (*Gynandropsis* spp.) and monocotyledoneous (maize) C4 plants. Within this group of 294 genes, 68 genes were transcriptionally up-regulated in the leaf blade and also had an up-regulated H3K9ac peak; 17 of these genes had an up-regulated R-SUP.

This is a significant enrichment of R-SUP genes within the list of potential C4 genes (data not shown; P [χ^2] = 0.01). Four of the six R-SUP genes encoding C4 enzymes (Fig. 6) were also found in this comparison. This included PEPC and pyruvate-inorganic phosphate-dikinase but also, both CA genes described above. One of the two CA genes was already described as C4 specific (GRMZM2G121878; Studer et al., 2014). The second CA gene (GRMZM2G414528) was mainly expressed in leaves (Sekhon et al., 2013), and transcripts were highly enriched in M cells compared with BS cells in both maize and *Gynandropsis* spp. (Aubry et al., 2014). Thus, this gene might contribute to C4-CA activity. This is in line with the minor growth reduction of C4-CA double mutants (Studer et al., 2014), suggesting that some remaining CA activity was present in these double mutants for the formation of C4 cycle substrates.

Another R-SUP-associated gene in the list by Aubry et al. (2014) was a sulfate transporter (GRMZM2G395114), which was specifically expressed in BS cells of maize and *Gynandropsis* spp. Transcription of this gene was 10-fold higher in the blade compared with the base of the leaf, but H3K9ac at the TIS was unaffected by the transcriptional up-regulation. Instead, we observed a 7-fold up-regulation of the R-SUP that was positioned at −1,600 bp relative to the TIS. This gene might also play a role in C4 photosynthesis, consistent with the confinement of sulfate metabolism to BS cells in C4 plants (Kopriva and Koprivova, 2005).

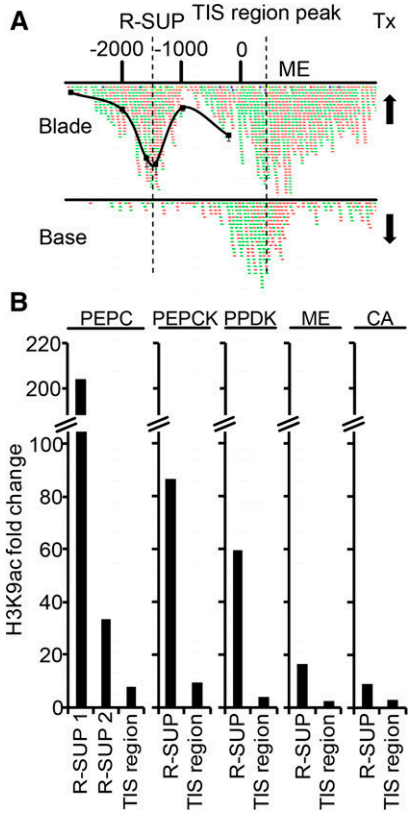


Figure 6. Regulation of R-SUPs and TIS region peaks on genes encoding C4 enzymes. A, H3K9ac profile on the ME gene at the leaf base and blade. Positions are relative to the TIS in base pairs. Red and green dots represent mapped forward and reverse ChIP-Seq reads. The black line represents previously obtained ChIP-qPCR data from the mature leaf (Heimann et al., 2013). Arrows represent high and low transcript levels. Tx, Transcription. B, H3K9ac fold changes between leaf base and blade for R-SUPs and TIS region peaks on genes encoding C4 enzymes. PEPCK, Phosphenolpyruvate carboxykinase; PPDK, pyruvate-inorganic phosphate-dikinase.

A BS-specific DNA binding with one finger (DOF) transcription factor (ZmDOF30; AC233935.1_FG005) that had previously been suggested as a candidate C4 regulator also had an up-regulated R-SUP. DOF transcription factors have been shown to control expression of genes involved in carbon metabolism (Yanagisawa, 2000) and specifically, the C4-PEPC gene (Yanagisawa and Sheen, 1998). Both activating and repressive functions have been suggested (Yanagisawa and Sheen, 1998). ZmDOF30 might, therefore, act as an activator of BS-specific gene expression and/or a repressor of M-specific genes in BS cells.

In total, we found 25 transcription factors from the GRASSIUS Database (Yilmaz et al., 2009) that were associated with an up-regulated R-SUP. We, therefore, also compared these transcription factors with a list of transcription factors that were preferentially expressed in M or BS cells in maize (Li et al., 2010). Nine transcription factor genes were found on both lists. Again, an ERF gene was among these candidates (maize ethylene response element binding protein17; GRMZM2G029323). Together with the multiple ERFs identified as candidate C4 regulators by Pick et al. (2011) and the ERE sequences enriched in R-SUP regions (see above), this underlines the potential role of EREs as C4 regulatory elements.

The ZmDOF30 (AC233935.1_FG005) transcription factor was again in the overlap. We furthermore identified two R-SUP-associated genes encoding BS-specific basic leucine zipper (bZIP) transcription factors (GRMZM2G092137 and AC203957.3_FG004). Transcription factors from the bZIP family have been shown to interact with DOF transcription factors (Zhang et al., 1995) and be involved in the regulation of histone acetylation and chromatin structure on maize genes (Casati and Walbot, 2008). Thus, complexes of ZmDOF30 and bZIP factors might be involved in the control of C4 gene expression.

CONCLUSION

During evolution of C4 photosynthesis, C4 genes evolved from C3 ancestors (Brown et al., 2010; Christin et al., 2012) but had to adopt new promoter elements to control higher and more regulated gene expression (Ku et al., 1996; Hibberd and Covshoff, 2010). The comparative analysis of regulated histone modifications in different developmental zones of the maize leaf in this study allowed for the identification of R-SUPs that are distinct from the often-described acetylation peaks in TIS regions. R-SUPs are often associated with EREs and might, therefore, constitute sites for binding of regulatory transcription factors. This binding would be facilitated by the strong acetylation of histones that increases accessibility of DNA in the chromatin context (Dion et al., 2005). Moreover, R-SUPs were identified on many photosynthetic and C4-related genes, suggesting that they might contribute to regulation of these genes. Our data suggest an evolutionary scenario, where parallel recruitment of upstream promoter elements from photosynthetic genes contributed to reprogramming of gene expression during establishment of C4 photosynthesis in maize. Comparative analyses on the chromatin structure of orthologs of C4 promoters in C3 grasses and transgenic studies with promoter deletion constructs will help to test this hypothesis.

MATERIALS AND METHODS

Plant Material and Growth Conditions

Maize (*Zea mays*) 'B73' was grown in growth chambers with a photoperiod of 16 h of light at 25°C and 8 h of dark at 20°C. The photon flux intensity was adjusted to 130 μmol m^{-2} s^{-1}. Water-soaked seeds were transferred to soil, and plants were harvested when the third leaf was fully expanded (13–14 d after sowing); 4 cm from the middle of the illuminated area of the third leaf were harvested for analysis of the leaf blade. Base samples were isolated from the same leaves. Plants were cut 0.5 cm above soil, and the lowest 1 to 2 cm of the leaf area without light access was harvested.

RNA Isolation and RNA Sequencing

Harvested leaf sections from six plants per biological replicate were pooled and ground in liquid nitrogen. About 50 to 60 mg of ground plant material was dissolved in 1 mL of Trizol and agitated for 15 min; 0.2 volumes of chloroform was added. Samples were agitated for 10 min, and phases were separated by centrifugation (16,100*g* at 4°C for 15 min); 400 μL of the aqueous phase was transferred to a new reaction tube and washed with 1 volume of chloroform. The RNA was precipitated by addition of 2 volumes of 96% (v/v) ice-cold ethanol, incubation at −20°C for 30 min, and centrifugation (16,100*g* at 4°C for 15 min). The precipitate was washed with 70% (v/v) ice-cold ethanol and dissolved in 30 μL of RNase-free water. The isolated RNA was treated with the RNase Free DNase Set (79254; Qiagen), purified with the RNeasy MiniElute Cleanup Kit (74204; Qiagen), and eluted two times in 14 μL of RNase-free water. Kits were used according to the manufacturer's instructions. The integrity of the RNA was analyzed by agarose gel electrophoresis, and the concentration was measured photometrically; 2 μg of RNA per biological replicate was submitted to sequencing on an Illumina HiSeq2000 Instrument with 50-bp single-end reads (GATC Biotech). The service included quality assessment on a 2100 Bioanalyzer, complementary DNA synthesis from fragmented poly(A)+ RNA, and library preparation.

ChIP and ChIP Sequencing

We generated two biological replicates for the analysis of the histone modifications, which is in accordance with the guidelines for ChIP-Seq experiments from the Encyclopedia of DNA Elements and Model Organism Encyclopedia of DNA Elements Consortia (Landt et al., 2012). Harvested leaf sections from 10 plants per biological replicate were pooled, and chromatin was cross linked as described (Horst et al., 2009). Chromatin isolation and immunoprecipitation were performed as described (Heimann et al., 2013), with the exception that protein A agarose was not blocked with salmon sperm DNA. Modifications were detected with 5 μL of antiacetyl H3K9 (07-352; Merck Millipore) or 2.5 μL of anti-H3K4me3 (04-745; Merck Millipore). After washing and reversion of the cross link, samples were treated for 2 h with RNase A (R4642; Sigma-Aldrich) according to the manufacturer's instructions. The precipitated DNA was purified with the MiniElute PCR Purification Kit (428004; Qiagen). For each biological replicate, eight ChIP reactions were pooled; 10 ng of precipitated DNA was submitted to sequencing on an Illumina HiSeq2000 (for H3K9ac) or Illumina HiSeq2500 (for H3K4me3) Instrument with 50-bp single-end reads (GATC Biotech). The service included quality assessment on a 2100 Bioanalyzer and library preparation.

Processing and Analysis of RNA-Seq Data

Quality of raw sequencing reads was checked with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and reads were processed with the Trimmomatic software (Bolger et al., 2014). The adapter clipper,

leading trimmer, trailing trimmer, and sliding window trimmer modules were used with their default settings to trim the raw sequences. Processed sequences were imported into the CLC Genomics Workbench 7.5.1 (http://www.clcbio.com/products/clc-genomics-workbench/; Qiagen). Read mapping was carried out with the RNA-Seq Analysis Tool. Reads were mapped with default settings only to regions that were covered by gene annotations of the maize reference genome (Zea_mays.AGPv3.21.dna.toplevel; ftp://ftp.ensemblgenomes.org/pub/plants/release-21/). Reads that mapped equally to multiple positions were ignored. Raw read counts were chosen as expression values. RPKM values were calculated for each gene to quantify absolute expression (Mortazavi et al., 2008). Statistical evaluation of differentially expressed genes was carried out with the Empirical Analysis of DGE Tool, which implements edgeR analysis (Robinson et al., 2010) and Benjamini-Hochberg correction of *P* values (Benjamini and Hochberg, 1995). For additional analyses, only genes annotated as protein coding and with an RPKM $>$ 0.1 in at least one of the developmental zones were selected. This threshold was determined based on the coefficient of variation between RPKM values of the biological replicates that strongly increased at RPKM $\leq$ 0.1.

Processing and Analysis of ChIP-Seq Data

Quality assessment and processing of raw ChIP-Seq sequencing reads were carried out as for RNA-Seq data. Processed reads were imported to the CLC Genomics Workbench. Duplicated reads were removed with the tool Remove Duplicate Reads using default settings. The reads were mapped to the maize reference genome (Zea_mays.AGPv3.21.dna.toplevel; ftp://ftp.ensemblgenomes.org/pub/plants/release-21/) using default parameters with two exceptions: reads were mapped without masking, and reads that mapped equally to multiple positions were ignored. The Peak Shape Tool from the CLC Genomics Workbench was used to call peaks. This tool uses the Hotelling observer to account for variable shapes of peaks (Hotelling, 1931). This approach was successfully used before to call peaks with varying shapes (Kumar et al., 2013). Peaks were initially called with a *P* value of 0.0001. Afterward, the Advanced Peak Shape Tools were used to optimize peak calling according to the manual. For each sample, a subset of 1,000 called peaks with the highest *P* values was used as positive regions to define a unique filter with the tool Learn Peak Shape Filter. The number of bins was calculated for each sample as window size per size of bins (35 bp). This optimized filter was used for the definite peak calling with the tool Apply Peak Shape Filter again with a *P* value of 0.0001.

Differential analysis of regulated peaks between the blade and base of the maize leaf was carried out with the R Bioconductor Package DiffBind following the manual (Stark and Brown, 2011). DiffBind merged overlapping peaks from different samples if peak coordinates overlapped in at least two samples. Reads within the coordinates of merged consensus peaks were counted and normalized for each sample, even if the peak was not called in this sample. Quantitative differences of the normalized reads in the merged peak areas were statistically evaluated using the implemented R Package edgeR (Robinson et al., 2010). Consensus peaks were annotated to genes using the PeakAnalyzer software (Salmon-Divon et al., 2010) with options TSS-Nearest Transcription Start Site and Coding and Non-Coding Genes. If several transcript variants were annotated for one gene, peak distances to the TIS were always calculated for the most 5′-located transcript variant. Only peaks that were located a maximum of 4,000 bp upstream of the TIS or within the transcribed region of the annotated gene were included. As described for RNA-Seq data, the analysis was focused on protein-coding genes with an average RPKM $>$ 0.1. We checked the core C4 genes manually and realized that a strong H3K9ac peak upstream of the CA gene (GRMZM2G121878) was wrongly annotated to a neighboring gene (GRMZM2G122025) that was not expressed (RPKM = 0) and annotated as a low-confidence gene. This peak was reannotated to GRMZM2G121878.

Enrichment Analysis of MapMan Processes

MapMan bins were used to assign genes to processes (Thimm et al., 2004). The χ^2 test was used to test for enriched processes. Only processes were taken into account if their expectation in the analyzed group was $\geq$5. The calculated *P* values were adjusted with the Benjamini-Hochberg correction to account for multiple testing (Benjamini and Hochberg, 1995).

Motif Identification

De novo identification of motifs was performed with MEME-ChIP (Machanick and Bailey, 2011; http://meme-suite.org). Genomic sequences that were covered by R-SUPs were used as input. MEME-ChIP uses 100 bp from the center of the sequences and performs de novo motif discovery using the algorithms MEME and DREME. Reported motifs were automatically compared with known motifs from the JASPAR CORE Plantae Database (Mathelier et al., 2014) with TOMTOM. MEME-ChIP was used with its default settings, except for limitation of maximum width to 10 nucleotides in MEME.

The original data sets from this article can be found in the National Institutes of Health Gene Expression Omnibus database under the accession number GSE67551.

Supplemental Data

The following supplemental materials are available.

Supplemental Figure S1. Read mapping statistics of ChIP-Seq and RNA-Seq libraries and comparison of RNA-Seq data from this study with data from Li et al., 2010.

Supplemental Figure S2. H3K9ac profiles on selected C4 genes at the leaf base and blade.

Supplemental Table S1. List of genes with R-SUPs.

Supplemental Table S2. List of R-SUP-associated genes with induction of both transcription and R-SUP acetylation in the leaf blade compared with the base and their functional annotations; additionally, overlaps between these genes and gene lists from two C4-related studies and induced R-SUP genes assigned to the photosynthesis bin are shown.

Received June 22, 2015; accepted June 24, 2015; published June 25, 2015.

LITERATURE CITED

Aubry S, Kelly S, Kümpers BM, Smith-Unna RD, Hibberd JM (2014) Deep evolutionary comparison of gene expression identifies parallel recruitment of trans-factors in two independent origins of C4 photosynthesis. PLoS Genet **10:** e1004365

Bailey TL (2011) DREME: motif discovery in transcription factor ChIP-seq data. Bioinformatics **27:** 1653–1659

Bailey TM, Elkan C (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc Int Conf Intell Syst Mol Biol. **2:** 28–36

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Series B Stat Methodol **57:** 289–300

Berger SL (2007) The complex language of chromatin regulation during transcription. Nature **447:** 407–412

Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics **30:** 2114–2120

Brown NJ, Palmer BG, Stanley S, Hajaji H, Janacek SH, Astley HM, Parsley K, Kajala K, Quick WP, Trenkamp S, et al (2010) C acid decarboxylases required for C photosynthesis are active in the mid-vein of the C species Arabidopsis thaliana, and are important in sugar and amino acid metabolism. Plant J **61:** 122–133

Brusslan JA, Bonora G, Rus-Canterbury AM, Tariq F, Jaroszewicz A, Pellegrini M (March 23, 2015) A genome-wide chronological study of gene expression and two histone modifications, H3K4me3 and H3K9ac, during developmental leaf senescence. Plant Physiol **168:** 1246–1261

Casati P, Walbot V (2008) Maize lines expressing RNAi to chromatin remodeling factors are similarly hypersensitive to UV-B radiation but exhibit distinct transcriptome responses. Epigenetics **3:** 216–229

Christin PA, Edwards EJ, Besnard G, Boxall SF, Gregory R, Kellogg EA, Hartwell J, Osborne CP (2012) Adaptive evolution of C(4) photosynthesis through recurrent lateral gene transfer. Curr Biol **22:** 445–449

Danker T, Dreesen B, Offermann S, Horst I, Peterhänsel C (2008) Developmental information but not promoter activity controls the methylation state of histone H3 lysine 4 on two photosynthetic genes in maize. Plant J **53:** 465–474

Dion MF, Altschuler SJ, Wu LF, Rando OJ (2005) Genomic characterization reveals a simple histone H4 acetylation code. Proc Natl Acad Sci USA **102:** 5501–5506

Du Z, Li H, Wei Q, Zhao X, Wang C, Zhu Q, Yi X, Xu W, Liu XS, Jin W, et al (2013) Genome-wide analysis of histone modifications: H3K4me2, H3K4me3, H3K9ac, and H3K27ac in Oryza sativa L. Japonica. Mol Plant **6:** 1463–1472

Gutterson N, Reuber TL (2004) Regulation of disease resistance pathways by AP2/ERF transcription factors. Curr Opin Plant Biol **7:** 465–471

Hart CM, Nagy F, Meins F Jr (1993) A 61 bp enhancer element of the tobacco β-1,3-glucanase B gene interacts with one or more regulated nuclear proteins. Plant Mol Biol **21:** 121–131

Hassan AH, Awad S, Al-Natour Z, Othman S, Mustafa F, Rizvi TA (2007) Selective recognition of acetylated histones by bromodomains in transcriptional co-activators. Biochem J **402:** 125–133

