Signal integration on photosynthetic promoters in the C4 grass species *Zea mays*, *Sorghum bicolor* and *Setaria italica*

Von der Naturwissenschaftlichen Fakultät
der Gottfried Wilhelm Leibniz Universität Hannover
zur Erlangung des Grades
Doktorin der Naturwissenschaften
Dr. rer. nat.

genehmigte Dissertation
von
M. Sc. Louisa Heimann
Geboren am 29.03.1984 in Hagen

2013

Referent : Prof. Dr. Christoph Peterhänsel
Koreferent : Prof. Dr. Hans-Peter Braun
Tag der Promotion: 23.08.2013

Summary

Expression of the C4-specific isoform of the phosphoenolpyruvate carboxylase gene (C4-*Pepc*) in maize is affected by multiple endogenous and exogenous stimuli. Regulation takes place primarily on the level of transcription so that most of these stimuli have to be integrated into a promoter response. The gene is therefore an excellent model to study signal integration on the chromatin level.

The first aim of this project was to compare the chromatin regulation of C4 gene expression in the millet species *Setaria italica* and *Sorghum bicolor* with the previously analyzed regulatory mechanisms in *Zea mays.* The hypothesis was that chromatin modification patterns found in maize are ubiquitous mechanisms for integration of information on promoters. *Sorghum bicolor* and *Setaria italica* are suitable candidates to answer this question, because *Sorghum* is closely related to maize whereas S*etaria* separated from the maize/*Sorghum* lineage approximately 25 Million years ago and evolved C4 photosynthesis independently. Modification profiling on two selected C4 genes in *Sorghum* and *Setaria* suggested that a histone code is used in independent C4 lineages and, thus, was probably recruited into C4 from an ancient mechanism already existing in C3 plants. A model summarizing our current knowledge about epigenetic gene regulation of C4 promoters is provided.

In the second part of this work, the current knowledge about histone modifications in C4 gene regulation was expanded by profiling the abundance of five so far non-characterized histone H3 modifications on C4-*Pepc*. The experiment revealed that each modification showed a specific distribution over the gene. Interestingly, all acetylations were regulated by light whereas methylation was always regulated in a cell-type specific manner.

Long noncoding RNAs are another factor influencing the activity of promoters. In the third part of this work, the aim was to characterize a long noncoding RNA (lncRNA) associated with the C4-*Pepc* promoter of maize. A lncRNAs homologous to 126 bp of promoter sequence of the gene in maize was identified. It was only detectable in tissues where the gene was expressed and followed the protein-coding transcript in abundance through light/dark treatment as well as in diurnal regulation. However, whereas the protein-coding transcript was completely suppressed by a treatment mimicking high sugar availability, the lncRNA remained unaffected by this stimulus. Pharmacological suppression of RNA polymerase II completely abandoned promoter activity, but did not affect lncRNA levels. Additionally it could be shown that the lncRNA is associated with chromatin.

Altogether these results expand our knowledge about the role of chromatin in signal integration in plants and specifically the complex regulation pattern of the C4-*Pepc* gene.

Zusammenfassung

Die Expression der C4-spezifischen Isoform des Phosphoenolpyruvatcarboxylase-Gens (C4-*Pepc*) wird durch vielfältige endogene und exogene Stimuli reguliert. Die Regulation erfolgt in erster Linie auf der Ebene der Transkription, so dass viele dieser Stimuli zu einer Promotorantwort integriert werden müssen. Das Gen ist daher ein hervorragendes System zur Studie der Signalintegration auf der Chromatinebene.

Das erste Ziel dieser Arbeit war es, das Vorhandensein eines allgemeinen Musters für die Genregulation in den Hirsesorten *Setaria italica* und *Sorghum bicolor* im Vergleich zu bereits erforschten regulatorischen Mustern in Mais zu prüfen. Die Arbeitshypothese war, dass Chromatinmodifikationsmuster generell zur Signalintegration auf Promotoren verwendet werden. *Sorghum bicolor* und *Setaria italica* stellen zur Untersuchung dieser Hypothese angemessene Kandidaten dar, denn *Sorghum* ist ein naher Verwandter von Mais, wohingegen sich *Setaria* vor etwa 25 Millionen Jahren von der Mais/*Sorghum* Linie getrennt und C4-Photosynthese unabhängig entwickelt hat. Die Modifikationsanalyse von zwei ausgewählten Genen in *Sorghum* und *Setaria* deutet daraufhin, dass dieser Code in unabhängigen C4 Organismen verwendet wird und demnach von einem Mechanismus der bereits in C3-Pflanzen etabliert war, in C4-Pflanzen rekrutiert wurde. Zusätzlich wurde ein Modell erstellt welches das bisherige Wissen zur epigenetischen Regulation von C4-Genen zusammenfasst.

Im zweiten Teil dieser Arbeit wurde das aktuelle Wissen über Histonmodifikationen in der C4-Genregulation durch die Analyse von fünf bis dahin noch nicht in Mais charakterisierten Modifikationen erweitert. Die Experimente zeigten, dass jede Modifikation eine einzigartige Verteilung über das Gen aufweist. Interessanterweise wurden alle untersuchten Acetylierungen durch Licht reguliert, wohingegen alle untersuchten Methylierungen gewebespezifisch reguliert wurden.

Ein weiterer Faktor, der die Promotoraktivität beeinflussen kann, sind lange nichtkodierende RNAs. Das Ziel des dritten Teils dieser Arbeit war es, eine lange nichtkodierende RNA (lncRNA) zu charakterisieren die mit dem C4-*Pepc* Promotor assoziiert ist. Es konnte festgestellt werden, dass die nichtkodierende RNA homolog zu 126 Basenpaaren der C4-*Pepc* Promotorsequenz aus Mais ist. Die lncRNA konnte nur in Geweben detektiert werden, in denen das Gen exprimiert wird. Ebenso zeigte die nichtkodiernede RNA während Hell/Dunkel Experimenten sowie in der diurnalen Regulation eine ähnlich hohe Abundanz wie das proteinkodierenden Transkript. Weiterhin konnte gezeigt werden, dass Stimuli wie hohe Zuckerverfügbarkeit, welche die Transkription des proteinkodierenden Transkripts hemmen, keinen Einfluss auf die Abundanz der nichtkodierenden RNA haben. Ebenfalls keinen Einfluss hatte die pharmakologische Suppression der RNA Polymerase II, welche die Promotoraktivität von C4-*Pepc* herabsetzt. Mit Hilfe von Chromatinanreicherung war es außerdem möglich eine Chromatinassoziation der lncRNA zu zeigen.

Zusammengefasst erweitern diese Ergebnisse unser bisheriges Wissen über die Rolle des Chromatins in der pflanzlichen Signalintegration und im Speziellen über die komplexe Regulation des C4-*Pepc* Gens.

Key words:

C4 photosynthesis, histone modifications, histone code

C4 Photosynthese, Histonmodifikationen, Histon Code

Table of contents

List of abbreviations

°C	Degree Celsius
µg	Micro gram
µl	Micro liter
3-PGA	3-phosphoglycerate
ac	Acetylation
ATP	Adenosine triphosphate
bp	Base pairs
B	Bundle sheath
cDNA	complementary DNA
ChIP	Chromatin Immunoprecipitation
CO_2	Carbon dioxide
DNA	Deoxyribonucleotide acid
DOG	Desoxyglucose
h	Hour
H_2O	Water
HAT	Histone acetyltransferase
HDAC	Histone deacatylase
hnRNA	Heterogeneous nuclear RNA
lncRNA	Long noncoding RNA
M	Mesophyll
ME	Malic enzyme
me	Methylation
me2	Dimethylation
me3	Trimethylation
min	Minute
miRNA	microRNA
M-MLV-RT	Reverse transcriptase (from Moloney Murine Leukemia virus)
C4-*Me*	Malic enzyme
mRNA	Messenger RNA
NADPH	Nicotinamide adenine dinucleotide phosphate (reduced)
NADP-ME	NADP-Malic Enzyme
O	Oxygen
PCR	Polymerase- Chain- reaction
PEP	Phosphoenolpyruvate (enzyme)
C4-*Pepc*	Phospho*enol*pyruvate carboxylase (gene)
PS	Photosystem
qPCR	Quantitative Real-time PCR
RT	Reverse transcriptase
Rubisco	Ribulose-1,5-bisphosphate carboxylase-oxygenase
siRNA	Short interfering RNA
TIS	Transcription initiation site
UTR	Untranslated regions

CHAPTER 1

General Introduction

1 Introduction

1.1 Photosynthesis

The process of photosynthesis is one important metabolic pathway in plants. It is the source for plant growth and development. Plants are photoautotrophic organisms that are able to use light energy for the conversion of inorganic compounds into organic compounds (Koussevitzky et al., 2007). Photosynthesis can be separated into two different reactions, the light-reactions and the light-independent reactions. Plants absorb light using the pigment chlorophyll. They cannot use the entire light spectrum but absorb light at specific wavelengths in the blue and red spectral range. Light-dependent reactions occur in the thylakoid membranes of the chloroplasts and the main function is to use light energy to synthesize adenosine triphosphate (ATP) and reduced nicotineamide adenine dinucleotide phosphate (NADPH). These energy and reducing equivalents can further be used in the light-independent reaction, the Calvin cycle, for the fixation of carbon dioxide (CO_2). During the process of photosynthesis sugars are synthesized from carbon dioxide and water. In all plants CO_2 is fixed by the enzyme Ribulose-1.5-bisphosphate carboxylase-oxygenase (Rubisco). Rubisco is the most abundant protein in the world and it is involved in the light-independent reaction of the Calvin cycle. One important characteristic of Rubisco, besides the carboxylase activity, is its oxygenase activity. It is able to fix CO_2 as well as O_2 and both compounds compete for the active site of Rubisco. Despite the fact that it has a higher specificity for CO_2 than for O_2, every fourth reaction is an oxygenation (Brooks, 1985).

Due to evolutionary changes in the atmosphere, CO_2 concentrations decreased and O_2 concentrations increased, Rubisco oxygenase activity increased as well. The fixation of oxygen results in one molecule each of 3-phosphoglycerate and 2-phosphoglycolate. Phosphoglycolate has no metabolic purpose (Gowik and Westhoff, 2011) and the process in which phosphoglycolate is disposed is called photorespiration. During this process, energy has to be invested and previously fixed carbon is released. Therefore, photorespiration leads to decreased carbohydrate production and thus leads to losses in both biomass and yield of plants (Peterhansel et al., 2010). In the Calvin cycle CO_2 is converted to two molecules of 3-phosphoglycerate (3-PGA). Subsequently 3-PGA is reduced to 3-glyceraldehyde-phosphate (GAP), a reaction that uses the NAPDH and ATP produced in the light reaction. In summary, three CO_2 fixation events are necessary to produce one reduced carbohydrate (Taiz L, 2007).

Overall the photosynthetic efficiency of many plants is reduced by the inability of Rubisco to suppress the reaction with oxygen. C4 plants overcome this limitation by increasing the concentration of carbon dioxide around the enzyme (Gowik and Westhoff, 2011).

1.1.1 C4-Photosynthesis

Major features of C4 plants are the high rates of photosynthesis as well as the efficient use of water and nitrogen resources. It is an adaptive trait that reduces photorespiration under low CO_2 and high temperature (Sage et al., 2012).

The high photosynthetic efficiency of C4 plants is due to their mechanism of carbon assimilation which results from several biochemical and anatomical modifications that allow plants to concentrate CO_2 at the site of Rubisco. As a consequence of this concentrating mechanism of CO_2, the competitive inhibition of Rubisco by oxygen is largely repressed and C4 plants show drastically reduced rates of photorespiration (Wyrich et al., 1998; Gowik and Westhoff, 2011). C4 leaves consist of mainly two cell types, mesophyll (M) and bundle sheath (B) cells. The leaf anatomy allows C4 plants a primary and secondary CO_2 fixation reaction. This peculiar anatomy is called Kranz anatomy (see also Figure 1-1).

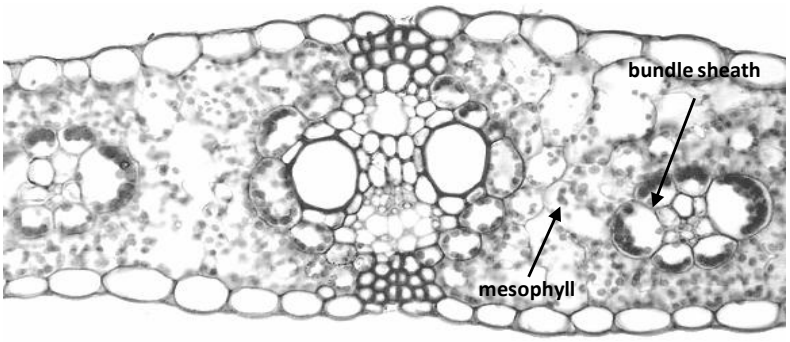


Figure 1-1: Kranz anatomy of C4 leaves.
Illustrated is the microscopic picture of a cross section of a leaf. Arrows point to bundle sheath cells and mesophyll cells, respectively (Hahnen S, 2004).

Figure 1-2 shows a schematic overview of NADP-ME dependent C4 photosynthesis. Different modifications of this pathway can be found beside the NADP-ME dependent C4 cycle, depending on the decarboxylating enzyme and the localization of this enzyme. Different from the model described in the next chapter the decarboxylation can also take place in the mitochondria or in the cytoplasm of the B cells (Kanai, 1999).

In C4 leaves, atmospheric CO_2 enters the cytoplasm of M cells through stomata, where it is initially fixed into the four-carbon compound, oxaloacetate by phosphoenolpyruvate carboxylase (PEPC). Oxalacetate is reduced to malate in the mesophyll chloroplasts by NADP dependent malate dehydrogenase. Malate diffuses from M cells to B cells, presumably through palsmodesmata, which are abundant at the interface of the two cell types (Gowik and Westhoff, 2011). In B cells chloroplasts, malate is decarboxylated by NADP-malic enzyme to release CO_2 near to Rubisco. The product formed from this reaction, pyruvate, is returned to the M cells where it is phosphorylated to phosphoenolpyruvate by Pyruvate-P_i-Dikinase (Langdale, 2011).

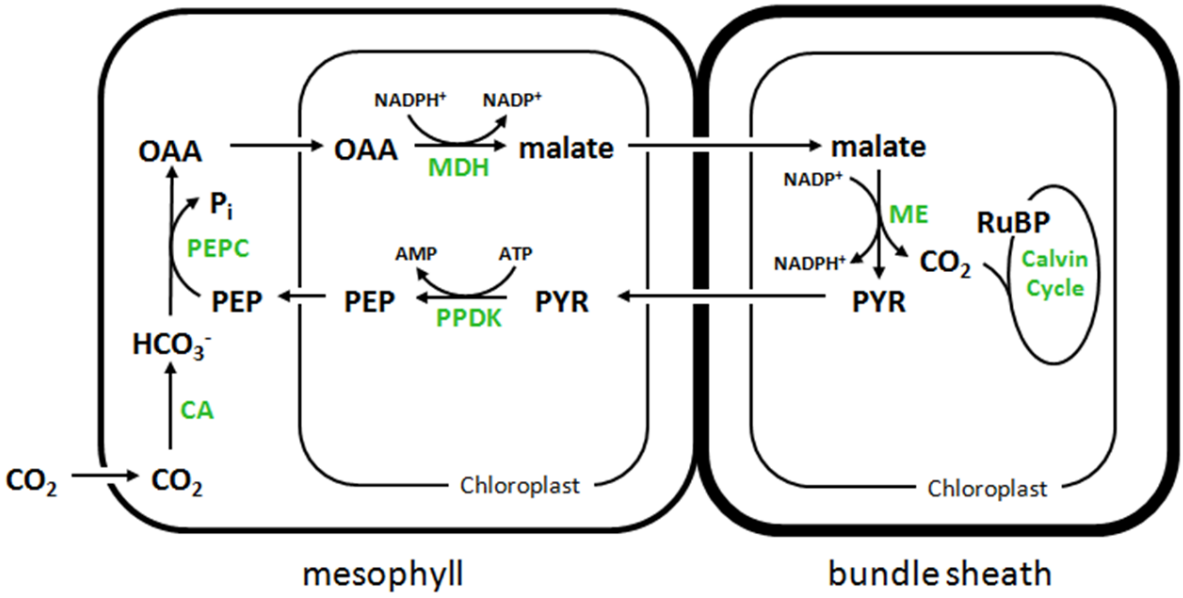


Figure 1-2: Schematic overview of the C4 photosynthesis of *Zea mays*.
Schematic overview of NAPD-ME type like C_4 photosynthesis like it is preformed in maize. To avoid CO_2 diffusion, the bundle sheath cells have a fortified cell wall that is denoted in the Figure. **CA:** Carboanhydrase; **PEPC:** Phospho*enol*pyruvate carboxylase; **MDH:** Malate dehydrogenase; **ME:** Malic enzyme; **PPDK:** Pyruvate-P_i-Dikinase; **HCO_3^-:** Bicarbonate; **PEP:** Phosphoenolpyruvate; **OAA:** Oxaloacetate; **PYR:** Pyruvate; **RuBP:** Ribulose1-5,bisphosphate; **$NAPDH^+$:** Nicotinamide adenine dinucleotide phosphate (reduced); **$NAPD^+$:** Nicotinamide adenine dinucleotide phosphate (oxidized); **AMP:** Adenosine monophosphate; **ATP:** Adenosine triphosphate. (Horst, 2009)

The energy costs of the C4 cycle are higher than the costs of C3 plants because every CO_2 molecule has to be fixed twice. Firstly CO_2 is converted into a 4-carbon organic acid and secondly it is fixed by Rubisco. Although the energy demands of C4 plants are higher, the energetic costs of the C4 pathway are balanced through the fact that CO_2 is concentrated at the site of Rubisco to prevent oxygen from competing for the active site of Rubisco (Ewing et al., 1998; Dai et al., 1993).
Thus, C4 plants are especially abundant under environmental conditions that would normally favour photorespiration such as high light intensities, high temperatures, and dryness. C4 plants are mostly found in grassland floras and in the tropical and subtropical regions of the earth (Edwards and Smith, 2010).

1.2 Phylogenetic relationship between maize, *Sorghum* and *Setaria*

C4 photosynthesis is a very good example for convergent evolution because the same biological trait evolved independently about 62 times in at least 17 plant families (Sage et al., 2012). A number of C4 plants are among the most productive crops in agriculture like *Zea mays*, or the millet species *Sorghum bicolor* and *Setaria italica* (common name: foxtail millet) (Wyrich et al., 1998, Christin et al., 2009a). Whereas *Sorghum* and maize share a common C4 origin, *Setaria* separated from the maize/*Sorghum* lineage approximately 25 Million years ago and evolved C4 photosynthesis independently (Vicentini et al., 2008). All of the three species belong to the PACMAD clade of *Poaceae* family (Christin et al., 2009a). The PACMAD clade contains the six subfamilies, Panicoideae, Arundinoideae, Chloridoideae, Micrairoideae, Aristidoideae, and Danthonioideae. The Panicoideae

subfamily contains the agronomically important crop plants maize, *Sorghum*, in the tribe Andropogoneae and *Setaria* in the tribe Paniceae (Mauro-Herrera et al., 2013; Li and Brutnell, 2011). Also C3 species are found in the Paniceae as well as in the Andropogoneae. Figure 1-3 exemplarily illustrates the phylogenetic relationship of some C4 and C3 species among the *Poaceae*.

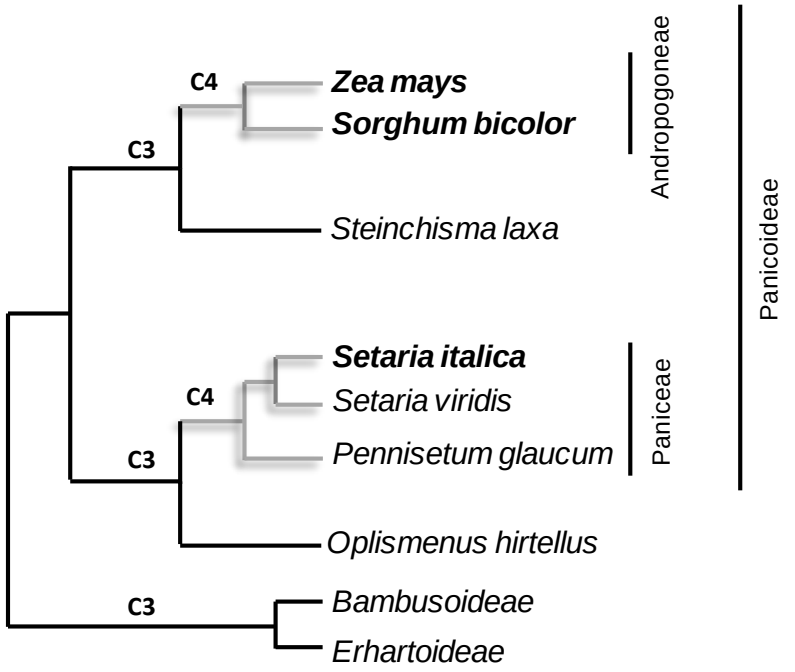


Figure 1-3: Phylogeny of the Poaceae.
The cladogram illustrates the phylogenetic relationship of C4 and C3 species among Poaceae. Highlighted are *Zea mays*, *Sorghum bicolor* and *Setaria italica.* Maize and *Sorghum* share a common C4-origin. *Setaria* evolved C4-metabolism independently. Species on the end of grey branches perform C4 photosynthesis, on black branches C3 photosynthesis, respectively (altered after Brutnell et al., 2010 and Ibrahim et al., 2008).

17 independent C4 origins are proposed by Christin et al., 2007 in the Panicoideae subfamily and none in the other five families (Langdale; 2011).

One major aspect and a general prerequisite of C4 evolution is based on the creation of new genes and the recruitment of pre-existing genes to encode the enzymes of the C4 pathway (Hibberd and Covshoff, 2010). These genetic redundancies have been acquired through the process of duplications of whole genomes, genome parts, or only single genes. Therefore, massive changes in gene regulation are needed and came along with these alterations (Gowik and Westhoff, 2011). Furthermore, leaves have been altered towards Kranz anatomy, a photorespiratory CO_2 pump was established, and finally a C4 cycle was created. There are just few exceptions, e. g. in the aquatic lineages and in two Chenopod lineages (*Binertia* and *Suaeda*), where the pathway takes place in a single cell (Langdale 2011).

1.2.1 The C4 specific phosphoenolepyruvate carboxylase of maize, sorghum and *Setaria italica*

The C4-*Pepc* gene family is very well studied especially in maize and the enzyme plays an important role in the C4 system (Wang et al., 2009; Schäffner and Sheen, 1992). C4-*Pepc* belongs to a multigene family which contains non-photosynthetic and C4 isoforms. In contrast to the specific C4 isoform, all the other genes are expressed on a constitutive basal level in different tissues. However the transcription of the C4 isoform is induced by light and the gene is exclusively expressed in mesophyll cells. Therefore the promoter had to acquire new functions such as high light-inducibility, and mesophyll-specific expression of the gene

The C4-*Pepc* gene is well analyzed for several regulatory mechanisms. Besides the light regulation (Horst, 2009), the nitrogen availability of the plant plays an important role (Sheen, 1999). High hexose concentrations were identified to act inhibitory on C4-*Pepc* transcription (Sheen, 1990) and it is also known that the C4-*Pepc* gene is regulated in the diurnal rhythm (Horst, 2009). Recent studies from our lab indicated that chromatin structure and histone modifications are important in the response of the promoter to all these signals (Mellor, 2005; Offermann et al., 2008). Because of its complex transcriptional regulation, C4-*Pepc* is an excellent model for studying the integration of environmental and developmental stimuli on a promoter and thus regulating of its activity. Figure 1-4 demonstrates the gene organization of the C4-*Pepc* genes found in maize, *Sorghum* and *Setaria.* It was altered after phytozome.org, where the gene structure and environment for each species was predicted based on genome sequence information.

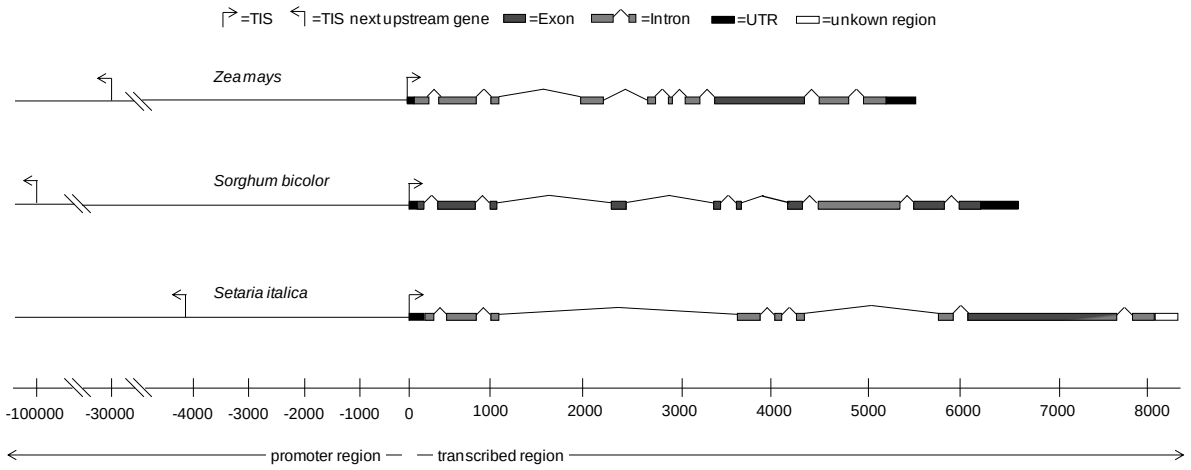


Figure 1-4: Gene organization of the C4-*Pepc* genes of *Zea mays, Sorghum bicolor and Setaria italica*.
Gene structure is shown as an intron-line, exon-block diagram. The *Zea mays* C4-*Pepc* gene is 5408 bp long and consists of 10 exons and 9 introns, the *Sorghum bicolor* C4-*Pepc* gene is 6536 bp long and consists of 10 exons and 9 introns and the *Setaria italica* C4-*Pepc* gene is 8433 bp long and consists of 9 exons and 8 introns. The three genes showed comparable intron-exon organization, but no detectable sequence homology on the putative promoters. 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 4 kb upstream of the *Setaria* C4-*Pepc* gene. The 3'UTR of *Setaria* C4-*Pepc* is so far unknown (sequence information was taken from phytozome.org).

The genome of eukaryotes is mainly organized in linear chromosomes. The localization of genes on the chromosomes is different for each species. The investigated C4-*Pepc* gene is located on different chromosomes in the three different species analyzed. It is located in the maize genome on chromosome 9, in the *Sorghum* genome on chromosome 10 and in the *Setaria* genome on chromosome 4 (Christin et al., 2009; Goodstein et al., 2012). The three genes differ in length but

showed comparable intron-exon organization but no sequence homology on the putative promoters (Figure 1-4). Furthermore the neighboring gene architecture upstream of the C4-*Pepc* genes was different in all three species (see also Chapter 2). The *Setaria* C4-*Pepc* is located nearby a *serine/threonine phosphatase* gene, which is transcribed in opposite direction to the C4-*Pepc* gene. In contrast, the maize and *Sorghum* C4-*Pepc* genes are not flanked by a gene in a comparable distance.

