Ghent University

Faculty of Sciences

Department of Plant Biotechnology and Bioinformatics

Enhancement of biomass production and accessibility of the cell wall for fermentation in *Brachypodium distachyon* as a model and *Zea mays* as a crop

Wannes Voorend

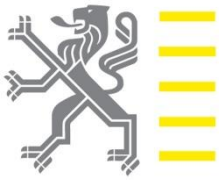
Thesis submitted as fulfillment to the requirements for the degree of
PhD in Sciences: Biotechnology & Biochemistry
Academic year 2013-2014

Promotor: Prof. Dr. Dirk Inzé[1,2]
Copromotor: Dr. Hilde Muylle[3]

[1] Department of Plant Systems Biology, VIB, Technologiepark 927, 9052 Gent, Belgium
[2] Department of Plant Biotechnology and Bioinformatics, Ghent University, Technologiepark 927, 9052 Gent, Belgium

[3] Plant Sciences Unit – Growth and Development, Institute for Agricultural and Fisheries Research (ILVO), Caritasstraat 21, 9090 Melle, Belgium

This work has been conducted at the University of Ghent, department of Plant Systems Biology in the Flanders Institute for Biotechnology (VIB) and at the Plant Sciences Unit, Growth and Development group in the Institute for Agricultural and Fisheries Research (ILVO). The research was funded by the agency for Innovation by Science and Technology (IWT), the European seventh framework program (SUNLIBB), Ghent University, the Flanders Institute for Biotechnology (VIB) and the Institute for Agricultural and Fisheries Research (ILVO).



agency for Innovation
by Science and Technology

Exam Commission

Promotor

Prof. Dr. **Dirk Inzé**
Email: dirk.inze@psb.vib-ugent.be
Department of Plant Biotechnology and Bioinformatics, Ghent University

Copromotor

Dr. **Hilde Muylle**
Email: hilde.muylle@ilvo.vlaanderen.be
Plant Sciences Unit – Growth and Development, Institute for Agricultural and Fisheries Research

Chair

Prof. Dr. **Wout Boerjan**
Department of Plant Biotechnology and Bioinformatics, Ghent University

Other members of the jury (* members of the reading committee)

Prof. Dr. **Frank Van Breusegem***
Department of Plant Biotechnology and Bioinformatics, Ghent University

Prof. Dr. **Geert Haesaert***
Department of Applied biosciences, Ghent University

Prof. Dr. **Tom Desmet***
Department of bioengineering, Ghent University

Dr. **David Caparrós-Ruiz ***
Centre for Research in Agricultural Genomics (CRAG), Barcelona, Spain

Prof. Dr. **Mieke Van Lijsebettens**
Department of Plant Biotechnology and Bioinformatics, Ghent University

Prof. Dr. **Isabel Roldán-Ruiz**
Plant Sciences Unit – Growth and Development, Institute for Agricultural and Fisheries Research

Dr. **Ruben Vanholme**
Department of Plant Biotechnology and Bioinformatics, Ghent University

Dr. **Hilde Nelissen**
Department of Plant Biotechnology and Bioinformatics, Ghent University

1. List of abbreviations

4CL: 4-coumarate:CoA ligase
ADF: acid detergent fiber
ADL: acid detergent lignin
AGP: Arabinogalactan protein
AMP: adenosine monophosphate
At : *Arabidopsis thaliana*
BAR: basta resistance
BC: backcross
Bd: *Brachypodium distachyon*
BK2: brittle stalk2
BLAST: Basic Local Alignment Search Tool
bm: *brown midrib*
bmr: *brown midrib*
bp: base pairs
bs: *brown stem*
C3H: p-coumarate 3-hydroxylase
C4H: cinnamate 4-hydroxylase
CAD: cinnamyl alcohol dehydrogenase
CaMVp35S: cauliflower mosaic virus 35S promoter
CCoAOMT: caffeoyl-CoA O-methyltransferase
CCR: cinnamoyl-CoA reductase
cDNA: copy-DNA
CEC: compact embryogenic callus
CESA: cellulose synthase
CHI: chalcone isomerase
CHS: chalcone synthase
CK: cytokinin
COMT: caffeic acid O-methyltransferase
CSE: caffeoyl shikimate esterase
CSL: CESA-like
CSPP: candidate substrate-product pair
CWR: cell wall residue
DAPI: 4',6-diamidino-2-phenylindole
DFR: dihydroflavonol 4-reductase
diFA: diferulic acid
DIMBOA: 2,4-dihydroxy-7-methoxy-2H-1,4-benzoxazin-3(4H)-one
DNA: Deoxyribonucleic acid
dsRNA: double stranded RNA
DW: dry weight
Ec: conductivity
EF1a: elongation factor 1-alpha
EMS: Ethyl Methanesulfonate

EST: expressed sequence tag
EtOH: ethanol
F3'H: flavonoid 3-hydroxylase
F3H: naringenin 3-dioxygenase
F5H: ferulate 5-hydroxylase
FA: ferulic acid
FAO: Food and Agriculture Organisation
FLS: flavonol synthase
G: guaiacyl
GA : gibberellic acid
GA20ox1: gibberellic acid 20 oxidase1
GAX: glucoarabinoxylan
GFP: Green Fluorescent protein
GO: gene ontology
GOD-POD: glucose oxidase-peroxidase
H: *p*-hydroxyphenyl
HBOA: 2-hydroxy-2*H*-1,4-benzoxazin-3(4*H*)-one
HCl: hydrogen chloride
HCT: hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase
HMBOA: 2-hydroxy-7-methoxy-1,4-benzoxazin-3-one
HPT: Hygromycin Phosphotransferase
IBI: International Brachypodium Initiative
ILVO: Instituut voor Landbouw en Visserij Onderzoek; Institute for Agricultural and Fisheries Research
INRA-IJPB : Institut National de Recherche Agronomique – Institut Jean-Pierre Bourgin
IVDMD: *in vitro* dry matter digestibility
IVNDFD: *in vitro* NDF digestibility
LAC: laccase
LED: leaf elongation duration
LER: leaf elongation rate
LERmax: maximal leaf elongation rate
Lm: final leaf length
Mb: mega base pairs
miRNA: microRNA
MLG: mixed-linkage glucan
Mu: *Mutator*
NaOH: sodium hydroxide
NDF: neutral detergent fiber
NIRS: near infrared reflectance spectroscopy
NMR: nuclear magnetic resonance
NT: non-transgenic
PAL: phenylalanine ammonia lyase
PAT: phosphoribosylanthranilate transferase