He G, Chen B, Wang X, Li X, Li J, He H, Yang M, Lu L, Qi Y, Wang X, et al (2013) Conservation and divergence of transcriptomic and epigenomic variation in maize hybrids. Genome Biol **14:** R57

Heimann L, Horst I, Perduns R, Dreesen B, Offermann S, Peterhansel C (2013) A Common histone modification code on C4 genes in maize and its conservation in Sorghum and *Setaria italica*. Plant Physiol **162:** 456–469

Hibberd JM, Covshoff S (2010) The regulation of gene expression required for C4 photosynthesis. Annu Rev Plant Biol **61:** 181–207

Horst I, Heimann L, Peterhansel C (2013) Signal integration on plant promoters: a case study in maize. Plant Signal Behav **8:** e25389

Horst I, Offermann S, Dreesen B, Niessen M, Peterhansel C (2009) Core promoter acetylation is not required for high transcription from the phosphoenolpyruvate carboxylase promoter in maize. Epigenetics Chromatin **2:** 17

Hotelling H (1931) The generalization of student's ratio. Ann Math Stat **2:** 360–378

Hussey SG, Mizrachi E, Groover A, Berger DK, Myburg AA (2015) Genome-wide mapping of histone H3 lysine 4 trimethylation in Eucalyptus grandis developing xylem. BMC Plant Biol **15:** 117

Kopriva S, Koprivova A (2005) Sulfate assimilation and glutathione synthesis in C4 plants. Photosynth Res **86:** 363–372

Kouzarides T (2007) Chromatin modifications and their function. Cell **128:** 693–705

Ku MS, Kano-Murakami Y, Matsuoka M (1996) Evolution and expression of C4 photosynthesis genes. Plant Physiol **111:** 949–957

Külahoglu C, Denton AK, Sommer M, Maß J, Schliesky S, Wrobel TJ, Berckmans B, Gongora-Castillo E, Buell CR, Simon R, et al (2014) Comparative transcriptome atlases reveal altered gene expression modules between two Cleomaceae C3 and C4 plant species. Plant Cell **26:** 3243–3260

Kumar V, Muratani M, Rayan NA, Kraus P, Lufkin T, Ng HH, Prabhakar S (2013) Uniform, optimal signal processing of mapped deep-sequencing data. Nat Biotechnol **31:** 615–622

Landt SG, Marinov GK, Kundaje A, Kheradpour P, Pauli F, Batzoglou S, Bernstein BE, Bickel P, Brown JB, Cayting P, et al (2012) ChIP-seq guidelines and practices of the ENCODE and modENCODE consortia. Genome Res **22:** 1813–1831

Langdale JA (1998) Cellular differentiation in the leaf. Curr Opin Cell Biol **10:** 734–738

Lauberth SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH, Roeder RG (2013) H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. Cell **152:** 1021–1036

Lauria M, Rossi V (2011) Epigenetic control of gene regulation in plants. Biochim Biophys Acta **1809:** 369–378

Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, Kebrom TH, Provart N, Patel R, Myers CR, et al (2010) The developmental dynamics of the maize leaf transcriptome. Nat Genet **42:** 1060–1067

Machanick P, Bailey TL (2011) MEME-ChIP: motif analysis of large DNA datasets. Bioinformatics **27:** 1696–1697

Majeran W, Cai Y, Sun Q, van Wijk KJ (2005) Functional differentiation of bundle sheath and mesophyll maize chloroplasts determined by comparative proteomics. Plant Cell **17:** 3111–3140

Majeran W, Friso G, Ponnala L, Connolly B, Huang M, Reidel E, Zhang C, Asakura Y, Bhuiyan NH, Sun Q, et al (2010) Structural and metabolic transitions of C4 leaf development and differentiation defined by microscopy and quantitative proteomics in maize. Plant Cell **22:** 3509–3542

Mathelier A, Zhao X, Zhang AW, Parcy F, Worsley-Hunt R, Arenillas DJ, Buchman S, Chen CY, Chou A, Ienasescu H, et al (2014) JASPAR 2014: an extensively expanded and updated open-access database of transcription factor binding profiles. Nucleic Acids Res **42:** D142–D147

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nat Methods **5:** 621–628

Nelissen H, Rymen B, Jikumaru Y, Demuynck K, Van Lijsebettens M, Kamiya Y, Inzé D, Beemster GT (2012) A local maximum in gibberellin levels regulates maize leaf growth by spatial control of cell division. Curr Biol **22:** 1183–1187

Offermann S, Danker T, Dreymüller D, Kalamajka R, Töpsch S, Weyand K, Peterhänsel C (2006) Illumination is necessary and sufficient to induce histone acetylation independent of transcriptional activity at the C_4-specific phosphoenolpyruvate carboxylase promoter in maize. Plant Physiol **141:** 1078–1088

Offermann S, Dreesen B, Horst I, Danker T, Jaskiewicz M, Peterhansel C (2008) Developmental and environmental signals induce distinct histone acetylation profiles on distal and proximal promoter elements of the C4-Pepc gene in maize. Genetics **179:** 1891–1901

Papait R, Cattaneo P, Kunderfranco P, Greco C, Carullo P, Guffanti A, Vigano V, Stirparo GG, Latronico MV, Hasenfuss G, et al (2013) Genome-wide analysis of histone marks identifying an epigenetic signature of promoters and enhancers underlying cardiac hypertrophy. Proc Natl Acad Sci USA **110:** 20164–20169

Pick TR, Bräutigam A, Schlüter U, Denton AK, Colmsee C, Scholz U, Fahnenstich H, Pieruschka R, Rascher U, Sonnewald U, et al (2011) Systems analysis of a maize leaf developmental gradient redefines the current C4 model and provides candidates for regulation. Plant Cell **23:** 4208–4220

Robinson MD, McCarthy DJ, Smyth GK (2010) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26:** 139–140

Ross-Innes CS, Stark R, Teschendorff AE, Holmes KA, Ali HR, Dunning MJ, Brown GD, Gojis O, Ellis IO, Green AR, et al (2012) Differential oestrogen receptor binding is associated with clinical outcome in breast cancer. Nature **481:** 389–393

Roudier F, Ahmed I, Bérard C, Sarazin A, Mary-Huard T, Cortijo S, Bouyer D, Caillieux E, Duvernois-Berthet E, Al-Shikhley L, et al (2011) Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. EMBO J **30:** 1928–1938

Salmon-Divon M, Dvinge H, Tammoja K, Bertone P (2010) PeakAnalyzer: genome-wide annotation of chromatin binding and modification loci. BMC Bioinformatics **11:** 415

Sekhon RS, Briskine R, Hirsch CN, Myers CL, Springer NM, Buell CR, de Leon N, Kaeppler SM (2013) Maize gene atlas developed by RNA sequencing and comparative evaluation of transcriptomes based on RNA sequencing and microarrays. PLoS One **8:** e61005

Song J, Angel A, Howard M, Dean C (2012) Vernalization - a cold-induced epigenetic switch. J Cell Sci **125:** 3723–3731

Stark R, Brown G (2011) DiffBind: differential binding analysis of ChIP-Seq peak data. http://bioconductor.org/packages/release/bioc/vignettes/DiffBind/inst/doc/DiffBind.pdf (March 2, 2015)

Studer AJ, Gandin A, Kolbe AR, Wang L, Cousins AB, Brutnell TP (2014) A limited role for carbonic anhydrase in C4 photosynthesis as revealed by a *ca1ca2* double mutant in maize. Plant Physiol **165:** 608–617

Tausta SL, Li P, Si Y, Gandotra N, Liu P, Sun Q, Brutnell TP, Nelson T (2014) Developmental dynamics of Kranz cell transcriptional specificity in maize leaf reveals early onset of C4-related processes. J Exp Bot **65:** 3543–3555

Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY, Stitt M (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J **37:** 914–939

To TK, Kim JM (2014) Epigenetic regulation of gene responsiveness in Arabidopsis. Front Plant Sci **4:** 548

Turner BM (2002) Cellular memory and the histone code. Cell **111:** 285–291

Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, van Schaik FM, Varier RA, Baltissen MP, Stunnenberg HG, Mann M, Timmers HT (2007) Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell **131:** 58–69

Wang L, Czedik-Eysenberg A, Mertz RA, Si Y, Tohge T, Nunes-Nesi A, Arrivault S, Dedow LK, Bryant DW, Zhou W, et al (2014) Comparative

analyses of C_4 and C_3 photosynthesis in developing leaves of maize and rice. Nat Biotechnol **32:** 1158–1165

Wang X, Elling AA, Li X, Li N, Peng Z, He G, Sun H, Qi Y, Liu XS, Deng XW (2009) Genome-wide and organ-specific landscapes of epigenetic modifications and their relationships to mRNA and small RNA transcriptomes in maize. Plant Cell **21:** 1053–1069

Yanagisawa S (2000) Dof1 and Dof2 transcription factors are associated with expression of multiple genes involved in carbon metabolism in maize. Plant J **21:** 281–288

Yanagisawa S, Sheen J (1998) Involvement of maize Dof zinc finger proteins in tissue-specific and light-regulated gene expression. Plant Cell **10:** 75–89

Yilmaz A, Nishiyama MY Jr, Fuentes BG, Souza GM, Janies D, Gray J, Grotewold E (2009) GRASSIUS: a platform for comparative regulatory genomics across the grasses. Plant Physiol **149:** 171–180

Zhang B, Chen W, Foley RC, Büttner M, Singh KB (1995) Interactions between distinct types of DNA binding proteins enhance binding to *ocs* element promoter sequences. Plant Cell **7:** 2241–2252

Zhang X (2008) The epigenetic landscape of plants. Science **320:** 489–492

Zhou J, Tang X, Martin GB (1997) The Pto kinase conferring resistance to tomato bacterial speck disease interacts with proteins that bind a cis-element of pathogenesis-related genes. EMBO J **16:** 3207–3218

Perduns et al. (2015) - Supplement

Supplemental figure S1: Read mapping statistics of ChIP-Seq and RNA-Seq libraries and comparison of RNA-Seq data from this study with data from Li et al., 2010.

A

RNA-Seq	Processed reads	Uniquely mapped reads	% (Uniquely mapped)
Base A	59,346,263	46,246,282	78%
Base B	33,432,364	24,687,777	74%
Blade A	74,724,999	53,086,298	71%
Blade B	27,582,965	18,668,247	68%

H3K9ac-ChIP	Processed reads	Uniquely mapped reads	% (Uniquely mapped)
Base A	29,553,783	23,724,782	80%
Base B	18,884,476	15,417,201	82%
Blade A	30,611,609	23,999,153	78%
Blade B	24,498,527	19,320,849	79%

H3K4me3-ChIP	Processed reads	Uniquely mapped reads	% (Uniquely mapped)
Base A	13,877,095	7,718,604	56%
Base B	7,915,293	6,624,821	84%
Blade A	16,204,778	10,212,086	63%
Blade B	25,950,063	21,664,983	83%

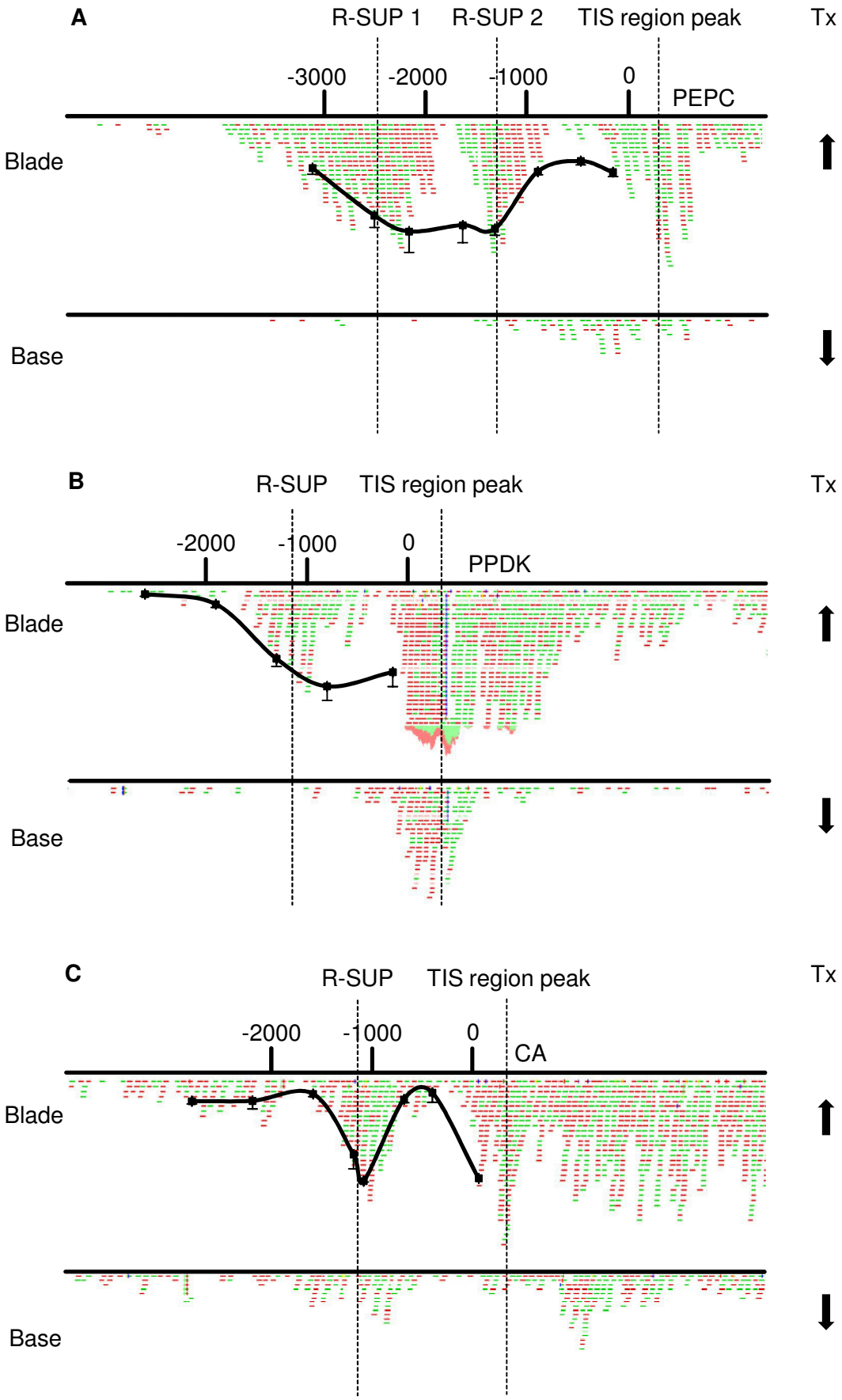
B

Base	Base A	Base B	Li et al., Base A	Li et al., Base B
Base A		0.9	0.94	0.89
Base B	0.9		0.9	0.93
Li et al., Base A	0.94	0.9		0.95
Li et al., Base B	0.89	0.93	0.95	

Blade	Blade A	Blade B	Li et al., Blade A	Li et al., Blade B
Blade A		0.93	0.96	0.96
Blade B	0.93		0.92	0.93
Li et al., Blade A	0.96	0.92		1
Li et al., Blade B	0.96	0.93	1	

Figure S1: Read mapping statistics of ChIP- and RNA-Seq libraries and comparison of RNA-Seq data from this study with data from Li et al., 2010. **A** Summary of processed reads and mapping statistics of ChIP- and RNA-Seq libraries for all replicates and samples. **B** Pearson correlation coefficients between analogous fragments from Li et al. (2010) and this study.

Perduns et al. (2015) - Supplement
Supplemental figure S2: H3K9ac profiles on selected C4 genes at the leaf base and blade.



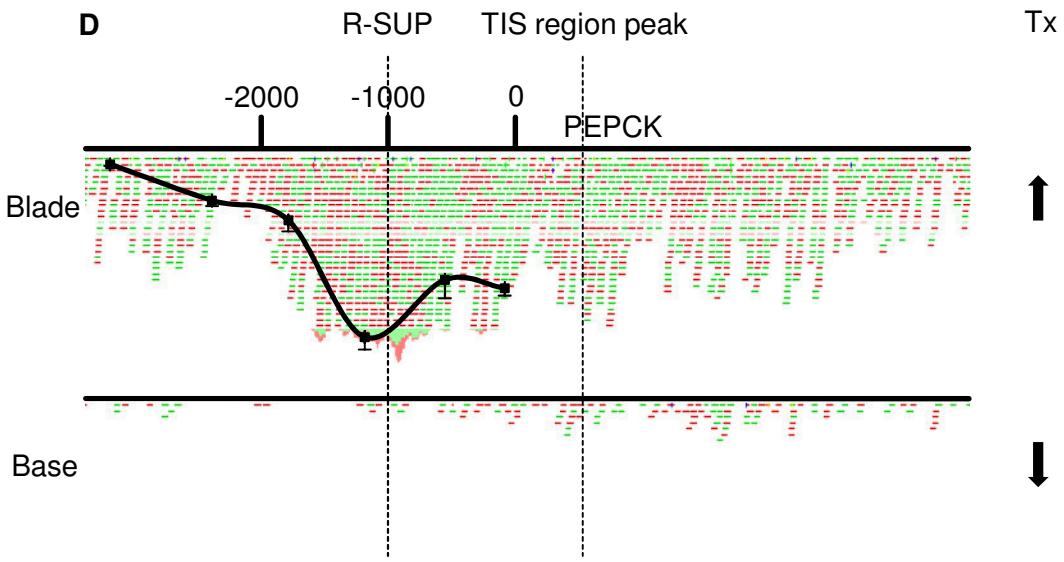


Figure S2: H3K9ac profiles on selected C4 genes at the leaf base and blade. Positions are relative to TIS in bp. Red and green dots represent mapped forward and reverse ChIP-Seq reads. The solid lines represent previously obtained ChIP-qPCR data from the mature leaf (Heimann et al., 2013). Arrows represent high and low transcript levels (Tx=Transcription). PEPC = phosphoenolpyruvate carboxylase, PPDK = pyruvate-Pi-dikinase, CA = carbonic anhydrase, PEPCK = phosphenolpyruvate carboxykinase.

Chapter 4 - Manuscript I

Gene regulation along the maize leaf gradient - a consequence of light?

Renke Perduns and Christoph Peterhansel

Leibniz University Hannover, Institute of Botany, 30419 Hannover, Germany

Type of authorship:	first author
Type of article:	research article
Share of the work:	95%
Contribution to the manuscript:	performed all experiments, analyzed the data, prepared all figures and basically wrote the manuscript

Gene regulation along the maize leaf gradient - a consequence of light?