1.3 Chromatin

The genome sizes of the species investigated in this study are about 2500 Mbp for maize (Schnable et al., 2009), 730 Mbp for *Sorghum* (Paterson et al., 2009) and 490 Mbp for *Setaria* (Doust et al., 2009). The maize genome consists of a chromosome set which would be up to 1.7 m uncompressed. This amount of DNA needs to be packed tightly together in order to be stored in the nucleus of a cell. For this purpose DNA is associated with certain proteins and this DNA-protein-complex is called chromatin. The core particle of chromatin is the nucleosome octamer which consists of each two of the histone proteins H2A, H2B, H3 and H4 and DNA that is wound around this core particle 1.65 times (Li et al., 2007, Kouzarides, 2007). Histone H1 acts as a stabilizer for this complex. This is the first step of DNA organization. In the second step multiple histones wrap into a fibre which has a density of 40-fold and contains nucleosomes arrays in their most compact form (Luger, 1997). The third state of packing is the chromosome structure, which can be seen under a light microscope. Since the DNA is wrapped around the histone proteins, chromatin is the real template for essential processes in the cell like replication, recombination and repair events (Li et al., 2007). Transcriptionally active chromatin is named euchromatin, whereas transcriptional inactive chromatin is called heterochromatin.

Although every cell of an individual contains the same amount and sequence of DNA, different sets of genes are transcribed in different cells. Different genes are expressed at different strength and under different conditions, as required for growth and development. This leads to a specific model of gene regulation. Gene regulation is possible on different levels like transcription, RNA processing, or translation. Nucleosomes can be part of the gene regulation on the transcriptional level.

1.3.1 The role of histone modifications in gene regulation

In principle chromatin structure is not accessible for transcription (Narlikar et al., 2002) so histones and their N-terminal tails are targets of several histone modifications, which are used to regulate transcription, DNA replication or DNA repair mechanisms (Lusser, 2001). Histones can be covalently modified in multiple ways. Acetylation and methylation of multiple lysines on the N-terminal tails of H3 and H4 have been best studied. Whereas acetylation is almost exclusively associated with active promoters, methylation can induce active or repressive states dependent on the residue that is methylated (Wang et al., 2009a; Pokholok et al., 2005).

Chromatin modifications are known to regulate chromatin structure by recruiting remodelling enzymes that can use the energy derived from ATP hydrolysis to rearrange nucleosomes and allow transcription or polymerase binding (Turner, 2002; Bannister and Kouzarides, 2011).

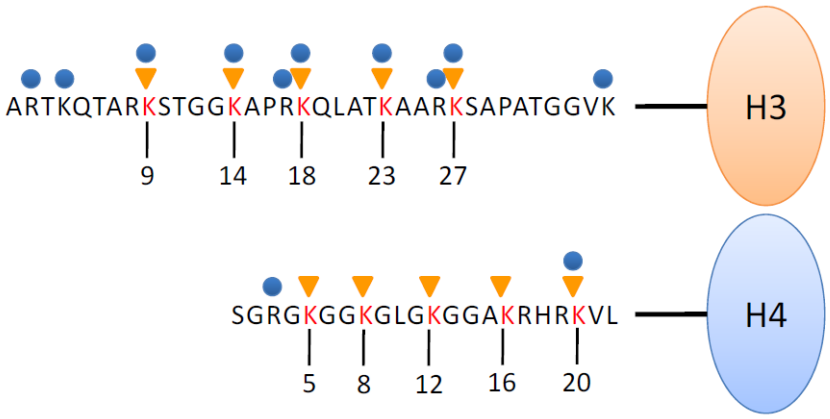


Figure 1-5: Postranslational acetylation and methylation of histone H3 and H4
Shown are the N-terminal end of histone H3 and histone H4 with potential acetylation (orange triangles) and methylation sites (blue dots) (altered after Benhamed et al., 2006; Chen and Tian, 2007).

Figure 1-5 exemplarily gives an overview about histone modifications found on histone H3 and histone H4. Methylation sites are illustrated as blue dots, whereas acetylation sites are presented as yellow triangles. Generally, acetylation takes place on lysine residues (Shahbazian and Grunstein, 2007) whereas methylation can take place on lysine and arginine residues (Pang et al., 2010). Lysine methylation is more extensively studied than arginine methylation. Lysines can be mono-, di- or tri-methylated (Dutnall, 2003). Beside acetylation and methylation, regulation can also take place by phosphorylation and ubiquitination (Yang and Seto, 2008), but these modifications are not discussed here because they are not relevant for this work.

Basically, illumination of plants leads to a stronger acetylation of the histones H3 and H4 in the promoter region and at the start of the coding region of genes being regulated by light. The steady-state equilibrium is controlled by the interplay of histone acetyltransferases (HAT) and histone deacetylases (HDACs) (Choi and Howe, 2009). It has been argued that, beside steady-state levels, the dynamic turn-over rate of acetylation is important in controlling transcription initiation (Clayton et al., 2006). Euchromatic areas are more often acetylated than heterochromatic areas. At the same time the frequency of specific acetylations differs strongly between plants and other organisms (Fuchs et al., 2006). Histone acetylation and thus the regulation of the maize C4-*Pepc* gene by this modification are already well studied (Offermann, 2006; Danker et al., 2008; Horst et al., 2009). Acetylation, especially histone H3 lysine 9 (H3K9) and histone H4 lysine 5 (H4K5) acetylation, is a modification that frequently correlates with actively transcribed genes (Offermann et al., 2008).

As well as acetylation, the equilibrium of histone methylation is controlled by certain enzymes, histonemethyltransferases and -demethylases (Bannister and Kouzarides, 2011; Biel et al., 2005; Shi et al., 2004). Histone H3 lysine 4 (H3K4) methylation is the most prominent methylation and H3K4 tri-methylation (H3K4me3) is also associated with euchromatic areas like acetylation. In contrast H3K9 methylation correlates with inactive gene areas (Martin and Zhang, 2005). In principle, promoters and coding regions can be associated with methylation, but whole genome studies revealed that the majority of methylation is concentrated around the transcription initiation site (Zhou et al., 2010).

1.3.2 Hypothesis of the histone code and charge neutralization model

Chromatin modifications are known to regulate chromatin structure by recruiting remodelling enzymes that can use the energy derived from ATP hydrolysis to rearrange nucleosomes (Wang et al., 2009a). They may also affect higher order chromatin structure by influencing the contact between nucleosomes and DNA. Histone modifications can lead to neutralization of positively charged histone tails, thereby weakening the interaction with the negatively charged DNA. This might allow better access for RNA polymerases and other transcription factors (Turner, 2002).

There are two competing hypotheses that explain these regulation patterns (see Figure 1-6). One possible way how histone modifications can be involved in gene regulation is due to charge neutralization (charge neutralization model). It is thought that some histone modifications can lead to neutralization of positively charged histone tails, and so weaken the interaction between proteins with the negatively charged DNA. This might lead to a better access for RNA polymerases and other transcription factors (Dion et al., 2005; Henikoff and Shilatifard, 2011). In this model certain signals are interpreted to an information integration which further leads to histone modifications. In this model the role of the histone modifications would be just to control the response of a certain integrator.

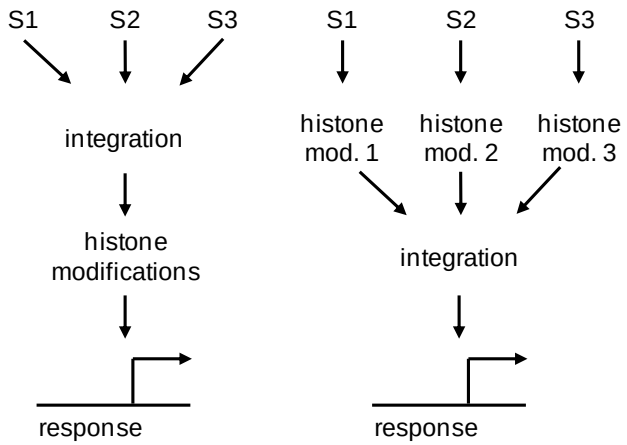


Figure 1-6: Model of the charge neutralization model and histone code model (see also Chapter 3).

The main idea of the histone code hypothesis is that genetic memory is not only stored in the DNA sequence itself, but also in tail modifications of histone proteins. The critical point is that histone modifications serve to recruit proteins by specific recognition of modified histones. These recruited proteins can act on chromatin structure and DNA accessibility to promote or prevent transcription. In this model, histone modifications are controlled by certain signals and used to integrate the signal and store the information on the promoter.

So the question arises, if histone modifications are used to store and amplify the signal or if they are just used to control the response through nucleosomes remodeling? Actually, it could be shown that both scenarios apply. The two hypotheses are discussed in detail in Chapter 3 ‘Signal integration on plant promoters: A case study in maize’.

1.4 Long non coding RNAs

The central dogma of molecular biology describes how genetic information is converted within a biological system. It is a classic view that DNA stores information and that this genetic information is transcribed into messenger RNA (mRNA) and finally translated into a protein (Francis Crick, 1970). In these days, a big part of the transcriptome was named as 'dark matter' (Yamada et al., 2003; van Bakel et al., 2010). The central has been extended. Beside transcripts coding for proteins (mRNAs), an ever increasing number of noncoding RNAs (ncRNAs) have been described in the transcriptome of eukaryotes. Nowadays noncoding RNAs (ncRNAs) have received increasing attention. With the help of microarray analysis and new sequencing technologies it has recently been shown that the majority of the genome is transcribed in eukaryotes (Lucia and Dean, 2011). As a consequence of these findings there is a need to discriminate between different noncoding RNAs. Beside long known examples such as ribosomal RNAs or transfer RNAs, Noncoding RNAs can include small RNAs, generally under 200 base pairs in length, and longer molecules, sometimes referred to as long noncoding RNAs (lncRNAs) (Kapranov et al., 2007; Jiao and Meyerowitz, 2010, Ørom et al., 2010).

Whereas small RNAs mostly act on the level of posttranscriptional gene regulation and RNA interference (Ghildiyal and Zamore, 2009), many possibilities have been considered of how long non coding RNAs might function. One possible function could be that lncRNAs guide certain proteins to specific genome regions (Baker, 2011), or that they act as scaffolds keeping protein complexes together. Another possibility is that lncRNAs act as scaffold and as guide at the same time. Further proposed possibilities are that lncRNAs act as byproducts to open the DNA and thereby to activate nearby genes, they act as effectors, allowing a protein to modify chromatin or otherwise regulate gene expression or they just act as enhancer or activator, for boosting gene transcription of genes encoding for proteins. Lee et al. (2012) were able to show that the XIST locus (*Xist* = X-inactive-specific transcript) encodes a 20 kbp RNA which is expressed from the inactive X chromosome. In plants, lncRNAs have been reported to function in directing chromatin-modifying activities to their targets and play an important role in development, such as flowering. Tsai et al. (2010) recently identified a 1.1 kbp noncoding RNA in *Arabidopsis thaliana* from the Flowering Locus C (FLC), called COLDAIR that recruits the chromatin-modifying complex, Polycomb repressive Complex 2 (PRC2) and establishes H3 lysine-27 trimethylation to mediate vernalization. Heo et al. (2011) recently showed that HOTAIR, a lncRNA derived from a Hox gene cluster, serves as a modular scaffold for PRC2.

Several lncRNAs have been identified with the ability to block transcription of their neighboring protein coding genes (Martens et al., 2004). On the other hand it was shown by Hirota et al. (2008) that transcription of certain promoter associated lncRNAs also can help to induce an open chromatin structure that allows the binding of activator proteins and transcription of the neighboring protein coding gene.

CHAPTER 2

A common histone modification code on C4 genes in maize and its conservation in Sorghum and *Setaria italica*

Authors

Louisa Heimann[1], Ina Horst[1], Renke Perduns, Björn Dreesen[2], Sascha Offermann, Christoph Peterhansel*

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

[1] both authors contributed equally to this work

[2] current address: Projektträger Jülich, 52428 Jülich

* corresponding author: Christoph Peterhansel, tel +49-511-7622632, fax +49-511-76219262, email: cp@botanik.uni-hannover.de

key words: C4 photosynthesis, histone modification, tissue-specificity, light-regulation, *Zea mays*, *Sorghum bicolor*, *Setaria italica*

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

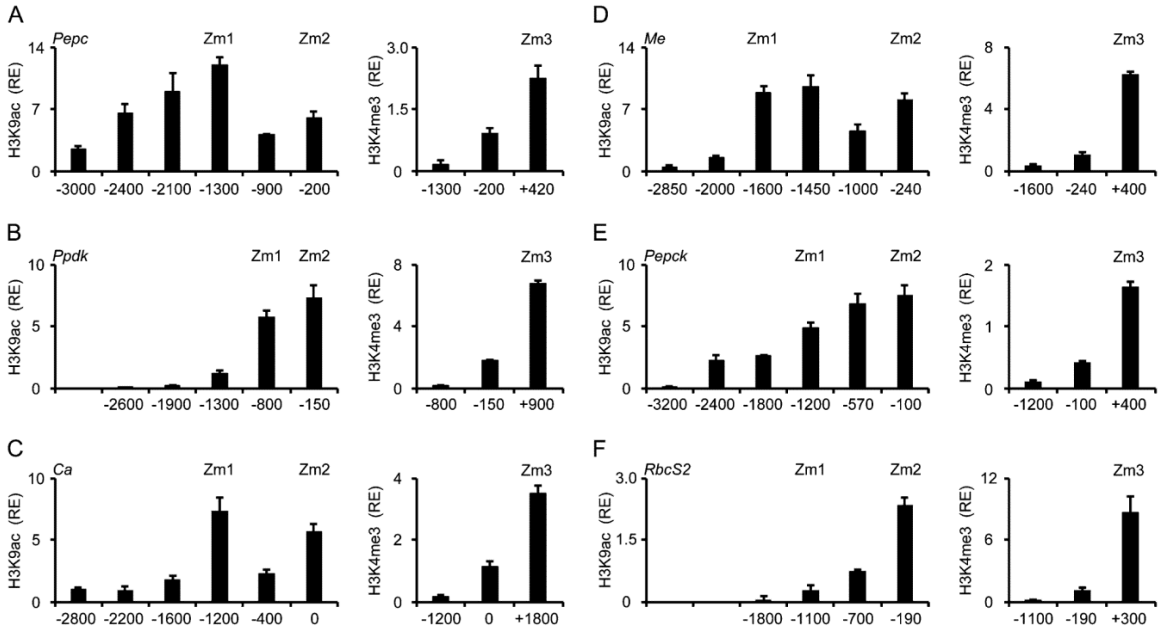


Figure 1. Histone modification profiles of six C4 genes from maize. Amounts of chromatin precipitated with an antibody specific for H3K9ac or H3K4me3 in illuminated leaves (4L) are shown. Numbers on the *x* axis indicate bp positions relative to TIS. Positions chosen for further analyses are designated 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

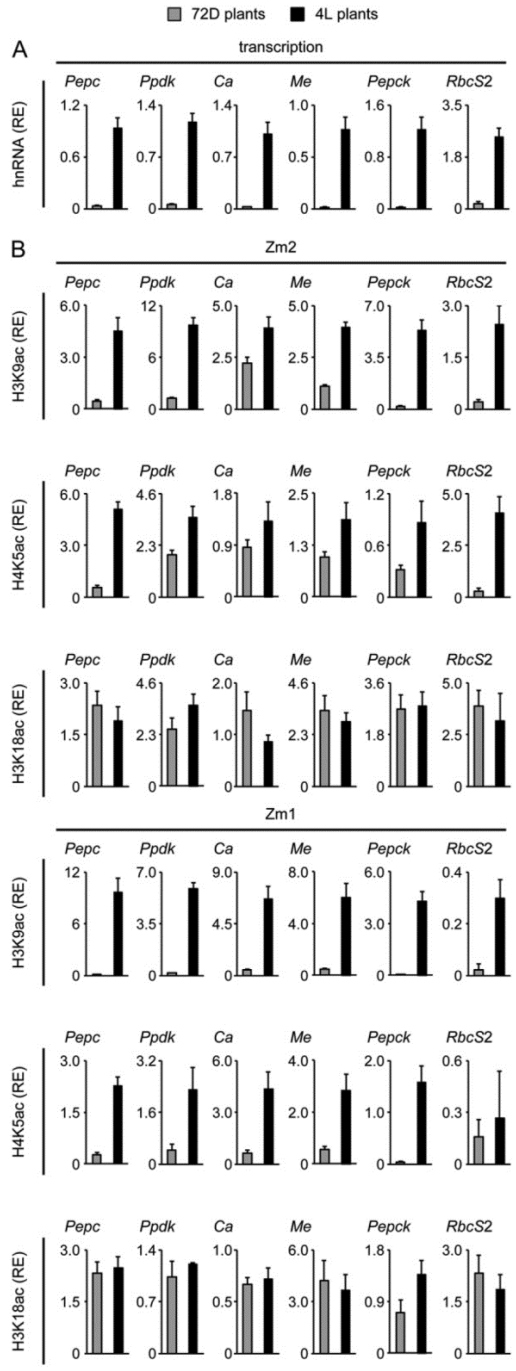


Figure 2. Light regulation of transcription and histone acetylation on six C4 genes in maize. A, Relative quantification of hnRNA expression levels of the six C4 genes in leaves from plants that were exposed to 72 h of darkness (72D; gray columns) and in plants that were illuminated for 4 h (4L; black 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.

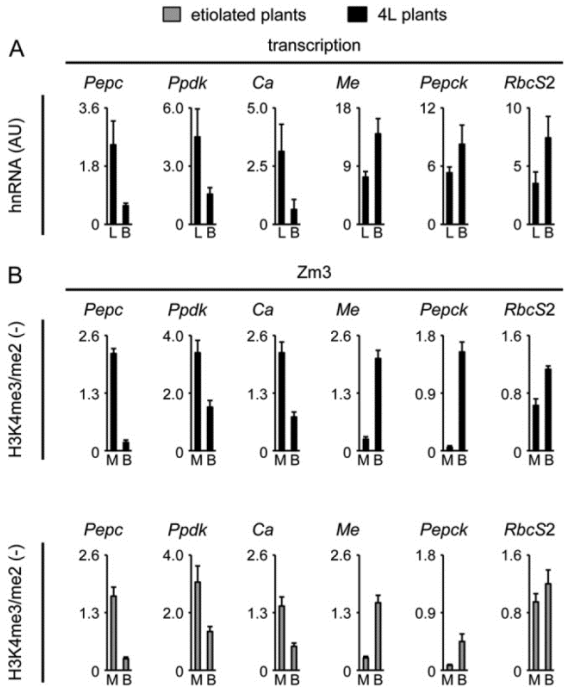


Figure 3. Cell type-specific transcription and histone methylation on six C4 genes in maize. A, Quantification of hnRNA expression levels of the six C4 genes in leaves (L) and B cells (B) isolated from plants that were illuminated for 4 h (4L plants). Values are arbitrary units (AU) derived from a cDNA 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).

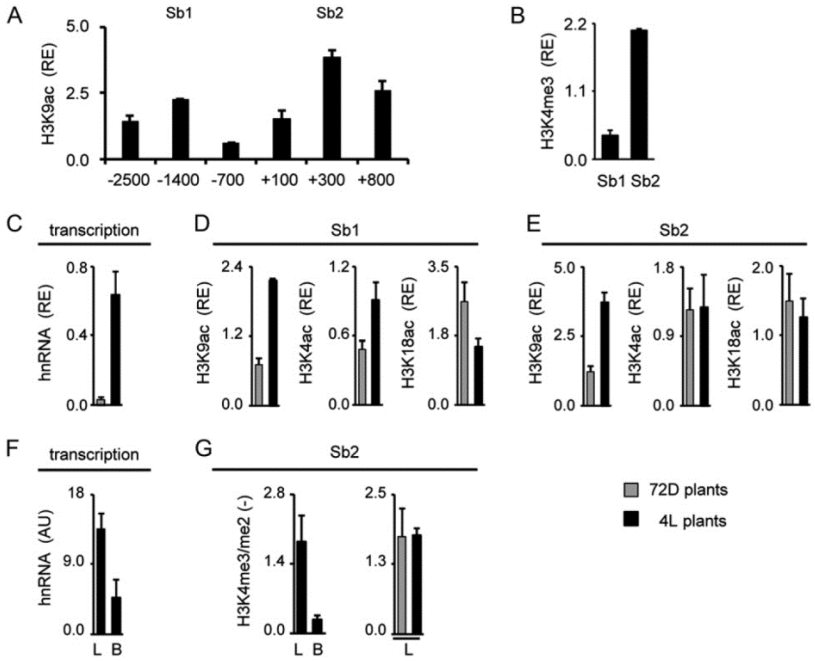


Figure 4. Histone modification profile of the C4-*Pepc* gene from sorghum. A and B, Amounts of chromatin precipitated with an antibody specific for H3K9ac and H3K4me3 in illuminated leaves (4L). Positions chosen for further analyses are designated Sb1 and Sb2. Values are presented as the relative enrichment (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

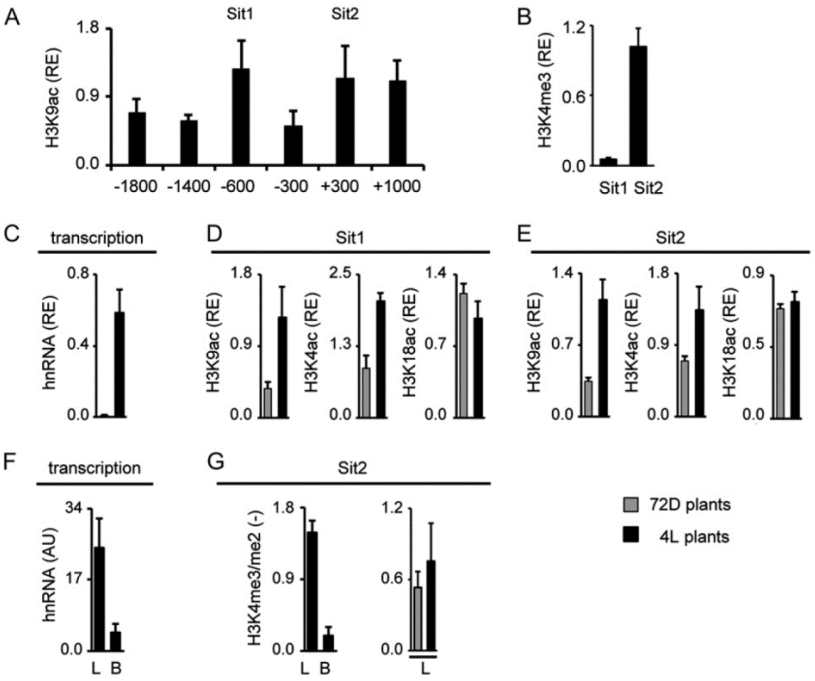


Figure 5. Histone modification profile of the C4-*Pepc* gene from *S. italica*. A and B, Amount of chromatin precipitated with an antibody specific for H3K9ac and H3K4me3 in illuminated leaves (4L). Positions chosen for further analyses are designated Sit1 and Sit2. Values are presented as the relative 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.

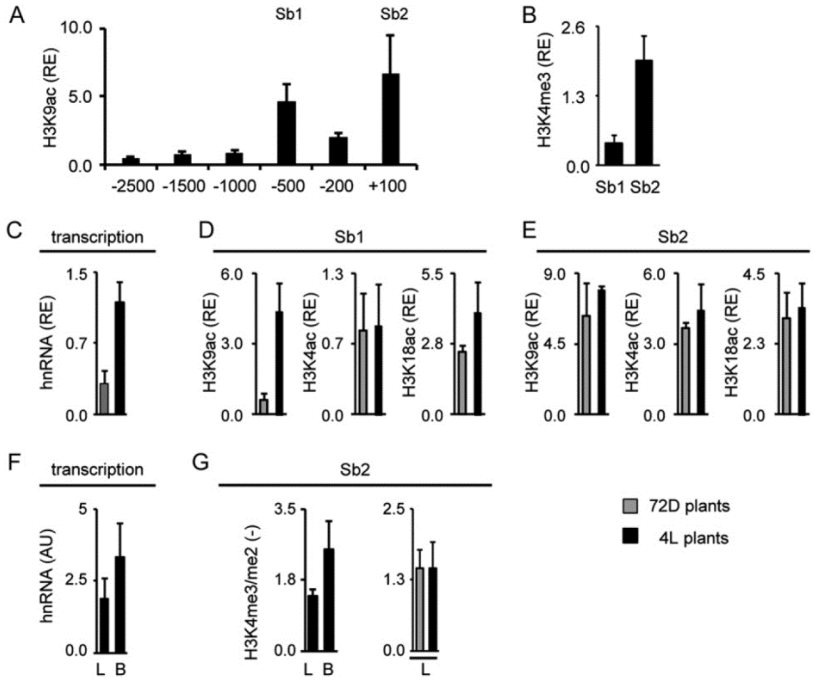


Figure 6. Histone modification profile of the C4-*Me* gene from sorghum. A and B, Amount of chromatin precipitated with an antibody specific for H3K9ac and H3K4me3 in illuminated leaves (4L). Positions chosen for further analyses are designated Sb1 and Sb2. Values are presented as the relative enrichment (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

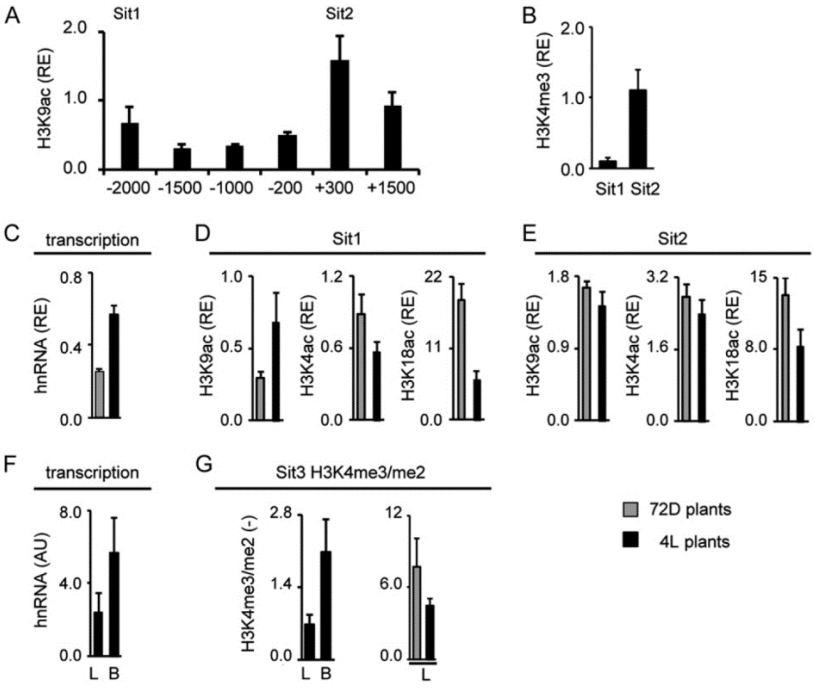


Figure 7. Histone modification profile of the C4-*Me* gene from *S. italica*. A and B, Amount of chromatin precipitated with an antibody specific for H3K9ac and H3K4me3 in illuminated leaves (4L). Positions chosen for further analyses are designated Sit1 and Sit2. Values are presented as the relative 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 tri-methylated 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

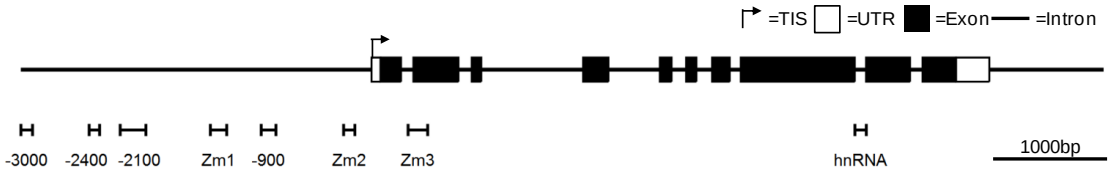
Yan C, Boyd DD (2006) Histone H3 acetylation and H3 K4 methylation define distinct chromatin regions permissive for transgene expression. Mol Cell Biol **26:** 6357–6371

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.