PC: principle component
*p*CA: *p*-coumaric acid
PCR : polymerase chain reaction
PER: peroxidase
PME: pectin methyl esterase
PSB: Plant Systems Biology
PTF: Plant Transformation Facility
pUBIL: Ubiquitin-long promoter
qRT-PCR: quantitative real-time PCR
RNA: ribonucleic acid
RNAi: RNA interference
rpm: rounds per minute
S: silking
S: syringyl
S+14d: fourteen days after silking
S+7d: seven days after silking
sRNA: small RNA
T: transgenic
t10%, t50%, t90%, te: time points at which the leaf reaches 10%, 50%, 90% and 100% of the final leaf length, respectively
t100: time point at which the leaf reaches 100 mm
TE: transposable element
TFA: trifluoro acetic acid
TILLING: Targeting Induced Local Lesions in Genomes
tm: time point at which the leaf reaches LERmax
UBI: ubiquitin
UPLC-MS: Ultra Performance Liquid Chromatography-Mass Spectroscopy
URGV : Unité de Recherches en Génétique Végétale
UTR: untranscribed region
VIB: Vlaams Instituut voor Biotechnologie; Flemisch Institute for Biotechnology
VIGS: virus-induced gene silencing
Vx: x visible leaf collars
Zm: *Zea mays*

2. Summary

In this PhD, the focus was set on the improvement of the lignocellulosic feedstock that can be used for the production of second generation bioethanol by improving biomass yield and biomass quality. Nevertheless, the insights gained in this work are also applicable to forage improvement. Two strategies to improve lignocellulosic biomass crops were explored: improving the amount of fermentable sugars that can be released from the biomass by lignin engineering and improving the dry matter yield, by boosting the innate growth potential of the crop. In both fields, fundamental insights have been gained using the dicot model plant *Arabidopsis thaliana*. Yet, a major potential feedstock for the production of second generation bioethanol and thus of the bio-based economy are fast growing grasses that are characterized by high yields per hectare and year and high resource use efficiency. Therefore, in this dissertation, a translational approach was taken from *A. thaliana* to *Brachypodium distachyon* (purple false brome), a recently proposed model plant for the family of the grasses and *Zea mays* spp. *mays* or maize, an economically important crop species providing a major source of biomass globally.

First, a transgenic approach was followed for growth enhancement and lignin perturbation in *B.distachyon* and maize. Overexpression of the *AtGA20ox1* gene in *B. distachyon* did not result in increased leaf length. This was unexpected since overexpression of the same gene in maize did result in increased leaf length and plant height. Transgenic approaches for improvement of the saccharification efficiency included downregulation of the lignin biosynthetic gene *cinnamyl alcohol dehydrogenase* (*CAD*), both in *B. distachyon* and maize. We were successful in generating transgenic Brachypodium and maize plants with expression of the silencing construct. However, the level of downregulation of the target genes *BdCAD1* for *B. distachyon* and *ZmCAD2* in maize was limited and dependent on the tissue and developmental stage. A phenotypic analysis of selected maize transgenics showed that the level of downregulation was insufficient to cause a reduction in lignin amount or increased saccharification.

Alternatively, the effect of mutations in lignin biosynthesis genes on cell wall characteristics was analyzed. In *B. distachyon*, several mutants in *4-coumarate:CoA Ligase1* (*Bd4CL1*) were identified by a TILLING approach. However, no mutations leading to a premature stop in the transcript sequence could be detected. Mutants with a predicted reduced activity of the Bd4CL1 protein were analyzed for lignin quantity and saccharification efficiency but no significant changes were found, possibly because of sufficient residual activity or redundancy in the 4CL family of *B. distachyon*. On the other hand, a TILLING mutant in *BdCAD1*, identified by a red-brown coloration of the stem, had increased saccharification efficiency. By means of an expression analysis, we showed that an auto regulatory mechanism might be present, regulating expression of *CAD* family members in *B. distachyon*.

In maize, two perturbations in lignin biosynthesis genes were analyzed for cell wall characteristics. The first was the disruption of *cinnamate-4-hydroxylase1* (*ZmC4H1*), coding for an enzyme that functions in an early step in the lignin biosynthetic pathway. This resulted in reduction of the amount of lignin which was compensated for by an increase of a comparable amount of hemicellulose in the cell wall of *zmc4h1* plants. Interestingly, the saccharification efficiency of *zmc4h1* plants was increased when the biomass was pretreated with acid but no increase in the saccharification efficiency was observed without pretreatment. The second was disruption of

ZmCAD2, of which the enzyme catalyzes the last step in the production of monolignols, the building blocks of lignin. These plants had a decreased amount of cell wall due to a decrease of lignin and cellulose amount, but not of hemicellulose. The stem biomass of *zmcad2* plants showed increased saccharification efficiency, both with and without acid pretreatment.

In addition, we followed a systems biology approach to interrogate the effects of lignin perturbation in maize, such as disruption of *C4H* and *CAD*, on gene expression levels to provide deeper inside into lignin biosynthesis and the complex metabolic network it is embedded in. Systems-wide gene expression profiling showed that expression of genes involved in cell wall modification and flavonoid biosynthesis was affected, besides lignin biosynthesis genes. We also revealed the presence of a general stress response. Phenolic profiling revealed metabolic shifts in these lignin mutants, such as altered abundance of benzoxazinoids, benzenoids, flavonoids, flavonolignols and oligolignols.

Lignin mutants have been associated with pleiotropic effects such as slower growth and reduced biomass yield. To be able to monitor the growth of grasses, we developed a user-friendly tool that can be used for fast, reliable and robust analysis of large leaf elongation datasets. This tool can be used to quantify leaf growth parameters at the organ and at the cellular level in response to genetic perturbation, ectopic expression or different environmental conditions or to identify inter-genotype differences.

A final goal of this PhD dissertation was to interrogate the connection between biomass quantity and biomass quality of lignocellulosic feedstock. Here we investigated whether increased biomass yield could have an effect on the biomass quality. Indeed, we found that overproduction of GA, resulting in higher growth rates and more stem biomass, has an effect on cell wall properties such as higher amount of lignin and reduced saccharification efficiency. These results show that biomass quantity and biomass quality are interconnected, which is important for future breeding strategies.

With this work, we showed that a biotechnology approach at the level of individual genes can be applied for improvement of lignocellulosic feedstock. However, single-gene based approaches most likely do not administer the major increase in biomass production and saccharification efficiency as would be necessary to make bioethanol production from lignocellulosic biomass economically profitable. Ultimately, combinations of improved plant growth traits with combinations of improved biomass degradation traits by gene stacking could form the basis of the next generation of bioenergy crops. Finally, the analysis of complex biological processes such as organ growth regulation and lignification of the secondary cell wall is best studied using a combination of data at the transcript, protein, metabolite and phenotype level, termed as ‘systems biology’. With the current progress in metabolic profiling of the maize stem, partly discussed in this work, I believe that maize can form a powerful model for systems biology in monocot species.