Authors

Renke Perduns and Christoph Peterhansel*

Leibniz University Hannover, Institute of Botany, Herrenhäuser Straße 2, 30419 Hannover, Germany

* Corresponding author: Christoph Peterhansel, tel: +49-511-7622632,
email: cp@botanik.uni-hannover.de

Key words: *Zea mays*, developmental gradient, light regulation, histone modification, histone acetyltransferases, histone deacetylases

Abstract

Various studies analyzed the developmental gradient of the maize (*Zea mays*) leaf to understand the gradual establishment of Kranz anatomy and C4 photosynthesis. We characterized the influence of light on genes that were differentially expressed from the base to the blade of the maize leaf by a comparative analysis of RNA-Seq data from three different tissues: base (division zone), illuminated blade (maturing zone) and blade exposed to prolonged darkness for 52 hours. We classified genes that were differentially expressed from the base to the blade into four classes; repressed (Rep) or induced (Ind) from the base to the blade in a light-dependent (Ld) or light-independent (Li) manner. The majority of Ind-genes were regulated light-dependently while Rep-genes were mainly regulated light-independently. Meta-analysis of classified genes showed that they were equally co-regulated with the histone modification histone 3 lysine 9 acetylation (H3K9ac). We also showed that regulated secondary upstream peaks (R-SUPs) of H3K9ac were not significantly enriched within the defined classes of regulation. Additionally, we identified histone acetyltransferases (HATs) and histone deacetylases (HDACs) as potential regulators of classified genes based on their expression profiles.

Introduction

The leaf gradient of monocotyledonous plants is a useful platform to study the establishment of various biological processes in a gradually developing system. The base of the growing monocot leaf is characterized by meristematic and undifferentiated cells that are continuously dividing (division zone; Sylvester et al., 1990; Nelissen et al., 2013). The size of the division zone strongly affects the final size of the mature leaf (Rymen et al., 2007; Nelissen et al., 2012). Undifferentiated cells expand and differentiate adjacent to the division zone; this zone was defined as expansion zone. When the expansion of cells is completed they are integrated into the mature zone where all structural and physiological parameters for a functional metabolism are established (Sylvester et al., 1990; Majeran et al., 2010; Li et al., 2010; Nelissen et al., 2013). Thus, all stages of various biological processes can be analyzed in a single leaf at a given time point of development.

Specifically, the analysis of the developmental gradient of the maize leaf is of interest to characterize the establishment and regulation of C4 photosynthesis. Kranz anatomy is essential to perform C4 photosynthesis and is established by the differentiation of cells into bundle sheath and mesophyll cells on the developmental gradient (Langdale, 1998; Majeran et al., 2005; Majeran et al., 2010). In recent years, numerous studies characterized the dynamics of metabolites, proteins, transcripts and anatomical structures along the developmental gradient of the maize leaf to understand processes and their regulation required to establish C4 photosynthesis (Li et al., 2010; Majeran et al., 2010; Pick et al., 2011; Tausta et al., 2014). Additionally, transcriptional profiles along the developmental gradient of maize were compared to the monocotyledonous C3 plant rice and the dicotyledonous C4 plant *Gynandropsis* to identify genes associated with C4 photosynthesis (Wang et al., 2014; Aubry et al., 2014). We previously extended the knowledge about gene regulation along the maize leaf gradient by analyzing the impact of the histone modifications histone 3 lysine 9 acetylation (H3K9ac) and histone 3 lysine 4 trimethylation (H3K4me3) on the transcription of associated genes on a genome-wide level. Wang et al. (2009) described previously that H3K9ac and H3K4me3 were found at transcription initiation sites of actively transcribed genes in shoots and roots of maize. We analyzed the dynamics of both modifications and the transcription of associated genes and observed a good correlation between transcriptional regulation of genes and associated H3K9ac and H3K4me3. Additionally, we identified and characterized regulated secondary upstream peaks (R-SUPs) as a unique feature of H3K9ac. R-SUPs were enriched on genes that were induced from the base to the blade and especially on C4 associated genes (Perduns et al., 2015).

The characterization of factors that are responsible for the regulation of gene expression along the developmental gradient has widely been overlooked so far. It was shown that the endogenous plant hormones cytokinin, auxin, castasterone and gibberellin were enriched in the division and expansion zone of the maize leaf and regulated leaf growth (Nelissen et al., 2012). Light, as an important environmental stimulus, is another promising candidate for the regulation of genes that are differentially expressed from the base to the blade of the maize leaf. The base of secondary maize leaves is shaded by older leaves while the maturing zone is illuminated. It has been described that light is required to establish functional photosynthesis and controls the expression of various genes through the perception by at least four different classes of photoreceptors: phytochromes, cryptochromes, phototropins and unidentified ultraviolet B (UVB) photoreceptors. These receptors use distinct pathways to interpret and transduce the perceived light signals to alter the gene expression for an appropriate adaptation of cellular functions and development (Sheen, 1999; Ma et al., 2001; Li et al., 2011). For example, it was shown that the expression of the small subunit of the ribulose-1,5-bisphosphate carboxylase-oxygenase (*RbcS*) is regulated in a phytochrome and cryptochrome dependent manner (Martinez-Hernandez et al., 2002; Oh and Montgomery, 2011). *RbcS* and other light regulated genes share conserved promoter elements, like G- and I-boxes, that are summarized as light-responsive *cis*-elements (LREs; Terzaghi and Cashmore, 1995). It has been shown that LREs are targeted by various transcription factors and that their structural arrangement affects the transcriptional responsiveness of corresponding genes to different light conditions (Menkens et al., 1995; Lopez-Ochoa et al., 2007). Several studies indicate that light regulation is mediated by histone modifications. Jang et al. (2011) observed that the transcription of *Arabidopsis* phytochrome A itself was tightly regulated by histone modifications. The regulatory impact of histone modifications on the light-dependent expression of phosphoenolpyruvate carboxylase (*Pepc*) of maize was characterized in detail in our lab (Offermann et al., 2006; Offermann et al., 2008; Horst et al., 2009). The observed histone modification profiles of *Pepc* were partially conserved on other C4 genes in maize and in the related C4 grasses *Sorghum* and *Setaria* (Horst et al., 2013; Heimann et al., 2013).

In this study, we aimed to characterize the influence of light on genes that were differentially expressed comparing base and blade of the maize leaf. In a first step, we compared RNA-Seq data from two developmental zones of the maize leaf, the base (division zone) and blade (maturing zone), to identify genes that were differentially expressed. In a second step, we used RNA-Seq data of blade samples of plants that were exposed to 52 hours of prolonged darkness (re-etiolated) to classify the regulation of genes from the base to the blade into light-dependent and -independent regulation. The resulting gene groups were

characterized and subsequently analyzed for their co-regulation with the histone modification H3K9ac and their association with R-SUPs in a meta-analysis. Additionally, the expression profiles of histone acetyltransferases (HATs) and histone deacetylases (HDACs) were screened to identify potential regulators of H3K9ac in the analyzed zones of the maize leaf.

Results and Discussion

Assessment of data quality and clustering of RNA-Seq samples

We isolated RNA from two developmental zones of the maize leaf: base (division zone) and blade (maturing zone). The zones correspond well to segments -4 and +4 from Li et al. (2010; Perduns et al., 2015). To investigate the influence of light on genes that are regulated from the base to the blade we furthermore isolated RNA from blades of maize plants that were exposed to 52 hours of prolonged darkness (re-etiolated). The abundance of mRNAs was analyzed by RNA-Seq. Re-etiolated blades were taken as a model for a leaf blade that developed in the absence of light. An alternative would have been blades from etiolated plants, because maize develops true leaves even when grown in complete darkness (Heimann et al., 2013). However, these blades differ in structure and shape from blades derived from plants grown in the light (Charron et al., 2009). Therefore, we decided to analyze blades from re-etiolated plants to investigate the influence of light on gene regulation from the base to the blade. Offermann et al. (2006) showed before that transcription of *Pepc* in re-etiolated maize plants was on a basal level, comparable to the expression in etiolated plants.

To assess quality and clustering of RNA-Seq data, we analyzed the correlation between replicates of the investigated samples (Table 1). Pearson correlation coefficients between the associated biological replicates were above 0.9 for all three samples. Interestingly, correlation coefficients between replicates from the blade and the re-etiolated blade ranged from 0.2 to 0.26 while the correlation coefficients between the base and the re-etiolated blade samples ranged from 0.08 to 0.15. Taken together, the transcriptomes of re-etiolated blade samples clustered closer to transcriptomes of the blade compared to base samples.

Thus, transcriptional changes caused by re-etiolation of the blade were smaller compared to changes of the transcriptomes from the base to the blade. The re-etiolation of the blade alone is not sufficient to reconstitute the transcriptome of the base within the differentiated blade, indicating that multiple factors are responsible for the transcriptional alterations from the base to the blade of the maize leaf. Additionally, transcriptional changes caused by re-etiolation of the blade might be different to changes caused by the initial illumination during leaf development.

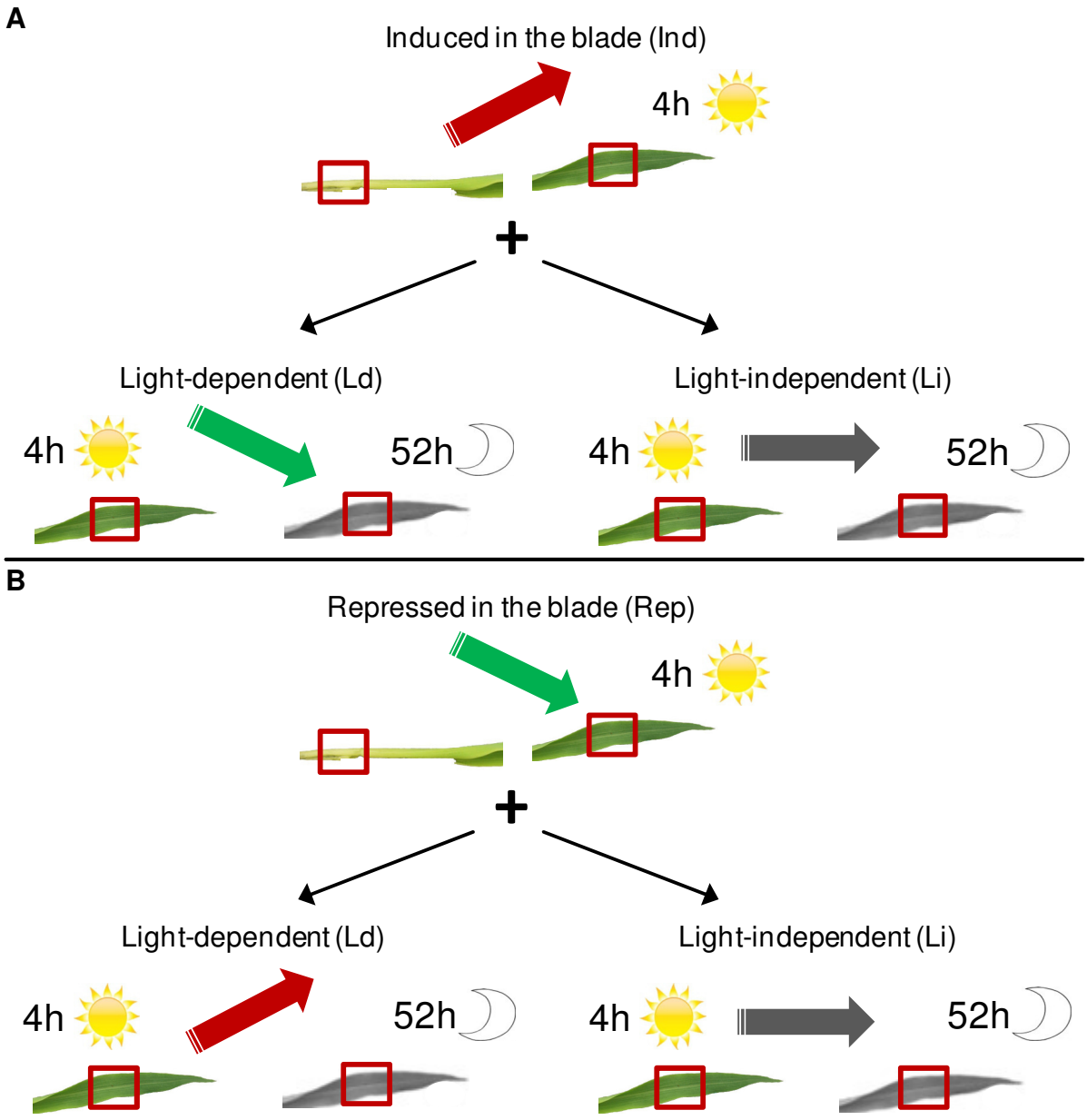
	Base_A	Base_B	Blade_A	Blade_B	Re-et. blade_A	Re-et. blade_B
Base_A		0.9	0.1	0.08	0.13	0.11
Base_B	0.9		0.14	0.15	0.1	0.08
Blade_A	0.1	0.14		0.93	0.26	0.25
Blade_B	0.08	0.15	0.93		0.22	0.2
Re-et. blade_A	0.13	0.1	0.26	0.22		0.98
Re-et. blade_B	0.11	0.08	0.25	0.2	0.98	

Table 1: Correlation of replicates from different tissues. Pearson correlation coefficients between all incorporated tissues (base, blade and re-etiolated [Re-et.] blade) and replicates (A/B) based on the expression values.

Classification of transcriptional regulation from the base to the blade of the maize leaf

To analyze the impact of light on the regulation of individual genes from the base to the blade, we classified the regulation into light-dependent and -independent groups with a stepwise approach (Figure 1). First of all, genes had to be significantly regulated from the base to the blade (fold change ≥ |10| with a false discovery rate [FDR] ≤ 0.05). For this step, only protein coding genes with reads per kilobase per million mapped reads (RPKM) > 0.1 in one of the analyzed zones were taken into consideration. Genes that were regulated from the base to the blade were further classified according to their transcriptional response upon re-etiolation of the blade in a second step. We defined groups of light-independently regulated genes that showed additional regulation between illuminated and re-etiolated blade according to Figure 1 (fold change ≥ |10| with an FDR ≤ 0.05). Genes that were not significantly regulated in the re-etiolated blade compared to the blade were classified as light-independent (FDR > 0.05). The classes of regulation are abbreviated according to Figure 1C.

We used comparably strict thresholds of fold changes for the definition of classes to ensure that the members were reliable. As a drawback, genes that were not regulated > |10| fold in one of the comparisons were not classified. To assess the influence of the used thresholds we analyzed the overlap between the defined classes when thresholds of a |10| or |2| fold regulation were used. Overall, 70% of the classified genes using the 2-fold threshold were still classified when using the 10-fold threshold (data not shown). We decided to use the |10| fold threshold in order to ensure that only truly regulated genes were included.



C

	Induced in blade vs base	Repressed in blade vs base
Light-dependent	Ind_Ld	Rep_Ld
Light-independent	Ind_Li	Rep_Li

Figure 1: Schematic overview of the definition from light-dependently and light-independently regulated genes that were differentially expressed between blade and base **A** Genes that were induced in the blade compared to the base and repressed in response to re-etiolation (Fold change ≥ |10|, FDR ≤ 0.05) were classified as light-dependent genes. Light-independently regulated genes were not regulated in response to re-etiolation (FDR > 0.05). Only protein coding genes with a mean RPKM > 0.1 in the blade or at the base were taken into consideration. **B** Genes that were repressed in the blade compared to the base and induced in response to re-etiolation (Fold change ≥ |10|, FDR ≤ 0.05) were classified as light-dependent genes. Light-independently regulated genes were not regulated in response to re-etiolation (FDR > 0.05). Only protein coding genes with a mean RPKM > 0.1 in the blade or at the base were taken into consideration. **C** Abbreviations of the defined classes.

Figure 2 summarizes numbers of genes in the defined classes. Overall 5938 genes were significantly regulated from the base to the blade. 2582 genes were induced in the blade compared to the base (Ind-genes) and 3356 genes were repressed (Rep-genes). 1191 Ind-genes were further classified as light dependent (Ind_Ld-genes) and 780 as light-independent genes (Ind_Li-genes), respectively. 265 Rep-genes were classified as light dependent (Rep_Ld-genes) and 2752 as light-independent (Rep_Li-genes). 611 Ind-genes and 339 Rep-genes showed an aberrant regulation upon re-etiolation and could not be classified clearly. For example, these genes were not regulated ≥ |10| fold (albeit significantly) in the re-etiolated blade compared to the blade.

The majority of Ind-genes were transcriptionally affected by light (60%), but, surprisingly, a notable amount of Ind-genes was not affected (40%). For Rep-genes the relation was inverted. 91% were classified as Rep_Li-genes and only 9% as Rep_Ld-genes. Thus, light was an important factor for the induction of genes during leaf development while Rep-genes were not affected by light. Several additional stimuli differ between the base and blade and could contribute to gene regulation during leaf development: for example, hormones, exposure to ambient air or the content of metabolic compounds (Pick et al., 2011; Nelson, 2011; Nelissen et al., 2012; Wang et al., 2014). Additionally, meristematic cells divide at the base of the leaf and differentiate into functional tissues along the leaf gradient (Nelissen et al., 2012). For example, the Kranz anatomy is gradually established and it was shown that many genes were specifically expressed in bundle sheath or mesophyll cells, respectively. The expression of such genes is potentially influenced by the developmental stage or the presence of the specific tissue (Majeran et al., 2010; Tausta et al., 2014; Chang et al., 2012). Thus, the light-independent regulation of the majority of Rep-genes and a considerable amount of Ind-genes is reasonable.