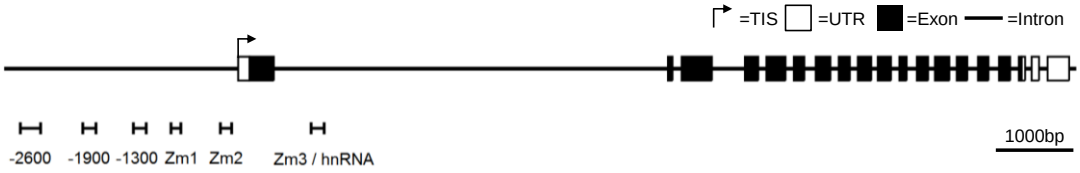
***Zea mays* – Phosphoenolpyruvate carboxylase (C4-*Pepc*)**	
Locus	GRMZM2G083841
Transcript variant	T01
Chromosome	9
Next upstream gene	~ 30 kb
Reference/Identification	genomic locus homologous to mRNA NM_001111948 (Wang et al., 2009) was identified by whole genome BLAST on www.phytozome.net (Goodstein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

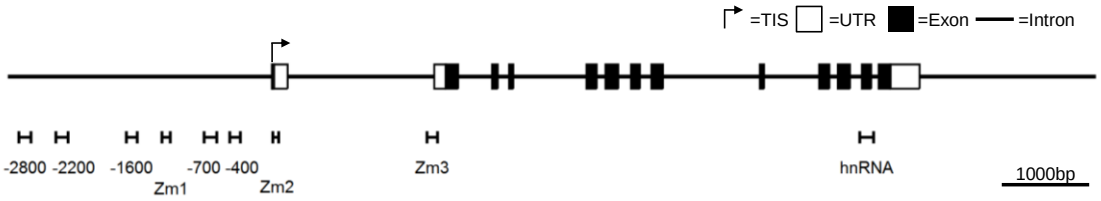
* oligonucleotide shows mismatches to database genome sequence, but efficiently amplifies DNA from the genotype used in this work

***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)



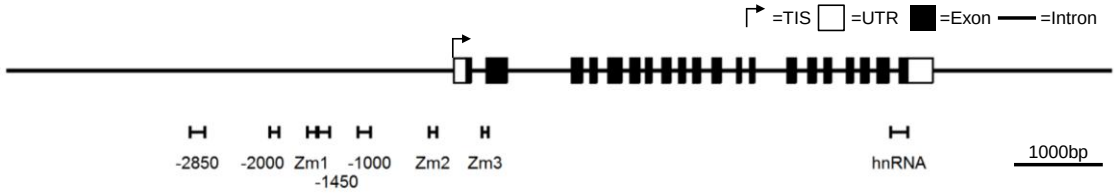
Name	bp relative to TIS	Sequence (5'→3')
	-2600	CGATCATCTCTCCAGTCAACTG
		CAAGCTCAGGGTGCTAAAATCAC
	-1900	GATTCATCAGTAGTTAGACTTAGTC
		CCTGGTAAGTCTTCATTCATAACC
	-1300	AGGGGTATTGTGAACAAGAGGATG
		CCAATTCCTCGCAAAGACACTTCAC
Zm1	-800	TGGAGGCGTTGGCTAAAGTAC
		AGAGGTAAATCAGATGACTACAAAAGAAAG
Zm2	-150	CACTATAGCCACTCGCCGCAAG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

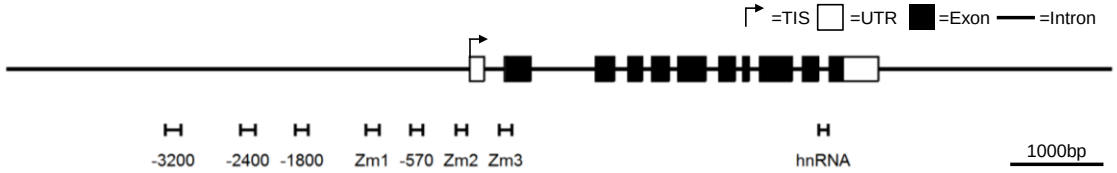
***Zea mays* – Malic Enzyme (C4-*Me*)**	
Locus	GRMZM2G085019
Transcript variant	T01
Chromosome	3
Next upstream gene	~ 13 kb
Reference/Identification	genomic locus homologous to mRNA NM_001111843 (Wang et al., 2009) was identified by whole genome BLAST on www.phytozome.net (Goodstein et al., 2012)



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 *

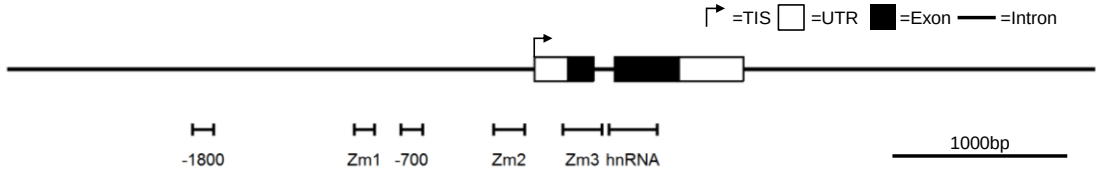
* oligonucleotide shows mismatches to database genome sequence, but efficiently amplifies DNA from the genotype used in this work

***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)

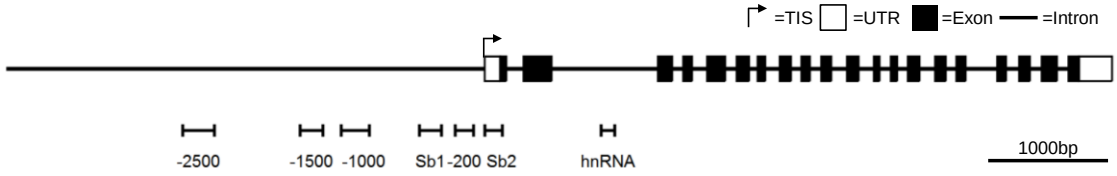
Name	bp relative to TIS	Sequence (5'→3')
	-100	TTTAAGGCTGCTGTACTGCTGTAGA
		CACTTTCTGCTCATGGTTTAAGG
mRNA	+170 **	CCTATCGTATGTGACAATGGCACT
		GCCTCATCACCTACGTAGGCAT

** 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)

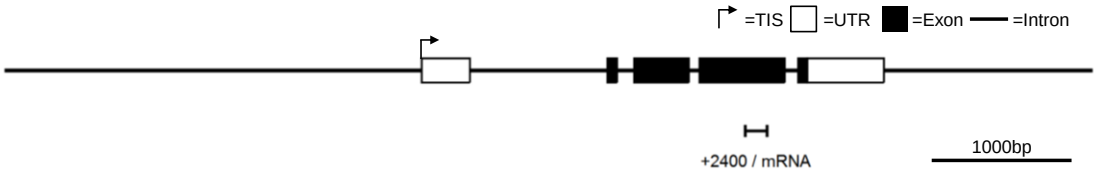
Name	bp relative to TIS	Sequence (5'→3')
	-2500	AGTTACTTCATTATCATAAATTTCTTGGCC
		GGTTTAGAAGATTTTGGCCATGAAGAC
Sb1	-1400	CGGGACATGTAATAAGGAGTTAGG
		GGTGGTGGTGAAGATATGCGG
	-700	GCCTTCCTCCAGCGCCATGCATCCTC
		CTGTTTGCAGTCAAGGCCGGATTCTGGGG
	-100	GCATGCCTTTCCAATCCCGCG
		CAATGCAGGGCGCCGGCC
Sb2	+300	GGACCTCCATGGCCCCAGCCTTCGCG
		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

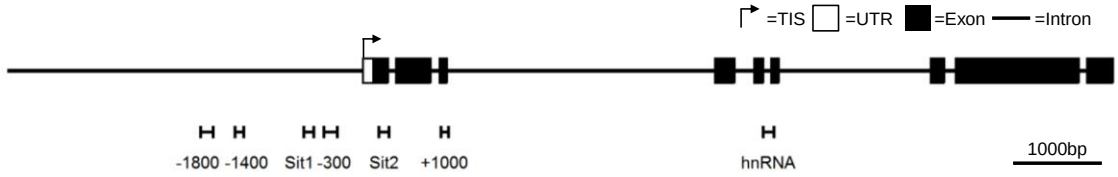
Sorghum bicolor* – *Actin	
Locus	Sb03g040880
Transcript variant	Sb03g040880.1
Chromosome	3
Next upstream gene	~ 4 kb
Reference/Identification	nearest homologue to *Actin*1 from *Zea mays* on www.phytozome.net (Goodstein et al., 2012)



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

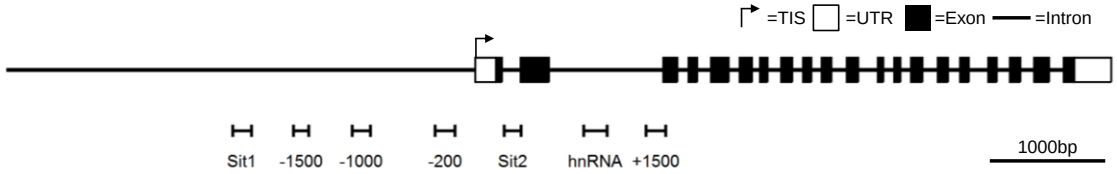
* oligonucleotide shows mismatches to database genome sequence, but efficiently amplifies DNA from the genotype used in this work

***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)



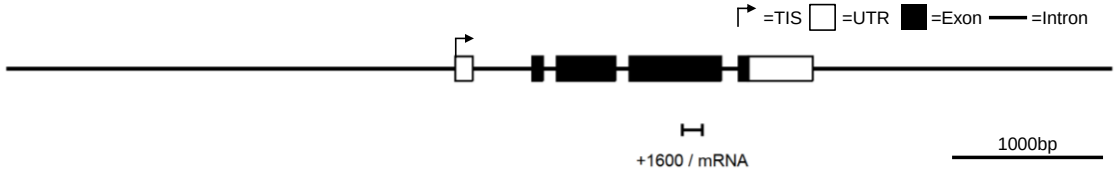
Name	bp relative to TIS	Sequence (5'→3')
	-1800	CGGAAGACAACATCATGCATGTGC
		ATGGATGCTCAGCCAGCTACC
	-1400	ATCAATGCCCTCGGCGCCAATC
		GATTATATATCGGCTCCATGTGTG
Sit1	-600	GAAAAAAACCTGGGAACAAGCC
		GTCTTCCTCTTCCCCTACCCG
	-300	GGGTATGTGGCTGGCTGTGCA
		GGACTCAAGCATGTGCTATATAGGAC
Sit2	+300	GGCAAGGTCTCCGAGGACGAC
		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

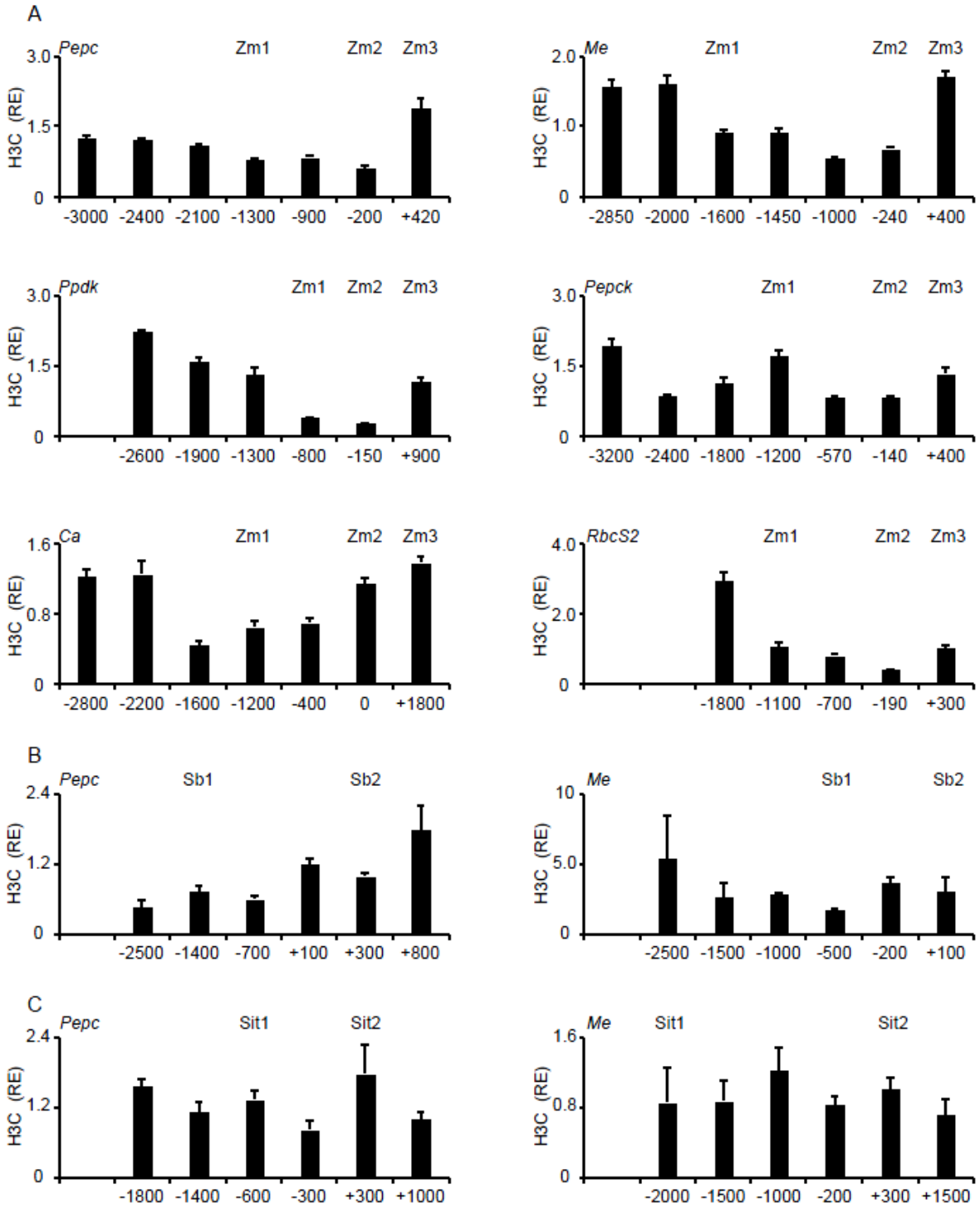
Sheen J (1991) Molecular mechanisms underlying the differential expression of maize pyruvate, orthophosphate dikinase genes. Plant Cell **3:** 225-245

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

Signal integration on plant promoters: A case study in maize

Authors

Ina Horst[1], Louisa Heimann[1] and Christoph Peterhansel*

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

[1] both authors contributed equally to this work

* corresponding author: Christoph Peterhansel, tel +49-511-7622632, fax +49-511-76219262, email: cp@botanik.uni-hannover.de

key words: histone modification, histone code, signal integration, C4 photosynthesis, *Zea mays*, *Sorghum bicolor*, *Setaria italica*

SHORT COMMUNICATION

Plant Signaling & Behavior 8:9, e25389; September 2013; © 2013 Landes Bioscience

Signal integration on plant promoters A case study in maize

Ina Horst†, Louisa Heimann†, and Christoph Peterhansel*

Leibniz University Hannover; Institute of Botany; Hannover, Germany

†These authors contributed equally to this work.

Keywords: histone modification, histone code, signal integration, C4 photosynthesis, *Zea mays*, Sorghum bicolor, *Setaria italica*

Gene promoters perceive numerous signals and integrate this information into a single response, the transcriptional activity of a gene. It was speculated that covalent modification of histones on the promoters might have an important function in storage and integration of signals. Using the genes for the core proteins of C4 metabolism in maize as a model, we associated the perception of specific signals with the establishment of individual histone modifications. Core elements of the histone code defined in these studies are conserved on all C4 genes and on other maize genes that respond to similar stimuli. Moreover, the code is used in independent C4 lineages. However, our data also advise caution because interpretation of histone modifications might differ dependent on the promoter position of the modification. The model provided here constitutes a starting point for genome-wide decoding of stimulus-modification pairs in epigenetic gene regulation.

In behavioral sciences, information integration theory describes the valuation and integration of information derived from different sources or signals into a response.[1] On a molecular level, information integration is a major function of gene promoters that translate numerous developmental or environmental inputs into a single output, the transcriptional rate. Eukaryotic genes are packed into chromatin. Research from the last two decades established chromatin as a key player in gene regulation.[2] The basic repeat unit of chromatin is the nucleosome that is made up of each two copies of the histone proteins H2A, H2B, H3 and H4 plus approximately 147 bp of DNA wound around this body.[3] Multiple amino acid residues on histone proteins can be modified in numerous ways with acetylation and methylation of the N-terminal domains of histones H3 and H4 being the best studied modifications.[4,5] Two different models for the function of histone modifications in gene regulation were proposed in the past (**Fig. 1A**): First, some histone modifications can lead to neutralization of positively charged histone tails, thereby weakening the interaction with the negatively charged DNA. This might allow better access for RNA polymerases and other transcription factors (charge neutralization model).[6,7] This model would imply that information integration takes place before chromatin is modified and that chromatin modifications would just be used to control the response function of the integrator. Alternatively, specific histone modifications are themselves recognized by transcription factors (histone code model).[8,9] In this model, histone modifications would be controlled by defined signals and used to store and integrate information on promoters. Thus, histone modification would act on the level of the integration function.

The C4 carbon concentrating mechanism in maize is an excellent system to study signal integration on chromatin, because the corresponding genes are highly transcribed and regulated in a similar manner by many different developmental and external signals. Major developmental stimuli include organ specificity (i.e., genes are only transcribed in leaves, but not in roots) and tissue specificity (i.e., within a leaf, most C4 genes are either transcribed in mesophyll cells or bundle sheath cells, but not in both tissues). Most important responses to external signals are a strong induction by light and a downregulation at low nitrogen availability or high leaf sugar levels.[10,11] In addition to the high degree of regulation of promoter activity, initial studies revealed that epigenetic factors are involved in the regulation of C4 genes.[12,13] It was shown that de-methylation of four specific cytosines in the upstream promoter region of the C4 gene encoding phosphoenolpyruvate carboxylase (C4-Pepc) occurred in a light- and tissue-specific manner. Using the same promoter as a model, we identified histone modifications associated with gene regulation (**Fig. 1B**). Among others, these analyses revealed that core promoter histone modifications were controlled by specific stimuli, e.g., acetylation of lysine 9 on histone H3 (H3K9ac) and H4K5ac were exclusively controlled by light,[14] whereas trimethylation of H3K4 (H3K4me3) potentiated the gene for activation

Correspondence to: Christoph Peterhansel; Email: cp@botanik.uni-hannover.de
Submitted: 05/16/13; Revised: 06/13/13; Accepted: 06/13/13
Citation: Horst I, Heimann L, Peterhansel C. Signal integration on plant promoters: a case study in maize. Plant Signal Behav 2013; 8:e25389
http://dx.doi.org/10.4161/psb.25389

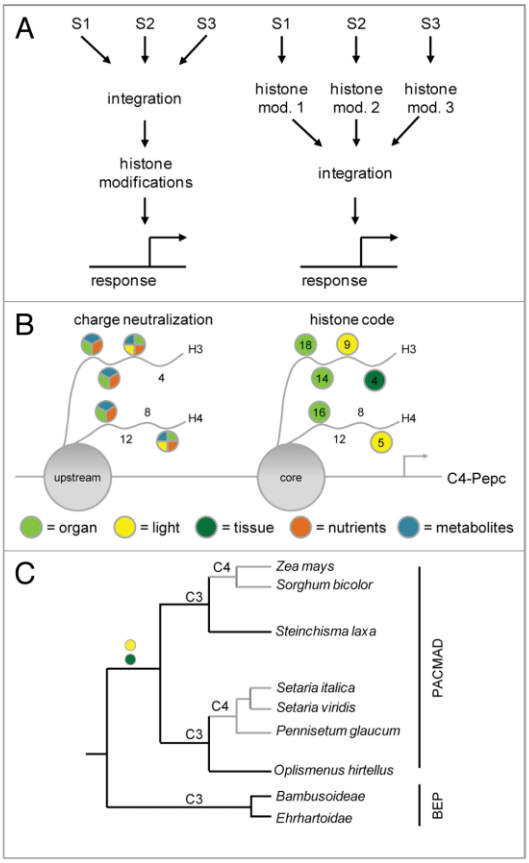


Figure 1. Histone modification models and phylogeny of the Poaceae. (**A**) Two different models for the function of histone modifications in signal integration. Charge neutralization model (left), histone code model (right). (**B**) Schematic model of the function of histone modifications in C4-Pepc gene regulation. Numbers represent lysine residues on the N-terminal tails of histones H3 and H4. The color represents as listed in the figure. Core promoter modifications follow the histone code model, upstream promoter modifications follow the charge neutralization model. (**C**) The cladogram illustrates the phylogenetic relationship of representative C4 and C3 species among the Poaceae. The PACMAD clade contains both C4 (gray branches) and C3 species (black branches), whereas the BEP clade contains only C3 species. The most recent possible phylogenetic origins of light-induced histone acetylation (yellow dot) and tissue-specific histone methylation (green dot) are indicated.

in mesophyll cells.[15] Such stimulus-modification pairs were compatible with a histone code model and argued for a function of histone modifications in signal integration. However, in more upstream promoter regions, histone modifications responded to all tested stimuli, including nitrogen availability and metabolite repression, in a similar and dose-dependent manner suggesting that histone modifications were only used to control the response function of the promoter as predicted in the charge neutralization model.14

We now analyzed whether other C4 genes in maize use the same histone code.16 Key features of the code such as light-dependent histone acetylation and tissue-specific histone methylation were highly reproducible on all genes encoding enzymes of the C4 core cycle.16 Here, we add data for the core promoters of four additional maize genes (**Fig. 2**). Glk1 encodes a kinase with possible function in mesophyll chloroplast development,17 whereas Cp24, Cp26 and Cp29 encode elements of the light harvesting complex II.18 All these genes are preferentially transcribed in mesophyll cells of illuminated maize leaves (data not shown and ref. 19). Accordingly, we found high H3K9ac only in illuminated leaves, but not in leaves from plants exposed to prolonged darkness. H3K4me3 levels were increased in mesophyll compared with bundle sheath cells, even when leaves were never illuminated. These additional results substantiate our hypothesis that a universal code is used for the control of promoters by light and tissue-specific signals in maize.

The question remains to which extent this histone code has been established during evolution of C4 metabolism or whether a previously existing code has been recruited into C4. Due to its recent evolutionary origin (approximately 25 million years), the C4 syndrome is an outstanding example for parallel evolution with more than 60 independent origins in different plant lineages.20 Recent analyses of DNA sequence elements responsible for C4-specific gene expression indicated that these elements were active in different C4 lineages21 and already found in the C3 orthologs of some of these genes.22,23 We therefore compared chromatin patterns on C4-Pepc and a second C4 gene, C4-malic enzyme (C4-Me) in maize, sorghum and Setaria italica.16 Whereas maize and sorghum share the same C4 origin, C4 photosynthesis in S. italica evolved independently (**Fig. 1C**, altered after refs. 24 and 25). All three species belong to the PACMAD clade of the Poaceae family that contains both C3 and C4 plants whereas the sister BEP clade exclusively contains C3 plants.26 The comparative chromatin analyses again revealed light induction of H3K9ac, but tissue-specific control of H3K4me3 in all three species.16 Thus, the two core features of the maize C4 histone code were retrievable in independent C4 lineages. These results indicate that elements of the histone code had been recruited into C4 from a preexisting mechanism. The most recent possible phylogenetic origin of this mechanism was after separation of the PACMAD and BEP clades (yellow and green dots in **Fig. 1C**). Further comparative analyses will show whether the origin can be dated back to even earlier time points.

In conclusion, data from previous work and the chromatin analyses on additional genes added here point to an important

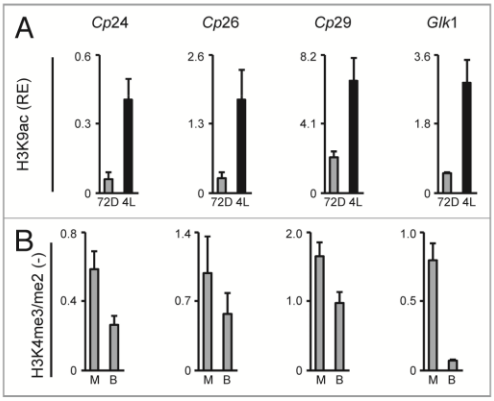


Figure 2. Light regulation of histone acetylation and cell-type specific histone methylation on four maize genes.(**A**) Light-dependent acetylation of histone H3 lysine 9 (H3K9ac) in leaves from plants that were exposed to 72 h darkness (72D, gray columns) and from plants that were illuminated for 4 h (4L, black columns). Values are presented as the relative enrichment (RE) of modifications per nucleosome over modifications per nucleosome found on the Actin1 promoter. (**B**) Ratio of the histone H3 lysine 4 trimethylation (H3K4me3) and histone H3 lysine 4 dimethylation (H3K4me2) in mesophyll (M) or bundle sheath (**B**) cells isolated form etiolated leaves. Cp24, Cp26 and Cp29 encode components of light harvesting complex II, Glk1 a kinase involved in mesophyll chloroplast development. All data points are based on at least four independent experiments. Vertical lines indicate standard errors.

role of histone modifications in the integration function of plant promoters. The histone code used to display the perception of specific stimuli seems to be highly conserved. Dependent on the modification and the position on the promoter, histone modifications might in addition help to implement the response function of promoters.