Table of Contents

Chapter 1: General introduction

Chapter 1: General introduction

1. Transition to a bio-based economy: from fossil fuel to renewable resources

Sustainable development is the organization of human life on our finite planet in a way that it does not only meet current but also future generations' needs and aspirations. This definition of sustainable development formed the basis of the Brundtland report "Our Common Future" (World Commission on Environment and Development 1987), which placed climate, social and economic issues on the political agenda in the form of striving for a sustainable development. Transforming our currently petroleum-based economy into a bio-based economy would greatly benefit the sustainable society (Ragauskas et al. 2006; Vanholme et al. 2013a). The bio-based economy is an economy based on materials, chemicals and energy from renewable resources instead of fossil resources. As part of an integrated approach of employing alternative resources such as solar, wind, geothermal, hydroelectric, and wave energy, liquid biofuels are developed to provide transportation energy. It is believed that plant biomass is the only suitable and renewable resource than can provide alternative transportation fuels such as bioethanol or biodiesel in the short-term (Sun and Cheng 2002; Hamelinck et al. 2005). The current so-called first generation biofuels are derived mainly from corn starch and edible vegetable oil, for the production of bioethanol and biodiesel respectively. Yet, first generation biofuels provide little or no greenhouse gas reduction once all impacts of cultivation, including indirect land use change (ILUC) which occurs when existing cropland is used for biomass cultivation or natural areas are converted into cropland, and processing are taken into account (European Academies Science Advisary Council 2012). Furthermore, first generation biofuels, produced from biomass sources that can be used for food and feed, are subject to the food versus fuel debate (Williams 2008; Valentine et al. 2012). Therefore, second-generation biofuels, including cellulosic ethanol were proposed as alternative. Cellulosic ethanol does not directly compete with the food chain since it can be produced from non-edible lignocellulosic plant parts (Valentine et al. 2012). A major potential feedstock for the production of second generation bioethanol and thus of the bio-based economy are fast growing grasses that are characterized by high yields per hectare and year and high resource use efficiency (Andersen et al. 2008; Carpita and McCann 2008; Corredor et al. 2009; Feltus and Vandenbrink 2012; Jung et al. 2012; van der Weijde et al. 2013). In this dissertation, *Brachypodium distachyon* (purple false brome), a recently proposed model plant for the family of the grasses and *Zea mays* spp. *mays* or maize, an economically important crop species providing a major source of biomass globally, formed the subjects for studying strategies to improve lignocellulose feedstock for the production of bioethanol.

2. The Poaceae

The grass family (*Poaceae*) provides a major portion of human nutrition in the form of cereal grains. In addition, domestic animals are raised on diets based partly or wholly on grasses. The three subfamilies that contain the species of highest economic relevance are 1) the *Panicoideae* including maize (*Zea mays*), sorghum (*Sorghum bicolor*), common millet (*Panicum miliaceum*), pearl millet (*Pennisetum glaucum*), Shama millet (*Echinochloa colona*), foxtail millet (*Setaria italica*), switchgrass (*Pabicum viragatum*) and *Miscanthus* spp., 2) the *Ehrhartoideae* including rice (*Oryza*

spp.) and 3) the *Pooideae*, including wheat (*Triticum* spp.), barley (*Hordeum vulgare*), oat (*Avena sativa*), rye (*Secale cereale*) and many pasture grasses such as fescue (*Festuca* spp.) and ryegrass (*Lolium* spp.) (Figure 1).

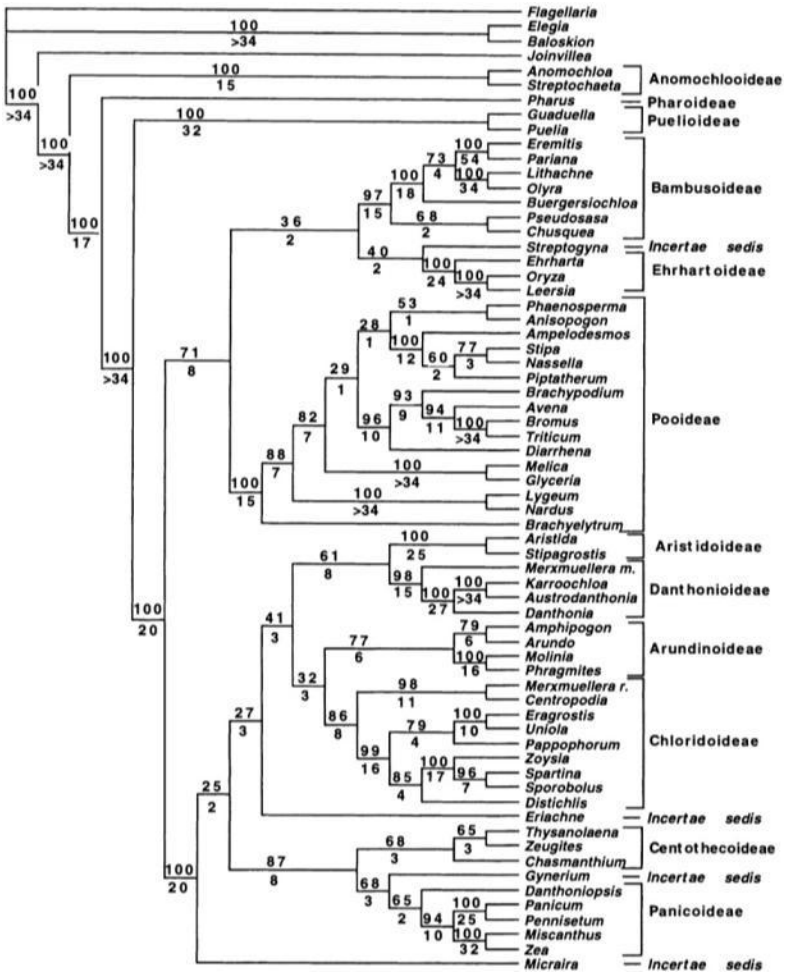


Figure 1. Phylogenetic relationships in the *Poaceae*. The image shown is a single most parsimonious tree for the grasses and relatives, based on eight sets of data (chloroplast restriction sites, *ndh*F, *rbc*I, *rbo*C2, phytochrome B, chloroplast restriction site polymorphisms, ITS and GBSSI). Figures above branches are bootstrap support values and figures below are Bremer support values. Brackets indicate the botanical classification for the *Poaceae*. From Grass Phylogeny Working Grp (2001),

On a global scale, 703 million hectares or 47% of the total arable area is used for the cultivation of cereals including rice, maize, wheat, barley, sorghum and millet. Another 45 million hectares, cultivated with maize, ryegrass, Sorghum and other grasses such as bluegrass (*Poa* spp.), Columbus grass (*Sorghum almum*), fescue, elephant grass (*Pennisetum purpureum*), orchard grass (*Dactylis glomerata*), Sudan grass (*Sorghum vulgare* var. *sudanense*) and Timothy (*Phleum pratense*) are harvested specifically for forage and silage (FAO Statistics Division 2013a).