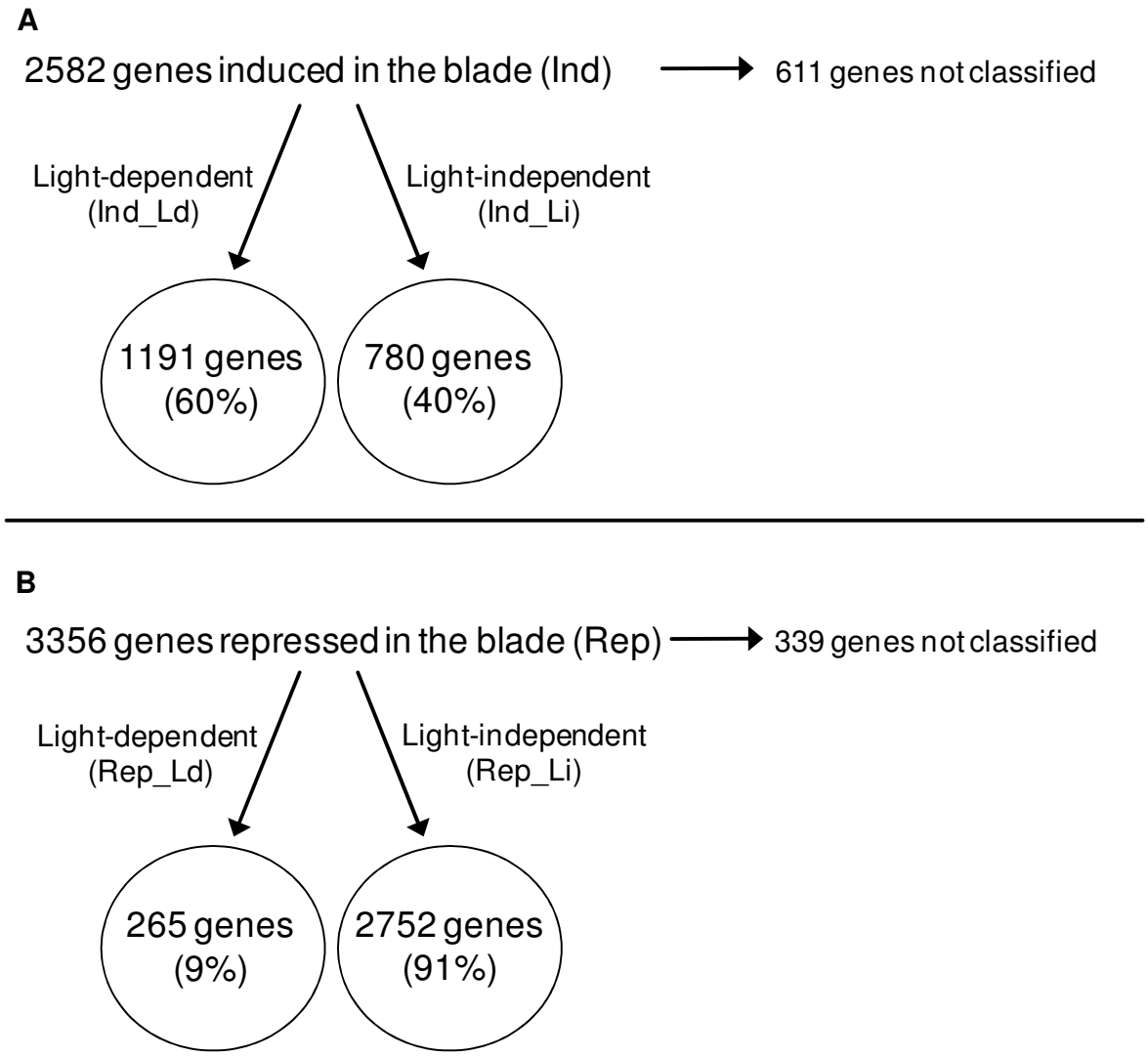


Figure 2: Quantities of genes within the defined classes of regulation. **A** Quantities of Ind-genes that were classified according to the definition from Figure 1A. Percentages were calculated relative to the sum of Ind-genes that could be classified further. **B** Quantities of Rep-genes that were classified according to the definition from Figure 1B. Percentages were calculated relative to the sum of Rep-genes that could be classified further.

Characterization of classified genes

To evaluate if the defined classes of regulation could be associated with distinct processes we tested whether genes from biological processes defined by MapMan bins were enriched (Figure 3; Thimm et al., 2004) in any of the classes. Four processes were significantly over- or underrepresented within both classes of Ind-genes.

The most striking process was *photosynthesis* which was enriched within Ind_Ld-genes and heavily underrepresented within Ind_Li-genes. It is known that the establishment of photosynthesis is dependent on light and that many photosynthetic genes are regulated by light (Sheen, 1999; Sakamoto et al., 2008; Heimann et al., 2013). Under- and overrepresentation of photosynthetic genes within Ind_Li- and Ind_Ld-genes was therefore a

reliable indicator for the validity of the defined classes. Additionally, genes from the process *RNA* were underrepresented and genes from the process *secondary metabolism* were enriched within Ind_Ld-genes. Within Ind_Li-genes, only the process *cell* was enriched.

C4 genes mainly belong to the process *photosynthesis*. We analyzed C4 genes defined by Chang et al. (2012) manually and observed that from 16 genes associated with C4 photosynthesis 14 were classified as Ind-genes and 12 were further classified as Ind_Ld-genes (Supplemental table S2). Pyruvate-P_i-dikinase (*Ppdk*) and malic enzyme (*Me*) were not classified as Ind-Ld-genes because they were repressed less than 10 fold (albeit significantly) between illuminated and re-etiolated blade. Transcript amounts were still on a high level in the re-etiolated blade with an RPKM of 2459 for *Ppdk* and 942 for *Me*. This might indicate that both genes are also participating in metabolic processes which are not light-dependent or that their mRNAs are more stable compared to other C4-associated genes.

Members of the process *secondary metabolism* were also enriched within Ind_Ld-genes. We analyzed the sub processes in detail and observed that genes associated with isoprenoid and phenylpropanoid metabolism were significantly enriched within Ind_Ld-genes. The ratio from observed to expected counts was 1.7 for both sub processes with p-values of the χ^2-test < 0.02 (data not shown). Metabolites of both classes are known to be involved in the absorption of photons (Dixon and Paiva, 1995; Ruiz-Sola and Rodríguez-Concepción, 2012; Singh and Sharma, 2015). Thus, the light-dependent induction of associated genes in the blade is expected.

Within Ind_Li genes, the process *cell* was enriched. This included the sub processes cell-organization, -division, -cycle and -motility. In general, the majority of associated genes are repressed along the developmental gradient (Li et al., 2010). Associated genes which are induced in the blade might be responsible for maintenance and organization of the cellular organization, including Kranz anatomy, at the blade of the leaf. Accordingly, Chang et al., (2012) found 173 genes, associated with the process *cell* that were enriched in bundle sheath (156) or mesophyll cells at the blade of the leaf (17).

Additionally, genes from the process *RNA* were underrepresented within Ind_Ld-genes. Genes from the process *RNA* are mainly transcription factors. Transcription factors are involved in the regulation of various processes. Their underrepresentation within Ind_Ld-genes indicates that Ind-Ld-genes might be regulated by common transcription factors or that responsible transcription factors are not transcriptionally co-regulated with comparable fold changes.

However, no process was significantly over- or underrepresented within Rep-genes. Sizes of both classes of Rep-genes hindered statistical enrichment analyses. The Rep_Ld class was

very small and a strong under- or overrepresentation was needed for significance. On the other hand, the class of Rep_Li genes nearly comprises all Rep-genes and within this group, no biological process was significantly enriched.

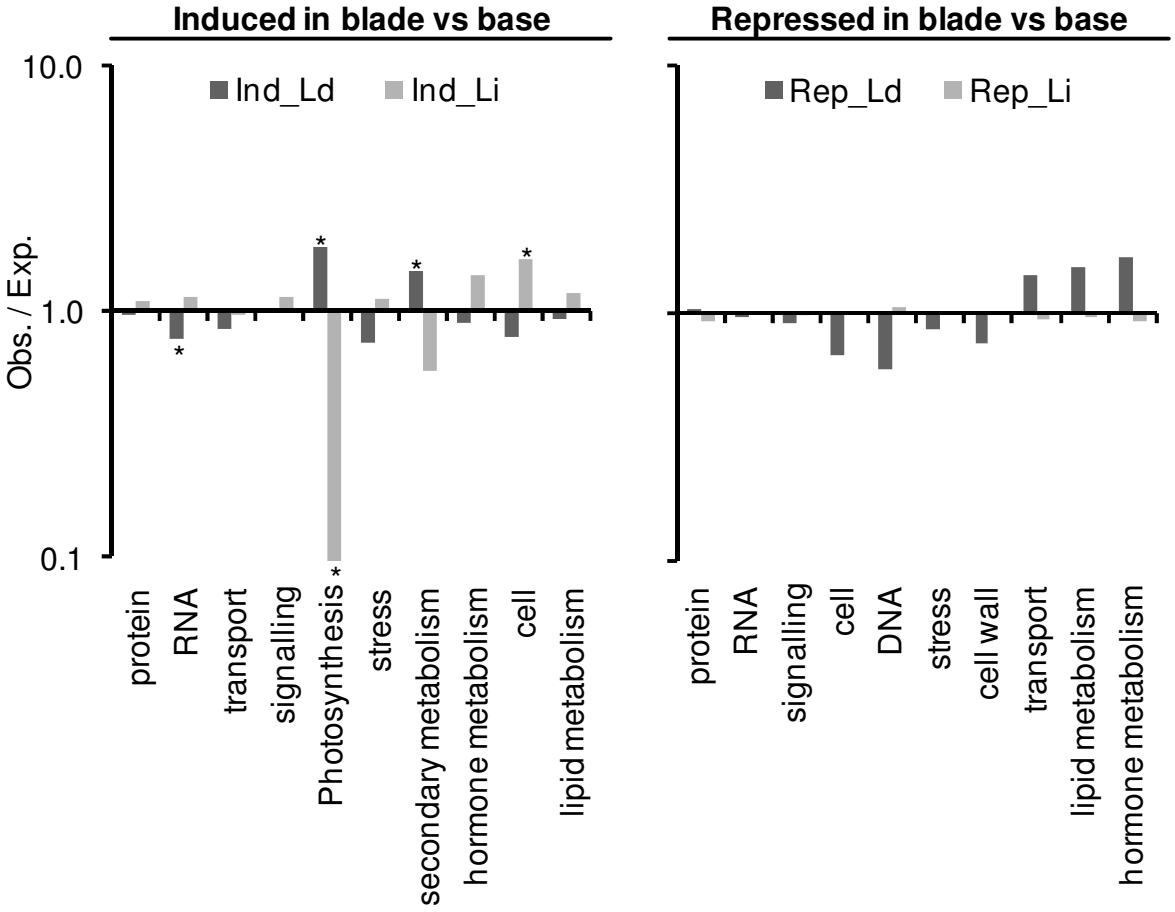


Figure 3: Enrichment analysis of processes within the defined classes of regulation. Processes were defined by MapMan bins. Only processes were analyzed if at least five counts were expected in the tested group. Ten processes with the highest expectations are shown. The degree of enrichment is given as ratios of expected (Exp.) and observed (Obs.) counts (Log_{10}-transformed y-axis). The significances of enrichments were evaluated by χ^2-test and derived p-values were adjusted using Benjamini Hochberg correction. *: FDR ≤ 0.05.

Meta-analysis of H3K9ac signals associated with defined classes of regulation

In a previous study, we analyzed the quantitative relationship between the histone modification H3K9ac and transcription of associated genes by comparing the dynamics of both parameters from the base to the blade of the maize leaf. We identified and characterized regulated secondary upstream peaks (R-SUPs) as a unique feature of the histone modification H3K9ac. We observed that R-SUP associated genes were enriched within genes that were induced in the blade compared to the base and especially on photosynthetic genes (Perduns et al., 2015). In this study, we showed that light is a crucial factor for the induction of photosynthetic genes from the base to the blade and we were

interested if the occurrence of R-SUPs can be linked to light regulation of associated genes. We therefore analyzed if R-SUP associated genes were enriched within the defined classes of regulation (Figure 4).

R-SUP associated genes were not significantly enriched in any of the defined classes, but tendencies for enrichment within Ind_Ld-genes could be observed. 81 R-SUP associated genes were expected and 92 were observed. The p-value for the enrichment was 0.09. Within Ind_Li-genes 55 R-SUP associated genes were expected and 46 observed (p-value = 0.1). For both classes of Rep-genes, no such tendencies could be observed.

We showed that the majority photosynthetic genes were upregulated on the gradient and downregulated by re-etiolation (Ind-Ld genes; Figure 3). Therefore, the tendencies for over- and underrepresentation of R-SUP associated genes within Ind_Ld- and Ind_Li-genes might be caused by the previously shown association of R-SUPs with photosynthetic genes. To test this hypothesis, we excluded all photosynthetic genes from Ind-genes and re-analyzed the enrichment within both classes. The degree of over- and underrepresentation declined and the p-values increased to 0.3 for both classes (data not shown), indicating that over- and underrepresentation of R-SUP associated genes within Ind_Ld- and Ind_Li-genes was mainly caused by photosynthetic genes. On the other hand, this observation underpins the specific association of R-SUPs with photosynthetic genes, independent of the class of regulation.

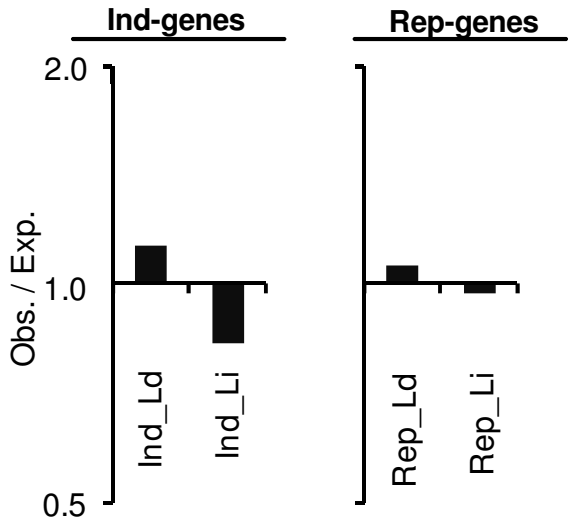


Figure 4: Enrichment analysis of R-SUP associated genes within the defined classes of regulation. R-SUP associated genes were identified in the study Perduns et al. (2015). All Ind- and Rep-genes associated with H3K9ac peaks, were chosen as reference for the indicated groups. The degree of enrichment is given as ratios of expected (Exp.) and observed (Obs.) counts (Log_{10}-transformed y-axis). The significances of enrichments were evaluated by χ^2-test.

In a previous study, we also showed a good correlation of fold changes from the base to the blade of H3K9ac and the transcription of associated genes (Perduns et al., 2015). We were interested whether this holds true for the different classes of regulation defined here (Figure 5). Percentages of genes modified with H3K9ac were similar between the defined classes of regulation. Around 80% of all classified genes were associated with H3K9ac. Also percentages of H3K9ac modifications that were co-regulated with transcription of associated genes were similar between the classes. Around 90% of all analyzed H3K9ac peaks were significantly co-regulated with the transcription of associated genes, while 3% to 4% were regulated opposite to the transcription in each case.

This suggests that genes that were regulated from the base to the blade of the maize leaf in a light-dependent or -independent manner used H3K9ac to regulate transcription. It has previously been shown that transcription was co-regulated with H3K9ac between shoots and roots from maize (He et al., 2013). This indicates that H3K9ac is tightly linked to the transcriptional activity of associated genes in general.

A

Class of regulation		Genes	H3K9ac modified genes
Ind	Ld	1192	961 (81%)
	Li	781	625 (80%)
Rep	Ld	265	220 (83%)
	Li	2752	2157 (78%)

B

Transcription	Ld-genes		Li- genes	
Ind-genes	4%	86%	3%	88%
Rep-genes	90%	4%	92%	3%
	Rep-H3K9ac	Ind-H3K9ac	Rep-H3K9ac	Ind-H3K9ac

H3K9ac

Figure 5: Relationship between the regulation of transcription and H3K9 acetylation. **A** Genes which were classified according to Figure 1 were analyzed for the frequencies of associated H3K9ac modification. The percentages were calculated relative to the number of genes within each class of regulation. **B** The frequencies of significant co-regulation between the transcription and H3K9ac were calculated. The numbers represent the percentages of regulated genes which were associated with significantly regulated H3K9ac according to the indicated regulation (FDR ≤ 0.05). The values of H3K9ac modified genes from Figure 5A were taken as 100% for each transcriptional class of regulation. Rep, repressed in the blade compared to the base; Ind, induced in the blade compared to the base.

We wanted to understand how H3K9ac is regulated between the base and the blade of the maize leaf. In general, histone acetylations are incorporated by histone acetyltransferases (HATs) and removed by histone deacetylases (HDACs; Pandey et al., 2002; Chen and Tian, 2007). It was shown that H3K9ac is associated with actively transcribed genes (Figure 5; Dion et al., 2005; Wang et al., 2009; Perduns et al., 2015). Therefore, HATs of H3K9ac can be regarded as positive regulators of transcription and HDACs of H3K9ac as negative regulators. We expected that the expression of HATs is correlated with potential target genes and that the expression of HDACs is negatively correlated with potential target genes. We classified HATs and HDACs of the maize genome manually to find candidate enzymes for the regulation of H3K9ac within the analyzed zones of the maize leaf (Table 2).
Three out of 16 annotated HDACs could be classified based on the scheme in Figure 1. *Hdt103*, 104 and *Hda119* were classified as Rep_Li-genes. These HDACs are candidates for the suppression of genes at the base that were induced in the blade. They were also not regulated by light according to our criteria. All other HATs and HDACs could not be classified. This is probably due to the very strict criteria for the classification.
However, the expression profiles of HATs and HDACs that were not classified also allow for the estimation of potential target genes. For example, *Hda101* and *Hda110* were mainly expressed in the re-etiolated blade and at the base. These might be potential HDACs for the regulation of Ind_ld-genes. *Hda120* was the only HDAC with enhanced expression in the blade and might be responsible for the suppression of genes at the blade.

Overall, HATs were expressed at lower levels and their regulation was weaker compared to HDACs. No HAT could be classified based on the scheme in Figure 1. *Ham102* was the highest expressed HAT and the only one with enhanced expression in the blade. It is likely that *Ham102* induces gene expression in the blade. *Hag101* and *Hag102* were strongly expressed at the base and only moderately in the re-etiolated blade, indicating a role in the induction of H3K9ac at the base of the leaf. Surprisingly, the expression of seven HATs was strongest in the re-etiolated blade. Potentially, HATs are more important for the regulation of the transcriptional adaptations caused by the absence of light compared to HDACs. Accordingly, it was shown that the expression of rice HATs is regulated by various stresses (Liu et al., 2012). However, transcription of HATs or HDACs alone does not necessarily indicate activity. There are several examples that individual HATs and HDACs interact with non-histone proteins to become activated or to be transported from the cytosol to the nucleus upon specific stimuli (Bhat et al., 2004; Gonzalez et al., 2007; Alinsug et al., 2012; Kuang et al., 2012). Conversely, the absence of a specific RNA indicates that the enzyme is inactive. Further experiments are needed to understand the role of individual HATs and HDACs in the regulation of gene expression along the developing maize leaf.