Disclosure of Potential Conflicts of Interest

No potential conflicts of interest were disclosed.

Acknowledgments

This work was supported by grant PE819/1-3 from the Deutsche Forschungsgemeinschaft to CP

References

1. Anderson NH. Integration theory and attitude change. Psychol Rev 1971; 78:171-206; http://dx.doi.org/10.1037/h0030834
2. Li B, Carey M, Workman JL. The role of chromatin during transcription. Cell 2007; 128:707-19; PMID:17320508; http://dx.doi.org/10.1016/j.cell.2007.01.015
3. Kouzarides T. Chromatin modifications and their function. Cell 2007; 128:693-705; PMID:17320507; http://dx.doi.org/10.1016/j.cell.2007.02.005
4. Chen ZJ, Tian L. Roles of dynamic and reversible histone acetylation in plant development and polyploidy. Biochim Biophys Acta 2007; 1769:295-307; PMID:17556080
5. Luo M, Liu X, Singh P, Cui Y, Zimmerli L, Wu K. Chromatin modifications and remodeling in plant abiotic stress responses. Biochim Biophys Acta 2012; 1819:129-36; PMID:21708299
6. Dion MF, Altschuler SJ, Wu LF, Rando OJ. Genomic characterization reveals a simple histone H4 acetylation code. Proc Natl Acad Sci USA 2005; 102:5501-6; PMID:15795371; http://dx.doi.org/10.1073/pnas.0500136102
7. Henikoff S, Shilatifard A. Histone modification: cause or cog? Trends Genet 2011; 27:389-96; PMID:21764166; http://dx.doi.org/10.1016/j.tig.2011.06.006
8. Berger SL. The complex language of chromatin regulation during transcription. Nature 2007; 447:407-12; PMID:17522673; http://dx.doi.org/10.1038/nature05915
9. Jenuwein T, Allis CD. Translating the histone code. Science 2001; 293:1074-80; PMID:11498575; http://dx.doi.org/10.1126/science.1063127
10. Hibberd JM, Covshoff S. The regulation of gene expression required for C4 photosynthesis. Annu Rev Plant Biol 2010; 61:181-207; PMID:20192753; http://dx.doi.org/10.1146/annurev-arplant-042809-112238
11. Sheen J. C4 gene expression. Annu Rev Plant Physiol Plant Mol Biol 1999; 50:187-217; PMID:15012208; http://dx.doi.org/10.1146/annurev.arplant.50.1.187
12. Langdale JA, Taylor WC, Nelson T. Cell-specific accumulation of maize phosphoenolpyruvate carboxylase is correlated with demethylation at a specific site greater than 3 kb upstream of the gene. Mol Gen Genet 1991; 225:49-55; PMID:2000091; http://dx.doi.org/10.1007/BF00282641
13. Tolley BJ, Woodfield H, Wanchana S, Bruskiewich R, Hibberd JM. Light-regulated and cell-specific methylation of the maize PEPC promoter. J Exp Bot 2012; 63:1381-90; PMID:22143916; http://dx.doi.org/10.1093/jxb/err367
14. Offermann S, Dreesen B, Horst I, Danker T, Jaskiewicz M, Peterhansel C. Developmental and environmental signals induce distinct histone acetylation profiles on distal and proximal promoter elements of the C4-Pepc gene in maize. Genetics 2008; 179:1891-901; PMID:18689888; http://dx.doi.org/10.1534/genetics.108.087411
15. Danker T, Dreesen B, Offermann S, Horst I, Peterhänsel C. Developmental information but not promoter activity controls the methylation state of histone H3 lysine 4 on two photosynthetic genes in maize. Plant J 2008; 53:465-74; PMID:18179650; http://dx.doi.org/10.1111/j.1365-313X.2007.03352.x
16. Heimann L, Horst I, Perduns R, Dreesen B, Offermann S, Peterhansel C. A common histone modification code on C4 genes in maize and Its conservation in sorghum and Setaria italica. Plant Physiol 2013; 162:456-69; PMID:23564230; http://dx.doi.org/10.1104/pp.113.216721
17. Wang P, Fouracre J, Kelly S, Karki S, Gowik U, Aubry S, et al. Evolution of GOLDEN2-LIKE gene function in C(3) and C (4) plants. Planta 2013; 237:481-95; PMID:22968911; http://dx.doi.org/10.1007/s00425-012-1754-3
18. Kovács L, Damkjaer J, Kereïche S, Ilioaia C, Ruban AV, Boekema EJ, et al. Lack of the light-harvesting complex CP24 affects the structure and function of the grana membranes of higher plant chloroplasts. Plant Cell 2006; 18:3106-20; PMID:17114352; http://dx.doi.org/10.1105/tpc.106.045641
19. Sawers RJ, Liu P, Anufrikova K, Hwang JT, Brutnell TP. A multi-treatment experimental system to examine photosynthetic differentiation in the maize leaf. BMC Genomics 2007; 8:12; PMID:17212830; http://dx.doi.org/10.1186/1471-2164-8-12
20. Sage RF, Sage TL, Kocacinar F. Photorespiration and the evolution of C4 photosynthesis. Annu Rev Plant Biol 2012; 63:19-47; PMID:22404472; http://dx.doi.org/10.1146/annurev-arplant-042811-105511
21. Brown NJ, Newell CA, Stanley S, Chen JE, Perrin AJ, Kajala K, et al. Independent and parallel recruitment of preexisting mechanisms underlying C_4 photosynthesis. Science 2011; 331:1436-9; PMID:21415351; http://dx.doi.org/10.1126/science.1201248

22. Brown NJ, Palmer BG, Stanley S, Hajaji H, Janacek SH, Astley HM, et al. 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 2010; 61:122-33; PMID:19807880; http://dx.doi.org/10.1111/j.1365-313X.2009.04040.x
23. Kajala K, Brown NJ, Williams BP, Borrill P, Taylor LE, Hibberd JM. Multiple Arabidopsis genes primed for recruitment into C_4 photosynthesis. Plant J 2012; 69:47-56; PMID:21883556; http://dx.doi.org/10.1111/j.1365-313X.2011.04769.x
24. Brutnell TP, Wang L, Swartwood K, Goldschmidt A, Jackson D, Zhu XG, et al. Setaria viridis: a model for C4 photosynthesis. Plant Cell 2010; 22:2537-44; PMID:20693355; http://dx.doi.org/10.1105/tpc.110.075309
25. Ibrahim DG, Burke T, Ripley BS, Osborne CP. A molecular phylogeny of the genus Alloteropsis (Panicoideae, Poaceae) suggests an evolutionary reversion from C4 to C3 photosynthesis. Ann Bot 2008; 103:127-36; PMID:18974099; http://dx.doi.org/10.1093/aob/mcn204
26. Buell CR. Poaceae genomes: going from unattainable to becoming a model clade for comparative plant genomics. Plant Physiol 2009; 149:111-6; PMID:19005087; http://dx.doi.org/10.1104/pp.108.128926

CHAPTER 4

Histone modification profiling on the C4- phosphoenolpyruvate carboxylase promoter in *Zea mays* reveals *a* light and tissue dependent regulation of so far non-characterized histone H3 modifications

Authors

Louisa Heimann and Christoph Peterhansel*

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

* corresponding author: Christoph Peterhansel, tel +49-511-7622632, fax +49-511-76219262, email: cp@botanik.uni-hannover.de

key words: histone modification, histone code, signal integration, crosstalk, C4 photosynthesis, *Zea mays*

Abstract

Recent studies revealed the importance of certain histone modifications and the co-regulation of these modifications. Here, we were able to show that the reversible light induced expression of the C4-phosphoenolpyruvate carboxylase (C4-*Pepc*) locus is accompanied by several changes in so far non-characterized histone H3 modifications (H3K4ac, H3K4me3, H3K27ac, H3K27me3, H3K36ac, H3K36me3) and that each modification, showed a gene-specific distribution of modifications. We found that C4-*Pepc* activation during light was accompanied by enrichment in H3K4ac, H3K27ac and H3K36ac in the promoter region whereas the tri-methylation of these lysines was mainly found at the start of the coding region and more downstream. We further extended our investigations to the cell-type specific response of these modifications of C4-*Pepc* and additionally of a second gene, malic enzyme (C4-*Me*). These studies revealed that all tested methylation did not respond to the light stimulus but were rather regulated in a cell-type specific manner and that H3K27me3 presumably acts as ‘off’-marker for C4-*Pepc* in bundle sheath cells The presence of activating and repressive histone marks suggests a mechanism for the rapid and reversible regulation of C4-*Pepc* by dark and light, and a specific code for cell-type specific expression.

Introduction

Light (L) is an important environmental factor that impacts on plant development. A group of genes that are strongly regulated by light are the genes that encode proteins of C4 photosynthesis such as the phosphoenolpyruvate carboxylase (C4-*Pepc*) or malic enzyme (C4-*Me*) gene (Sheen, 1999; Kausch et al., 2001). C4-*Pepc* is a key enzyme of C4 photosynthesis because it catalyzes primary CO_2 fixation. In addition to light dependent induction it is also regulated by the nutrient availability and the metabolic state of the cell (Sheen, 1999). C4-*Me* is involved in the carbon concentrating mechanism near Rubisco. It catalyzes the oxidation reaction of malate to pyruvate and CO_2 and is also strongly regulated by light (Gowik and Westhoff, 2011).

When C4 plants are germinated in darkness or exposed to prolonged darkness, leaves turn to be yellowish because of the missing chlorophyll (Yoshida et al., 2001). In parallel to the L-dependent morphological changes, changes in gene expression could be observed (Jiao and Meyerowitz, 2010). These changes in gene expression have been frequently correlated with changes in specific histone marks. The pattern and distribution of histone modifications are highly complex, because of the high number of residues that potentially can be modified and the multiple combinatorial modifications (Zhang, 2008). Some modifications directly alter chromatin structure, whereas others serve as binding platforms to recruit additional effectors. The most prominent modifications to date are acetylation and methylation, which in combination can modulate chromatin conformation (Barrand et al., 2010).

C4-*Pepc* is an excellent model for studying the function of different histone modifications. Recent studies form our lab indicated that C4-*Pepc* has an extend promoter and that chromatin structure and histone modifications are important for the regulation of this model gene (Offermann et al., 2006; Offermann et al., 2008).

Additionally, there are reports among the literature where histone modifications are involved in L-dependent gene expression. An important example and well studied modification is histone H3 lysine 4 tri-methylation (H3K4me3). H3K4me3 is generally understood as ‚on'-marker for transcription and is usually found at the start of the coding region of genes (Heintzman et al., 2007; Pokholok et al. 2005). Jang et al. (2011) recently showed that acetylation of H3K9 and H3K27 as well as H3K4me3 enrichment is correlated with the activation of the *Phytochrome A* locus during deetiolation of *Arabidopsis thaliana* seedlings. In comparison to methylation, acetylation of histone H3 (H3K4ac) is not very well studied. Nevertheless, Guillemette et al. (2011) were able to show that H3K4ac is enriched on actively transcribed genes and can often be found upstream of H3K4me3.

Light is not the only stimulus that strongly effects gene expression of C4 genes. As a result of the spatial separation of the C4 cycle in mesophyll (M) and bundle sheath (B) cells, there needs to be a certain code that leads to M or B specific expression. We previously reported that trimethylation of H3K4 (H3K4me3) marks the C4-*Pepc* and C4-*Me* gene for activation in the corresponding tissue (Danker et al., 2008). In contrast to H3K4me3 the H3K27me3 was found to be associated in Arabidopsis with euchromatic regions of silenced genes (Barrand et al., 2010). Additionally it has previously been shown by Zhang et al. (2008) that H3K27me3 can primarily be found on genes which are silenced in a tissue-depended manner. Acetylation of the same lysine on histone H3 (H3K27ac) instead is again widely understood as an 'on'-marker (Creyghton et al., 2010). Another important histone modification is H3K36me3. Barrand et al. (2010) has shown that H3K36me3 can be found in

the coding region of actively transcribed genes and also up to 6 kbp downstream the transcription initiation site (TIS). In contrast H3K36ac is mainly found in promoters of transcribed genes (Morris et al., 2006). Additionally it has been supposed that H3K36ac plays a role in the so called ‘exon-marking’ and is found on the borders between exons and introns to link epigenetic information, transcription and splicing together (Barrand et al., 2010).

Modifications may act alone or influence each other to open or repress chromatin. This communication across modifications is called ‘crosstalk’ (Fischle et al., 2003; Suganuma and Workman, 2008). Over the years several examples showed the importance of crosstalk. A very well studied example is the COMPASS complex where H2B monoubiquitination is needed to trigger H3K4 methylation and H3 lysine 79 (H3K79) methylation (Suganuma and Workman, 2008).

In this study we wanted to understand global chromatin changes associated with the light/dark transition (L/D transition) on C4-*Pepc* and in accordance to cell-type specificity on C4-*Pepc* and C4-*Me* in maize.

Results

We used chromatin immunoprecipitation from illuminated leaves harvested 4h after onset of illumination (4L plants) and plants that were exposed to prolonged darkness (72D plants) to identify global chromatin changes associated with the L/D transition on C4-*Pepc*. We compared the changes of 6 histone marks over 7 regions for D/L transition a of the C4-*Pepc* locus (see Figure 0-1).

These regions encompass the upstream and core promoter regions, as well as several regions within the coding region of the gene. Next to light induced methylation of H3K4me3, antibodies specific for methylation and acetylation on histone H3 (H3K4ac, H3K27ac, H3K27me3 H3K36ac and H3K36me3) were investigated to create a complete picture of the changes in histone modifications due to the light stimulus and the cell type specificity. Since we already investigated the distribution of H3K4me3 on the C4-*Pepc* gene we used this antibody as internal control to ensure the quality of material preparation, especially the B cells preparation (Danker et al., 2008). As internal control for repressed chromatin the highly conserved TY1 class *copia* LTR retrotransposon was used. The expression of a *copia*-like transposon (Copia) was shown to be repressed in rice (Liu et al., 2004), and real-time PCR experiments suggested that it is repressed in maize as well (Haring et al., 2007).

Normally precipitation efficiencies for the promoter of the housekeeping gene *Actin*-1 or *Gapdh* (Glyceraldehyde 3-phosphate dehydrogenase) were used to correct the data for possible variations in the quality of chromatin preparations from different samples (Haring et al., 2007). Since the levels of *Actin*-1 or *Gapdh* were not always comparable between 4L and 72D plants for every tested antibody we refrained from correcting measured data for a housekeeping gene.

Figure 0-1 A shows a model of the C4-*Pepc* gene with all measured positions indicated as black bars under the gene structure. To detect nucleosome occupancy on the gene, we precipitated chromatin with an antibody directed to the invariant domain of the C-term of histone H3 (H3C) (Figure 0-1 A). Nucleosome occupancy did not change significantly upon illumination. In both 72D and 4L plants, the most obvious characteristic was a 2-fold increase at the start of the transcribed region, as previously shown by Offermann et al. (2006). The left panel shows the analyzed methylations and the right panel analyzed acetylations. H3K4me3 peaked at the start of the coding sequence and was weak more upstream of the TIS, but did not show significant differences between L and D plants as described before (Danker et al., 2008). H3K4me3 of C4-*Pepc* showed significantly higher modifications on position (-200 bp to 4300 bp) than Copia, which is in accordance with previously shown data (Danker et al., 2008) . In comparison to that, H3K27me3 only showed a higher signal than Copia on Positions +500 bp and +4300 bp. H3K36me3 signals were enriched mainly in the coding region (+500 bp and +4300 bp) but also in the core promoter region. Signal intensities were always lower or in the range of Copia. It is striking that almost no modification was observed with all tested antibodies on position +1900 bp which was the only region investigated, covering mainly an exon. No significant changes between the D/L samples were observed with all tested methylation antibodies (Figure 0-1 B-G).

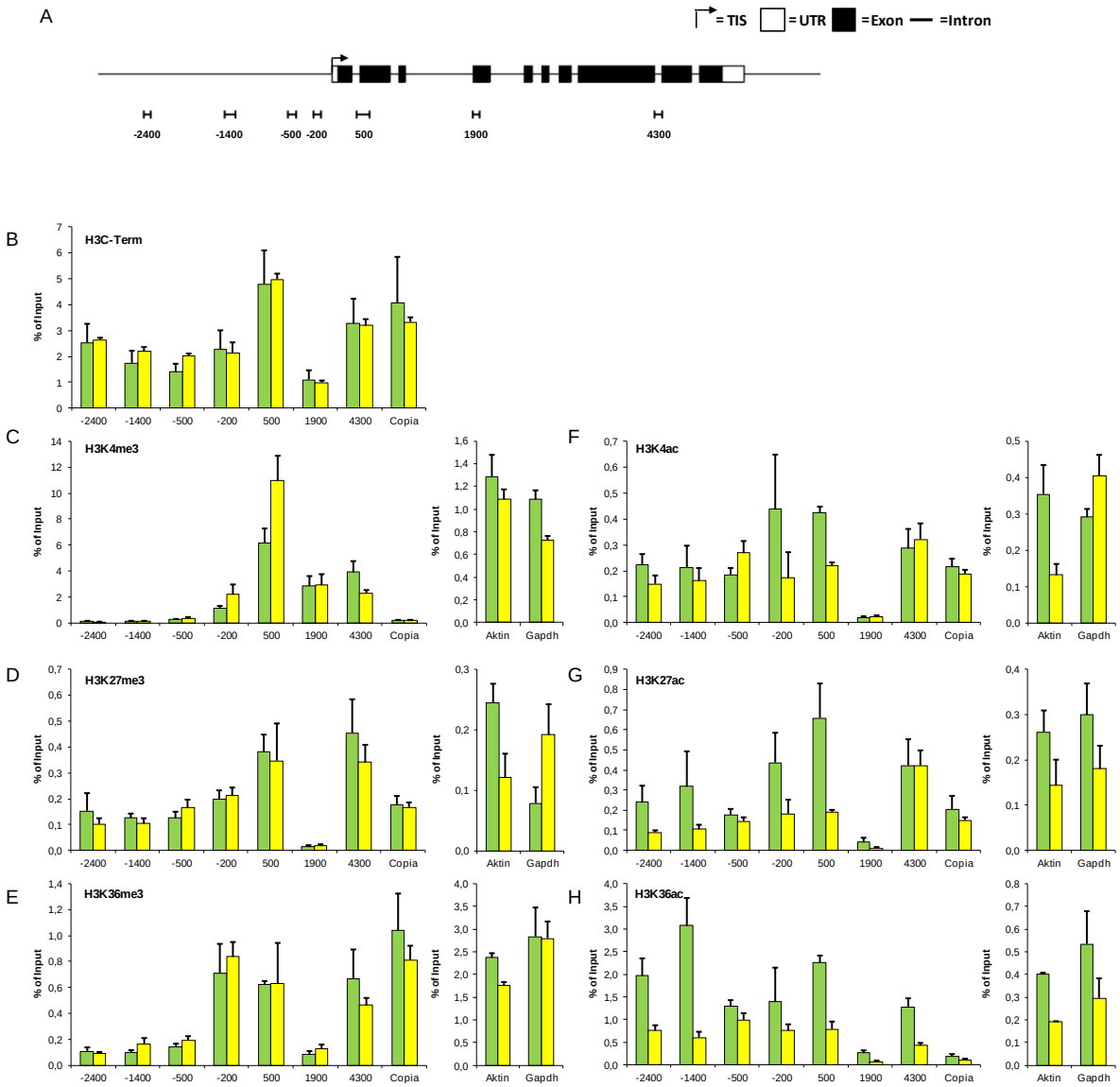


Figure 0-1: Distribution of the histone acetylation and tri-methylation over the C4-*Pepc* promoter and coding region in 4L and 72D plants.

Presented are the acetylation and tri-methylation levels of plants harvested 4 hours after illumination (green bars) and plants that were harvested after prolonged darkness (yellow bars) over the promoter and coding region of the C4-*Pepc* gene. A, model of the C4-*Pepc* gene with all measured positions indicated with black bars; Exons, 5' UTR and 3'UTR presented as black and white boxes, respectively. The arrow indicates the transcription initiation site. B, shows the amount precipitated with an invariant domain of the C-terminal part of histone H3C, C, shows the H3K4 tri-methylation; D, H3K27 tri-methylation; E, H3K36 tri-methylation; F, H3K4 acetylation; G, H3K27 acetylation; H, H3K36 acetylation.

A 2-fold increase of acetylation was detected for H3K4ac in the core promoter and at the start of the coding region (-200 bp, +500 bp), but only a slight increase could be detected on the upstream promoter (-2400 bp) but all upstream positions anyways peaked in the range of the Copia signal. An enrichment of acetylation in 4L plants in the core promoter and coding region was also found for H3K27ac and H3K36ac. But in contrast to H3K4ac the increase of acetylation covered the whole promoter region.

For H3K27ac the enrichment was always 2- to 3-fold, except on position 500 bp and the highest peak was observed in the core promoter region and at the start of the coding region. The distribution of H3K36ac differs from H3K4ac and H3K27ac. The highest acetylation peak was found in the far

upstream promoter region (-1400 bp) with a 4-fold increase in 4L plants compared to 72D plants. Also acetylation sites seemed to be regulated in the coding region of the gene and all signals were significantly higher than on Copia. The methylation state of C4-*Pepc* is in contrast to the general global increase of acetylation during the L period. Figure 0-1 B-G also reveals that light activation was clearly associated with enrichment of acetylation of all tested modifications, mainly in the promoter region.

In a second experiment the cell-type specificity of certain histone modification were tested between leaves an B cells on the M cell specific gene (C4-*Pepc*) over 6 regions and the B cell specific gene (C4-*Me*) over 5 regions. Since the upstream promoter position -2400 bp showed only signals under background in B cells we refrained to show this position.

Figure 0-2 and Figure 4-3 illustrate the distribution of the histone modifications over the C4-*Pepc* and C4-*Me* genes in 4L samples and B cells derived from 4L plants. Figure 4-2 shows the distribution of the histone acetylation and tri-methylation over the C4-*Pepc* gene in 4L and B cells. The left panel shows again the analyzed methylations and the right panel the analyzed acetylations. Again precipitation with an antibody directed to an invariant domain of the C-terminal part of histone H3C was measured (Figure 4-2 A) but the nucleosomes occupancy differed a lot between 4L samples and B cells.

Figure 4-2 B shows that H3K4me3 is clearly enriched in 4L plants compared to B cells. In both samples the modification peaked at the start of the coding sequence and was weak more upstream of the TIS. H3K27me3 was found to be 2-3-fold enriched within the gene (+500 bp and +4300 bp) in B cells, compared to 4L plants and H3K36me3 signals were enriched mainly in the coding region (+500 bp and +4300 bp) but also in the core promoter region in 4L plants. In B cells the same distribution was observable with the exception that H3K36me3 showed no peak on position +4300 bp. In contrast to the methylation, H3K4ac peaked as already observed in the first experiment in the promoter region, but was hardly detectable in B cells. H3K27ac on the contrary showed comparable high acetylation levels in both samples and peaked mainly in the core promoter (-200 bp) and upstream promoter (-1400 bp) as well as in the end of the gene (+4300 bp). H3K36ac level were mainly enriched in the promoter region. Whereas 4L plants showed a continuously increase towards the upstream promoter, the acetylation levels in B cells do not change.

Figure 4-3 shows the distribution of the histone acetylation and tri-methylation over the C4-*Me* gene in 4L and bundle sheath cells. The left panel shows again the analyzed methylations and the right panel analyzed acetylations. Again precipitation with an antibody directed to an invariant domain of the C-terminal part of histone H3C was measured (Figure 4-2 A) but the nucleosomes occupancy differed a lot between 4L samples and B cells.

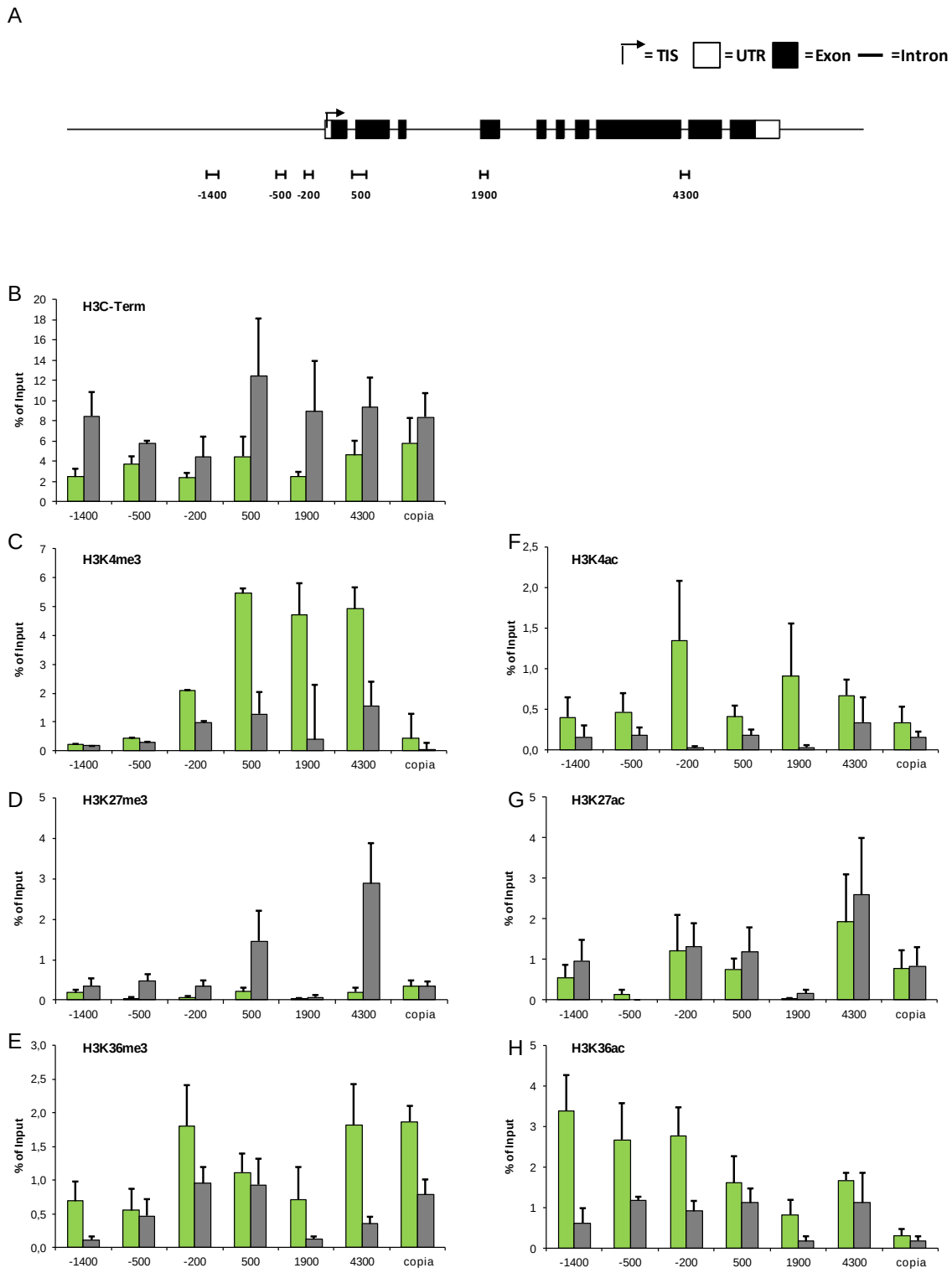


Figure 0-2: Distribution of the histone acetylation and tri-methylation over the C4-*Pepc* promoter and coding region in 4L plants and bundle sheath cells.