Of special interest for this PhD dissertation are the species *Zea mays* ('maize' in what follows), a crop plant of high economic value, and *Brachypodium distachyon* ('Brachypodium' in what follows)

a plant of no agricultural value but emerging as alternative model system for the grasses, belonging to the Pooideae subfamily.

Maize, also referred to as corn, is cultivated on, a total of 177 million hectares, making it the world's second most cultivated crop species after wheat in terms of cultivated area (FAO Statistics Division 2013b). A geographic distribution of global maize cultivation is shown in Figure 2. At an average productivity of 4.9 tonnes per hectare, 875 million tonnes of maize is produced yearly. Together with rice and wheat, these three crops provide 89% of the total cereal production (FAO Statistics Division 2013b).

Brachypodium is native to dry, open habitats in southern Europe and is now established in California and Australia where it grows in dry, disturbed areas on sandy or rocky soils (Schippmann 1991; Rivas-Martinez et al. 1999). Importantly, it was proposed as a model plant for the *Poaceae* in 2001 (Draper et al. 2001). Due to its reduced stature, small, diploid genome and close phylogenetic relationship to economically important temperate cereals and forage grasses, it forms a suitable model to study traits important for sustainable food and feed production (Vogel et al. 2010; Bevan et al. 2010). Brachypodium was also proposed as a model for bioenergy grasses to study important traits for the production of cellulosic bioethanol such as biomass yield and cell wall composition (Brkljacic et al. 2011).

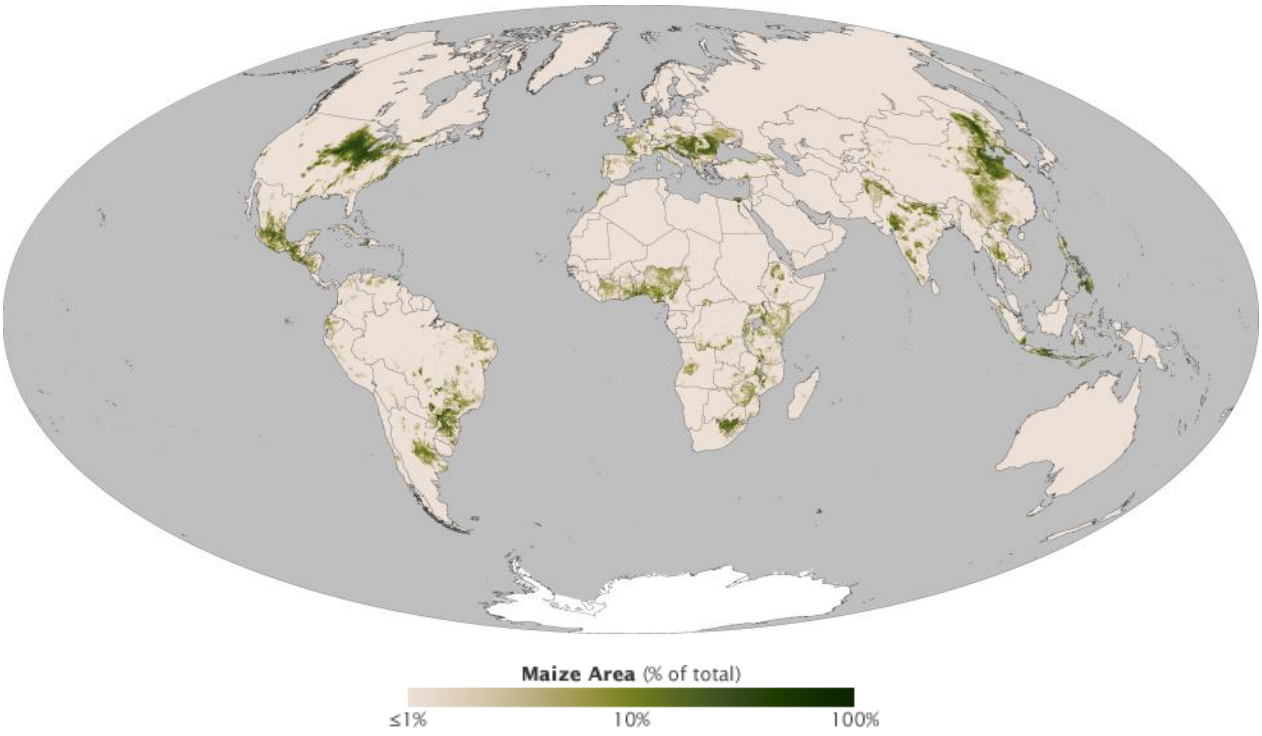


Figure 2. NASA Earth Observatory image showing the distribution of maize cultivation worldwide, created by Jesse Allen using crop data from 2000 published in Monfreda et al. (2008). From http://visibleearth.nasa.gov/view.php?id=47250.

3. Anatomy of the grass plant

The grass plant is instantly recognizable by its tubular, erect stems divided into sections by solid nodes bearing alternate and opposite leaves (except genera *Micraira* and *Arundaclaytonia*) with narrow blades, and supporting terminal spikes, racemes or panicles of spikelets (Figure 3). However, since the family of the grasses is one of the largest in terms of species diversity, a range of morphological variation is present, the most noticeable being annual or perennial habit; tufted, cane-like or even tree-like life forms, herbaceous or woody stems and the design of the inflorescence (Renvoize 2002). Although each grass species displays a somewhat different morphology, a schematic drawing of a typical grass is shown in Figure 3.

The aerial part of the grass plant is generally made up of several shoots or tillers and the individual tillers are formed by the successive formation of phytomers from the apical meristem. A phytomer consists of a node, an axillary bud, an internode and a leaf (Moore and Moser 1995; Langdale 2005). Moving up a tiller from the bottom, each phytomer is younger than the one below (Kallenbach 2012). The node is a solid region on the culm to which the leaf is attached. It also bears an axillary meristem from which secondary tillers can form (Pautler et al. 2013). Moreover, in several grass species, adventitious roots can grow from the aboveground nodes. These are rather apparent in e.g. maize, sugarcane and rice (McSteen 2010). The piece of stem between two consecutive nodes is the internode. The leaf has an upper part, called the blade and a lower part called the sheath and they are separated by a ligule (Figure 3). The leaf blade and the sheath are important for photosynthesis, respiration, and photoperception. The function of the ligule however, is not so clear. The presence of the ligule at the blade-sheath junction may provide mechanical strength to support the weight of the leaf blade and protection by excluding water, dust, harmful spores and pests from the 'interior' of the plant. The presence of auricles, present in a large number of grasses (Figure 3), would provide strength to the leaf by folding around the stem and act as dripping points for water removal. However, ligules might also be important as a secretory tissue (Chaffey 2000). In the vegetative stage of plant development, the leaf sheaths may be rolled into a structure called a pseudostem, apparent in e.g. maize, rice and banana.