A

HDAC	Locus	Base	Blade	Re-et. blade	Class of regulation
Hdt101	GRMZM2G057044	1.5	0.2	0.8	-
Hdt102	GRMZM2G100146	140.7	44.7	17.6	-
Hdt103	GRMZM2G159032	145.3	5.0	14.6	Rep_Li
Hdt104	GRMZM5G898314	23.9	0.5	0.6	Rep_Li
Srt101	GRMZM2G058573	7.6	4.8	12.5	-
Hda101	GRMZM2G172883	41.5	28.7	50.9	-
Hda102	GRMZM2G119703	14.0	3.5	8.4	-
Hda108	GRMZM2G136067	29.3	9.6	11.0	-
Hda109	GRMZM2G457889	21.7	15.4	15.8	-
Hda110	GRMZM2G107309	10.6	3.4	9.7	-
Hda115	GRMZM2G456917	Not expressed			
Hda116	GRMZM2G081474	8.6	0.6	2.3	-
Hda117	GRMZM2G046824	4.3	1.4	13.6	-
Hda118	GRMZM2G456473	1.2	1.2	0.2	-
Hda119	GRMZM2G163572	0.5	0.0	0.1	Rep_Li
Hda120	GRMZM2G008425	0.9	6.6	2.5	-

B

HAT	Locus	Base	Blade	Re-et. blade	Class of regulation
Hag101	GRMZM2G046021	18.8	4.2	8.4	-
Hag102	GRMZM5G851405	17.5	1.4	1.4	-
Hag103	GRMZM2G100872	11.3	3.3	18.8	-
Hag107	GRMZM2G359735	6.2	1.9	10.4	-
Ham101	GRMZM2G044126	9.0	8.3	11.7	-
Ham102	GRMZM2G140288	11.1	22.9	9.8	-
Hac101	GRMZM2G069886	8.2	4.3	14.6	-
Hac111	GRMZM2G470556	5.2	7.5	17.1	-
Hac113	GRMZM2G139977	9.1	9.2	14.9	-
Hac115	GRMZM2G335438	Not expressed			
Haf101	GRMZM2G394722	5.7	4.9	15.4	-

Table 2: Expression values and defined classes of regulation of annotated HDACs (**A**) and HATs (**B**). The color code indicates expression strength (dark = high) and is applied for each gene independently. Re-et., re-etiolated.

Material and methods

Plant material and growth conditions

Maize (B73) plants were cultivated in growth chambers under long-day conditions. The photoperiod was adjusted to 16 h of light with a photon flux intensity of 130 µmol s^{-1} m^{-2} and a temperature of 25°C. During the dark period, maize plants were grown at 20°C for 8 h. Water soaked seedlings were transferred to soil (VM, Einheitserde, Sinntal) and cultivated until the third leaf was fully expanded (13 to 14 days). Blade and base samples were harvested from the third leaf of the identical plants 4 h after the onset of light. 4 cm from the middle of the illuminated part of the leaf were harvested as blade samples. Plants were cut 0.5 cm above the soil and the lowest 1-2 cm of the third leaf with no access to light were harvested for base samples. Plants for the analysis of the influence of light on gene expression along the leaf gradient were transferred to darkness just before the onset of light and kept in prolonged darkness for 52 h (re-etiolated) so that they were harvested at identical time points of the day compared to the blade and base samples.

RNA isolation

Harvested leaf sections from six plants per biological replicate were pooled and ground in liquid nitrogen. About 50 to 60 mg of ground plant material was dissolved in 1 mL of Trizol and agitated for 15 min; 0.2 volumes of chloroform were added. Samples were agitated for 10 min, and phases were separated by centrifugation (16,100g at 4°C for 15 min); 400 µL of the aqueous phase was transferred to a new reaction tube and washed with 1 volume of chloroform. The RNA was precipitated by addition of 2 volumes of 96% (v/v) ice-cold ethanol, incubation at -20°C for 30 min, and centrifugation (16,100g at 4°C for 15 min). The precipitate was washed with 70% (v/v) ice-cold ethanol and dissolved in 30 µL of RNase-free water. The isolated RNA was treated with the RNase Free DNase Set (79254; Qiagen, Venlo), purified with the RNeasy Mini-Elute Cleanup Kit (74204; Qiagen, Venlo), and eluted two times in 14 µL of RNase-free water. Kits were used according to the manufacturer's instructions. The integrity of RNA was analyzed by agarose gel electrophoresis, and the concentration was measured photometrically; 2 µg of RNA per biological replicate was submitted to sequencing on an Illumina HiSeq2000 Instrument with 50 bp single-end reads (GATC Biotech). The service included quality assessment on a 2100 Bioanalyzer, complementary DNA synthesis from fragmented poly(A)+ RNA, and library preparation.

Processing and analysis of RNA-Seq data

Quality of raw sequencing reads was checked with FastQC (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and reads were processed with

the Trimmomatic software (Bolger et al., 2014). The adapter clipper, leading trimmer, trailing trimmer, and sliding window trimmer modules were used with their default settings to trim the raw sequences. Processed sequences were imported into the CLC Genomics Workbench 7.5.1 (http://www.clcbio.com/products/clc-genomics-workbench/; Qiagen, Venlo). Read mapping was carried out with the RNA-Seq Analysis Tool. Reads were mapped with default settings only to regions that were covered by gene annotations of the maize reference genome (Zea_mays.AGPv3.21.dna.toplevel; ftp://ftp.ensemblgenomes.org/pub/plants/release-21/). Reads that mapped equally to multiple positions were ignored. Raw read counts were chosen as expression values. RPKM values were calculated for each gene to quantify absolute expression (Mortazavi et al., 2008). Statistical evaluation of differentially expressed genes was carried out with the Empirical Analysis of DGE Tool, which implements edgeR analysis (Robinson et al., 2009) and Benjamini-Hochberg correction of P values (Benjamini and Hochberg, 1995). Only genes annotated as protein coding and with an RPKM > 0.1 in at least one of the developmental zones were selected for the first step of the analysis: comparison of base and blade samples. This threshold was determined based on the coefficient of variation between RPKM values of the biological replicates that strongly increased at RPKM ≤ 0.1. Genes which were regulated between the base and blade with a fold change ≥ |10| and an FDR ≤ 0.05 were further analyzed for their regulation in the re-etiolated blade compared to the blade from illuminated plants.

Enrichment analysis of MapMan processes

MapMan bins were used to assign genes to processes (Thimm et al., 2004). The χ^2-test was used to test for significantly enriched processes. Only processes were taken into account if ≥ 5 counts were expected within the analyzed group. The calculated p-values were adjusted with the Benjamini Hochberg correction to account for multiple testing (Benjamini and Hochberg, 1995).

Meta-analysis of H3K9ac ChIP-Seq data within defined classes of regulation

ChIP-Seq data from Perduns et al. (2015) was used to analyze the defined classes of regulation for the enrichment of R-SUP associated genes and the co-regulation with H3K9ac signals.

HAT and HDAC sequences

Sequences of maize HATs and HDACs were downloaded from ChromDB (Gendler et al., 2007). The sequences were used to identify the current IDs of corresponding genomic loci with the BLAST tool from Phytozome (http://phytozome.jgi.doe.gov/pz/portal.html#).

Supplemental data

Supplemental table S1: Read mapping statistics of RNA-Seq libraries

Supplemental table S2: Overview of the regulation of C4 genes

Literature cited

Alinsug MV, Chen FF, Luo M, Tai R, Jiang L, Wu K, Wu Q (2012) Subcellular Localization of Class II HDAs in Arabidopsis thaliana: Nucleocytoplasmic Shuttling of HDA15 Is Driven by Light. PLoS ONE **7:** e30846.

Aubry S, Kelly S, Kümpers BMC, Smith-Unna RD, Hibberd JM (2014) Deep Evolutionary Comparison of Gene Expression Identifies Parallel Recruitment of Trans-Factors in Two Independent Origins of C4 Photosynthesis. PLoS Genetics **10:** e1004365.

Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: a Practical and Powerful Approach to Multiple Testing. Journal of the Royal Statistical Society B **57:** 289–300.

Bhat RA, Borst JW, Riehl M, Thompson RD (2004) Interaction of maize Opaque-2 and the transcriptional co-activators GCN5 and ADA2, in the modulation of transcriptional activity. Plant Molecular Biology **55:** 239–252.

Bolger A, Lohse M, Usadel B (2014) Trimmomatic: A flexible trimmer for Illumina Sequence Data. Bioinformatics **30:** 2114–2120.

Chang Y, Liu W, Shih AC, Shen M, Lu C, Lu MJ, Yang H, Wang T, Chen SC, Chen SM, Li W, Ku MSB (2012) Characterizing Regulatory and Functional Differentiation between Maize Mesophyll and Bundle Sheath Cells by Transcriptomic Analysis. Plant Physiology **160:** 165–177.

Charron JF, He H, Elling AA, Deng XW (2009) Dynamic Landscapes of Four Histone Modifications during Deetiolation in Arabidopsis. The Plant Cell **21:** 3732–3748.

Chen J, Tian L (2007) Roles of dynamic and reversible histone acetylation in plant development and polyploidy. Biochimica et Biophysica Acta **1769:** 295–307.

Dion MF, Altschuler SJ, Wu LF, Rando OJ (2005) Genomic characterization reveals a simple histone H4 acetylation code. Proceedings of the National Academy of Sciences USA **102:** 5501–5506.

Dixon R, Paiva N (1995) Stress-Induced Phenylpropanoid Metabolism. The Plant Cell **1995:** 1085–1097.

Gendler K, Paulsen T, Napoli C (2007) ChromDB: The Chromatin Database. Nucleic Acids Research **36:** D298-D302

Gonzalez D, Bowen AJ, Carroll TS, Conlan RS (2007) The Transcription Corepressor LEUNIG Interacts with the Histone Deacetylase HDA19 and Mediator Components MED14 (SWP) and CDK8 (HEN3) To Repress Transcription. Molecular and Cellular Biology **27:** 5306–5315.

He G, Chen B, Wang X, Li X, Li J, He H, Yang M, Lu L, Qi Y, Wang X, Deng XW (2013) Conservation and divergence of transcriptomic and epigenomic variation in maize hybrids. Genome Biology **14:** R57.

Heimann L, Horst I, Perduns R, Dreesen B, Offermann S, Peterhansel C (2013) A Common Histone Modification Code on C4 Genes in Maize and Its Conservation in Sorghum and Setaria italica. Plant Physiology **162:** 456–469.

Horst I, Heimann L, Peterhansel C (2013) Signal integration on plant promoters. a case study in maize. Plant Signaling Behavior **8:** e25389.

Horst I, Offermann S, Dreesen B, Niessen M, Peterhansel C (2009) Core promoter acetylation is not required for high transcription from the phosphoenolpyruvate carboxylase promoter in maize. Epigenetics Chromatin **2:** 17.

Jang I, Chung PJ, Hemmes H, Jung C, Chua N (2011) Rapid and Reversible Light-Mediated Chromatin Modifications of Arabidopsis Phytochrome A Locus. The Plant Cell **23:** 459–470.

Kuang J, Chen J, Luo M, Wu K, Sun W, Jiang Y, Lu W (2012) Histone deacetylase HD2 interacts with ERF1 and is involved in longan fruit senescence. Journal of Experimental Botany **63:** 441–454.

Langdale JA (1998) Cellular differentiation in the leaf. Current Opinion in Cell Biology **10:** 734–738.

Li J, Li G, Wang H, Deng WX (2011) Phytochrome Signaling Mechanisms. The Arabidopsis Book **9:** e0148.

Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, Kebrom TH, Provart N, Patel R, Myers CR, Reidel EJ, Turgeon R, Liu P, Sun Q, Nelson T, Brutnell TP (2010) The developmental dynamics of the maize leaf transcriptome. Nature Genetics **42:** 1060–1067.

Liu X, Luo M, Zhang W, Zhao J, Wu K, Tian, L Duan J (2012) Histone acetyltransferases in rice (Oryza sativa L.): phylogenetic analysis, subcellular localization and expression. BioMed Central Plant Biology **12:** 145.

Lopez-Ochoa L, Acevedo-Hernandez G, Martinez-Hernandez A, Arguello-Astorga G, Herrera-Estrella L (2007) Structural relationships between diverse cis-acting elements are critical for the functional properties of a rbcS minimal light regulatory unit. Journal of Experimental Botany **58:** 4397–4406.

Ma L, Li J, Qu L, Hager J, Chen Z, Zhao H, Deng XW (2001) Light Control of Arabidopsis Development Entails Coordinated Regulation of Genome Expression and Cellular Pathways. The Plant Cell **13:** 2589–2607.

Majeran W, Cai Y, Sun Q, van Wijk KJ (2005) Functional Differentiation of Bundle Sheath and Mesophyll Maize Chloroplasts Determined by Comparative Proteomics. The Plant Cell **17**: 3111–3140.

Majeran W, Friso G, Ponnala L, Connolly B, Huang M, Reidel E, Zhang C, Asakura Y, Bhuiyan NH, Sun Q, Turgeon R, van Wijk KJ (2010) Structural and Metabolic Transitions of C4 Leaf Development and Differentiation Defined by Microscopy and Quantitative Proteomics in Maize. The Plant Cell **22:** 3509–3542.

Martinez-Hernandez A, Lopez-Ochoa L, Arguello-Astorga G, Herrera-Estrella L (2002) Functional Properties and Regulatory Complexity of a Minimal RBCS Light-Responsive Unit Activated by Phytochrome, Cryptochrome, and Plastid Signals. Plant Physiology **128:** 1223–1233.

Menkens AE, Schindler U, Cashmore AR (1995) The G-box: a ubiquitous regulatory DNA element in plants bound by the GBF family of bZIP proteins. Trends in Biochemical Sciences **20:** 506–510.

Mortazavi A, Williams BA, McCue K, Schaeffer L, Wold B (2008) Mapping and quantifying mammalian transcriptomes by RNA-Seq. Nature Methods **5:** 621–628.

Nelissen H, Rymen B, Coppens F, Dhondt S, Fiorani F, Beemster GTS (2013) Kinematic Analysis of Cell Division in Leaves of Mono- and Dicotyledonous Species: A Basis for Understanding Growth and Developing Refined Molecular Sampling Strategies. Methods Molecular Biology **959:** 247–264.

Nelissen H, Rymen B, Jikumaru Y, Demuynck K, van Lijsebettens M, Kamiya Y, Inzé D, Beemster GTS (2012) A Local Maximum in Gibberellin Levels Regulates Maize Leaf Growth by Spatial Control of Cell Division. Current Biology **22:** 1183–1187.

Nelson T (2011) The grass leaf developmental gradient as a platform for a systems understanding of the anatomical specialization of C4 leaves. Journal of Experimental Botany **62:** 3039–3048.

Offermann S, Danker T, Dreymüller D, Kalamajka R, Töpsch S, Weyand K, Peterhänsel C (2006) Illumination is necessary and sufficient to induce histone acetylation independent of transcriptional activity at the C4-specific phosphoenolpyruvate carboxylase promoter in maize. Plant Physiology **141:** 1078–1088.

Offermann S, Dreesen B, Horst I, Danker T, Jaskiewicz M, Peterhansel C (2008) Developmental and environmental signals induce distinct histone acetylation profiles on distal and proximal promoter elements of the C4-Pepc gene in maize. Genetics **179:** 1891–1901.

Oh S, Montgomery BL (2011) Identification of Proteins Associated with Spatial-Specific Phytochrome-Mediated Light Signaling in Arabidopsis thaliana by Liquid Chromatography-Tandem Mass Spectrometry. Gravitational and Space Biology **25:** 22–32.

Pandey R, Mueller A, Napoli CA, Selinger DA, Pikaard CS, Richards EJ, Bender J, Mount DW, Jorgensen RA (2002) Analysis of histone acetyltransferase and histone deacetylase families of Arabidopsis thaliana suggests functional diversification of chromatin modification among multicellular eukaryotes. Nucleic Acids Research **30:** 5036–5055.

Perduns R, Horst-Niessen I, Peterhansel C (2015) Photosynthetic Genes and Genes Associated with the C4 Trait in Maize Are Characterized by a Unique Class of Highly Regulated Histone Acetylation Peaks on Upstream Promoters. Plant Physiology **168:** 1378–1388.

Pick TR, Brautigam A, Schluter U, Denton AK, Colmsee C, Scholz U, Fahnenstich H, Pieruschka R, Rascher U, Sonnewald U, Weber APM (2011) Systems Analysis of a Maize Leaf Developmental Gradient Redefines the Current C4 Model and Provides Candidates for Regulation. The Plant Cell **23:** 4208–4220.

Robinson MD, McCarthy DJ, Smyth GK (2009) edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics **26:** 139–140.

Ruiz-Sola MÁ, Rodríguez-Concepción M (2012) Carotenoid Biosynthesis in Arabidopsis: A Colorful Pathway. The Arabidopsis Book **10:** e0158.

Rymen B, Fiorani F, Kartal F, Vandepoele K, Inze D, Beemster GTS (2007) Cold Nights Impair Leaf Growth and Cell Cycle Progression in Maize through Transcriptional Changes of Cell Cycle Genes. Plant Physiology **143:** 1429–1438.

Sakamoto W, Miyagishima S, Jarvis P (2008) Chloroplast Biogenesis: Control of Plastid Development, Protein Import, Division and Inheritance. The Arabidopsis Book **6:** e0110.

Sheen J (1999) C4 GENE EXPRESSION. Annual Review of Plant Biology **50**: 187–217.

Singh B, Sharma RA (2015) Plant terpenes: defense responses, phylogenetic analysis, regulation and clinical applications. 3 Biotech **5:** 129–151.

Sylvester AW, Cande ZW, Freeling M (1990) Division and differentiation during normal and liguleless-1 maize leaf development. Development **110:** 985–1000.

Tausta SL, Li P, Si Y, Gandotra N, Liu P, Sun Q, Brutnell TP, Nelson T (2014) Developmental dynamics of Kranz cell transcriptional specificity in maize leaf reveals early onset of C4-related processes. Journal of Experimental Botany **65:** 3543–3555.

Terzaghi WB, Cashmore AR (1995) Light-Regulated Transcription. Annual Review of Plant Physiology and Plant Molecular Biology **46:** 445-474.

Thimm O, Bläsing O, Gibon Y, Nagel A, Meyer S, Krüger P, Selbig J, Müller LA, Rhee SY, Stitt M (2004) Mapman: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. The Plant Journal **37:** 914–939.