Presented are the acetylation and tri-methylation levels of plants harvested 4 hours after illumination (green bars) and from bundle sheath cells (grey bars) over the promoter and coding region of the C4-*Pepc* gene. A, model of the C4-*Pepc* gene with all measured positions indicated with black bars; Exons, 5' UTR and 3'UTR presented as black and white boxes, respectively. The arrow indicates the transcription initiation site. B, shows the amount precipitated with an invariant domain of the C-terminal part of histone H3C, C, shows the H3K4 tri-methylation; D, H3K27 tri-methylation; E, H3K36 tri-methylation; F, H3K4 acetylation; G, H3K27 acetylation; H, H3K36 acetylation.

The most obvious characteristic was that H3K4me3 levels of C4-*Me* were comparably high in B cells, compared to the 4L samples, except on position -1400 bp. Comparable to the distribution on the C4-*Pepc* gene; highest H3K4me3 was found at the start of the coding region and was weak more upstream. On the contrary H3K27me3 peaked in the promoter region and at the start of the coding region in 4L plants (-1400 bp to +500 bp), whereas in B cells the highest peaks were observed in the upstream promoter region (-1400 bp) and at the end of the gene (+1900 bp and +4300 bp). H3K36me3 showed a comparable distribution like H3K4me3 on C4-*Me.* Highest H3K36me3 levels were detected within the gene for 4L plants (+500 bp and +1900 bp) and on positions -200 bp and +500 bp for B cells. Acetylation levels were found to be more abundant in the promoter region for H3K4ac. The highest peak was observed on position -1400 bp. From there the acetylation continuously decreases towards the end of the gene. In contrast to 4L plants H3K4ac was always 2-4 fold lower in B cells.

H3K27ac showed again comparable high acetylation levels in 4L plants and B cells. Highest peaks were observed on position -1400 bp and at the start of the coding region +500 bp. Also H3K36ac levels showed comparable high acetylation levels in 4L plants and B cells in the promoter region. 4L plants also peaked at the start of the coding region and in both samples acetylation levels decreases towards the end of the gene.

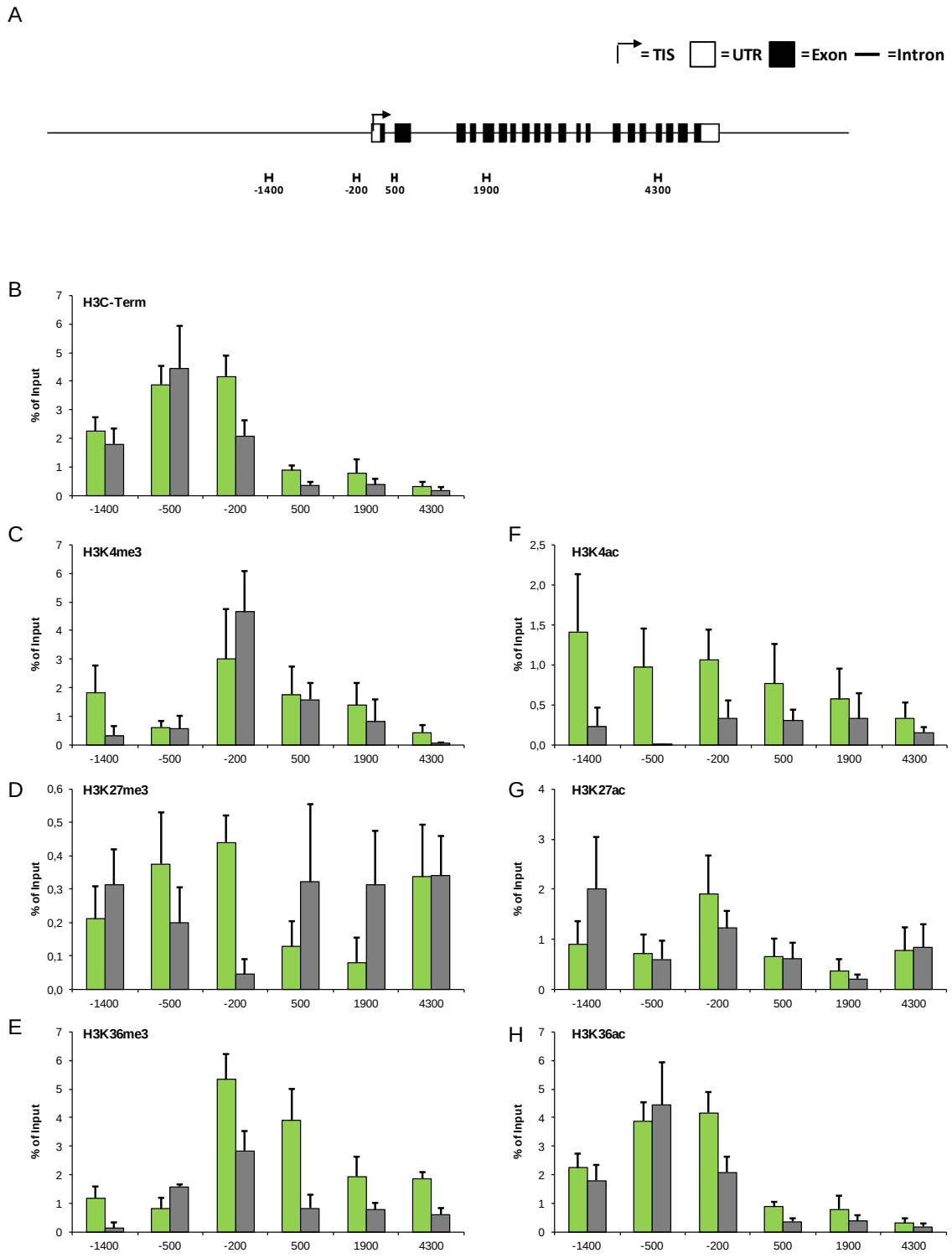


Figure 0-3: Distribution of the histone acetylation and tri-methylation over the C4-*Me* promoter and coding region in 4L plants and bundle sheath cells.

Presented are the acetylation and tri-methylation levels of plants harvested 4 hours after illumination (green bars) and from bundle sheath cells (grey bars) over the promoter and coding region of the C4-*Me* gene. A, model of the C4-*Me* gene with all measured positions indicated with black bars; Exons, 5' UTR and 3'UTR presented as black and white boxes, respectively. The arrow indicates the transcription initiation site. B, shows the amount precipitated with an invariant domain of the C-terminal part of histone H3C, C, shows the H3K4 tri-methylation; D, H3K27 tri-methylation; E, H3K36 tri-methylation; F, H3K4 acetylation; G, H3K27 acetylation; H, H3K36 acetylation.

Discussion

In this study we have focused on investigating chromatin changes on the C4-*Pepc* and C4-*Me* locus for two reasons: firstly, both genes are activated by light and expressed in cell-type specific manner; and secondly, C4-*Pepc* and C4-*Me* are well investigated loci which allow a more detailed investigations of changes in histone modifications along different genomic and promoter regions.

We were able to show that the reversible light induced expression of the C4-*Pepc* locus is accompanied by several changes in so far non-characterized histone H3 modifications and that each modification, showed a gene-specific distribution of modifications. We chose to study three pairs of modifications, due to the assumption, that on positions were methylation can be detected no acetylation would occur.

We found that acetylation of all tested modifications of C4-*Pepc* peaked in the promoter region, whereas methylation was mainly found at the start of the coding region and more downstream of the gene. Additionally this data revealed that all tested methylations are not influenced by the light stimulus. We observed a striking gap in x and y at position 1900 on the C4-Pepc gene (see Figure 4-1). However, this was not reproducible in the 2nd experiment (see fig 4-2) and therefore not taken into consideration for data analysis. These findings are in accordance to the literature. Jang et al. (2011) were able to show in *Arabidopsis thaliana* seedlings that acetylation of H3K9 and H3K27 as well as H3K4me3 enrichment correlated with gene activation during deetiolation on the *Phytochrome A* locus. Additionally using genome-wide chromatin Immunoprecipitation, Guillemette et al. (2011), were able to show that H3K4ac was enriched at promoters of actively transcribed genes and located upstream of H3K4me3. This pattern was found to be conserved in human cells and it is presumably that these modifications crosstalk. H3K36me3 was shown to be associated with the transcription of active genes with the distinct pattern that H3K36me3 increases towards the 3' end of genes. Down-regulation experiments of the H3K36 methyltransferase SetD2 revealed global and gene-specific H3K36 demethylation as well as global H3K27 hypermethylation. It was also shown that H3K27me3 levels on promoters were not affected (Wagner and Carpenter, 2012; Barrand et al., 2010). Another evidence for a crosstalk was given by (Nightingale et al., 2006). They were able to show with the help of mass spectrometry analysis that H3K4me3 was highly associated with high acetylations on histone H3 (K9, K14, K18, K23 and K27). Since the H3K4 acetylation levels for C4-*Pepc* only reaches the Copia detection limit on two positions in both experiments, it is just speculation whether a crosstalk can be observed here too. H3K4ac is enriched in the core promoter region on 4L plants whereas H3K4me3 is enriched at the coding region and this pattern was reproducible in both experiments (Figure 0-1 B and C, Figure 0-2 B and C). A clearer picture was received for C4-*Me* in 4L plants where a clear increase of H3K4ac in 4L plants could be observed on the promoter in comparison to Copia (see Figure 0-3 B and C).

A crosstalk between H3K4me3 and H3K27ac can only be assumed for B cells on C4-*Pepc* because H3K4me3 peaked in the core promoter and at the start of the coding region, while H3K27ac was highly enriched within the gene. Further a crosstalk of H3K4me3 and H3K27ac could also be assumed for C4-*Me* B cells. Whereas in B cells the highest H3K27me3 peak was observed within the upstream promoter region, it was found for H3K4me3 at the start of the coding region. It has recently be shown by Zhang et al. (2007) that H3K27me3 is mainly found on genes associated with tissue-dependent

transcription. In addition to that it was recently reported that modifications like H3K27ac and H3K9ac are correlated with the activation of photosynthetic genes.

Further the results from Charron et al. (2009) and Jang et al. (2011) show that H3K27me3 and H3K27ac are regulated in an inverse pattern. We could not observe an inverse regulation in these experiments , but our results indeed support the theory that H3K27me3 as an 'off'-marker is associated with tissue-dependent transcription, given the fact, that C4-*Pepc* is a M cell specific gene and high H3K27me3 levels were found on this gene in B cells, where it should be inactive (see Figure 0-2 D).

On the contrary signal intensities of H3K27me3 on C4-*Me* are very low and hardly reached the detection limit, whereas H3K27ac are enriched on the upstream promoter and at the start of the coding region. The finding that the 'off'-marker H3K27me3 is low on C4-*Me* in B cells is in accordance to a tissue-dependent transcription, because the gene should be active in B cells.

H3K36ac is mainly be found in promoters of Pol II transcribed genes from yeast (Morris et al., 2006), an inverse position pattern in comparison to H3K36me3. The distribution of H3K36ac correlates with patterns found for other histone H3 acetylation sites like H3K9ac and H3K14ac (Barrand et al., 2010; Krogan et al., 2003). For C4-*Pepc* H3K36me3 never reaches the detection limit of Copia in all tested samples. So we can assume that C4-*Pepc* is not regulated by this modification. C4-*Me* instead showed enriched H3K36me3 signals in 4L plants as well as in B cells in the coding region of the gene. These findings are in agreement with Barrand et al. (2010) and Krogan et al. (2003), that modification for 4L plants were found at the beginning of the coding region and further downstream. Additionally it was assumed that H3K36me3 plays a role in so called exon-marking and that this exon marking links chromatin modifications and mRNA processing. It was proposed and shown that H3K36me3 associated with the elongating form of Pol II and that the absence of H3K36me3 affects productive elongation, resulting in RNAPII stalling at the 3' end of genes. This assumption could not be confirmed, at least not for C4-*Pepc* but we observed a general increase in H3K36me3 occupancy on C4-*Me* exons (+500bp, +1900bp, +4300 bp).

We were able to detect so far not described activating and repressive histone modifications on C4-*Pepc* and C4-*Me*. Since the maize genome encodes a number of histone acetyltransferases/deacetylases and methyltransferases/demethylases, it further needs to be investigated which role they play in global chromatin changes and therefore for the gene expression in response to D/L and cell-type specificity.

Material and Methods

Plant material and growth conditions

Maize (*Zea mays* cv. Montello), *Sorghum* (*Sorghum bicolor* BTx623) and *Setaria* (*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, Sinntal-Jossa, Germany), with a photon flux density of 120-180 µmol m^{-2} s^{-1} until the third leaf was fully expanded. 72D plants were grown in the normal light rhythm, but darkened for three days before harvest.

Chromatin immunoprecipitation (ChIP)

As described previously by Horst et al. (2009), 6 g leaves from 10- to 12-d-old maize seedlings were harvested and crosslinked. Chromatin immunoprecipitation was performed as described by Haring et al. (2007). The material was ground, resuspended in extraction buffer (10 mM Na-Butyrate, 400 mM sucrose, 10 mM Tris-HCl, pH 8.0, 5 mM beta-mercaptoethanol, 0.1 mM phenylmethylsulfonyl fluoride (PMSF) and 1x Complete (Roche Applied Science, Mannheim, Germany)) and incubated for 15 min at 4°C. Afterwards, the solution was filtered through 4 layers of Miracloth (VWR, Darmstadt, Germany) and the residue was washed with purification buffer 1 (10 mM Na-Butyrate, 250 mM sucrose, 10 mM Tris-HCl, pH 8.0, 5 mM beta-mercaptoethanol, 0.1 mM PMSF, 10 mM $MgCl_2$, 1% w/v Triton X-100 and 1x Complete) and afterwards with purification buffer 2 (10 mM Na-Butyrate, 1.64 M sucrose, 10 mM Tris-HCl, pH 8.0, 5 mM beta-mercaptoethanol, 0.1 mM PMSF, 2 mM $MgCl_2$, 0.15% w/v Triton X-100 and 1x 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 1x Complete).

Chromatin was sheared with a Bioruptor (Diagenode, Liège, Belgium) 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, 0.1% w/v Triton X-100) and precleared with 40 µL protein A agarose (Roche Applied Science, Mannheim, Germany). 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 protein A agarose and modified histones were detected with 2.5 µL anti-trimethyl H3K4 (04-745, Millipore, Billerica, USA), 5 µL anti-acetyl H3K4 (07-539, Millipore, Billerica, USA), 5 µL anti-acetyl H3K27 (07-360, Millipore, Billerica, USA), 5 µL anti-acetyl H3K36 (07-540, Millipore, Billerica, USA), 5 µL anti-trimethyl H3K27 (07-449, Millipore, Billerica, USA),), 5 µL anti-trimethyl H3K36 (ab9050, Abcam, Cambridge, UK) and 1 µL anti-H3 C-term (ab1791, Abcam, Cambridge, UK).

The control serum for determination of background precipitation was derived from rabbits immunized with an unrelated protein from potato. After washing, the antibody-bound complexes were released and de-crosslinked 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 co-precipitated DNA was purified using the MSB Spin PCRapace kit (Invitek, Berlin, Germany).Typically, 2 µL of eluted DNA were 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 (NCS, see above). The NCS signal was never more than 10% of the signal obtained with a specific antibody.

qPCR

Quantitative PCR was performed on an ABI PRISM 7300 sequence detection system (Life Technologies, Darmstadt, Germany) using SYBR Green fluorescence (Platinum SYBR Green QPCR Mix, Life Technologies, Darmstadt, Germany) for detection. Oligonucleotides were purchased from Metabion (Martinsried, Germany). Oligonucleotide sequences are given in Supplemental data 1. Amplification conditions were 2 min of initial denaturation at 95°C, followed by 40 cycles of 15 s at 95°C, 1 min at 60°C. Afterwards 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.

References

Barrand S, Andersen IS and Collas P (2010). Promoter-exon relationship of H3 lysine 9, 27, 36 and 79 methylation on pluripotency-associated genes. Biochenm Bioph Res Co 401: 611–617.

Charron J-BF, He H, Elling AA and Deng XW (2009). Dynamic landscapes of four histone modifications during deetiolation in Arabidopsis. Plant Cell 12: 3732–3748.

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA and Jaenisch R (2010). From the Cover: Histone H3K27ac separates active from poised enhancers and predicts developmental state. PNAS 107: 21931–21936.

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.

Fischle W, Wang Y and Allis CD (2003). Histone and chromatin crosstalk. Curr Opin Cell Biol 15: 172–183.

Gowik U and Westhoff P (2011). The path from C3 to C4 photosynthesis. Plant Physiol 155: 56–63.

Guillemette B, Drogaris P, Lin H-HS., Armstrong H, Hiragami-Hamada K, Imhof A, Bonneil É, Thibault P, Verreault A, Festenstein RJ and Madhani HD (2011). H3 lysine 4 Is acetylated at active gene promoters and is regulated by H3 lysine 4 methylation. PLOS Genet 3: e1001354.

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

Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE and Ren B (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet 39: 311–318

Jang I-C, Chung PJ, Hemmes H, Jung C and Chua N-H (2011). Rapid and Reversible Light-Mediated Chromatin Modifications of Arabidopsis Phytochrome A Locus. Plant Cell 2: 459–470.

Jiao Y and Meyerowitz EM (2010). Cell-type specific analysis of translating RNAs in developing flowers reveals new levels of control. Mol Syst Biol 6.

Kausch AP, Owen TP, Zachwieja SJ, Flynn AR, Sheen J (2001). Mesophyll-specific, light and metabolic regulation of the C4 PPCZm1 promoter in transgenic maize. Plant Mol Biol 45:1-15.

Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C, Shilatifard A, Buratowski S and Greenblatt J (2003). Methylation of Histone H3 by Set2 in *Saccharomyces cerevisiae* Is Linked to Transcriptional Elongation by RNA Polymerase II. Mol Cell Biol 23: 4207–4218.

Liu B, Fedak G, Liu ZL, Han FP, Tan M, Shan XH, Dong YZ, Wang XZ and Hao S (2004). Activation of a rice endogenous retrotransposon Tos17 in tissue culture is accompanied by cytosine demethylation and causes heritable alteration in methylation pattern of flanking genomic regions. Theor Appl Genet 109: 200–209.

Morris SA, Rao B, Garcia BA, Hake SB, Diaz RL, Shabanowitz, J, Hunt DF, Allis CD, Lieb JD and Strahl BD (2006). Identification of Histone H3 Lysine 36 Acetylation as a Highly Conserved Histone Modification. J Biol Chem 10: 7632–7640.

Nightingale KP, Gendreizig S, White DA, Bradbury C, Hollfelder F and Turner BM (2006). Cross-talk between Histone Modifications in Response to Histone Deacetylase Inhibitors. J Biol Chem 7: 4408–4416.

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 Physiol 141:1078-1088.

Offermann S, Dreesen, B, Horst I, Danker, T, Jaskiewicz M, Peterhänsel 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.

Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolsheimer E, Zeitlinger J, Lewitter F, Gifford DK, Young RA (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell 122:517-527.

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

Suganuma T and Workman JL (2008). Crosstalk among Histone Modifications. Cell 4: 604–607.

Wagner EJ and Carpenter PB (2012). Understanding the language of Lys36 methylation at histone H3. Nat Rev Mol Cell Biol 2: 115–126.

Yoshida S, Ito M, Nishida I and Watanabe A (2001). Isolation and RNA gel blot analysis of genes that could serve as potential molecular markers for leaf senescence in Arabidopsis thaliana. Plant Cell Physiol 2: 170–178.

Zhang X (2008). The epigenetic landscape of plants. Science 5875: 489–492.

CHAPTER 5

A diurnally regulated and tissue-specific long noncoding RNA is associated with the C4 phosphoenolypruvate promoter in maize

Authors

Louisa Heimann, Désirée Waidmann, Ina Horst, Renke Perduns, Sascha Offermann and Christoph Peterhansel*

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

* corresponding author: Christoph Peterhansel, tel +49-511-7622632, fax +49-511-76219262, email: cp@botanik.uni-hannover.de

key words: long noncoding RNAs, chromatin, C4 photosynthesis, diurnal, *Zea mays*

Abstract

Long noncoding RNAs (lncRNAs) have been implicated in regulation of transcription in eukaryotes. We detected an abundant lncRNA homologous to 126 bp of promoter sequence of the C4-specific phosphoenolpyruvate carboxylase gene (C4-*Pepc*) in maize. C4-*Pepc* is one of the most highly expressed genes in maize and strongly regulated by diverse environmental and developmental stimuli. The lncRNA was only detected in leaf mesophyll cells where C4-*Pepc* expression is highest. It followed the protein-coding transcript in abundance through light/dark treatments and diurnal regulation. However, whereas the protein-coding transcript was completely suppressed by feeding of desoxyglucose to leaves, the lncRNA remained unaffected by this treament. Pharmacological suppression of RNA polymerase II activity with α-Amanitin also completely abolished synthesis of the protein-coding transcript, but lncRNA levels remained high. Using subcellular fractionation, we show that the lncRNA is associated with chromatin. A possible function of the lncRNA in increasing accessibility of C4-*Pepc* promoter chromatin to transcription factors is discussed.

Introduction

With the help of microarray analysis and new sequencing technologies, it has been shown that the majority of the genome in eukaryotes is transcribed into RNA (Lucia and Dean 2011). A major function of RNA is to transfer genomic code information to protein synthesis (messenger RNA = mRNA). However, it is known since long that other RNAs rather have structural or regulatory functions. The family of noncoding RNAs (ncRNAs) contains such diverse members as transfer RNA (tRNA), ribosomal RNA (rRNA), microRNAs (miRNA), small interfering RNA (siRNAs) and long noncoding RNAs (lncRNA). rRNAs and tRNAs are abundant and well-described constituents of the translation machinery (Ghildiyal and Zamore 2009). miRNAs are small cytoplasmic RNAs of 20-25 bases in length that are involved in posttranscriptional gene silencing and induce degradation or repression of productive translation of homologous mRNAs (Brennecke et al. 2005; Cui et al. 2006). Small RNAs of similar size were described to control the formation of heterochromatin on transposons and repetitive DNA elements in the nucleus (casiRNA) (Ghildiyal and Zamore 2009). lncRNAs are also nuclear RNAs and associated with chromatin, but clearly longer (> 100 nucleotides) (Rinn and Chang 2012). The best studied example for a lncRNA is the “X inactive specific transcript” *Xist* that binds in *cis* to one of the two X chromosomes in females, recruits the polycomb repressive complex 2 (PRC2), and induces the repression of this chromosome for dosage compensation (Lee et al. 2012; Rinn and Chang 2012; Wutz 2011). Similarly, the lncRNA HOTAIR, that is itself derived from a Hox gene cluster, can suppress other Hox gene clusters in *trans* in a PRC2-dependent manner (Tsai et al. 2010). In plants, Heo and Sung (2011) recently identified a 1.1 kbp lncRNA derived from an intron of the Flowering Locus C (FLC) called COLDAIR. The RNA is involved in establishment of a repressive chromatin structure on the gene that is induced by vernalization and required for flowering.

A common feature of lncRNAs characterized so far is their possible function in chromatin regulation. Chromatin is not only a passive barrier to transcription, but contributes actively to gene regulation. The core particle of chromatin is the nucleosome octamer which consists of each two of the histone proteins H2A, H2B, H3 and H4 and DNA that is wound around this core particle nearly two times (Kouzarides et al. 2007). Histones can be covalently modified in multiple ways. Acetylation and methylation of multiple lysines on the N-terminal tails of H3 and H4 have been best studied. Whereas acetylation is almost exclusively associated with active promoters, methylation can induce active or repressive states dependent on the lysine that is methylated and the number of methyl groups that are transferred to the terminal amino group (Wang et al. 2009a; Pokholok et al. 2005). Chromatin modifications are known to regulate chromatin structure by recruiting remodelling enzymes that can use the energy derived from ATP hydrolysis to rearrange nucleosomes (Bannister and Kouzarides 2011).

The gene encoding C4-specific phosphoenolpyruvate carboxylase (C4-*Pepc*), a key enzyme of C4 photosynthesis that catalyzes primary CO_2 fixation in this photosynthetic subtype, is among the most highly transcribed genes in maize (Sheen and Bogorad 1987; Kausch et al. 2001). Because of its complex transcriptional regulation, C4-*Pepc* is an excellent model for studying the integration of environmental and developmental stimuli into a promoter response. C4-*Pepc* is only transcribed in leaves, but not in other plant organs. Within the leaf, promoter activity is restricted to mesophyll cells. In bundle sheath cells, the second photosynthetic cell type that cooperates with mesophyll cells in C4

photosynthesis, the gene remains inactive (Sheen 1999). C4-*Pepc* transcription is highly inducible by light (Horst et al. 2009), but suppressed by metabolic stimuli such as high sugar availability (Sheen 1999). Recent studies from our lab indicated that chromatin structure and histone modifications are important in the response of the promoter to all these signals (Horst et al. 2013; Offermann et al. 2008).

In this study, we report the identification of a transcript homologous to the core promoter region of C4-*Pepc*. Characterization of the abundance and localization of this transcript in response to diverse stimuli suggests that it is important for the establishment of an active chromatin structure on the C4-*Pepc* core promoter.

Results

Transcription of the C4-Pepc *promoter*

The promoters of many highly transcribed genes in humans are covered with non-coding transcripts (Wilusz et al. 2009; Kanhere et al. 2010). We wanted to determine whether C4-*Pepc* promoter regions were also transcribed. To this end, we isolated RNA from illuminated leaves of maize harvested 4h after onset of illumination (4L plants), prepared cDNA and determined the abundance of transcripts by quantitative real-time PCR (qPCR). The absence of genomic DNA contamination was tested by parallel cDNA synthesis reactions that did not contain reverse transcriptase (-RT control, see also Materials and Methods). With none of the primer systems, DNA was amplified from -RT controls (data not shown). Figure 5-1a shows a truncated diagram of C4-*Pepc* gene structure and the positions of the amplicons on the promoter. We tested 14 promoter regions of 100-200 bp in length (primer systems see online resource 1) and, for comparison, a gene region near to the 3' end of the gene (region 1). The latter amplicon covered an exon-intron border and, therefore, detected unspliced nascent transcripts (hnRNA) from cDNA. We and others have shown before that the abundance of hnRNA transcripts provides a good estimate for promoter activity (Elferink and Reiners 1996, Offermann et al. 2006; Wu et al. 2009). Relative abundances of RNA homologous to the 15 tested promoter regions are shown in Figure 5-1b and Figure 5-1c on two different scales. We detected various amounts of transcripts in 13 of the 15 tested promoter regions. By far strongest signals were obtained from regions 2 and 3. The corresponding PCR systems are centered at positions +10 and -100 relative to the predicted transcription initiation site (TIS), respectively. The abundances of the ncRNAs homologous to these regions were comparable to the abundance of hnRNA detected in region 1 (Figure 5-1b). A second much weaker peak was observed in region 12 (-1700 bp relative to TIS, Figure 5-1c). These results were obtained in the maize hybrid line Montello, but were completely reproducible in B73, an unrelated maize inbred line (see online resource 2). Thus, significant amounts of promoter transcripts were detected near to the TIS and in the core promoter region of the C4-*Pepc* gene.