From architectural point of view, a main difference between Brachypodium and maize is the number of tillers that they produce. Brachypodium generally produces numerous tillers that grow from the basal nodes (Figure 4E). Only when grown under very long day conditions (23h light/1h dark), Brachypodium plants produce a single tiller and flower within one month after germination (Figure 4F). Most maize plants produce only one tiller, the result of the domestication of its wild progenitor teosinte (Matsuoka et al. 2002). The height of the maize (and Brachypodium) stem depends both on the genotype and on the environment. A typical maize plant, grown in central United States, will have a stem height of 2-3 m and carry 16 to 22 leaves.

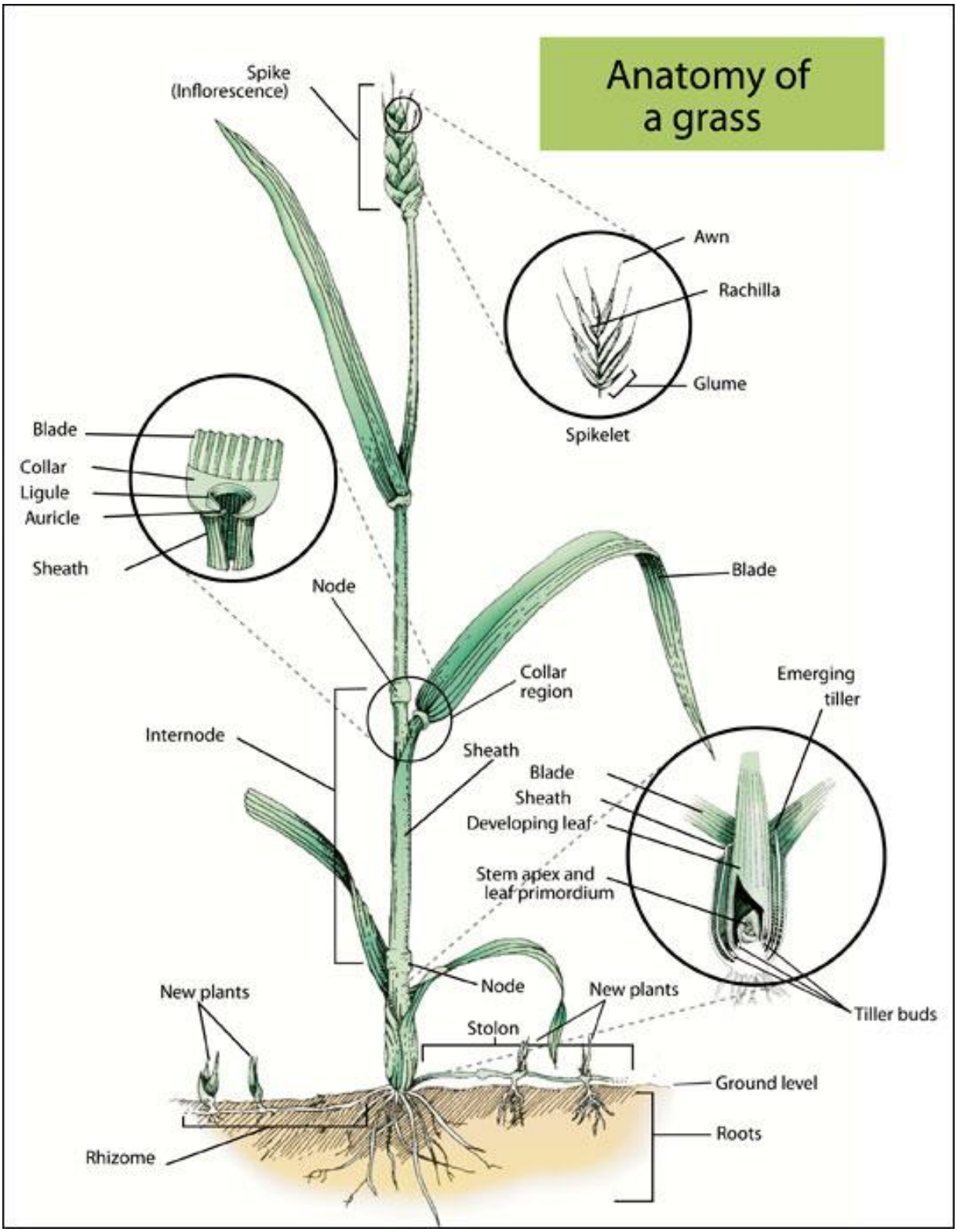


Figure 3. Anatomy of a typical grass plant; Although all grasses look somewhat different, this stylized drawing shows the major parts of a grass plant. From Kallenbach (2012).

4. Development of the grass plant

The growth of a grass plant can be divided into three main developmental phases. The first is the vegetative stage where the SAM initiates new leaves. In this stage the focus is on producing enough leaf biomass to allow for sufficient photosynthesis. Then, the vegetative SAM converts to a generative inflorescence meristem (IM) that will give rise to flowers. This transition is dependent on environmental and genotypic factors (Pautler et al. 2013). The major factors in controlling this transition is the day length and the temperature (Higgins et al. 2010). Different species of grasses have different sensitivities and thresholds for daylength-dependent flowering. For example, rice is considered a photoperiod-sensitive species, with a facultative short-day requirement. On the other hand, floral induction in maize can be either day-neutral (most temperate maize inbred lines) or respond to short-day inductive cues (tropical lines) (Colasanti and Coneva 2009). Other temperate grasses, such as wheat and barley, have a long-day requirement with a vernalization switch. Ultimately the flowers of the grass plant will become visible this point is usually referred to as "heading" or mostly called "flowering" (Eckardt 2000).

For this introduction, the timing of leaf emergence in grasses is further discussed in more detail. During the vegetative growth phase in grasses, the new leaves appear in an orderly and predictable way from the pseudostem. The time between the emergence of two successive leaves on the same plant is called the phyllochron (Klepper et al. 1982). The phyllochron of one species is found to be remarkably constant for leaves of similar hierarchical order and can be found back in many grasses (Craufurd and Bidinger 1988; Birch et al. 1990; Rickman and Klepper 1995; Clerget et al. 2008; Van Minnebruggen et al. 2013). The phyllochron depends on the simultaneous processes of leaf initiation, leaf elongation and construction of the dimension of the pseudostem. It has not yet been elucidated how this regularity emerges from a complex dynamic system. The regularity may be genetically defined by the rate of initiation of the primordium from the apex. On the other hand, it might be a self-regulating dynamic system in which the emergence of a leaf controls the formation and thus the timing of emergence of the younger ones (Fournier et al. 2005). The length of newly formed leaves depends on the length of previously formed leaves which can vary in development or can be altered by grazing or cutting. Accordingly, when the pseudostem is artificially extended, newly formed leaves will grow longer, supporting the view of a self-regulating dynamic system of leaf initiation (Verdenal et al. 2008; Verdenal et al. 2009; Verdenal et al. 2012). Furthermore, there is also evidence for coordinated leaf and tiller development. A study of leaf and tillering in three consecutive nodes in tall fescue (Skinner and Nelson 1994) showed that major transitions in leaf and tiller development were synchronized. At the oldest node, cessation of cell division in the leaf sheath was accompanied by initiation of cell division and elongation in the associated tiller bud. At the next younger node the ligule was being initiated, while at the youngest node cell division commenced in the leaf primordium, as elongation of a new leaf blade began.