Wang L, Czedik-Eysenberg A, Mertz RA, Si Y, Tohge T, Nunes-Nesi A, Arrivault S, Dedow LK, Bryant DW, Zhou W, Xu J, Weissmann S, Studer A, Li P, Zhang C, LaRue T, Shao Y, Ding Z, Sun Q, Patel RV, Turgeon R, Zhu X, Provart NJ, Mockler TC, Fernie AR, Stitt M, Liu P, Brutnell TP (2014) Comparative analyses of C4 and C3 photosynthesis in developing leaves of maize and rice. Nature Biotechnology **32:** 1158–1165.

Wang X, Elling AA, Li X, Li N, Peng Z, He G, Sun H, Qi Y, Liu XS, Deng XW (2009) Genome-Wide and Organ-Specific Landscapes of Epigenetic Modifications and Their Relationships to mRNA and Small RNA Transcriptomes in Maize. The Plant Cell **21:** 1053–1069.

Supplement

Supplemental table S1: Read mapping statistics of RNA-Seq libraries

	Processed reads	Mapped reads (unique)	% (unique mapped)
Base A	59,346,263	46,246,282	78%
Base B	33,432,364	24,687,777	74%
Blade A	74,724,999	53,086,298	71%
Blade B	27,582,965	18,668,247	68%
Re-et. Blade A	36,531,501	26,038,059	71%
Re-et. Blade A	37,431,260	26,734,597	71%

Supplemental table S1: Read mapping statistics of RNA-Seq libraries from all incorporated samples. Re-et., re-etiolated.

Supplemental table S2: Overview of the regulation of C4 genes

Gen	Locus	Blade vs Base	Blade vs re-et. Blade	Class of regulation
Pepc	GRMZM2G083841	1221.6	180.5	Ind_Ld-genes
Ppdk	GRMZM2G306345	765.0	6.7	-
Me	GRMZM2G085019	1321.9	8.2	-
Ca	GRMZM2G121878	1852.9	46.8	Ind_Ld-genes
Pepck	GRMZM2G001696	3468.1	4470.1	Ind_Ld-genes
Mdh	GRMZM2G129513	36.2	17.7	Ind_Ld-genes
RbsS2	GRMZM2G113033	81.3	979.2	Ind_Ld-genes
RbsS4	GRMZM2G098520	39.2	421.4	Ind_Ld-genes
Mep3	GRMZM2G138258	807.8	4684.5	Ind_Ld-genes
Asp-at1	GRMZM5G836910	13.9	32.1	Ind_Ld-genes
Tpt	GRMZM2G070605	83.6	226.3	Ind_Ld-genes
Ppt	GRMZM2G174107	16.0	82.1	Ind_Ld-genes
Dct2	GRMZM2G086258	34.9	141.8	Ind_Ld-genes
Sut1	GRMZM2G034302	84.9	11.7	Ind_Ld-genes
Omt	GRMZM2G383088	9.9	38.6	-
Dit1	GRMZM2G040933	8.4	21.8	-

Supplemental table S2: Overview of the regulation and defined classes of regulation of C4 genes. C4 genes were defined by Chang et al. (2012). Numbers represent significant fold changes of the indicated comparisons (FDR ≤ 0.05). Re-et., re-etiolated; *Pepc*, phosphoenolpyruvate carboxylase; *Ppdk*, pyruvate, P_i dikinase; *Me*, malic enzyme; *Ca*, carbonic anhydrase; *Pepck*, phosphoenolpyruvate carboxykinase; *Mdh*, malate dehydrogenase, *RbcS2/4*, ribulose-1,5-bisphosphate carboxylase-oxygenase2/4; *Mep3*, proton/pyruvate symporter3; *Asp-at1*, aspartate aminotransferase1; *Tpt*, triose-phosphate/phosphate antiporter; *Ppt*, phosphoenolpyruvate/phosphate translocator; *Dct2*, dicarboxylate transporter2; *Sut1*, sucrose transporter1; *Omt*, 2-oxoglutarate/malate transporter; *Dit1*, dicarboxylate transporter1

Chapter 5 - General discussion

5.1 Conservation of C4 specific histone modification patterns

The expression of many genes in eukaryotes is highly regulated by various signals of different origin. Different signals with potentially contrary impact on gene expression are simultaneously present and have to be integrated into one transcriptional output of the respective gene. Signals are integrated at the promoters of genes. The structures of promoters are very diverse and they are mainly located upstream of the transcription initiation site with varying lengths. Promoters are characterized by their DNA sequence that provides binding sites for various transcription factors and RNA-polymerases to regulate transcription (Gagniuc and Ionescu-Tirgoviste, 2012; Porto et al., 2014). The DNA of eukaryotes is tightly packed by the interaction with nucleosomes into chromatin that influences the accessibility of the DNA. Histone modifications are one mechanism of chromatin to regulate the expression of associated genes (Kouzarides, 2007; Li et al., 2007). Some histone modifications regulate gene expression by loosening the interaction of DNA and histone proteins to facilitate binding of transcription factors and RNA polymerases (charge neutralization model) or they recruit specific factors that interact with certain histone modifications (histone code model; Chapter 1.1.1). Comprehensive studies on the *Pepc* promoter in maize revealed that potentially both models act site by site to integrate different signals that regulate the expression of *Pepc* (Offermann, 2006; Offermann et al., 2008; Danker et al., 2008; Horst et al., 2009)

In the present study, observations from previous analyses on the promoter of *Pepc* were confirmed and extended: Analyzed histone modifications were found at two distinct sites at the *Pepc* promoter. The core promoter region was located around -200 bp and the upstream promoter region around -1300 bp upstream of the TIS (Chapter 2). In a recent study, Horst et al. (2013) summarized the current knowledge of signal integration at the *Pepc* promoter by histone modifications and the proposed models for both detected regions (Figure 5-1). Shortly, analyzed histone modifications in the upstream promoter region were regulated by all indicated signals in a dosage-dependent manner and they were highly co-regulated with the transcription of the *Pepc* gene. These modifications might therefore potentiate the transcription of *Pepc* by improving the access of DNA binding proteins to the DNA. Such regulation would argue for the charge neutralization model for this region of the *Pepc* promoter. This hypothesis is supported by the absence of the investigated histone methylation H3K4me3 in that region that would not influence the charge of histones.

H3K4me3 was only found at the core promoter region of the *Pepc* promoter (Chapter 2; Danker et al., 2008). Histone modifications behaved different in the core promoter region. H3K9ac and H4K5ac were induced by light, independent of the transcription rate. For example, in the light, H3K9ac and H4K5ac were also found in the core promoter region when the transcription of *Pepc* was repressed by metabolic signals (low nitrogen and high sugar concentrations; Offermann et al., 2008). The signal of the tissue was represented by H3K4me3 which was found only in the M cells at the core promoter of *Pepc*, independent of the light signal (Chapter 2; Danker et al., 2008). Such pairs of specific histone modifications and distinct stimuli that are independent of the expression of the associated gene argue for the histone code model as suggested by Turner (2002).

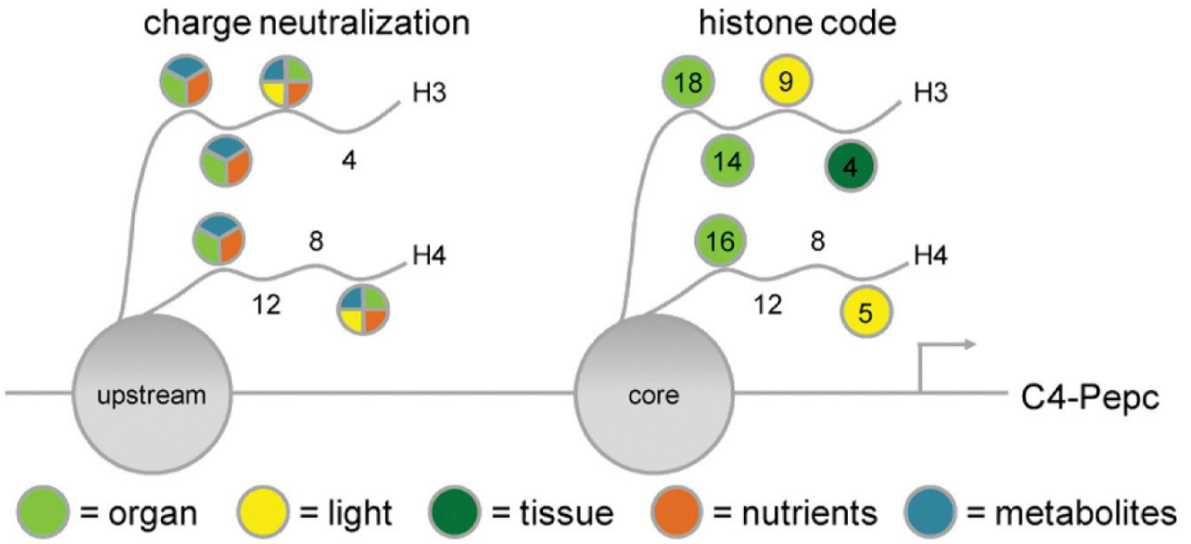


Figure 5-1: Histone modification model of C4-*Pepc* from *Zea mays*. Numbers indicate lysine residues at the N-terminal tails of histones H3 and H4. The color code symbolizes the signals listed at the bottom of the figure. The grey arrow at the right indicates the TIS. The two sites of enriched histone modifications are shown as distinct regions: core and upstream promoter regions, labeled with the proposed model of the regulation of gene expression by histone modifications (Horst et al., 2013).

Additionally, the influence of histone modifications on the expression of other C4 associated genes was investigated to analyze the conservation of the observations at the *Pepc* promoter (Chapter 2). The distribution and regulation of H3K9ac, H3K18ac, H4K5ac and H3K4me3 were analyzed at the promoters of carbonic anhydrase (*Ca*), malic enzyme (*Me*), pyruvate-P_i-dikinase (*Ppdk*), phosphoenolpyruvate carboxykinase (*Pepck*) and the small subunit of ribulose-1,5-bisphosphate carboxylase-oxygenase (*RbcS2*) in maize. H3K18ac was constitutively found at the promoters of investigated genes, independent of light. Offermann et al. (2008) analyzed H3K18ac in roots of maize and detected decreased intensities compared to the leaf in both regions of the *Pepc* promoter. It was supposed that H3K18ac represents the organ where *Pepc* is expressed. This hypothesis might also be valid for the

other investigated C4 genes. Similar to *Pepc*, H3K4me3 was restricted to the core promoter regions of the analyzed genes, only in the tissue where the corresponding genes were expressed. The proposed histone code model for H3K18ac and H3K4me3 in the core promoter region of *Pepc* might also be true for the analyzed C4 genes. H3K18ac and H3K4me3 might lead to a poised chromatin structure that allows rapid activation when the organ or tissue is illuminated.

Interestingly, the results in Chapter 2 showed that the promoters of the analyzed C4 genes also had increased H3K9ac and H4K5ac signals at the TIS and further upstream that correlated with the induction of associated genes by light. For *Ca* and *Me*, it was shown that H3K9ac in the upstream region was distinguishable from the core promoter region by a drop of signal between both regions; comparable to what had been reported for the *Pepc* promoter. At the promoters of *Ppdk*, *Pepck* and *RbcS2* strong H3K9ac peaks at the TIS were observed that slowly declined towards the upstream regions. In Chapter 3, the genome-wide distribution of H3K9ac was analyzed by ChIP-Seq. Manual inspection of the ChIP-Seq data of analyzed C4 genes revealed that also *Pepck* and *Ppdk* were associated with distinct H3K9ac peaks in the upstream region in the blade samples, while no distinct peaks could be detected by the qPCR analyses (Supplemental figure S2 of Chapter 3). Thus, the overlay of data from ChIP-qPCR experiments with ChIP-Seq data showed that the results from both methods were in agreement, but that the resolution of the ChIP-Seq data outperformed the resolution of the ChIP-qPCR data. This is reasonable, because the size and the density of used primer systems limits the resolution of ChIP-qPCR data (Haring et al., 2007). Thus, the data indicates that the distances between chosen primer systems in Chapter 2 was too large to detect the drop of H3K9ac between the peaks at the TIS and in the upstream region of *Ppdk* and *Pepck*. The genome-wide occurrence of upstream peaks is discussed in Chapter 5.2.1

Furthermore, it was shown that the promoters of orthologous *Me* and *Pepc* genes of the related C4 grasses *Sorghum bicolor* and *Setaria italica* also showed two distinct regions of enriched H3K9ac signals in the light (Chapter 2). Interestingly, the locations of both distinct regions were different compared to maize. The promoters of *Me* and *Pepc* in *Setaria* showed no H3K9ac signals in the expected core promoter region at the TIS but in the beginning of the coding region at +200 bp to +300 bp downstream of the TIS. The upstream promoter regions of enhanced H3K9ac signals shifted towards the TIS compared to maize. One reasonable explanation is the smaller genome size of *Setaria* (490 Mbp; Doust et al., 2009) compared to maize (2300 Mbp; Schnable et al., 2009) which might lead to condensed promoters. Accordingly, H3K9ac was tightly restricted to the TIS in *Arabidopsis* with a genome size of 125 Mbp (The Arabidopsis Genome Initiative, 2000; Zhou et al., 2010).

The regulation of the analyzed histone modifications in both promoter regions have to be analyzed in biological situations where the transcription of the according genes is modulated to understand the interplay of transcription and histone modifications in more detail. However, the described similarities of the analyzed histone modifications and the existence of modified upstream regions at promoters of prominent C4 genes in related species indicated a deep evolutionary conservation of the histone code used to integrate signals on the promoters of C4 genes.

5.2 Comparative analysis of H3K9ac and H3K4me3 on a genome-wide scale

In recent years, genome-wide analyses of histone modifications generated comprehensive knowledge about the global distribution of histone modifications in static situations (Chapter 1.1.2). In Chapter 3, it was aimed to extend the knowledge about signal integration on plant promoters and to focus on the regulation of histone modifications and the correlation with the transcription of associated genes between two biological situations. Additionally, the occurrence of upstream promoter regions with increased H3K9ac signals at C4 genes, like described in Chapter 5.1, should be investigated on a genome-wide scale to analyze their specificities. The developmental gradient of the maize leaf offers a useful tool to study regulatory features of C4 photosynthesis (Chapter 1.3.1). Therefore, the dynamics of transcription and the histone modifications H3K9ac and H3K4me3 were analyzed between two developmental zones of the maize leaf: the base and the blade. The base of the leaf is characterized by undifferentiated cells that develop to mature photosynthetic cells at the blade (Nelson, 2011). This system has been studied comprehensively to understand the establishment of C4 photosynthesis in maize. High throughput technologies have been used to analyze transcriptomes, proteomes and metabolomes at different stages of development (Li et al., 2010; Majeran et al., 2010; Pick et al., 2011; Ponnala et al., 2014). These studies offered the opportunity to compare the generated data set with published data on the developmental gradient and integrate the findings. Accordingly, obtained RNA-Seq data were correlated with the RNA-Seq data from Li et al. (2010) and it was shown that the analyzed zones (base and blade) correlated well with the segments -4 and +4 from Li et al. (2010). Therefore, RNA-Seq data from Li et al. (2010) could be used to ensure the identification of significantly regulated genes between the base and blade to enhance the reliability of the analyses. The ChIP-Seq data were generated from identical zones and were initially screened for previously described characteristics as quality control. In Chapter 3, it was shown that H3K9ac and H3K4me3 were associated with actively transcribed genes and mainly located at the TIS regions, independent of the developmental zone. The overlap between both modifications was substantial: 80% of transcribed genes were associated with

H3K9ac, 79% with H3K4me3 and 75% with both modifications. Both histone modifications have been described to be enriched at the TIS regions of actively transcribed genes in *Arabidopsis*, rice and maize in different biological situations and tissues (Wang et al., 2009; Zhou et al., 2010; Roudier et al., 2011; Du et al., 2013; He et al., 2013).

Genome-wide data of histone modifications in static situations were frequently linked to genome-wide transcriptomic data from identical biological situations to correlate intensities of histone modifications with transcription levels of associated genes. For such analyses, the genes were grouped according to their transcription level and the average signal intensities of the associated histone modifications were calculated for the according groups. It was shown that groups of highly expressed genes were associated with high levels of H3K9ac and H3K4me3 while weakly expressed genes were associated with low levels of H3K9ac and H3K4me3 (Wang et al., 2009; Roudier et al., 2011). In Chapter 3, the correlation of signal intensities of individual genes in a static view was compared with the correlation of the fold changes of histone modifications and the transcription of individual genes between the base and blade. The Pearson correlation coefficients for the correlation of individual signal intensities in a static view were calculated as 0.2 for H3K9ac and 0.3 for H3K4me3 while they were higher for the correlation of fold changes of individual genes: 0.7 for both H3K9ac and H3K4me3. This might indicate a disadvantage of such correlations between individual genes in a static view. The genomic environment might influence the impact of histone modifications on the expression of individual genes. Additionally, different condensation levels of the chromatin might technically influence the accessibility of the histone modifications for the specific antibodies within the cross linked samples. For example, we showed in Chapter 2 that expression of analyzed C4 genes varied only by a factor of 2 within the illuminated blade while the intensities of H3K9ac at the TIS varied by a factor of 4. At the upstream promoter positions the variation between different C4 genes was even more than 10 fold. The analyses of individual genes in dynamic situations could potentially lower the effects of such biases because the genomic context of the correlated parameters remains constant. The dynamics of histone modifications have rarely been studied in the plant field. In maize, only one study analyzed H3K9ac and H3K4me3 from roots and shoots in a comparative manner (He et al., 2013). Similar to the results in Chapter 3, it was also shown that the fold changes of H3K9ac and H3K4me3 between shoots and roots correlated well with the fold changes of the transcription of associated genes. However, no comparison to the correlation of signal intensities on the bases of individual genes was drawn like shown in Chapter 3. The data indicated that the intensities of H3K9ac and H3K4me3 were not suitable parameters for the estimation of the transcription of associated genes in static situations while the fold changes between two biological situations could be used to estimate the fold changes in transcription of associated genes.