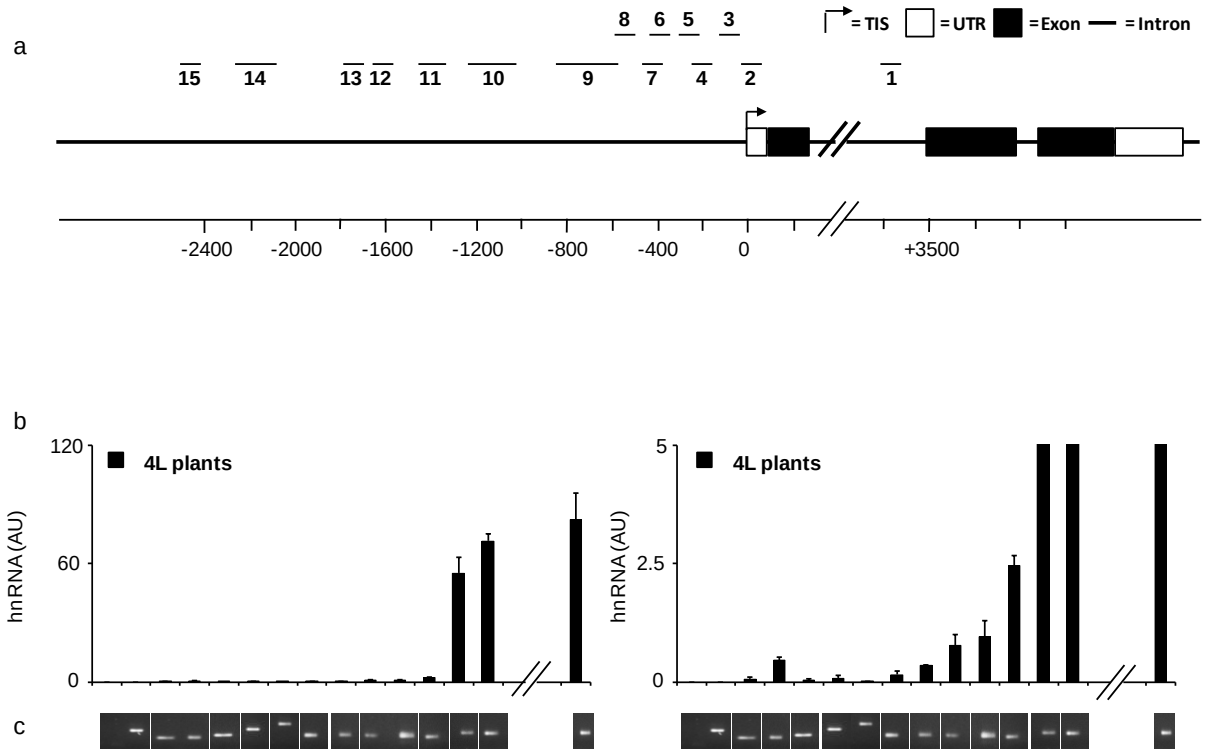


Figure 0-1: Overview about promoter transcripts found on the *C4-Pepc* promoter.
a, Truncated gene-model of C4-*Pepc.* Exons, 5' UTR and 3'UTR presented as black and white boxes, respectively. The arrow indicates the transcription initiation site. The amplification sites for qPCR analysis are presented as numbers with black bars above the map. b, Promoter transcript expression levels on two different Y-scales. c, Ethidium bromide-stained agarose gel of amplified promoter transcripts from the C4-*Pepc* loci shown in a. All data points are based on at least three independent experiments. Vertical lines indicate standard errors. Values are arbitrary units derived from a cDNA standard dilution series. hnRNA expression levels were determined by quantitative RT-PCR with a primer system specific for an intron (primer systems see online resource 1).

To determine whether region 2 and region 3 transcripts formed part of the protein-coding transcript or whether they were transcribed independently, we used PCR primer systems homologous to the core promoter (upstream primers of region 2 [Fw1] and region 3 [Fw2]) and to the start of the coding sequence (Rev1). Amplicons with Rev1 as the reverse primer cover an intron, thus, amplification from genomic DNA and RNA could be discriminated (Figure 5-2a). As shown in Figure 5-2b, all tested primer combinations amplified fragments from genomic DNA. In none of the cases, amplification from –RT controls was observed. Amplification from cDNA (+RT lanes) was only possible with Fw2 as a forward primer. Sequencing of the PCR products confirmed that the Fw2/Rev1 product was obtained from cDNA and not from contaminating genomic DNA, because the amplified sequence did not contain the intron (data not shown). These data suggested that the region 2 amplicon forms part of the protein-coding transcript and that the TIS was incorrectly mapped in earlier studies. The region 3 transcript was seemingly separately transcribed from the protein-coding transcript. To determine the length of the region 3 transcript, we also performed PCR testing a set of primer systems covering the core promoter region every 5 bp (Figure 5-2a, gray arrows). The longest product that could be obtained was 126bp in length (Fw2/Rev2). All other tested primer systems did not detect significant amounts of promoter transcripts.

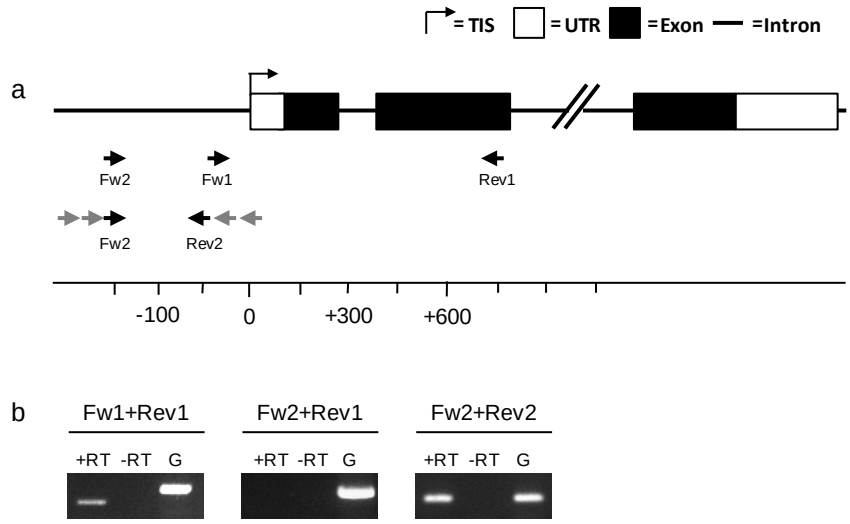


Figure 0-2: Determination of possible protein-coding transcripts and the size of region 3 transcript.
a, Truncated gene-model of C4-*Pepc.* Exons, 5' UTR and 3'UTR are represented as black and white boxes, respectively. The amplification sites for PCR and sequencing analysis are presented as black arrows below the map. Forward primers are designated with Fw, reverse primer with Rev. Grey arrows indicate primer systems covering the region around the noncoding RNA product. b, Ethidium bromide-stained PCR products obtained with indicated primer systems, +RT represents reverse transcription, -RT represents the control for DNA contaminations, G represents amplified products from DNA to determine the size.

Regulation of region 3 transcripts

In order to analyze possible functions of region 3 transcripts, we determined the abundance of these transcripts in different tissues and under various growth conditions. For comparison, abundance of C4-*Pepc* hnRNA (region 1 transcript) was always measured. A key feature of C4 gene regulation is gene induction by light. Region 1 and region 3 transcript levels were therefore measured in plants that were exposed to 72h darkness (72D plants) and compared to plants in a normal day night rhythm, harvested 4h after onset of illumination (4L plants) (Figure 3a). 72D plants contained 80-fold less region 1 transcripts than 4L plants. Region 3 transcripts were only reduced by approximately 50%. We further followed diurnal regulation of region 1 and region 3 transcripts at different time points after illumination (Figure 5-3b). Both transcripts showed a comparable induction in the morning, a reduction towards the end of the illumination phase, and constant levels during the night.

We further studied organ and cell-type specificity of the region 1 and 3 transcripts. It had been reported before that *C4-Pepc* is only transcribed in mesophyll cells (M cells) of leaves. We did not isolate M cells, because preparation of M cell protoplasts from leaves is a lengthy procedure during which C4 gene transcription is often strongly suppressed. Instead, we compared region 1 and region 3 transcript levels in total leaves and bundle sheath strands that can be rapidly isolated form leaves (see Materials and Methods). As expected, region 1 transcripts were 14-fold depleted from isolated BS cells compared to total leaves (Figure 5-3c). Such depletion was also observed for region 3 transcripts, albeit to a slightly lower extent.

Figure 5-3d shows amplification products from RT-PCRs with RNA derived from different plant organs (coleoptiles, roots, pollen, and seeds). In all organs, *Actin*1 hnRNA was detectable from RNA preparations. Amplification was dependent on reverse transcription indicating absence of DNA contamination in these preparations. Neither region 1 transcripts nor region 3 transcripts were

detectable in RNA from any of the tested organs. However, successful amplification was possible with genomic DNA derived from the same tissues as a template.

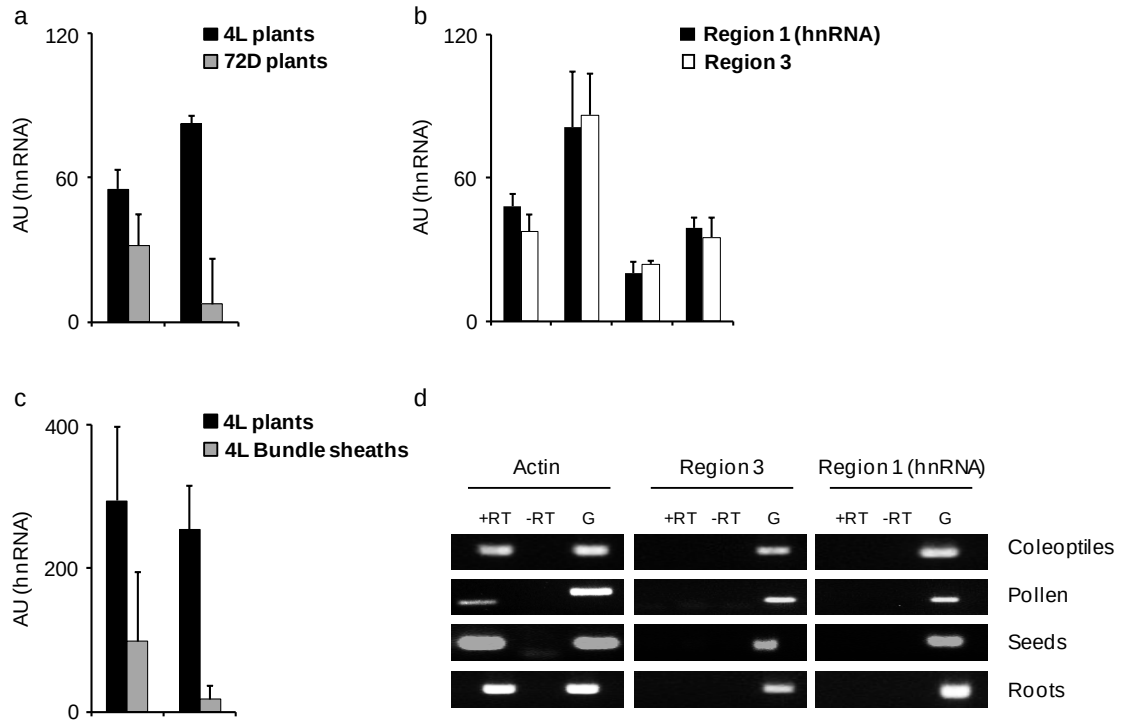


Figure 0-3: Light and tissue specific transcription of C4-*Pepc* and region 3 transcripts.
a, Relative quantification of C4-*Pepc* hnRNA and region 3 transcripts expression levels from plants that were exposed to 72h darkness (72D, grey columns) and in plants that were illuminated for 4h (4L, black columns). b, Relative quantification of C4-*Pepc* hnRNA (black columns) and region 3 transcripts (white columns) expression levels in a diurnal approach. Plants were harvested after 0h, 4h, 14, and 20h, respectively. c, Relative quantification of C4-*Pepc* hnRNA and region 3 transcripts expression levels from plants that were illuminated for 4h (4L, black columns) and bundle sheath cells (grey columns). d, Ethidium bromide-stained PCR products derived from different tissues, coleoptiles, pollen, seeds and roots. As control, always *Actin*-1 transcription was measured. a-d, Values are arbitrary units derived from a cDNA standard dilution series. hnRNA expression levels were determined by quantitative RT-PCR with a primer system specific for an intron (for details see online resource 1). All data points are based on at least three independent experiments. Vertical lines indicate standard errors.

We additionally tested a metabolic stimulus that was known to suppress C4-*Pepc* transcription. Detached leaves were fed with different concentrations of desoxyglucose (DOG) via the transpiration stream and transcript levels were compared (Figure 5-4a). The amounts of region 1 transcripts were reduced with increasing concentrations of DOG. Region 2 transcripts from the 5' end of the protein-coding transcript (see above) were also tested and showed a similar dose-dependent suppression by DOG. However, amounts of region 3 transcripts remained completely unaffected. Leaves were also treated for 4 and 8h with α-Amanitin, a potent inhibitor of RNA polymerase II, (Figure 5-4b). Region 1 transcripts and region 2 transcripts were undetectable in these samples. However, region 3 transcripts were easily detectable and found at similar levels as in control plants that were also detached, but not treated with α-Amanitin ("-" samples in Figure 5-4b). In these control samples, also region 2 and region 3 transcripts were detected and levels of all three amplicons were comparable.

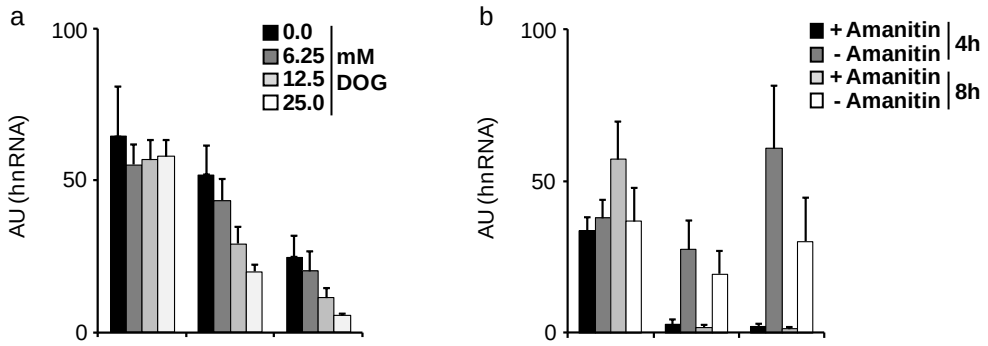


Figure 0-4: Metabolite and α-Amanitin dependent transcription of C4-*Pepc* and region 3 transcripts.
a, Relative quantification of C4-*Pepc* hnRNA and region 2 and 3 transcripts expression levels from plants that were treated with increasing amounts of 2-deoxyglucose (DOG) via the transpiration stream, from 0 mM to 25 mM, respectively. b, Relative quantification of C4-*Pepc* hnRNA and region 2 and 3 transcripts expression levels from plants that were treated with α-Amanitin via the transpiration stream for 4 and 8 hours. As control always transcription levels of non-treated plants were measured. a-b, Values are arbitrary units derived from a cDNA standard dilution series. hnRNA expression levels were determined by quantitative RT-PCR with a primer system specific for an intron (for details see online resource 1). All data points are based on at least three independent experiments. Vertical lines indicate standard errors.

Chromatin association of region 3 transcripts

We hypothesized that region 3 transcripts remain associated with the promoter after synthesis. We therefore isolated chromatin from 4L and 72D plants. Nucleic acids and proteins were covalently crosslinked with formaldehyde before start of the purification to ensure that loosely associated macromolecules were not lost. Figure 5-5a shows the abundance of region 3 transcripts compared to the abundance of C4-*Pepc* mRNA in chromatin preparations. Corresponding abundances of the transcripts in total leaves are shown in Figure 5-5b. Only region 3 transcripts, but not C4-*Pepc* mRNA were detected in the chromatin preparations. Also *Actin1* mRNA and region 1 transcripts were not detected in chromatin (data not shown) indicating that chromatin preparations were largely free of cytosolic contamination. In total leaves, both region 3 transcripts and mRNAs were detected and transcript abundance was dependent on illumination (as shown in Figure 5-3). The data indicate that region 3 transcripts are associated with chromatin.

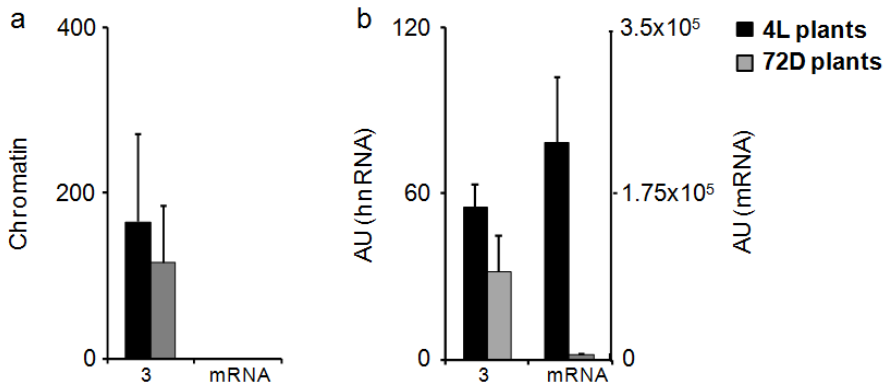


Figure 0-5: Chromatin association of region 3 transcript.
a, Relative quantification of region 3 transcripts and mRNA levels detected in chromatin extract in plants that were exposed to 72h darkness (72D, grey columns) and in plants that were illuminated for 4h (4L, black columns). b, Relative quantification of C4-*Pepc* hnRNA (data taken from figure 3a) and mRNA levels from plants that were exposed to 72h darkness (72D, grey columns) and in plants that were illuminated for 4h (4L, black columns). All data points are based on at least three independent experiments. Vertical lines indicate standard errors.

Discussion

We detected diverse RNAs when screening by RT-PCR for transcripts that were homologous to the C4-*Pepc* promoter in maize. Surprisingly, two of these transcripts derived from the core promoter near to the predicted TIS were present at similar levels as nascent unspliced C4-*Pepc* hnRNA (Figure 5-1). hnRNAs differ in abundance by two or three orders of magnitude from the corresponding mRNAs and are often hardly detectable when genes are not highly transcribed (Reed 2003; Offermann et al. 2006). One possible interpretation of the result was that the TIS was not correctly mapped in previous work (Yanagisawa and Izui 1989) or differed between the cultivar used here and in the previous study. Indeed, PCR analyses showed that the more downstream promoter transcript formed part of the protein-coding transcript (Figure 5-2) and, consequently, showed an identical regulation as C4-*Pepc* hnRNA over all tested treatments (Figure 5-4). This was reproducible in the model inbred line B73, the most used maize line for research purposes (Schnable et al. 2009). Thus, the major TIS of C4-*Pepc* was located at least 48 bp more upstream in both Montello and B73 than previously annotated. Our data do not exclude that other initiation sites are used with lower efficiency.

Ample evidence suggested that the more upstream region 3 transcript was transcribed independent from the protein-coding transcript. First, different from region 2 transcripts, amplification as a continuum with parts of the protein-coding transcript failed (Figure 5-2). Second, regulation was similar to the protein-coding transcript, but several important differences including sensitivity to α-Amanitin and DOG existed (Figures 5-3 and 5-4). Third, region 3 transcripts were associated with chromatin whereas the coding transcript was not (Figure 5-5). These results were different for the recently described promoter transcript on the GLDPA gene in the C4 plant *Flaveria trinervia* that was derived from an alternative transcription initiation site in the promoter and that was merged with the major protein-coding sequence by splicing (Wiludda et al. 2012). It is therefore probable that region 3 promoter transcripts belong to the class of lncRNAs. Their size of 126 nucleotides is at the lower end of what is known for such RNAs, but clearly longer than other chromatin-associated RNAs that are derived from the miRNA pathway (Lee 2012; Carthew and Sontheimer 2009).

Whereas diverse functions in gene regulation have been assigned to lncRNAs derived from introns or intergenic sequences (compare Introduction), only few reports exist that described the function of lncRNAs homologous to core promoter sequences. Kapranov et al. (2007) described a class of short transcripts in humans that were less than 200 nt in length and that were found on promoters of highly transcribed genes. In yeast, transcripts of heterogeneous length spanning active promoters were also observed (Davis and Ares 2006). Conversely, Kanhere et al. (2010) reported that ncRNAs, about 50 to 200 bp in length, transcribed downstream of the promoter region of Polycomb genes, can recruit the repressor protein PRC2 and prevent formation of the protein-coding transcript. Thus, promoter-associated lncRNAs can act positively or negatively on transcription. In all studies, formation of these transcripts was assigned to RNA polymerase II (RNAPII) activity. It was speculated that these transcripts functioned as a scaffold for recruiting proteins or simply as a consequence of promoter stalling of RNAPII (Baker 2011). However, in this study, α-Amanitin treatment did not reduce abundance of region 3 transcripts, whereas the protein-coding transcript was completely abolished by the treatment (Figure 5-4). This indicates that region 3 transcripts were not formed by RNAPII. Plants, different from animals, express specific RNAPs (RNAP IV+V) that are insensitive to α-Amanitin for the

synthesis of small RNAs implied in gene silencing (Haag and Pikaard 2011) and that might also synthesize the region 3 transcript. Alternatively, region 3 transcripts might be relatively stable and therefore simply not synthesized during the 8h of treatment. α-Amanitin only inhibits formation of new transcripts, but does not impact on existing transcripts (Brueckner and Cramer 2008). However, this is in disagreement with the clear fluctuation of region 3 transcript levels in the diurnal rhythm (Figure 5-2b). Discrimination between ncRNAs with an actual function or ncRNAs that are simply nonfunctional byproducts is complicated because of the high number of randomly transcribed genome regions (Ponjavic et al. 2007). Because of the lack of efficient transformation systems in maize, we were unable to manipulate the amounts of region 3 transcripts by overepxression or RNA interference. Instead, we used co-regulation as evidence for functional association. Region 3 transcripts were regulated by light, tissue-specific signals and diurnal stimuli (Figure 5-3) in a very similar manner as the protein-coding transcript. We have shown before that all these stimuli also impact on activating histone modifications such as acetylation (Offermann et al. 2006; Danker et al. 2008; Horst et al. 2009) or H3K4 trimethylation (Patel and Berry, 2007) in the core promoter region where the region 3 transcript was formed. Moreover, as well histone modifications on the core promoter (Offermann et al. 2008) as region 3 transcript abundance (Figure 5-4) remained unaffected by high sugar availability. Based on this tight correlation over many treatments of (i) promoter activity, (ii) region 3 transcript abundance and (iii) activating histone modifications, we hypothesize that continuous formation of region 3 transcripts increases the accessibility of C4-*Pepc* promoter chromatin for transcription factors and RNAP II. Consistent with this idea, region 3 transcripts were shown to be associated with chromatin (Figure 5-5). Furthermore, accessibility of C4-*Pepc* promoter chromatin increased after illumination of dark-grown plants (Kalamajka et al. 2003), a treatment that also increased the abundance of region 3 transcripts. Further analyses will identify potential interactors of region 3 transcripts on the promoter that mechanistically link transcript formation to chromatin modification (Tsai et al. 2010).

In conclusion, analysis of the promoter region of C4-*Pepc* in maize identified a chromatin-associated lncRNA that is transcribed independently from the protein-coding transcript, but highly co-regulated. The most probable function of this transcript is to increase core promoter accessibility of the gene.

Materials and Methods

Plant material and growth conditions

Maize (*Zea mays* cv. Montello and B73), was 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, Sinntal-Jossa, Germany) with a photon flux density of 120-180 µmol m^{-2} s^{-1} until the third leaf was fully expanded. 72D plants were grown in the normal light rhythm, but darkened for three days before harvest.

Isolation of bundle sheath cells

For gene expression analyses, bundle sheath strands were isolated mechanically as described before by Hahnen et al. (2003), but without diethylether treatment. Leaves were washed extensively in ice cold water and homogenized in a Waring Blendor for 3 x 3 s. 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, Darmstadt, Germany) and the residue was washed extensively with ice-cold water. The isolated bundle sheath strands were shortly dried on paper and frozen in liquid nitrogen.

Plant treatment and tissue preparation

Plants were treated with α-Amanitin or 2-deoxyglucose (DOG) as previously described by Offermann et al. (2008). After 3 h light, 10- to 12-day-old leaves were detached under water 1 cm above the laminar joint and incubated for 4 h and 8 h in solutions containing 10 µM α-amanitin or DOG at varying concentrations as indicated in the figure, in combination with 5 µM trans-zeatin (all Sigma-Aldrich, Schnelldorf, Germany) and 16 mM KNO_3 (Li et al. 2010) in tap water.

RNA isolation and reverse transcription

Total RNA isolation was performed by phenol-chloroform extraction as described in Offermann et al. (2006). About 25-30 mg ground plant material was dissolved in 1 ml Trizol and agitated for 15 min. After addition of 0.2 volumes chloroform and agitation for 10 minutes, 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 2 volumes 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 H_2O. The quality of the isolated RNA was controlled by electrophoresis and the concentration was determined photometrically.
One unit of DNAseI (Fermentas, St. Leon Roth, Germany) per µg 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. Control

reactions in the absence of reverse transcriptase were used to exclude amplification from residual DNA contamination. 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, Mannheim, Germany) and 1mM dNTPs in reaction buffer as specified by the manufacturer. hnRNAs were amplified from cDNA using primer systems specific for introns (online resource 1). A dilution series of cDNA from illuminated leaves was used as a standard.

qPCR

Quantitative PCR was performed on an ABI PRISM 7300 sequence detection system (Life Technologies, Darmstadt, Germany) using SYBR Green fluorescence (Platinum SYBR Green QPCR Mix, Life Technologies, Darmstadt, Germany) for detection. Oligonucleotides were purchased from Metabion (Martinsried, Germany). Oligonucleotide sequences are given in online resource 1. Amplification conditions were 2 min of initial denaturation at 95°C, followed by 40 cycles of 15 s at 95°C, 1 min at 60°C. Afterwards, 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. A dilution series of genomic DNA was used to determine relative abundance between different PCR systems (arbitrary units [AU]). All standard curves used for comparison had similar slopes, intercepts, and correlation coefficients.