4.1 Developmental stages in Brachypodium

As a model plant, Brachypodium is small in size (approximately 30 cm, Figure 4A-F) and has a short life cycle of four months, under 16h light/8h dark regime at 23°C (Draper et al. 2001). However, for Bd21, the genotype for which the genome sequence was first available, the life cycle can be reduced

to the half by simply prolonging the day length to 20h (Molinari et al. 2013), 23h (Bevan et al. 2010) or even 24h (Garvin et al. 2008) at 24°C. The most commonly used Brachypodium inbred lines Bd21 and Bd21-3 do not need vernalization in order to flower when grown under long day conditions, in contrast to other accessions (Schwartz et al. 2010). Approximately one week after germination, plants have one to two leaves (Figure 4A). Two weeks after germination, plants are still vegetative and have one or two additional tillers, depending on the day length (Figure 4B). Approximately one week later, the shoot apical meristem converts into a flowering meristem and starts to produce flowers, which is generally referred to as the transition phase (Figure 4C). When spikelets are emerging from the pseudostem, approximately one month after germination, the plant is in the flowering phase (Figure 4D). Shortly after, seeds become mature and can be harvested. Harvested seed need to be stored dry for at least two weeks. Later they can be sown to start a next cycle of reproduction.

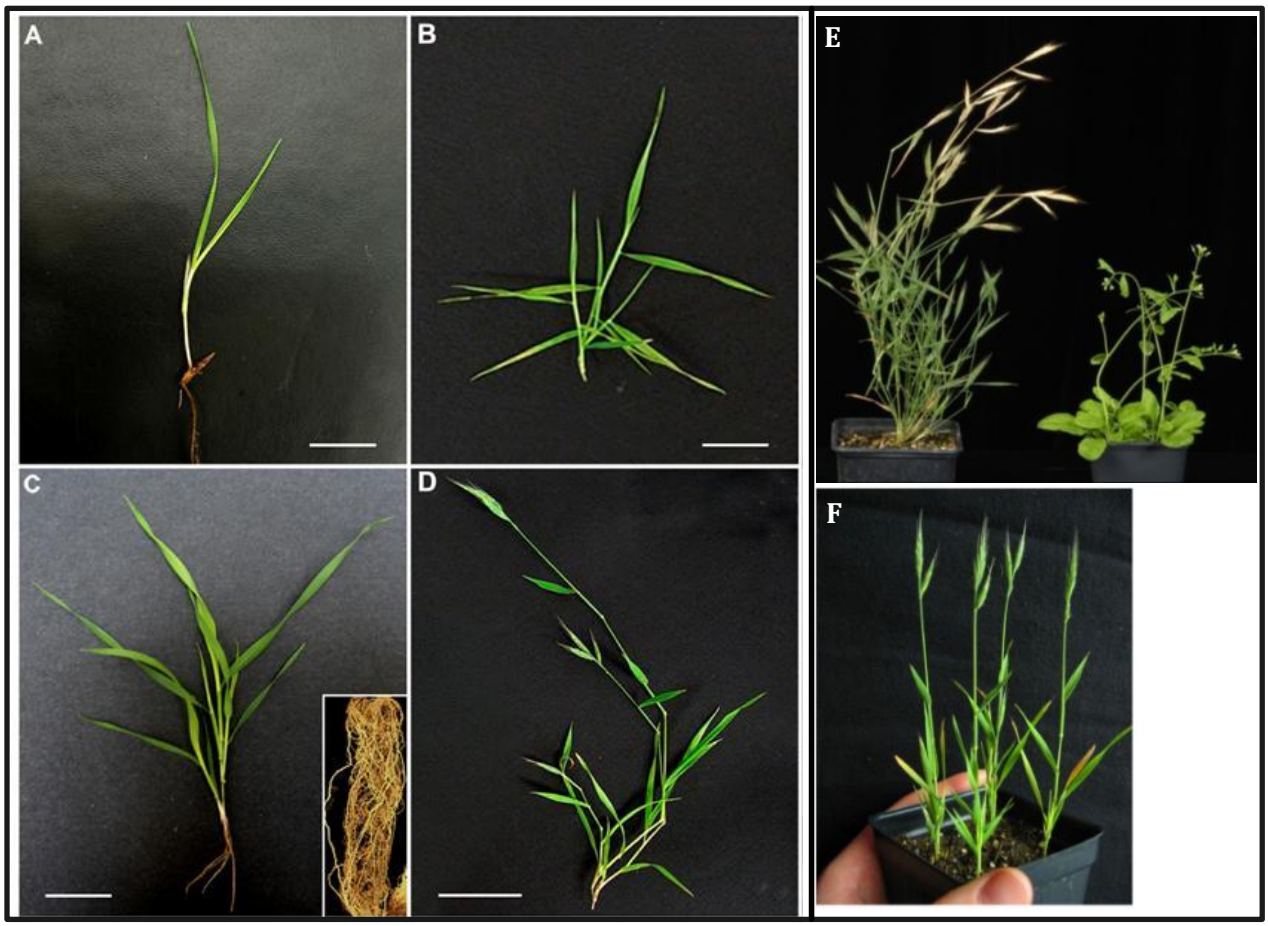


Figure 4. The small grass *Brachypodium distachyon* in different developmental stages, grown in standard glasshouse conditions (16h light/8h dark, 21°C). (A) Early vegetative phase, bar = 1cm; (B) late vegetative phase, bar = 2cm; (C) transition phase, bar = 2cm; (D) flowering phase, bar = 3cm. The figure is modified from Molinari et al. (2013); (E) *Brachypodium distachyon* (Bd21) compared to *Arabidopsis thaliana*, modified from mocklerlab.org/research; (F) Four *Brachypodium distachyon* plants (Bd21), 25 days after sowing grown under 23h day length, modified from Bevan et al. (2010).

4.2 Developmental stages in maize

Understanding the development of the maize plant is important for the study of the genetic and metabolic mechanisms underlying biomass production and cell wall biosynthesis. For instance, precise knowledge of the process of maize stem elongation is required to determine appropriate moments for internode sampling. For this purpose, maize growth is separated in relevant growth stages. The definition of unambiguous and easily recognizable growth stages in maize development aids in using a common terminology in research. The most common way to describe maize development is to divide growth into vegetative (V) and reproductive (R) phases. However, for maize, the definition of the moment of transition from vegetative to reproductive is deviating from earlier described terminology about phase changes in grasses. The commonly used terminology for maize regards the plant as vegetative until the first silks appear on the ear, after which the plant is considered reproductive. This is in contrast to conventional terminology for phase change in grasses: the moment at which the SAM converts to an IM, namely around the sixth leaf stage (V6) (O'Keeffe 2009).