In Chapter 5.1 it was discussed that H3K9ac represents the light signal at the core promoters of prominent C4 associated genes, like it was shown for the core promoter region of *Pepc* (Offermann et al., 2008). The results in Chapter 3 and 4 indicate that this seems not to be valid for all H3K9ac-modified genes on a genome-wide scale. In Chapter 4, RNA-Seq data of the re-etiolated blade were used to investigate the influence of light on gene regulation from the base to the blade. Groups of genes were defined that were induced or repressed from the base to the blade in a light-dependent or -independent manner. Interestingly, around 60% of all classified genes that were induced in the blade compared to the base were light regulated while only 9% of the repressed genes where regulated by light. Thus, light might be an important activating factor in the leaf blade, but only a weak repressive factor. The degree of co-regulation of transcription and H3K9ac was calculated within the groups and around 90% of H3K9ac modified genes that were significantly regulated from the base to the blade were associated with co-regulated H3K9ac, independent of the influence of light. For instance, also genes that were repressed in the blade compared to the base in a light-independent manner were associated with co-regulated H3K9ac. Additionally, He et al. (2013) showed that genes that were induced in the root compared to the shoot were highly co-regulated with H3K9ac. Therefore, the representation of the light signal by H3K9ac at the core promoter region of *Pepc* seemed to be a specific case that might be valid for other light regulated genes. Further experiments, such as discussed in Chapter 5.1, are needed to investigate the representation of the light signal by H3K9ac at other light-dependently regulated genes than *Pepc*. H3K9ac might also represent different signals at different genes. The analysis of H3K9ac in biological situations where the expression of light-independently regulated genes is modulated could help to understand the roles of H3K9ac.

5.2.1 Identification and characterization of R-SUPs

In Chapter 2, it was shown that two distinct H3K9ac peaks were found at the promoters of prominent C4 genes. One aim of the genome-wide analysis of histone modifications in Chapter 3 was to investigate whether specific groups of genes are characterized by such strong H3K9ac in upstream promoter regions. On a genome-wide level, no obvious enrichment of H3K9ac or H3K4me3 peaks in upstream promoter regions could be detected for all identified H3K9ac and H3K4me3 peaks that were associated with transcribed genes. However, when focusing on peaks that were significantly regulated from the base to the blade an increase of H3K9ac peaks in the upstream region was detected while the distribution of regulated H3K4me3 peaks was similar to all peaks. To exclude that these regulated upstream peaks were caused by false annotated TIS it was analyzed whether associated genes also showed an H3K9ac peak at the annotated TIS. Most of such genes

were additionally associated with a TIS peak that was distinguishable from the regulated upstream peak and, therefore, the upstream peak was a secondary peak. Those upstream peaks were accordingly termed “regulated secondary upstream peaks” (R-SUPs). The distribution of H3K9ac has been described in various studies in static situations but R-SUPs have not been identified to our knowledge. In Chapter 3, the analysis of regulated H3K9ac peaks enabled the identification of R-SUPs. In maize, only He et al. (2013) also analyzed the regulation of H3K9ac between roots and shoots of maize but no observation of R-SUPs was reported. In that study, precipitated DNA from three biological replicates were pooled and sequenced once, so that the biological variation could not be considered for the statistical evaluation. Therefore, He et al. (2013) included only H3K9ac peaks if the peak was found in both samples (shoots and roots) to ensure that the peaks were valid. Thus, genomic regions without a detected peak in one of the tissues were excluded from the analyses. A different approach was used for the statistical evaluation of the ChIP-Seq data in Chapter 3 that included H3K9ac peaks if the peak was found in at least two replicates that were analyzed independently. It was shown that R-SUPs had lower intensities compared to TIS peaks but their regulation was stronger. Also a manual inspection of C4 genes showed that R-SUPs were not detected at the base of the leaf whereas the TIS peaks were already observed with low intensities (Chapter 3). Potentially, He et al. (2013) missed R-SUPs because they were excluded from the analysis due to the absence in one of the investigated tissues.

To characterize the occurrence of R-SUPs and the associated genes their regulation from the base to the blade was analyzed (Chapter 3). It was shown that R-SUPs and transcription of associated genes were preferentially induced in the blade compared to TIS peaks and genes associated with a TIS peak only. To analyze if genes of certain biological processes were preferentially associated with R-SUPs the group of R-SUP associated genes that showed induction of both the R-SUP and transcription in the blade compared to the base was screened for enriched processes. Enrichment was only found for genes associated with photosynthesis. Among these genes, 5 genes were found that were known to be associated with C4 photosynthesis (*Pepc*, *Ppdk*, *Me*, *Ca* and *Pepck*; Hibberd and Covshoff, 2010). Another link to C4 photosynthesis was added by the analysis of DNA sequences that were associated with R-SUPs. As described in Chapter 1.1.1, it is supposed that acetylation of histones might diminish the interaction between DNA and histone proteins to facilitate the access of DNA-binding proteins, like transcription factors (Dion et al., 2005). Therefore, two algorithms were used to analyze whether distinct DNA motifs were enriched within the sequences associated with R-SUPs. Both algorithms identified similar sequences that were related to the motif of ethylene responsive factors (ERFs; Chapter 3). Although ERF transcription factors were frequently associated with biotic and abiotic stresses (Gutterson and Reuber, 2004; Mishra et al., 2015), Pick et al. (2011) identified 18 transcription factors

that were co-expressed with prominent C4 genes of which 9 belonged to the ERF family. The described characteristics indicated a link between R-SUPs and C4 photosynthesis. Therefore, it was analyzed whether R-SUPs were enriched within potential C4 genes that were previously identified based on their expression profiles along the leaf gradient of maize and between mesophyll and bundle sheath cells (Li et al., 2010; Aubry et al., 2014). It was shown that R-SUP associated genes were enriched within these lists and interesting candidates for C4 related genes were found based on their annotation and previous studies (Chapter 3). For example, an additional *Ca* gene was found that was mainly expressed in mesophyll cells (Aubry et al., 2014) and that might contribute to the initial hydration of CO_2. Another example was a DNA binding with one finger (DOF) transcription factor that was expressed in bundle sheath cells (Li et al., 2010; Aubry et al., 2014). DOF transcription factors have been shown to be involved in the regulation of the expression of *Pepc* in maize (Yanagisawa and Sheen, 1998) and light responsive genes in general (Kushwaha et al., 2011).

R-SUPs were identified as unique feature of H3K9ac and it was shown that R-SUPs were enriched at the promoters of known and potential C4 associated genes. R-SUPs might be a consequence of the adaptation of the expression of ancestral C3 genes required for the evolution of C4 photosynthesis (Hibberd and Covshoff, 2010). Changes in gene expression profiles might be acquired by addition or mutation of upstream promoter elements. For example, it was shown that two upstream promoter regions of the *Pepc* gene of *Flaveria trinervia* (C4) were mutated compared to the closely related C3 species *Flaveria pringlei.* The upstream promoter regions conferred mesophyll specific expression, a hallmark for *Pepc* in C4 species, when fused to the promoter of *Pepc* from *Flaveria pringlei* (Gowik et al., 2004). Such mutated upstream elements might be associated with altered histone modifications in those regions. Accordingly, the occurrence of R-SUPs could be used to narrow lists of potential C4 genes down, to focus on genes that fulfill the expectations of potential C4 genes on the transcriptional level and are additionally associated with an R-SUP. However, the described association of R-SUPs with C4 photosynthesis was mainly based on enrichment analyses. Additional wet lab experiments are needed to understand the role of R-SUPs in the regulation of C4 genes. The analyses of H3K9ac profiles of orthologous genes in C3 plants could help to investigate the specificity of R-SUPs for C4 genes. Promoter-deletion mutants with or without R-SUP regions could be studied in different biological situations to extend the understanding of R-SUPs. ChIP-Seq experiments to analyze binding sites of potential ERF1 transcription factors that were identified by Pick et al. (2011) could also help to analyze whether R-SUPs provide binding sites for distinct transcription factors.

5.2.2 Perspectives for the utilization of next generation sequencing data

Next generation sequencing (NGS) based approaches are nowadays the method of choice for the genome-wide analysis of RNA und DNA associated processes. RNA-Seq was first used by Weber et al. (2007) to re-sequence the transcriptome of *Arabidopsis thaliana*. Since then, NGS methods have been massively improved and current protocols enable the generation of millions of sequencing reads at affordable costs. RNA-Seq has been used frequently to analyze gene expression in non-model organisms, because the RNA-Seq data itself can be used for the *de novo* assembly of transcriptomes and, thus, no knowledge of the transcriptome is required (Martin et al., 2013). This is exemplified by the 1KP project that aims to sequence the transcriptomes of 1000 plant species (Matasci et al., 2014). The RNA-Seq data of published experiments are provided at public repositories and can be used by researchers to re-analyze them in further contexts. For example, Yu et al. (2015) used 22 RNA-Seq datasets from various developmental stages of maize to predict transcription factors (TFs) and corresponding transcription factor bindings sites (TFBs) by computational approaches. Promoter sequences of highly co-expressed genes from the same gene ontology (GO) were screened for conserved motifs within maize and orthologous genes of five additional species. Potential TFs that might bind to these motifs were identified by complex approaches to screen datasets of known TF-TFBs pairs of various species with the predicted TFBs of maize. This comprehensive analysis extended the database of known TF-TFBs pairs by far. Previously, only 30 TF-TFBs pairs were known in maize. Yu et al. (2015) predicted 253 TF-TFBs pairs. 12 newly identified TF-TFBs pairs were validated experimentally. RNA-Seq data generated in Chapters 3 and 4 could contribute to similar approaches that rely on co-expression analyses including various situations.

ChIP-Seq data have not been used for computational analyses to improve gene annotations or predict genes, although the enrichment of H3K9ac or H3K4me3 was frequently shown to be associated with the TIS of transcribed genes. One example from our lab for the improvement of individual gene annotations by ChIP-Seq data is shown in Figure 5-2. We aimed to annotate homologous histone methyltransferases (HMTs) in maize based on the *Arabidopsis* sequences. Two distinct regions of one HMT from *Arabidopsis* were found to be homologous to two genes in maize that were annotated consecutively (*Sdg115.1* and *.2*). The RNA-Seq data showed that both genes were equally expressed indicated by comparable levels of RNA-Seq reads in the exon regions of both genes. Additionally, the ChIP-Seq data showed that only *Sdg115.1* was associated with H3K9ac and H3K4me3 in the TIS region, suggesting that *Sdg115.1* and .2 belong to the same gene. A PCR from cDNA with primers spanning the gap between both genes validated this assumption (Ina Horst-Nießen; unpublished data).

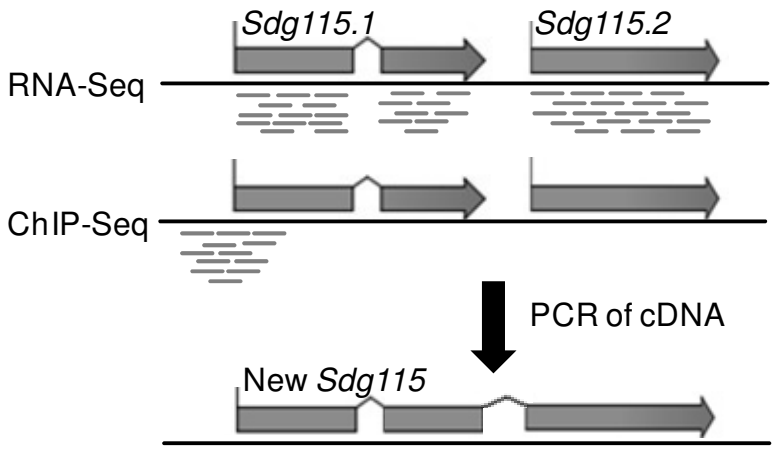


Figure 5-2: Simplified scheme for the usage of next generation sequencing data to improve the annotation of *Sdg115*. The arrows in the upper part symbolize the annotation of *Sdg115* as two distinct genes. Grey lines under the annotation illustrate mapped reads from the experiments indicated at the left. Both analyzed modifications (H3K9ac and H3K4me3) were enriched only at the start of *Sdg115.1*. The lower part shows the altered annotation that was verified by PCR from cDNA.

Another observation suggests the genome-wide use of ChIP-Seq and RNA-Seq data to validate the expression of genes that were annotated by *in silico* predictions only. Genes that were annotated as *low confidence* genes were excluded from the analyses in Chapters 3 and 4 to enhance the reliability. Such annotations were based on computational predictions (http://www.maizegdb.org/gene_center/gene). It was described that H3K9ac and H3K4me3 were associated with actively transcribed genes and mainly found in TIS regions. In the unfiltered dataset, 786 *low confidence* genes were found that were linked to both H3K9ac and H3K4me3 in the TIS region and that were transcribed in at least one of the analyzed developmental zones. Thus, the expression of these genes was validated and they were additionally associated with two histone modifications that were frequently described as activating modifications. Accordingly, the combination of ChIP-Seq and RNA-Seq data could also be used for *de novo* annotation of genes on a genome-wide scale.

References

Aubry S, Kelly S, Kümpers BMC, Smith-Unna RD, Hibberd JM (2014) Deep Evolutionary Comparison of Gene Expression Identifies Parallel Recruitment of Trans-Factors in Two Independent Origins of C4 Photosynthesis. PLoS Genetics **10:** e1004365.

Benhamed M, Bertrand C, Servet C, Zhou D (2006) Arabidopsis GCN5, HD1, and TAF1/HAF2 Interact to Regulate Histone Acetylation Required for Light-Responsive Gene Expression. The Plant Cell **18:** 2893–2903.

Brautigam A, Kajala K, Wullenweber J, Sommer M, Gagneul D, Weber KL, Carr KM, Gowik U, Mass J, Lercher MJ, Westhoff P, Hibberd JM, Weber APM (2011) An mRNA Blueprint for C4 Photosynthesis Derived from Comparative Transcriptomics of Closely Related C3 and C4 Species. Plant Physiology **155:** 142–156.

Brusslan JA, Rus Alvarez-Canterbury AM, Nair NU, Rice JC, Hitchler MJ, Pellegrini M, Dalal Y (2012) Genome-Wide Evaluation of Histone Methylation Changes Associated with Leaf Senescence in Arabidopsis. PLoS ONE **7:** e33151.

Candaele J, Demuynck K, Mosoti D, Beemster GTS, Inze D, Nelissen H (2014) Differential Methylation during Maize Leaf Growth Targets Developmentally Regulated Genes. Plant Physiology **164:** 1350–1364.

Chang Y, Liu W, Shih AC, Shen M, Lu C, Lu MJ, Yang H, Wang T, Chen SC, Chen SM, Li W, Ku MSB (2012) Characterizing Regulatory and Functional Differentiation between Maize Mesophyll and Bundle Sheath Cells by Transcriptomic Analysis. Plant Physiology **160:** 165–177.

Chen JF, Tian L (2007) Roles of dynamic and reversible histone acetylation in plant development and polyploidy. Biochim Biophys Acta **1769:** 295–307.

Christin P, Salamin N, Kellogg EA, Vicentini A, Besnard G (2009) Integrating Phylogeny into Studies of C4 Variation in the Grasses. Plant Physiology **149:** 82–87.

Danker T, Dreesen B, Offermann S, Horst I, Peterhänsel C (2008) Developmental information but not promoter activity controls the methylation state of histone H3 lysine 4 on two photosynthetic genes in maize. The Plant Journal **53:** 465–474.

Dion MF, Altschuler SJ, Wu LF, Rando OJ (2005) Genomic characterization reveals a simple histone H4 acetylation code. Proceedings of the National Academy of Sciences USA **102:** 5501–5506.

Doust AN, Kellogg EA, Devos KM, Bennetzen JL (2009) Foxtail Millet: A Sequence-Driven Grass Model System. Plant Physiology **149:** 137–141.

Du Z, Li H, Wei Q, Zhao X, Wang C, Zhu Q, Yi X, Xu W, Liu XS, Jin W, Su Z (2013) Genome-Wide Analysis of Histone Modifications: H3K4me2, H3K4me3, H3K9ac, and H3K27ac in Oryza sativa L. Japonica. Molecular Plant **6:** 1463–1472.

Efroni I, Eshed Y, Lifschitz E (2010) Morphogenesis of Simple and Compound Leaves: A Critical Review. The Plant Cell **22:** 1019–1032.

Ehleringer JR, Cerling TE, Helliker BR (1997) C4 photosynthesis, atmospheric CO_2, and climate. Oecologia **112:** 285–299.

Evert RF, Russin WA, Bosabalidis AM (1996) Anatomical and ultrastructural changes associated with sink-to-source transition in developing maize leaves. International Journal of Plant Sciences **157:** 247–261.

Gagniuc P, Ionescu-Tirgoviste C (2012) Eukaryotic genomes may exhibit up to 10 generic classes of gene promoters. BioMed Central Genomics **13:** 512.

Gowik U, Burscheidt J, Akyildiz M, Schlue U, Koczor M, Streubel M, Westhoff P (2004) cis -Regulatory Elements for Mesophyll-Specific Gene Expression in the C 4 Plant Flaveria trinervia , the Promoter of the C 4 Phosphoenolpyruvate Carboxylase Gene. The Plant Cell **16:** 1077–1090.

Gowik U, Westhoff P (2011) The Path from C3 to C4 Photosynthesis. Plant Physiology **155:** 56–63.

Groth A, Rocha W, Verreault A, Almouzni G (2007) Chromatin Challenges during DNA Replication and Repair. Cell **128:** 721–733.

Gutterson N, Reuber TL (2004) Regulation of disease resistance pathways by AP2/ERF transcription factors. Current Opinion in Plant Biology **7:** 465–471.

Ha M (2013) Understanding the chromatin remodeling code. Plant Science **211:** 137–145.

Haring M, Offermann S, Danker T, Horst I, Peterhansel C, Stam M (2007) Chromatin immunoprecipitation: optimization, quantitative analysis and data normalization. Plant Methods **3:** 11.

Hassan AH, Awad S, Al-Natour Z, Othman S, Mustafa F, Rizvi TA (2007) Selective recognition of acetylated histones by bromodomains in transcriptional co-activators. Biochemical Journal **402:** 125–133.

He G, Chen B, Wang X, Li X, Li J, He H, Yang M, Lu L, Qi Y, Wang X, Deng XW (2013) Conservation and divergence of transcriptomic and epigenomic variation in maize hybrids. Genome Biology **14:** R57.

Hibberd JM, Covshoff S (2010) The Regulation of Gene Expression Required for C 4 Photosynthesis. Annual Review of Plant Biology **61:** 181–207.

Ho JWK, Bishop E, Karchenko PV, Nègre N, White KP, Park PJ (2011) ChIP-chip versus ChIP-seq: Lessons for experimental design and data analysis. BioMed Central Genomics **12:** 134.