Chromatin preparation

As described previously by Horst et al. (2009), 6 g leaves from 10- to 12-d-old maize seedlings were harvested and vacuum infiltrated with crosslink buffer (10 mM Na-Butyrate, 400 mM sucrose, 10 mM Tris-HCl, pH 8.0, 5 mM beta-mercaptoethanol, 0.1 mM phenylmethylsulfonyl fluoride (PMSF) and 3% v/v formaldehyde). The crosslink was stopped after 10 min by addition of glycine to a final concentration of 0.2 M. The material was ground, resuspended in extraction buffer (10 mM Na-Butyrate, 400 mM sucrose, 10 mM Tris-HCl, pH 8.0, 5 mM beta-mercaptoethanol, 0.1 mM phenylmethylsulfonyl fluoride (PMSF) and 1x Complete (Roche Applied Science, Mannheim, Germany)) and incubated for 15 min at 4°C. Afterwards, the solution was filtered through 4 layers of Miracloth (VWR, Darmstadt, Germany) and the residue was washed with purification buffer 1 (10 mM Na-Butyrate, 250 mM sucrose, 10 mM Tris-HCl, pH 8.0, 5 mM beta-mercaptoethanol, 0.1 mM PMSF, 10 mM $MgCl_2$, 1% w/v Triton X-100 and 1x Complete) and afterwards with purification buffer 2 (10 mM Na-Butyrate, 1.64 M sucrose, 10 mM Tris-HCl, pH 8.0, 5 mM beta-mercaptoethanol, 0.1 mM PMSF, 2 mM $MgCl_2$, 0.15% w/v Triton X-100 and 1x 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 1x Complete). Chromatin was sheared with a Bioruptor (Diagenode, Liège, Belgium) for 10 min (setting: high, interval 30/30 s) under constant cooling. Afterwards chromatin was de-crosslinked 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 precipitated RNA was dissolved in 1 ml Trizol and agitated for 15 min. Further preparation steps were carried out as described in section RNA isolation.

References

Baker M (2011). Long noncoding RNAs: the search for function. Nat Meth 8:379-383.

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

Brennecke J, Stark A, Russell RB, Cohen SM (2005). Principles of microRNA-target recognition. PLOS Biol 3: 404-418.

Brueckner F and Cramer P (2008). Structural basis of transcription inhibition by α-amanitin and implications for RNA polymerase II translocation. Nat Struct Mol Biol 15:811-818.

Carthew RW and Sontheimer EJ (2009). Origins and mechanisms of miRNAs and siRNAs. Cell 136:642-655.

Cui Q, Yu Z, Purisima EO, Wang E (2006). Principles of microRNA regulation of a human cellular signaling network. Mol Syst Biol 2:1-7.

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.

Davis CA and Ares M (2006). Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA. 103:3262-3267.

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

Ghildiyal M and Zamore PD (2009). Small silencing RNAs: an expanding universe. Nat Rev Genet 10:94-108.

Haag JR and Pikaard CS (2011). Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. Nat Rev Mol Cell Biol 12:483-492.

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. Photosyn Res 75:183-192.

Heo JB and Sung S (2011). Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science 331:76-79.

Horst I, Heimann L, Peterhänsel C (2013). Signal integration on plant promoters: A case study in maize. Plant Signalling & Behavior, *in press*.

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

Kalamajka R, Hahnen S, Cavalar M, Töpsch S, Weier D, Peterhänsel C (2003). Restriction accessibility in isolated nuclei reveals light-induced chromatin reorganization at the PEPC promoter in maize. Plant Mol Biol 52:669-678.

Kanhere A, Viiri K, Araújo CC, Rasaiyaah J, Bouwman RD, Whyte WA, Pereira CF, Brookes E, Walker K, Bell GW, Pombo A, Fisher AG, Young RA, Jenner RG (2010). Short RNAs are transcribed from repressed polycomb target genes and interact with Polycomb Repressive Complex-2. Mol Cell 38:675-688.

Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, Stadler PF, Hertel J, Hackermüller J, Hofacker IL, Bell I, Cheung E, Drenkow J, Dumais E, Patel S, Helt G, Ganesh M, Ghosh S, Piccolboni A, Sementchenko V, Tammana H, Gingeras TR (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science 316:1484-1488.

Kausch AP, Owen TP, Zachwieja SJ, Flynn AR, Sheen J (2001). Mesophyll-specific, light and metabolic regulation of the C4 PPCZm1 promoter in transgenic maize. Plant Mol Biol 45:1-15.

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

Lee JT (2012). Epigenetic regulation by long noncoding RNAs. Science 338:1435-1439.

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. Nat Genet 42:1060-1067.

Lucia F and Dean C (2011). Long non-coding RNAs and chromatin regulation. Curr Opin Plant Biol 14:168-173.

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 Physiol 141:1078-1088.

Offermann S, Dreesen, B, Horst I, Danker, T, Jaskiewicz M, Peterhänsel 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 and Berry JO (2007). Rubisco gene expression in C4 plants. J Exp Bot 59:1625-1634.

Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolsheimer E, Zeitlinger J, Lewitter F, Gifford DK, Young RA (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell 122:517-527.

Ponjavic J, Ponting CP, Lunter G. (2007). Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. Genome Research 17:556-565.

Reed R (2003). Coupling transcription, splicing and mRNA export. Curr Opin in Cell Biol 15:326-331.

Rinn JL and Chang HY (2012). Genome regulation by long noncoding RNAs. Annu Rev Biochem 81:145-166.

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.

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

Sheen JY and 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.

Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY (2010). Long noncoding RNA as modular scaffold of histone modification complexes. Science 329:689-693.

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

Wiludda C, Schulze S, Gowik U, Engelmann S, Koczor M, Streubel M, Bauwe H, Westhoff P (2012). Regulation of the photorespiratory GLDPA gene in C4 Flaveria: An intricate interplay of transcriptional and posttranscriptional processes. Plant Cell 24:137-151.

Wilusz JE, Sunwoo H, Spector DL (2009). Long noncoding RNAs: functional surprises from the RNA world. Gene Dev 23:1494-1504.

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.

Wutz A (2011). Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. Nat Rev Genet 12: 542-553.

Yanagisawa S and Izui K (1989). Maize phosphoenolpyruvate carboxylase involved in C4 photosynthesis: nucleotide sequence analysis of the 5' flanking region of the gene. J Biochem 106:982-987.

Online Resource 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')
1	+4300	GTATGCTGCCATTGCCCATTGC
		TAGCCTGATAGTGAGTGACGCACA
2	+10	AACGACTCCCCATCCCTATTTGAAC
		AGCAGGGAAGCGAGACGGTTG*
3	-100	CCGGTTCCGTTGCGGTTACC
		CAAATAGGGATGGGGAGTCGTTGG
4	-200	CGATTGCCGCCAGCAGT
		GAACCGGCTGTGGCTGAG *
5	-300	GCACCGAGCCAAGCCAAAAGG
		GATGTGGAGAGGGGTGTCTGCT
6	-350	TTAACTGCTAAGGGACGCCCTCTC
		GTCAACAGCACCGAGCCAAGC
7	-400	CCCTCTCCACATCCTGCAAAGC
		ATTCCGTTGGCTAATTGGGTAGCA
8	-500	CTCTTAGCCACAGCCGCCTCA
		TGCAGCCTACTTGCTAACAGACG
9	-700	TGGCACCCTTATCCCTACAATAGC
		GTCTGTTTGCAGGATGTGGTTGAG
10	-1100	GTGTTAGGACACGTGGTTAGC
		CACTTGGCAGCGGTGAAGATAC
11	-1400	GTACAAATGAGGTGCCGGATTGATG
		CGGCCATGGCATGATACAATTCTCA
12	-1600	CCAAACAGACCCTAAAAATGTGTG
		GCAGTTGATCTATTCCAGCCTCTTA
13	-1700	AGAAACAAAAGCAAGGTCAAGGTG
		GGTTCTGTTTTCCTGCTTCTAAAAGT
14	-2100	GTCACAATTGAAGATTCGTGCAAGG
		CAGTTTGAACTAAACGACTTCCAAC
15	-2400	TATCCTTCTGCCTAGGTTGAGTAGCT
		TGTTGACACCAAATCCTAACCAAA

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

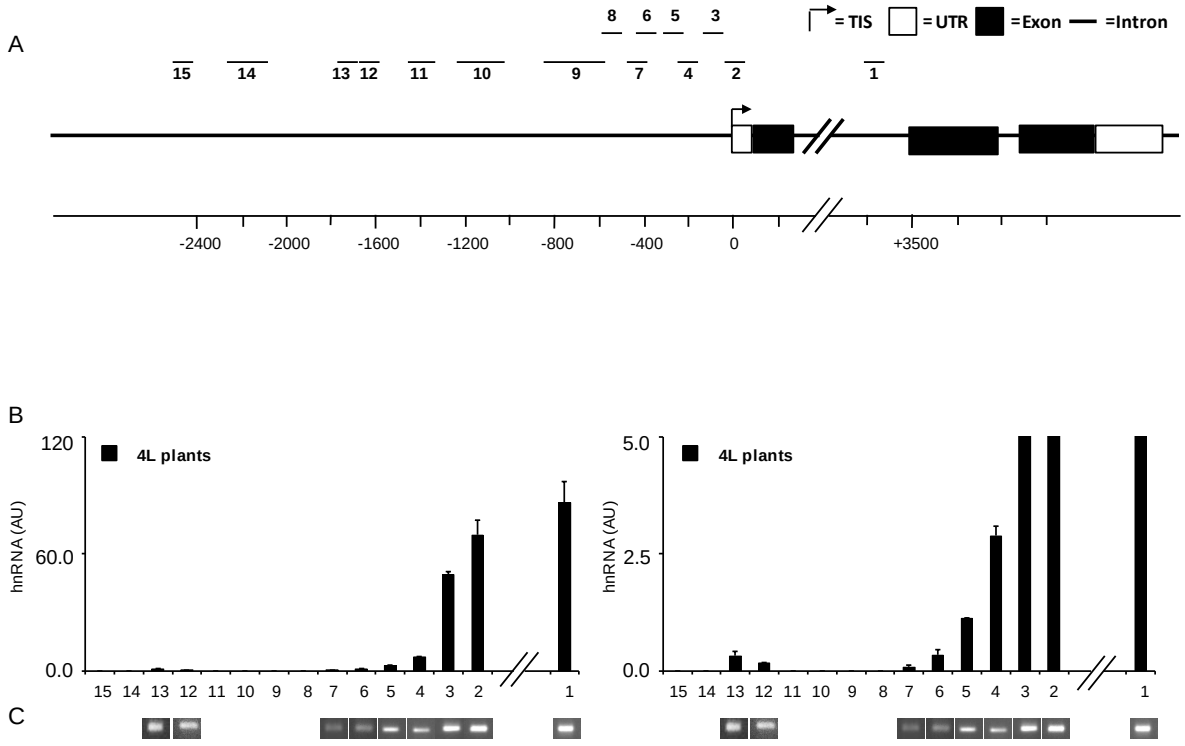
Literature cited

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

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

Online resource 2 Overview about promoter transcripts found on the *C4-Pepc* promoter in B73.

a, Truncated gene-model of C4-*Pepc.* Exons, 5' UTR and 3'UTR presented as black and white boxes, respectively. The arrow indicates the transcription initiation site. The amplification sites for qPCR analysis are presented as numbers with black bars above the map. b, Promoter transcript expression levels on two different Y-scales. c, Ethidium bromide-stained agarose gel of amplified promoter transcripts from the C4-*Pepc* loci shown in a. All data points are based on at least three independent experiments. Vertical lines indicate standard errors. Values are arbitrary units derived from a cDNA standard dilution series. hnRNA expression levels were determined by quantitative RT-PCR with a primer system specific for an intron (see online Resource 1).



CHAPTER 6

General discussion

6 General discussion

6.1 Modified model of signal integration on the C4-*Pepc* promoter

The results obtained in the present study helped to extend and confirm the existing model for signal integration on the C4-*Pepc* promoter from maize. It was shown in previous studies of the C4-*Pepc* promoter, that histones underlie a highly complex regulation of different modifications. In the following figure all results that are known to date are summarized. The model consists mainly of two theories after which histone modifications and in this context signal integration and promoter response can be regulated. The two theories were extensively described in Chapter 3. Figure 6-1 illustrates how information of nutrient and metabolic regulation, organ and tissue specificity, as well as gene induction by light is stored on the chromatin level in leaves. In previous years, using the promoter of the C4-*Pepc* gene, histone modifications were identified which are associated with each of these signals. However, beside the so far tested modifications, histone lysine residues can have many additional modifications that may be regulated by different stimuli. Some residues can be even both acetylated and methylated. For this reason three acetylation/methylation sites were selected (H3K4, H3K27, H3K36) and tri-methylation and acetylation from the same samples was recorded.

It was previously shown by Offermann et al. (2008) that light induces acetylation in the core promoter on the lysine residues H3K9 and H4K5 (Figure 6-1, light inducibility presented as yellow dots). These acetylations are removed when the plants are exposed to prolonged darkness, suggesting that information about previous illumination is lost. It was also shown that H3K14ac, H4K16ac and H3K18ac did not respond to the light stimulus (Figure 6-1, organ specificity presented as light green dots). It is probable that these modifications are not involved in potentiating a transcriptional response of the C4-*Pepc* gene since they were also found in tissues where C4-*Pepc* was only expressed at a basal level. They rather are involved in a poised chromatin state that allows transcriptional activation by illumination. This in contrast to the modifications investigated in this study. All investigated acetylations responded strongly to the light stimulus (see Fig. 4-1) in the core promoter region, H3K27ac and H3K36ac did also respond in the upstream promoter region. This is also in accordance with a study on the light induction of Phytochrome A from Jang et al. (2011). They were able to show that H3K9/14ac, H4K5ac, and H3K27ac are involved in light induction in the promoter and coding region of the Arabidopsis *Phytochrome A* locus. Offermann et al. (2006) were able to show that core promoter hyperacetylation also occurred in B cells and that histone acetylation was induced by light even though the transcription was not induced. It could be shown in this study that H3K27ac and H3K36ac also cover the promoter in B cells and that these modifications are induced by light in total leaves. However it seems to be that these modifications are presumably differentially interpreted in B cells than in leaves and do not induce transcription of C4-*Pepc.* A possibility might be that these modifications cannot be interpreted because an important information is missing e.g. a modification that marks the tissue. Badeaux et al. (2013) proposed a model, where chromatin modifications are interpreted by ‘reader’ proteins and that the role of histone modifications is to recruit these ‘reader’ proteins directly or with the help of crosstalk between histone modifications.

However, histone acetylation alone is not sufficient for promoter activation of C4-*Pepc*. Through extensive studies of the tri-methylation of histone H3 lysine 4 it could be shown that this modification acts as an 'on'-marker for transcriptionally active states (Figure 6-1, tissue-specificity presented as green dots). After separation of the two photosynthetic active tissues M and B cells it could be shown that C4-*Pepc* showed high tri-methylation of H3K4 in M cells and high di-methylation in B cells. An inverse pattern was observed for C4-*Me.* Based on this, the conclusion could be made that tri-methylation of H3K4 marks the tissue in which the corresponding gene is activated through light (Danker et al., 2008).

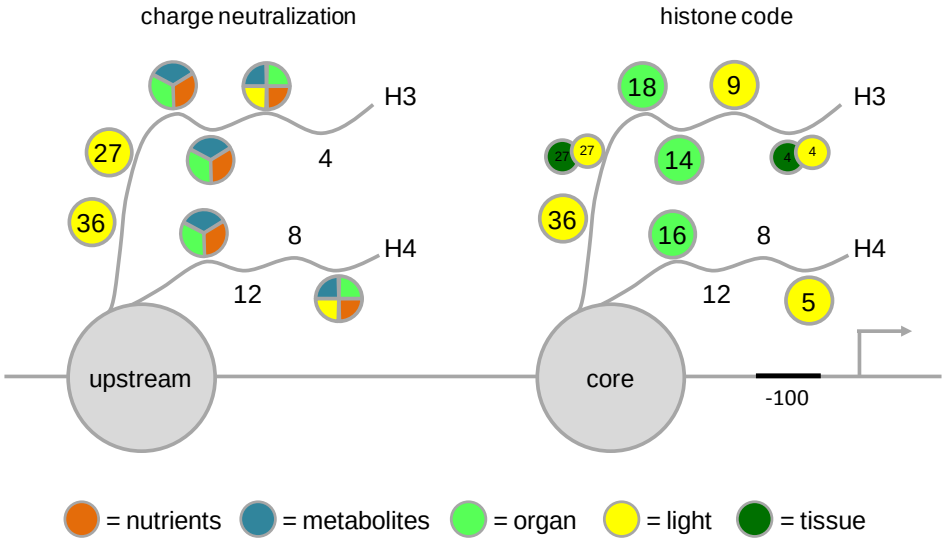


Figure 6-1: Extended histone modification model of C4-*Pepc* from *Zea mays*.
Numbers represent lysine residues on the N-terminal tails of histones H3 and H4. The colors represent signals as listed in the Figure. Core promoter modifications follow the histone code model; upstream promoter modifications follow the charge neutralization model.

Additionally to the findings of Danker et al. (2008) this study revealed that H3K27me3 also seemed to act as B cell specific modification. As an 'off'-marker it was found to be strongly enriched on C4-*Pepc* in B cells around the TIS and within the gene and not regulated by light (Chapter 4, Figure 4-2). It is possible that tissue-specific methylations poise the promoter for light-dependent activation.

Next to the analysis of histone modifications in this study, more chromatin based processes associated with C4-*Pepc* regulation in maize are found in the literature. Tolley et al. (2012) showed how DNA methylation is affected by illumination of the C4-*Pepc* promoter. They identified 4 cytosine residues in the C4-*Pepc* promoter in leaves where they could show that illumination leads to de-methylation of these cytosines, and that de-methylation correlates with C4-*Pepc* expression. Additionally de-methylation of these cytosines was only found to appear in M cells. Surprisingly, they could show that in the 0.6 kb promoter region, which was shown to be required for M-specific expression, no cytosine methylation was found. They rather proposed that unmethylated CpG islands recruit certain proteins which direct H3K4 tri-methylation. It could be shown that H3K4me3 is enriched around the TIS.

Another example of how DNA methylation can influence chromatin changes is DNA methyltransferase-1 (DNMT1). When the event of replication takes place, the state of methylated DNA

is preserved. Recruitment of DNMT1 is essential to prevent the uncontrolled transcription of otherwise silenced genes. The activity of DNMT1 is regulated by two proteins (SET7 and AKT1) which can either methylate or phosphorylate DNMT1 to decrease or increase the stability of the protein. These findings reveal a crosstalk of modifications directly on the DMNT1 protein (Estève et al., 2011, Badeaux and Shi, 2013).

Since it was shown, that all acetylations are light regulated and all methylations are regulated in a tissue-dependent matter, it can be speculated whether this is due to the half-life of these modifications. Cell-type specificity of a modification is defined independently of illumination and can therefore be assumed as long-term memory which is in accordance to the fact that, methylation are supposed to have a longer half-life than acetylation. The turnover rate of a methylation e.g. H3K27m3 was shown to be 3.128 days, whereas the turnover rate of acetylation of histone H3 was measured to be 3 to 30 min. In contrast to that, DNA-methylation can stay a life-time and appears to be mitotically stable (Barth and Imhof, 2010, Badeaux and Shi, 2013).

Less is known about the regulation of histone modifications in the transcribed region of C4 genes. One possible role could be that acetylation in the coding region of the gene is the continuation of the signaling cascade that starts in the promoter region and helps to integrate the signals that finally leads to transcription. Another possible role is that they act along with other modifications in a positive or negative feedback loop via 'crosstalk'. In Chapter 4, a possible role of the so called 'crosstalk' between certain modifications is discussed. The question that remains after the extended study of so far non-characterized histone modifications in maize is, if the charge neutralization vs. histone code hypothesis is still valid! It could be shown that light induced all acetylations in the upstream and core promoter, as well as in the start of the coding region, whereas methylation was not affected by this stimulus and concentrated around the TIS. There is no evidence that the described hypothesis in chapter 3 should not be valid, but to come to a better overall understanding of these modifications, more experiments are needed. When manipulating nitrogen and sugar availability, Offermann et al (2008) were able to show that these treatments only modulated the acetylation state of the upstream promoter region and that this regulation of modification most certainly followed the charge neutralization model. That still needs to be proven for the newly tested modifications. Nevertheless, this study revealed that H3K27ac and H3K36ac are regulated by light also in the upstream promoter region.

The regulation studies of the promoter of maize C4-*Pepc* were further extended to the RNA level. Another factor that might influence C4-*Pepc* expression that was found in the core promoter region was a long noncoding RNA. This lncRNA is regulated by the same stimuli that influence the core promoter like light or tissue specificity, but it showed no response to nutrients and metabolites. This is highly reminiscent of the control of histone acetylations on the C4-*Pepc* promoter. It seems that core promoter histone modification correlates with the presence of the lncRNA. Since we were able to detect the lncRNA in crude chromatin extracts, one possible option for the function is that it keeps the promoter open for the transcription machinery and/or prepares an euchromatic state. Another possibility is that it interacts with other e.g. enhancer elements and transcription is initialized or held because of this interaction. Kalamajka et al. (2003) reported an increase in the accessibility of

promoter chromatin after transcriptional activation. The most prominent example to date in plants is the ‘Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA’ of the Flowering locus C (FLC) described by Heo et al (2011) .The The authors were able to identify a 1.1 kb lncRNA derived from an intron of the Flowering Locus C (FLC) called COLDAIR. It is a good example of how an lncRNA is involved in a negative feedback loop associated with histone modifications. When the FLC gene is actively transcribed COLDAIR levels also increases. In a next step, COLDAIR recruits the chromatin-modifying complex, Polycomb repressive Complex 2 (PRC2) which establishes H3 lysine-27 trimethylation as an ‘off’-marker, by this reducing transcription of the FLC gene. The concomitant reduction of COLDAIR then again results in lower H3K27me3 levels.

Mondal et al. (2010) characterized an intergenic noncoding RNA that also regulated gene expression of the neighboring genes FANK1 and ADAM12 in two different tissues (placenta and brain) and that it was associated with chromatin. The biological relevance of the lncRNA found on the C4-*Pepc* promoter is supported by the strong correlation between hnRNA and lncRNA transcription. However, the actual function still remains unclear.

6.2 Regulation of C4-*Pepc* from *Sorghum bicolor* and *Setaria italica*

To answer the question whether the control of specific modifications is conserved on the orthologous C4-*Pepc* genes of the C4 grasses *Sorghum bicolor* and *Setaria italica*, first of all the Chromatin Immunoprecipitation protocol already established for maize (Haring et al., 2007) needed to be altered towards the requirements of *Sorghum* and *Setaria.* Chromatin Immunoprecipitation (ChIP) is an experimental technique to isolate chromatin and to investigate protein-DNA interactions. The type of ChIP used in my work is called cross-linked ChIP (xChIP). It uses reversibly cross-linked chromatin sheared by sonication. The crosslink is achieved by using formaldehyde as a reversible crosslink agent. Protocol steps that had to be modified were the sort and length of formaldehyde fixation and the chromatin fragmentation (data not shown). After the successful adaptation and establishment of the xChIP protocol, we wanted to analyze promoter histone modifications of photosynthetic genes from *Sorghum* and *Setaria* to answer the question to which extent the histone code had been established during development of C4 metabolism or whether a previously existing code had been recruited into C4. We therefore compared chromatin patterns on C4-*Pepc*. A challenge in these comparative studies was that, whereas the maize promoters had been well defined in previous studies, the *Sorghum* and *Setaria* promoters were just annotated based on automated genome sequence analysis. The definition of a promoter by such technologies is problematic, because promoters lack conserved consensus sequences. Traditionally, promoters were defined by simple sequence elements such as the TATA box (Dikstein, 2011), but a more comprehensive analysis of many promoters in eukaryotes revealed that they often lack TATA boxes and other expected elements (Mencía et al., 2002, Basehoar et al., 2004). The definition of the promoter region is rather complicated, e.g. the size of a promoter most certainly depends on the genome size of the species that is investigated. It is likely to say, that promoters in *Setaria* (genome size 490 Mbp) are shorter than promoters in maize with a genome size of 2300 Mbp (Schnable et al., 2009, Doust et al., 2009) simply due to the limited space between genes. Which set of experiments are suitable to define a promoter region? The size of the promoter region of *Sorghum* and *Setaria* is a point that still needs to be exactly evaluated. The best guess was made by nucleosome occupancy and histone acetylation assays (see Chapter 2). These results indicate that *Sorghum* has an intermediate promoter size compared to maize C4-*Pepc* with highest acetylation peak found -1400 bp of the TIS, whereas *Setaria* has the smallest promoter size of the three investigated species with highest acetylation peak already found -600 bp of the TIS. Still no exact size of the promoter could be defined yet. Further approaches to narrow down the border of a promoter are the investigation on the chromatin level. It was previously shown in maize that the ‘end’ of certain C4 gene promoter was associated with high H3K9 dimethylation and low H3K9 acetylation levels (data not shown). H3K9me2 has been reported to be mainly located in heterochromatic areas in *Arabidopsis thaliana*, while H3K9ac is associated with euchromatic regions (Zhou et al., 2010). Additionally, other genome-wide studies revealed that acetylated histones are found in the promoter and actively transcribed regions whereas H3K4me3 is concentrated around the TIS (Zhou et al., 2010, Barth and Imhof, 2010, Wang et al., 2009a).

The following figure (Figure 6-2) presents the model for the regulation of histone modifications and thus promoter response that could be established for *Sorghum* and *Setaria* in this study.

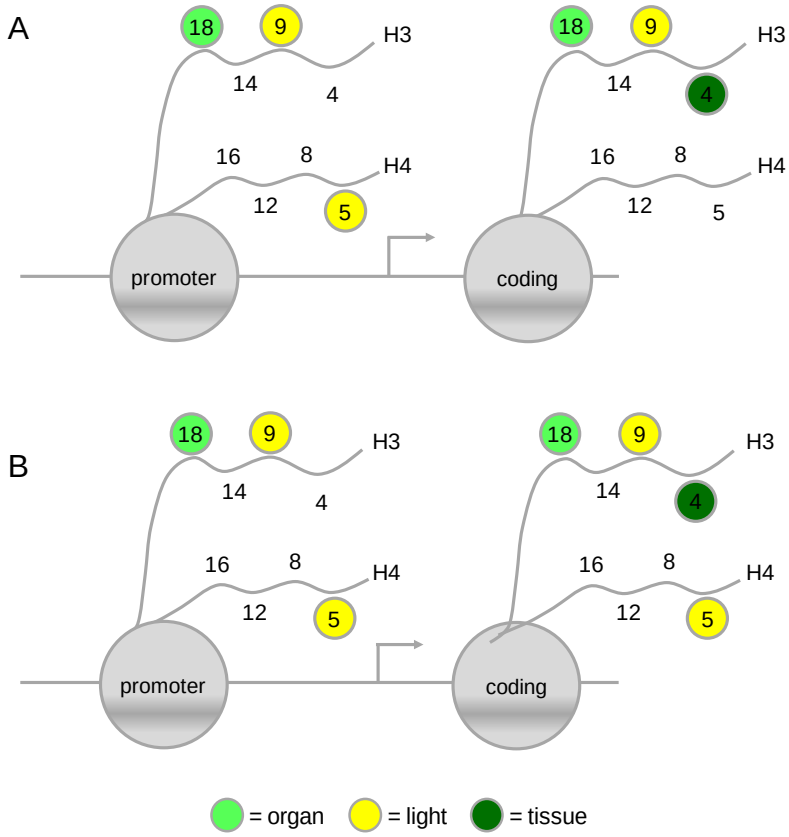


Figure 6-2: Histone modification model of C4-*Pepc Sorghum bicolor* and C4-*Pepc Setaria italica*.
Numbers represent lysine residues on the N-terminal tails of histones H3 and H4. The colors represent signals as listed in the Figure. A, Histone modification model of C4-*Pepc Sorghum bicolor*. B, Histone modification model of C4-*Pepc Setaria italica*.