The first V stage is VE for 'emergence' and subsequent growth stages are defined by the number of leaves in V, and characteristics of the ear or kernel in R. The number of leaves is generally counted as the number of leaves for which the leaf collar has become visible (see Figure 6B). Accordingly, the V stages are designated as V1, V2, V3, etc. through V(n), where (n) represents the last leaf stage before male flowering or tasseling (VT). For maize, the transition from vegetative phase to reproductive phase is characterized by the appearance of two inflorescences. One inflorescence meristem located at the apex of the plant, develops into a branched male reproductive organ, the tassel, which emerges first. Another inflorescence meristem develops into the unbranched female reproductive organ, the ear, which arises in the axil of a vegetative leaf (Thompson and Hake 2009; Tanaka et al. 2013). The R stages are numbered R1 through R6 and are named after appearance of the ear or kernel: silk, blister, milk, dough, dent and mature respectively. Figure 5 outlines the different stages in maize plant development following Ritchie et al. (1992), Nafziger (2008) and O'Keeffe (2009) and are discussed below.

Maize developmental phases or maturity are often expressed as by growing - degree days (GDD). GDD result from summation of mean daily temperatures above the base temperature, below which the plant does not grow or grows very slowly. For maize the base temperature is set at 10°C (Nafziger 2009). The rate of growth thus increases as temperature increases above the base temperature. The GDD concept is frequently used since plant growth and development are more closely related to accumulated mean daily temperature above a base value in the absence of other limiting conditions (Neild and Newman 1990). The approximate number of GDD required to reach a particular developmental stage as indicated in the text, refers to a maize hybrid that requires 2700 GDD to reach maturity in the corn belt region of the Midwestern United States (Neild and Newman 1990).

4.2.1 Germination and emergence (VE)

Under favorable conditions (moist, warm soil), the kernel absorbs water (imbibition) and initiates growth. The radicle elongates first (2-3 days after sowing), followed by the coleoptile which is a protective structure for the enclosed plumule (embryonic plant consisting of the first leaves) that

directs the shoot to the surface. Then, three to four seminal roots are formed (Figure 6) and by rapid elongation of the mesocotyl, a white, tubular, stem-like part sometimes considered as the first internode, the coleoptile reaches the surface (VE). The coleoptile emerges from the soil 6 to 10 days after planting (equivalent to approximately 115 GDD). As soon as the coleoptile tip emerges from the soil, it splits at the tip and two true leaves unfold. By means of mesocotyl elongation the height of the crown or growing point, located just above the mesocotyl, is adjusted to 2 – 3 cm below the soil surface. The roots that are formed at this point are referred to as seminal roots and sustain the young plants only for a short period of time (approximately three weeks). Soon after VE, nodal roots appear take over the function of the seminal roots.

4.2.2 V1-V5

Each new leaf will grow from the apex upward inside the pseudostem, a cylindrical structure that is formed by the sheaths of previously formed leaves. After the second leaf has emerged, a new leaf will unfold at the rate of about one leaf every three days, or about one leaf for each 65 GDD. A V2 plant is shown in Figure 6A. Two weeks after the plant emerges (appr. 315 GDD after sowing), the V3 stage begins (Figure 6B). By now, the nodal roots form the major part of the root system. The stem apex (growing point) is still below the soil surface. The first five leaves are already formed as part of the embryo in the maize seed. Therefore, these leaves are called embryonic leaves. These leaves display a different morphology than later formed leaves and are for this reason also called 'juvenile' leaves (Sylvester 2001). Already by V5 (475 GDD after sowing), all leaves and ear shoots that the plant will eventually produce are initiated. Even the tassel, still microscopically small, is present in the stem apex.

4.2.3 V5-V8

From V6 (555 GDD after sowing), the growing point or shoot apex is above the soil surface and stem elongation begins. During these stages (V5-V8) in its life cycle, the maximum number of rows around the ear is determined. A developing ear has a meristematic dome at the tip that is producing new rows of ovules along its length. These ovules eventually divide to produce a pair of rows of ovules from each single row, explaining why a maize ear always has an even number of kernel rows. Establishment of the number of kernel rows around the ear is susceptible to environmental stress. When resources are limited, ovules near the tip will be sacrificed to allow the maize plant to adequately support the remaining viable ovules. At this point in development, also the tassel is clearly visible (V7, 635 GDD after sowing). In V8 (715 GDD after sowing), nodal roots are elongating from the fourth node and several ear shoots are present. A potential ear shoot is formed at every above-ground node except at the upper six to eight nodes. However, only the upper one or two ear shoots eventually form harvestable ears. Degeneration and loss of the two lowest leaves may have already occurred by the V8 stage.

4.2.4 V9-V11

Approximately 5 weeks after emergence (845 GDD after sowing), the V10 stage begins. As the stem is now elongating at a rapid pace, the maize plant quickly accumulates dry matter. In this stage, a new leaf appears every 2 or 3 days. Demand for nutrients and water is high in order to meet the plant needs.

4.2.5 V12-V14

The potential number of kernel rows was already established at V5, but by the V12 stage, six weeks after emergence (845 GDD after sowing), the final number of ovules (potential kernels) on each ear is being determined. Brace roots, that are used for support and extracting additional water and nutrients from the soil, are developing from the fifth node. The plant reaches V14 at seven weeks after emergence (1045 GDD after sowing).

4.2.6 V15-V17

These stages are by far the most critical in terms of seed yield. As a V15 (1095 GDD after sowing) plant is only two weeks away from reproductive growth (R1, silking). Any injury or nutrient, biotic or abiotic stress may seriously reduce the number of kernels that develop. The tassel is near full size but still enclosed by the pseudostem. A new leaf stage can be formed every 1 or 2 days. By V17 (1180 GDD after sowing) the upper ear shoots may have grown enough so that their tips are visible (without dissection) at the top of the leaf sheaths that surround them. The tip of the tassel may also be visible at V17.

4.2.7 V18 or VT

Simultaneously or shortly after the collar of the 18th leaf is visible (1220 GDD after sowing), the tassel appears from the pseudostem. VT or tasseling stage (1350 GDD after sowing) is characterized by a fully expanded male flower. The maize plant reaches its full height. One week from VT, the silks will appear from the husk leaves and pollen shedding will start. In the absence of stress, the tip of the tassel can be seen at about the same time that the tip of the emerging ear becomes visible.

4.2.8 R1-R6

The only function of the tassel is to produce enough pollen to fertilize the ovules in the female flower. A vigorous maize tassel can produce between 2 million and 5 million pollen grains. Pollen shed generally last for several days (typically 5 to 10), with peak production on about the third day. However, it stops when the tassel is too wet or too dry and begins again when temperature and moisture conditions are favorable. Pollination in the greenhouse therefore usually takes place in the morning, when temperatures and humidity are favorable.