Horst I, Heimann L, Peterhansel C (2013) Signal integration on plant promoters: A case study in maize. Plant Signaling & Behavior **8:** e25389.

Horst I, Offermann S, Dreesen B, Niessen M, Peterhansel C (2009) Core promoter acetylation is not required for high transcription from the phosphoenolpyruvate carboxylase promoter in maize. Epigenetics & Chromatin **2:** 17.

Jang CJ, Sheen J (1994) Sugar Sensing in Higher Plants. The Plant Cell **6:** 1665–1679.

Kornberg RD (1977) Structure of Chromatin. Annual Review of Biochemistry **46:** 931–954.

Kouzarides T (2007) Chromatin Modifications and Their Function. Cell **128:** 693–705.

Kushwaha H, Gupta S, Singh VK, Rastogi S, Yadav D (2011) Genome wide identification of Dof transcription factor gene family in sorghum and its comparative phylogenetic analysis with rice and Arabidopsis. Molecular Biology Reports **38:** 5037–5053.

Langdale JA (2011) C4 Cycles: Past, Present, and Future Research on C4 Photosynthesis. The Plant Cell **23:** 3879–3892.

Lauberth SM, Nakayama T, Wu X, Ferris AL, Tang Z, Hughes SH, Roeder RG (2013) H3K4me3 Interactions with TAF3 Regulate Preinitiation Complex Assembly and Selective Gene Activation. Cell **152:** 1021–1036.

Lauria M, Rossi V (2011) Epigenetic control of gene regulation in plants. Biochimica et Biophysica Acta **1809:** 369–378.

Li B, Carey M, Workman JL (2007) The Role of Chromatin during Transcription. Cell **128:** 707–719.

Li P, Ponnala L, Gandotra N, Wang L, Si Y, Tausta SL, Kebrom TH, Provart N, Patel R, Myers CR, Reidel EJ, Turgeon R, Liu P, Sun Q, Nelson T, Brutnell TP (2010) The developmental dynamics of the maize leaf transcriptome. Nature Genetics **42:** 1060–1067.

Luger K, Mäder AW, Richmond RK, Sargent DF, Richmond TJ (1997) Crystal structure of the nucleosome core particle at 2.8 Angstrom resolution. Nature **389:** 251–260.

Majeran W, Friso G, Ponnala L, Connolly B, Huang M, Reidel E, Zhang C, Asakura Y, Bhuiyan NH, Sun Q, Turgeon R, van Wijk KJ (2010) Structural and Metabolic Transitions of C4 Leaf Development and Differentiation Defined by Microscopy and Quantitative Proteomics in Maize. The Plant Cell **22:** 3509–3542.

Martin L, Fei Z, Giovannoni JJ, Rose JKC (2013) Catalyzing plant science research with RNA-seq. Frontiers in Plant Science **4**: 66.

Matasci N, Hung LH, Yan Z, Carpenter EJ, Wickett NJ, Mirarab S, Nguyen N, Warnow T, Ayyampalayam S, Barker M, Burleigh JG, Gitzendanner MA, Wafula E, Der JP, Pamphilis CW de, Roure B, Philippe H, Ruhfel BR, Miles NW, Graham SW, et al. (2014) Data access for the 1,000 Plants (1KP) project. GigaScience **3:** 17.

Mishra S, Phukan UJ, Tripathi V, Singh DK, Luqman S, Shukla RK (2015) PsAP2 an AP2/ERF family transcription factor from Papaver somniferum enhances abiotic and biotic stress tolerance in transgenic tobacco. Plant Molecular Biology **89**: 173–186.

Nelissen H, Rymen B, Coppens F, Dhondt S, Fiorani F, Beemster GTS (2013) Kinematic Analysis of Cell Division in Leaves of Mono- and Dicotyledonous Species: A Basis for Understanding Growth and Developing Refined Molecular Sampling Strategies. Methods Molecular Biology **959:** 247–264.

Nelissen H, Rymen B, Jikumaru Y, Demuynck K, van Lijsebettens M, Kamiya Y, Inzé D, Beemster GTS (2012) A Local Maximum in Gibberellin Levels Regulates Maize Leaf Growth by Spatial Control of Cell Division. Current Biology **22:** 1183–1187.

Nelson T (2010) Development of Leaves in C4 Plants: Anatomical Features That Support C4 Metabolism. Advances in Photosynthesis and Respiration **32**: 147–159.

Nelson T (2011) The grass leaf developmental gradient as a platform for a systems understanding of the anatomical specialization of C4 leaves. Journal of Experimental Botany **62:** 3039–3048.

Nelson T, Langdale JA (1989) Patterns of Leaf Development in 04 Plants. The Plant Cell **1:** 3–13.

Offermann S, Danker T, Dreymüller D, Kalamajka R, Töpsch S, Weyand K, Peterhänsel C (2006) Illumination is necessary and sufficient to induce histone acetylation independent of transcriptional activity at the C4-specific phosphoenolpyruvate carboxylase promoter in maize. Plant Physiology **141:** 1078–1088.

Offermann S, Dreesen B, Horst I, Danker T, Jaskiewicz M, Peterhansel C (2008) Developmental and Environmental Signals Induce Distinct Histone Acetylation Profiles on Distal and Proximal Promoter Elements of the C4-Pepc Gene in Maize. Genetics **179:** 1891–1901.

Peterhansel C, Horst I, Niessen M, Blume C, Kebeish R, Kürkcüoglu S, Kreuzaler F (2010) Photorespiration. The Arabidopsis Book **8:** e0130.

Pick TR, Brautigam A, Schluter U, Denton AK, Colmsee C, Scholz U, Fahnenstich H, Pieruschka R, Rascher U, Sonnewald U, Weber APM (2011) Systems Analysis of a Maize Leaf Developmental Gradient Redefines the Current C4 Model and Provides Candidates for Regulation. The Plant Cell **23:** 4208–4220.

Poethig SR (1987) Clonal Analysis of Cell Lineage Patterns in Plant Development. American Journal of Botany **74:** 581–594.

Poethig SR, Szymkowiak EJ (1995) Clonal analysis of leaf development in maize. Maydica **40:** 67–76.

Ponnala L, Wang Y, Sun Q, van Wijk KJ (2014) Correlation of mRNA and protein abundance in the developing maize leaf. The Plant Journal **78:** 424–440.

Porto MS, Pinheiro MP, Batista VG, dos Santos RC, Filho Pde A, de Lima LM (2014) Plant Promoters: An Approach of Structure and Function. Molecular Biotechnology **56:** 38–49.

Roudier F, Ahmed I, Bérard C, Sarazin A, Mary-Huard T, Cortijo S, Bouyer D, Caillieux E, Duvernois-Berthet E, Al-Shikhley L, Giraut L, Després B, Drevensek S, Barneche F, Dèrozier S, Brunaud V, Aubourg S, Schnittger A, Bowler C, Martin-Magniette M, Robin S, Caboche M, Colot V (2011) Integrative epigenomic mapping defines four main chromatin states in Arabidopsis. The EMBO Journal **30:** 1928–1938.

Rymen B, Fiorani F, Kartal F, Vandepoele K, Inze D, Beemster GTS (2007) Cold Nights Impair Leaf Growth and Cell Cycle Progression in Maize through Transcriptional Changes of Cell Cycle Genes. Plant Physiology **143:** 1429–1438.

Sage RF, Sage TL, Kocacinar F (2012) Photorespiration and the Evolution of C 4 Photosynthesis. Annual Review of Plant Biology **63:** 19–47.

Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, Minx P, Reily AD, Courtney L, Kruchowski SS, Tomlinson C, Strong C, Delehaunty K, Fronick C, Courtney B, Rock SM, Belter E, et al. (2009) The B73 Maize Genome: Complexity, Diversity, and Dynamics. Science **326:** 1112–1115.

Sharman BC (1942) Developmental Anatomy of the Shoot of Zea mays L. Annual Botany **9:** 245–284.

Shu H, Wildhaber T, Siretskiy A, Gruissem W, Hennig L (2012) Distinct modes of DNA accessibility in plant chromatin. Nature Communications **3:** 1281.

Sugiharto B, Miyata K, Nakamoto H, Sasakawa H, Sugiyama T (1990) Regulation of Expression of Carbon-Assimilating Enzymes by Nitrogen in Maize Leaf. Plant Physiology **92:** 963–969.

Sugiharto B, Suzuki I, Burnell JN, Sugiyama T (1992) Glutamine Induces the N-Dependent Accumulation of mRNAs Encoding Phosphoenolpyruvate Carboxylase and Carbonic Anhydrase in Detached Maize Leaf Tissue. Plant Physiology **100:** 2066–2070.

Suzuki I, Crétin C, Omata T, Sugiyama T (1994) Transcriptional and Posttranscriptional Regulation of Nitrogen-Responding Expression of Phosphoenolpyruvate Carboxylase Gene in Maize. Plant Physiology **105**: 1223–1229.

Sylvester AW, Cande ZW, Freeling M (1990) Division and differentiation during normal and liguleless-1 maize leaf development. Development **110:** 985–1000.

Tausta SL, Li P, Si Y, Gandotra N, Liu P, Sun Q, Brutnell TP, Nelson T (2014) Developmental dynamics of Kranz cell transcriptional specificity in maize leaf reveals early onset of C4-related processes. Journal of Experimental Botany **65:** 3543–3555.

The Arabidopsis Genome Initiative (2000) Analysis of the genome sequence of the flowering plant Arabidopsis thaliana. Nature **408:** 796–8015.

Turner BM (2002) Cellular Memory and the Histone Code. Cell **111:** 285–291.

van Attikum H, Fritsch O, Hohn B, Gasser SM (2004) Recruitment of the INO80 Complex by H2A Phosphorylation Links ATP-Dependent Chromatin Remodeling with DNA Double-Strand Break Repair. Cell **119:** 777–788.

van den Bergh E, Külahoglu C, Bräutigam A, Hibberd JM, Weber APM, Zhu X, Schranz ME (2014) Gene and genome duplications and the origin of C4 photosynthesis: Birth of a trait in the Cleomaceae. Current Plant Biology **1:** 2–9.

Vermeulen M, Mulder KW, Denissov S, Pijnappel W, van Schaik FM, Varier RA, Baltissen MP, Stunnenberg HG, Mann M, Timmers HT (2007) Selective Anchoring of TFIID to Nucleosomes by Trimethylation of Histone H3 Lysine 4. Cell **131:** 58–69.

Wang L, Czedik-Eysenberg A, Mertz RA, Si Y, Tohge T, Nunes-Nesi A, Arrivault S, Dedow LK, Bryant DW, Zhou W, Xu J, Weissmann S, Studer A, Li P, Zhang C, LaRue T, Shao Y, Ding Z, Sun Q, Patel RV, Turgeon R, Zhu X, Provart NJ, Mockler TC, Fernie AR, Stitt M, Liu P, Brutnell TP (2014) Comparative analyses of C4 and C3 photosynthesis in developing leaves of maize and rice. Nature Biotechnology **32:** 1158–1165.

Wang X, Elling AA, Li X, Li N, Peng Z, He G, Sun H, Qi Y, Liu XS, Deng XW (2009) Genome-Wide and Organ-Specific Landscapes of Epigenetic Modifications and Their Relationships to mRNA and Small RNA Transcriptomes in Maize. The Plant Cell **21:** 1053–1069.

Weber APM, Weber KL, Carr K, Wilkerson C, Ohlrogge JB (2007) Sampling the Arabidopsis Transcriptome with Massively Parallel Pyrosequencing. Plant Physiology **144:** 32–42.

West PT, Li Q, Ji L, Eichten SR, Song J, Vaughn MW, Schmitz RJ, Springer NM, Sullivan BA (2014) Genomic Distribution of H3K9me2 and DNA Methylation in a Maize Genome. PLoS ONE **9:** e105267.

Wysocka J, Swigut T, Xiao H, Milne TA, Kwon SY, Landry J, Kauer M, Tackett AJ, Chait BT, Badenhorst P, Wu C, Allis CD (2006) A PHD finger of NURF couples histone H3 lysine 4 trimethylation with chromatin remodelling. Nature **442**:86–90.

Yanagisawa S, Sheen J (1998) Involvement of Maize Dof Zinc Finger Proteins in Tissue-Specific and Light-Regulated Gene Expression. The Plant Cell **10:** 75–89.

Yu C, Chen SC, Chang Y, Liu W, Lin H, Lin J, Chen HJ, Lu Y, Wu Y, Lu MJ, Lu C, Shih AC, Ku MS, Shiu S, Wu S, Li W (2015) Transcriptome dynamics of developing maize leaves and genomewide prediction of cis elements and their cognate transcription factors. Proceedings of the National Academy of Sciences USA **112:** E2477.

Zhang X, Bernatavichute YV, Cokus S, Pellegrini M, Jacobsen SE (2009) Genome-wide analysis of mono-, di- and trimethylation of histone H3 lysine 4 in Arabidopsis thaliana. Genome Biology **10:** R62.

Zhang X, Clarenz O, Cokus S, Bernatavichute YV, Pellegrini M, Goodrich J, Jacobsen SE (2007) Whole-Genome Analysis of Histone H3 Lysine 27 Trimethylation in Arabidopsis. PLoS Biology **5:** e129.

Zhou J, Wang X, He K, Charron JF, Elling AA, Deng XW (2010) Genome-wide profiling of histone H3 lysine 9 acetylation and dimethylation in Arabidopsis reveals correlation between multiple histone marks and gene expression. Plant Molecular Biology **72:** 585–595.

Appendix

List of abbreviations

ADP	Adenosine diphosphate
AMP	Adenosine monophosphate
ATP	Adenosine triphosphate
ac	Acetylation
bp	Base pairs
BS	Bundle sheath
CA	Carbonic anhydrase (enzyme)
Ca	Carbonic anhydrase (gene)
cDNA	Complementary DNA
ChIP	Chromatin immunoprecipitation
CO_2	Carbon dioxide
DNA	Deoxyribonucleotide acid
DOF	DNA binding with one finger
DZ	Division zone
ERF1	Ethylene responsive factor
EZ	Expansion zone
GO	Gene ontology
H	Histone
HCO_3-	Bicarbonate
HMT	Histone methyltransferases
Mbp	Mega base pairs
MDH	Malate dehydrogenase (enzyme)
ME	Malic enzyme (enzyme)
Me	Malic enzyme (gene)
me	Methylation
me2	Dimethylation
me3	Trimethylation
MZ	Maturing Zone
K	Lysine

NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)
NADP	Nicotinamide adenine dinucleotide phosphate (oxidized)
NGS	Next generation sequencing
OAA	Oxalacetate
PCR	Polymerase chain reaction
PEPC	Phosphoenolpyruvate carboxylase (enzyme)
Pepc	Phosphoenolpyruvate carboxylase (gene)
PEPCK	Phosphoenolpyruvate carboxykinase (enzyme)
Pepck	Phosphoenolpyruvate carboxykinase (gene)
P_i	Inorganic phosphate
PPDK	Pyruvate-P_i-dikinase (enzyme)
Ppdk	Pyruvate-P_i-dikinase (gene)
qPCR	Quantitative Real-time PCR
RNA	Ribonucleic acid
R-SUP	Regulated secondary upstream peaks
RubisCO	Ribulose-1,5-bisphosphate carboxylase-oxygenase (gene)
RbcS2	Small subunit of RubisCO (gene)
Sdg	SET domain group
Seq	Sequencing
TE	Transposable element
TF	Transcription factor
TFB	Transcription factor binding site
TIS	Transcription initiation site

Danksagung

Ich danke Prof. Dr. Peterhänsel für die Möglichkeit die Doktorarbeit am Institut für Botanik anfertigen zu können. Auch möchte ich mich dafür bedanken, dass Du es mir ermöglichst hast, die sehr kostspieligen NGS-basierten Methoden zu erlernen und in der Arbeitsgruppe zu etablieren. Vielen Dank für das entgegengebrachte Vertrauen. Außerdem möchte ich mich für die anregenden wissenschaftlichen Diskussionen und die umfassenden Hilfestellungen bedanken. Deine sachlichen Kommentare haben mir dabei geholfen meine Gedanken und Ausführungen auf das Wesentliche zu lenken.

Ich danke Prof. Dr. Debener und Prof. Dr. Braun für die Übernahme des Korreferates beziehungsweise des Vorsitzes für meine Promotion.

Ina danke ich für die Gespräche über Wissenschaftliches und Privates. Deine Einschätzungen haben mir sehr dabei geholfen die komplexen und alltäglichen Probleme im Labor oder am Auswerterechner zu lösen oder zumindest zu umschiffen. Wenn es gerade mal keine Probleme gab, hatten wir immer noch den Fußball.

Den Arbeitsgruppen Peterhänsel, Huchzermeyer und Offermann möchte ich für die angenehme Arbeitsatmosphäre und die gebotenen Hilfestellungen danken.

Nicht zuletzt möchte ich meiner Lebensgefährtin Fabienne, meiner Familie und meinen Freunden für die umfangreiche Unterstützung danken. Wenn ich am Boden war, habt Ihr mich wieder aufgerichtet und mir dabei geholfen meine (Selbst)zweifel zu zerstreuen. Und wie es aussieht, hattet Ihr am Ende Recht.

Lebenslauf

Name	Renke Perduns
Geburtsdatum	03.06.1984
Geburtsort	Wittmund
Staatsangehörigkeit	Deutsch

Akademischer Werdegang

Seit 03.2011	Wissenschaftlicher Mitarbeiter / Promotion an der Leibniz Universität Hannover, Institut für Botanik (Prof. Dr. Peterhänsel)
11.2008 – 11.2010	Masterstudium Pflanzenbiotechnologie an der Leibniz Universität Hannover. Masterarbeit am Institut für Pflanzengenetik (Prof. Dr. Braun): Charakterisierung pathogen-induzierter Veränderungen des mitochondrialen Proteoms von *Medicago truncatula*
10.2005 – 10.2008	Bachelorstudium Pflanzenbiotechnologie an der Leibniz Universität Hannover. Bachelorarbeit am Institut für Pflanzengenetik (Prof. Dr. Braun): Wirkung des „Harpin“ Effektorproteins auf das oxidative Phosphorylierungssystem pflanzlicher Mitochondrien

Zivildienst

2004 – 2005	Zivildienst als Haustechniker Kreiskrankenhaus Wittmund

Schulischer Werdegang

1997 – 2004	Abitur, Alexander-von-Humboldt-Schule in Wittmund
1995 – 1997	Orientierungsstufe in Wittmund
1991 – 1995	Grundschule in Ardorf