The functional significance of certain promoter region is suggested by the strong reaction to illumination of certain histone modifications and the concomitant changes in gene transcription (see Chapter 2). The comparative analyses revealed light induction of H3K9ac and H4K5ac in the promoter and coding regions, but also tissue-specific control of H3K4me3 in both species. The only striking difference between the two histone modification models shown in Figure 6-2 is that in *Sorghum*, H4K5ac do not respond to the light stimulus in the coding region. Light did also not induce any change in acetylation levels of H3K18, so an organ specific regulation can be assumed based on the data obtained from maize (see above). To support this assumption, since only leaf tissue was investigated, suitable experiments are e.g. measuring acetylation in different tissues like roots or coleoptiles.

Surprisingly and in contrast to maize, the regulation of modification was not mainly found to be located in the promoter region but also shifted to the start of the coding region. This is probably due to the smaller genome size of *Sorghum* and *Setaria*. A similar distribution of modifications was also found in genome wide studies of Arabidopsis (Wang et al., 2009a)

It further needs to be investigated whether the histone acetylation in *Sorghum* and *Setaria* is only light-dependent or affected by other stimuli. Nitrogen- or sugar repression experiments could help to identify if a certain modification is following the histone code or charge neutralization model. In a definition by Turner et al. (2002), histone modifications which follow the histone code theory need to be set independently of the transcription event which could be a good guideline for further investigations.

We observed a high similarity of regulation of certain histone modifications between maize, *Sorghum* and *Setaria* on C4-*Pepc.* The histone code used to generate a response to specific stimuli seems to be highly conserved. Dependent on the modification and the position on the promoter, histone modifications can help to integrate different stimuli into a promoter response function. Together with data from previous work and the chromatin analyses on additional genes, this leads to the suggestion that histone modifications play an important role in signal integration. These results expand our knowledge about the role of chromatin in signal integration in plants and the complex regulation of photosynthesis in C4 plants.

References

Badeaux AI and Shi Y (2013). Emerging roles for chromatin as a signal integration and storage platform. Nat Rev Mol Cell Biol 14: 211–224.

Baker M (2011). Long noncoding RNAs: the search for function. Nat Meth 8:379-383.

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

Barrand S, Andersen IS and Collas P (2010). Promoter-exon relationship of H3 lysine 9, 27, 36 and 79 methylation on pluripotency-associated genes. Biochenm Bioph Res Co 401: 611–617.

Barth TK and Imhof A (2010). Fast signals and slow marks: the dynamics of histone modifications. Trends Biochem Sci 35: 618–626.

Basehoar AD, Zanton SJ and Pugh BF (2004). Identification and distinct regulation of yeast TATA box-containing genes. Cell 116: 699–709.

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

Biel M, Wascholowski V and Giannis A (2005). Epigenetics-an epicenter of gene regulation: histones and histone-modifying enzymes. Angew Chem Int Edit 44: 3186–3216.

Brennecke J, Stark A, Russell RB, Cohen SM (2005). Principles of microRNA-target recognition. PLOS Biol 3: 404-418.

Brooks A, Farquhar GD (1985). Effect of temperature on the CO_2/O_2 specificity of ribulose-1.5-bisphosphate carboxylase/oxygenase and the rate of respiration in the light. Planta 165: 397–406.

Brueckner F and Cramer P (2008). Structural basis of transcription inhibition by α-amanitin and implications for RNA polymerase II translocation. Nat Struct Mol Biol 15:811-818.

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

Carthew RW and Sontheimer EJ (2009). Origins and mechanisms of miRNAs and siRNAs. Cell 136: 642-655.

Charron J-BF, He H, Elling AA and Deng XW (2009). Dynamic landscapes of four histone modifications during deetiolation in Arabidopsis. Plant Cell 12: 3732–3748.

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

Choi JK and Howe LJ (2009). Histone acetylation: truth of consequences? Biochem Cell Biol 87: 139–150.

Christin PA, Petitpierre B, Salamin N, Buchi L and Besnard G (2009). Evolution of C4 phosphoenolpyruvate carboxykinase in grasses, from genotype to phenotype. Mol Biol Evol 26: 357–365.

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

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

Clayton AL, Hazzalin CA and Mahadevan LC (2006). Enhanced histone acetylation and transcription: A dynamic perspective. Mol Cell 23: 289–296.

Creyghton MP, Cheng AW, Welstead GG, Kooistra T, Carey BW, Steine EJ, Hanna J, Lodato MA, Frampton GM, Sharp PA, Boyer LA, Young RA and Jaenisch R (2010). From the Cover: Histone H3K27ac separates active from poised enhancers and predicts developmental state. PNAS 107: 21931–21936.

Cui Q, Yu Z, Purisima EO, Wang E (2006). Principles of microRNA regulation of a human cellular signaling network. Mol Syst Biol 2:1-7.

Dai Z, Ku MS and Edwards GE (1993). C4 photosynthesis (The CO_2-concentrating mechanism and photorespiration). Plant Physiol 103: 83–90.

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.

Davis CA and Ares M (2006). Accumulation of unstable promoter-associated transcripts upon loss of the nuclear exosome subunit Rrp6p in *Saccharomyces cerevisiae*. Proc Natl Acad Sci USA. 103:3262-3267.

Dikstein R (2011). The unexpected traits associated with core promoter elements. Transcription 2: 201–206.

Dion MF, Altschuler SJ, Wu LF and Rando OJ (2005). Genomic characterization reveals a simple histone H4 acetylation code. PNAS 102: 5501–5506.

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

Dutnall RN (2003). Cracking the Histone Code One, Two, Three Methyls, You're Out! Mol Cell 12: 3–4.

Edwards EJ and Smith SA (2010). Phylogenetic analyses reveal the shady history of C4 grasses. PNAS 107: 2532–2537.

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

Estève PO, Chang Y, Samaranayake M, Upadhyay AK, Horton JR, Feehery GR, Cheng X and Pradhan S (2011). A methylation and phosphorylation switch between an adjacent lysine and serine determines human DNMT1 stability. Nat Struct Biol 18: 42–48.

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

Fischle W, Wang Y and Allis CD (2003). Histone and chromatin crosstalk. Curr Opin Cell Biol 15: 172–183.

Francis Crick (1970). Central Dogma of molecular Biology. Nature, 1970: 561–563.

Fuchs J, Demidov D, Houben A and Schubert I (2006). Chromosomal histone modification patterns – from conservation to diversity. Trends Plant Sci 11: 199–208.

Ghildiyal M and Zamore PD (2009). Small silencing RNAs: an expanding universe. Nat Rev Genet 10:94-108.

Gibb EA, Brown CJ and Lam WL (2011). The functional role of long non-coding RNA in human carcinomas. Mol Cancer 10: 38.

Goodstein DM, Shu S, Howson R, Neupane R, Hayes RD, Fazo J, Mitros T, Dirks W, Hellsten U, Putnam N and Rokhsar DS (2012). Phytozome: a comparative platform for green plant genomics. Nucleic Acids Res. 40 (Database issue): D1178-86.

Gowik U and Westhoff P (2011). The path from C3 to C4 photosynthesis. Plant Physiol 155: 56–63.

Guillemette B, Drogaris P, Lin H-HS., Armstrong H, Hiragami-Hamada K, Imhof A, Bonneil É, Thibault P, Verreault A, Festenstein RJ and Madhani HD (2011). H3 lysine 4 Is acetylated at active gene promoters and is regulated by H3 lysine 4 methylation. PLOS Genet 3: e1001354.

Haag JR and Pikaard CS (2011). Multisubunit RNA polymerases IV and V: purveyors of non-coding RNA for plant gene silencing. Nat Rev Mol Cell Biol 12:483-492.

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. Photosyn Res 75:183-192.

Hahnen S (2004) C3- und C4-Photosynthese in Mais (*Zea mays*): Das Hüllblatt der weiblichen Infloreszenz als Modellsystem zur Analyse photosynthetischer Genexpression.: Dissertation, RWTH Aachen, Institut für Biologie I.

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

Heintzman ND, Stuart RK, Hon G, Fu Y, Ching CW, Hawkins RD, Barrera LO, van Calcar S, Qu C, Ching KA, Wang W, Weng Z, Green RD, Crawford GE and Ren B (2007). Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome. Nat Genet 39: 311–318.

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

Heo JB and Sung S (2011). Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science 331:76-79.

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

Hirota K, Miyoshi T, Kugou K, Hoffman CS, Shibata T and Ohta K (2008). Stepwise chromatin remodelling by a cascade of transcription initiation of non-coding RNAs. Nature 456: 130–134.

Horst I (2009). Die C4-spezifische Phosphoenolpyruvatcarboxylase in Zea mays L.: Modellsystem zur Signalintegration auf Chromatinebene in Pflanzen. Dissertation, RWTH Aachen, Institut für Biologie I.

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

Horst I, Heimann L, Peterhänsel C (2013). Signal integration on plant promoters: A case study in maize. Plant Signal Behav; 8 (9). pii: e25389

Ibrahim DG, Burke T, Ripley BS and Osborne CP (2008). A molecular phylogeny of the genus Alloteropsis (Panicoideae, Poaceae) suggests an evolutionary reversion from C4 to C3 photosynthesis. Annals of Botany 103: 127–136.

Jang I-C, Chung PJ, Hemmes H, Jung C and Chua N-H (2011). Rapid and Reversible Light-Mediated Chromatin Modifications of Arabidopsis Phytochrome A Locus. Plant Cell 2: 459–470.

Jiao Y and Meyerowitz EM (2010). Cell-type specific analysis of translating RNAs in developing flowers reveals new levels of control. Mol Syst Biol 6: 419.

Kalamajka R, Hahnen S, Cavalar M, Töpsch S, Weier D, Peterhänsel C (2003). Restriction accessibility in isolated nuclei reveals light-induced chromatin reorganization at the PEPC promoter in maize. Plant Mol Biol 52:669-678.

Kanai R and Edwards GE (1999). The Biochemistry of C4 Photosynthesis.: C4 Plant Biology. R. F. Sage and R. K. Monson. San Diego, Academic Press: 49–87.

Kanhere A, Viiri K, Araújo CC, Rasaiyaah J, Bouwman RD, Whyte WA, Pereira CF, Brookes E, Walker K, Bell GW, Pombo A, Fisher AG, Young RA, Jenner RG (2010). Short RNAs are transcribed from repressed polycomb target genes and interact with Polycomb Repressive Complex-2. Mol Cell 38:675-688.

Kapranov P, Cheng J, Dike S, Nix DA, Duttagupta R, Willingham AT, Stadler PF, Hertel J, Hackermüller J, Hofacker IL, Bell I, Cheung E, Drenkow J, Dumais E, Patel S, Helt G, Ganesh M, Ghosh S, Piccolboni A, Sementchenko V, Tammana H, Gingeras TR (2007). RNA maps reveal new RNA classes and a possible function for pervasive transcription. Science 316:1484-1488.

Kausch AP, Owen TP, Zachwieja SJ, Flynn AR, Sheen J (2001). Mesophyll-specific, light and metabolic regulation of the C4 PPCZm1 promoter in transgenic maize. Plant Mol Biol 45:1-15.

Khalil AM, Guttman M, Huarte M, Garber M, Raj A, Rivea Morales D, Thomas K, Presser A, Bernstein BE, van Oudenaarden A, Regev A, Lander ES and Rinn JL (2009). Many human large intergenic noncoding RNAs associate with chromatin-modifying complexes and affect gene expression. PNAS 106: 11667–11672.

Kim ED and Sung S (2012). Long noncoding RNA: unveiling hidden layer of gene regulatory networks. Trends Plant Sci 17: 16–21.

Koussevitzky S, Nott A, Mockler TC, Hong F, Sachetto-Martins G, Surpin M, Lim J, Mittler R and Chory J (2007). Signals from chloroplasts converge to regulate nuclear gene expression. Science 316: 715–719.

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

Krogan NJ, Kim M, Tong A, Golshani A, Cagney G, Canadien V, Richards DP, Beattie BK, Emili A, Boone C, Shilatifard A, Buratowski S and Greenblatt J (2003). Methylation of Histone H3 by Set2 in *Saccharomyces cerevisiae* Is Linked to Transcriptional Elongation by RNA Polymerase II. Mol Cell Biol 23: 4207–4218.

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

Lee JT (2012). Epigenetic regulation by long noncoding RNAs. Science 338:1435-1439.

Li B, Carey M and Workman JL (2007). The role of chromatin during transcription. Cell 128: 707–719.

Li P and Brutnell TP (2011). *Setaria viridis* and *Setaria italica*, model genetic systems for the Panicoid grasses. J Exp Bot 62: 3031–3037.

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. Nat Genet 42:1060-1067.

Liu B, Fedak G, Liu ZL, Han FP, Tan M, Shan XH, Dong YZ, Wang XZ and Hao S (2004). Activation of a rice endogenous retrotransposon Tos17 in tissue culture is accompanied by cytosine demethylation and causes heritable alteration in methylation pattern of flanking genomic regions. Theor Appl Genet 109: 200–209.

Lucia F and Dean C (2011). Long non-coding RNAs and chromatin regulation. Curr Opin Plant Biol 14:168-173.

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

Lusser A, Kölle D, Loidl P (2001). Histone acetylation: lessons from the plant kingdom. Trends Plant Sci 6: 59–65.

Martens JA, Laprade L and Winston F (2004). Intergenic transcription is required to repress the *Saccharomyces cerevisiae* SER3 gene. Nature 429: 571–574.

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

Mauro-Herrera M, Wang X, Barbier H, Brutnell TP, Devos KM and Doust AN (2013). Genetic Control and Comparative Genomic Analysis of Flowering Time in Setaria (Poaceae). G3 Genes, Genomes, Genetics 2: 283–295.

Mellor J (2005). The dynamics of chromatin remodeling at promoters. Mol Cell 2: 147–157.

Mencía M, Moqtaderi Z, Geisberg JV, Kuras L and Struhl K (2002). Activator-specific recruitment of TFIID and regulation of ribosomal protein genes in yeast. Mol Cell 4: 823–833.

Mondal T, Rasmussen M, Pandey GK, Isaksson A and Kanduri C. (2010). Characterization of the RNA content of chromatin. Genome Res 7: 899–907.

Morris SA, Rao B, Garcia BA, Hake SB, Diaz RL, Shabanowitz, J, Hunt DF, Allis CD, Lieb JD and Strahl BD (2006). Identification of Histone H3 Lysine 36 Acetylation as a Highly Conserved Histone Modification. J Biol Chem 10: 7632–7640.

Narlikar GJ, Fan H-Y and Kingston RE (2002). Cooperation between complexes that regulate chromatin structure and transcription. Cell 4: 475–487.

Nightingale KP, Gendreizig S, White DA, Bradbury C, Hollfelder F and Turner BM (2006). Cross-talk between Histone Modifications in Response to Histone Deacetylase Inhibitors. J Biol Chem 7: 4408–4416.

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 Physiol 141:1078-1088.

Offermann S, Dreesen, B, Horst I, Danker, T, Jaskiewicz M, Peterhänsel 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.

Ørom UA, Derrien T, Beringer M, Gumireddy K, Gardini A, Bussotti G, Lai F, Zytnicki M, Notredame C, Huang Q, Guigo R and Shiekhattar R (2010). Long noncoding RNAs with enhancer-like function in human cells. Cell 1: 46–58.

Pang C, Gasteiger E and Wilkins MR (2010). Identification of arginine- and lysine-methylation in the proteome of Saccharomyces cerevisiae and its functional implications. BMC Genomics 1: 92.

Patel M and Berry JO (2007). 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, Schmutz J, Spannagl M, Tang H, Wang X, Wicker T, Bharti AK, Chapman J, Feltus FA, Gowik U, Grigoriev IV, Lyons E, Maher CA, Martis M, Narechania A, Otillar RP, Penning BW, Salamov AA, Wang Y, Zhang L, Carpita NC, Freeling M, Gingle AR, Hash CT, Keller B, Klein P, Kresovich S, McCann MC, Ming R, Peterson DG, Mehboob-ur-Rahman, Ware D, Westhoff P, Mayer KFX, Messing J and Rokhsar DS (2009). The Sorghum bicolor genome and the diversification of grasses. Nature 7229: 551–556.

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

Pokholok DK, Harbison CT, Levine S, Cole M, Hannett NM, Lee TI, Bell GW, Walker K, Rolfe PA, Herbolsheimer E, Zeitlinger J, Lewitter F, Gifford DK, Young RA (2005). Genome-wide map of nucleosome acetylation and methylation in yeast. Cell 122:517-527.

Ponjavic J, Ponting CP, Lunter G. (2007). Functionality or transcriptional noise? Evidence for selection within long noncoding RNAs. Genome Res 17:556-565.

Reed R (2003). Coupling transcription, splicing and mRNA export. Curr Opin in Cell Biol 15:326-331.

Rinn JL and Chang HY (2012). Genome regulation by long noncoding RNAs. Annu Rev Biochem 81:145-166.

Sage RF, Sage TL and Kocacinar F (2012). Photorespiration and the Evolution of C4 Photosynthesis. Annu Rev Plant Biol 1: 19–47.

Schäffner AR and Sheen J (1992). Maize C4 photosynthesis involves differential regulation of phosphoenolpyruvate carboxylase genes. Plant J 2: 221–232.

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.

Shahbazian MD and Grunstein M (2007). Functions of Site-Specific Histone Acetylation and Deacetylation. Annu Rev Biochem 1: 75–100.

Sheen J (1990). Metabolic repression of transcription in higher plants. Plant Cell 10: 1027–1038.

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

Sheen JY and 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.

Shi Y, Lan F, Matson C, Mulligan P, Whetstine JR, Cole PA, Casero RA and Shi Y (2004). Histone demethylation mediated by the nuclear amine oxidase homolog LSD1. Cell 7: 941–953.

Suganuma T and Workman JL (2008). Crosstalk among Histone Modifications. Cell 4: 604–607.

Sugiharto B and T Sugiyama (1992). Effects of nitrate and ammonium on gene expression of phosphoenolpyruvate carboxylase and nitrogen metabolism in maize leaf tissue during recovery from nitrogen stress. Plant Phys 98: 1403–1408.

Taiz L and Zeiger E (2007). Plant Physiology. Spektrum Akademischer Verlag, 4. Edition.

Tolley BJ, Woodfield H, Wanchana S, Bruskiewich R and Hibberd JM (2012). Light-regulated and cell-specific methylation of the maize PEPC promoter. J Exp Bot 3: 1381–1390.

Tsai MC, Manor O, Wan Y, Mosammaparast N, Wang JK, Lan F, Shi Y, Segal E, Chang HY (2010). Long noncoding RNA as modular scaffold of histone modification complexes. Science 329:689-693.

Tsao DC, Park NJ, Nag A and Martinson HG (2012). Prolonged α-amanitin treatment of cells for studying mutated polymerases causes degradation of DSIF160 and other proteins. RNA 2: 222–229.

Turner BM (2002). Cellular memory and the histone code. Cell 3: 285–291.

van Bakel H, Nislow C, Blencowe BJ and Hughes TR (2010). Most "dark matter" transcripts are associated with known genes. PLOS Biol 5: e1000371.

Vicentini A, Barber JC, Aliscioni SS, Giussani LM and Kellogg EA (2008). The age of the grasses and clusters of origins of C 4 photosynthesis. Glob Change Biol 12: 2963–2977.

Wagner EJ and Carpenter PB (2012). Understanding the language of Lys36 methylation at histone H3. Nat Rev Mol Cell Biol 2: 115–126.

Wang X, Elling AA, Li X, Li N, Peng Z, He G, Sun H, Qi Y, Liu XS and 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 4: 1053–1069.

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

Wiludda C, Schulze S, Gowik U, Engelmann S, Koczor M, Streubel M, Bauwe H, Westhoff P (2012). Regulation of the photorespiratory GLDPA gene in C4 Flaveria: An intricate interplay of transcriptional and posttranscriptional processes. Plant Cell 24:137-151.

Wilusz JE, Sunwoo H, Spector DL (2009). Long noncoding RNAs: functional surprises from the RNA world. Gene Dev 23:1494-1504.

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.

Wutz A (2011). Gene silencing in X-chromosome inactivation: advances in understanding facultative heterochromatin formation. Nat Rev Genet 12: 542-553.

Wyrich R, Dressen U, Brockmann S, Streubel M, Chang C, Qiang D, Paterson AH and Westhoff P (1998). The molecular basis of C4 photosynthesis in sorghum: isolation, characterization and RFLP mapping of mesophyll- and bundle-sheath-specific cDNAs obtained by differential screening. Plant Mol Biol 2: 319–335.

Yamada K, Lim J, Dale JM, Chen H, Shinn P, Palm CJ, Southwick AM, Wu HC, Kim C, Nguyen M, Pham P, Cheuk R, Karlin-Newmann G, Liu SX, Lam B, Sakano H, Wu T, Yu G, Miranda M, Quach HL, Tripp M, Chang CH, Lee JM, Torium, M, Chan MMH, Tang CC, Onodera CS, Deng JM, Akiyama K, Ansari Y, Arakawa T, Banh J, Banno F, Bowser L, Brooks S, Carninci P, Chao Q, Choy N, Enju A, Goldsmith AD, Gurjal M, Hansen NF, Hayashizaki Y, Johnson-Hopson C, Hsuan VW, Iida K, Karnes M, Khan S, Koesema E, Ishida J, Jiang PX, Jones T, Kawai J, Kamiya A, Meyers C, Nakajima M, Narusaka M, Seki M, Sakurai T, Satou M, Tamse R, Vaysberg M, Wallender EK, Wong C, Yamamura Y, Yuan,S, Shinozaki K, Davis RW, Theologis A and Ecker JR (2003). Empirical analysis of transcriptional activity in the Arabidopsis genome. Science 5646: 842–846.

Yang X-J and Seto E (2008). Lysine acetylation: codified crosstalk with other posttranslational modifications. Mol Cell 4: 449–461.

Yanagisawa S and Izui K (1989). Maize phosphoenolpyruvate carboxylase involved in C4 photosynthesis: nucleotide sequence analysis of the 5' flanking region of the gene. J Biochem 106: 982-987.

Yoshida S, Ito M, Nishida I and Watanabe A (2001). Isolation and RNA gel blot analysis of genes that could serve as potential molecular markers for leaf senescence in Arabidopsis thaliana. Plant Cell Physiol 2: 170–178.

Zhang X (2008). The epigenetic landscape of plants. Science 5875: 489–492.

Zhou J, Wang X, He K, Charron J-BF, Elling AA and 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 Mol Biol 6: 585–595.

APPENDIX

Publications and manuscripts

The manuscript entitled ‘A common histone modification code on C4 genes in maize and its conservation in Sorghum and *Setaria italica*’ by Louisa Heimann, Ina Horst, Renke Perduns, Björn Dreesen, Sascha Offermann and Christoph Peterhänsel has been published in Plant Physiology, May 2013, Vol. 162, pp.456-469. Samples for Figure 3 A were prepared and data were evaluated by Louisa Heimann. Except from data obtained by Ina Horst and Renke Perduns all laboratory work as well as data evaluation for *Sorghum* and *Setaria* samples was performed by Louisa Heimann as well as preparation of Figure 4-7. All authors took part in writing the article, while Christoph Peterhänsel was the major contributor.

The manuscript entitled ‘Signal integration on plant promoters: A case study in maize’ by Ina Horst, Louisa Heimann and Christoph Peterhänsel has been submitted to Plant Singaling & Behavior and is currently under revision. Figure 1C was created by Louisa Heimann. All authors took part in writing while Christoph Peterhänsel was the major contributor.

The manuscript of Chapter 4 is entitled ‘Histone modification profiling on the C4- phosphoenolpyruvate carboxylase promoter in *Zea mays* reveals a light and tissue dependent regulation*’.*
Sample preparation as well as data evaluation for all Figures was done by Louisa Heimann. All authors took part in writing the article, while Louisa Heimann was the major contributor.

The manuscript of Chapter 5 is entitled ‘ A long noncoding RNA associated with C4-phosphoenolpyruvate carboxylase promoter chromatin in *Zea mays’* by Louisa Heimann, Desirée Waidmann, Sascha Offermann, Ina Horst and, Christoph Peterhänsel. Samples for Figure 1, Figure 2, Figure 3 A, C and Figure 5 were prepared by Louisa Heimann. Additionally all Figures were created and all data were evaluated by Louisa Heimann. Louisa Heimann and Christoph Peterhänsel took part in writing, while Louisa Heimann was the major contributor.

List of figures

List of publications

Heimann L, Horst I, Perduns R, Dreesen B, Offermann S, Peterhänsel 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

Horst I, **Heimann L**, Peterhänsel C (2013). Signal integration on plant promoters: A case study in maize. Plant Signal Behav; 8 (9). pii: e25389

Danksagung

Herrn Prof. Dr. Christoph Peterhänsel danke ich für die Möglichkeit diese Doktorarbeit an seinem Institut erstellen zu können, für die wissenschaftliche Betreuung und für die Übernahme des Hauptreferates der vorliegenden Arbeit.

Herrn Prof. Dr. Hans-Peter Braun und Prof. Dr. Jutta Papenbrock danke ich für die Übernahme des Zweitgutachtens und des Vorsitzes.

Lebenslauf

Persönliche Angaben	
Name	Louisa Heimann
Geburtsdatum	29.März 1984
Geburtsort	Hagen
Familienstand	ledig
Hochschulausbildung	
Seit 12/2009	Wissenschaftliche Angestellte am Institut für Botanik der Leibniz Universität Hannover (Promotion)
07/2007-08/2009	Master Studium „Chemische Biologie“ an der Technischen Universität Dortmund
10/2003-07/2007	Bachelor Studium „Chemische Biologie“ an der Technischen Universität Dortmund
Schulausbildung	
1994-2003	Theodor-Heuss-Gymnasium in Hagen
1990-1994	Grundschule in Hagen