In the absence of stress, the first silks produced on a plant emerge from the enclosing husks (R1, 1400 GDD after sowing) one to three days after pollen shed has begun. Under favorable growing conditions, all silks will emerge and be ready for pollination within three to five days. Fertilization occurs when viable pollen lands on receptive silks. After fertilization, carbon and nutrients are shuttled towards the ear. Both ear and kernels will undergo maturation and different stages are named after the appearance of the kernels in the particular stages. R2 (1660 GDD after sowing), 'blister' stage as kernels appear as watery blisters; R3 (1925 GDD after sowing), 'milk' stage after the milky substance high in sugars as the beginning of starch production; R4 (2190 GDD after sowing), 'dough' stage after the sticky, gummy dextrin rich substance in the kernel; R5 (2450 GDD after sowing), 'dent' stage after the dentation or the sides of the kernels that are pushed slightly inside and R6 (2700 GDD after sowing), 'mature' stage when the kernel reaches physiological maturity with maximum dry weight accumulation (Figure 5).

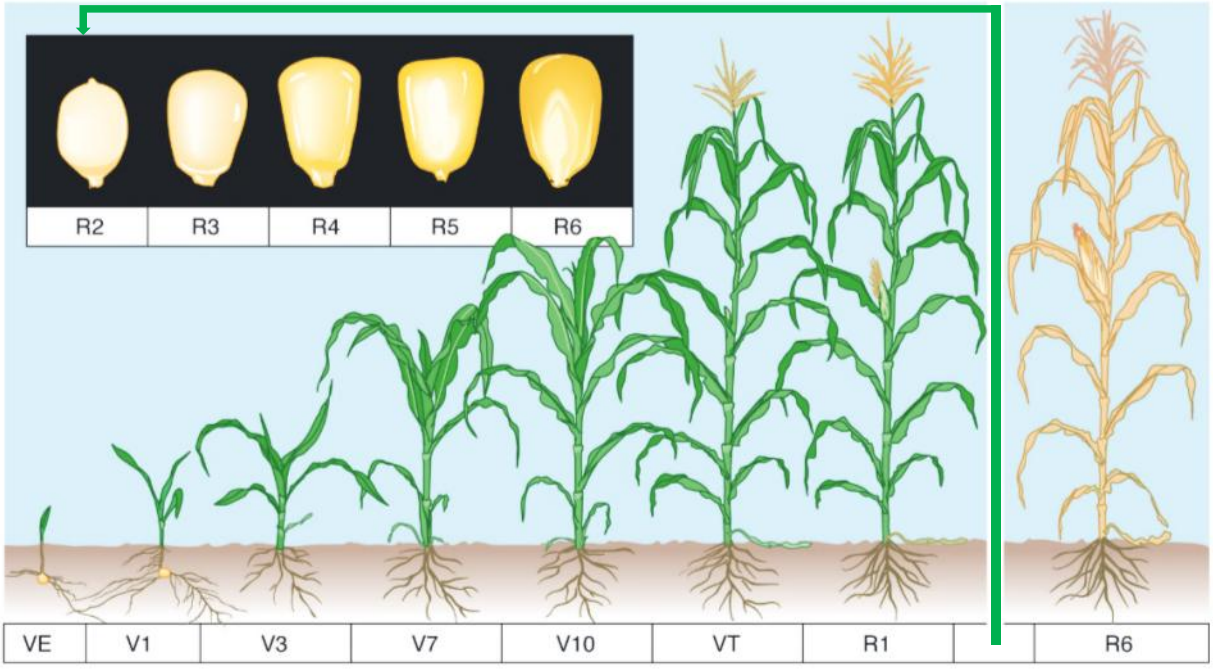


Figure 5. *Zea Mays* or maize in different developmental stages (Nafziger 2009)

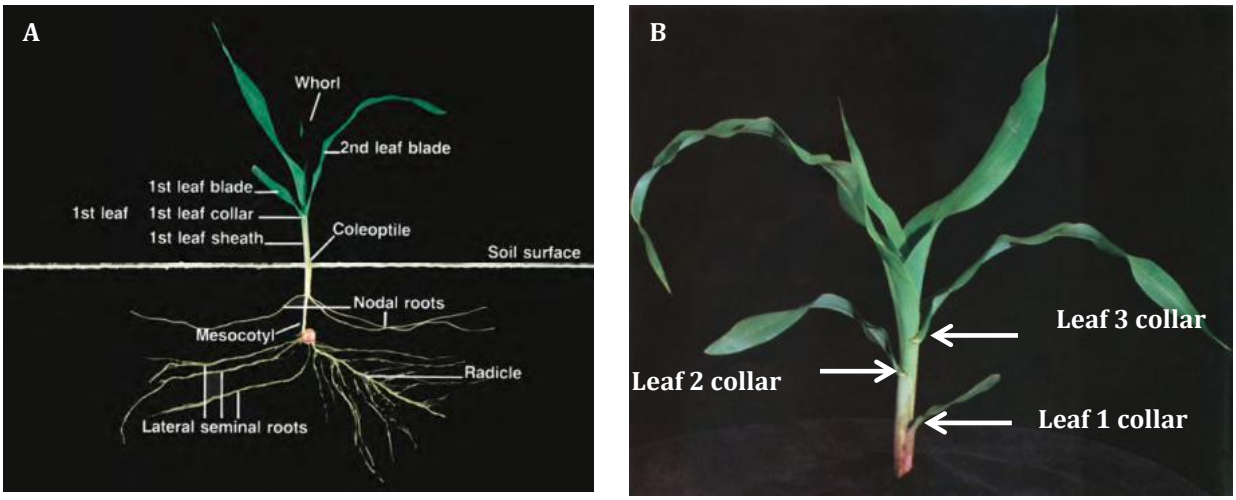


Figure 6. Different parts of the vegetative maize plant in V2 stage (A) and illustration of the number of visible leaf collars that is used as nomenclature to indicate the developmental stages of the vegetative maize plant, here a V3 plant (B). Images are taken from Ritchie et al. (1992).

5. The grass leaf

The leaves of grasses are generally narrow and long with a parallel venation pattern. However, as the veins typically diverge at the base of the leaf blade and converge and fuse toward the apex, the term 'striate' would be more appropriate than 'parallel' (Nelson and Dengler 1997). Maize and Brachypodium leaves, like those of other grasses, have longitudinal veins of different size. The

largest vein, the midvein, is thickened by embedment in parenchyma to form the midrib. On either side of the midvein several large veins run in parallel, and in between intermediate and small veins occur. At the smallest scale of venation, the longitudinal veins are interconnected, so the pattern is essentially reticulate (Nelson and Dengler 1997). As the leaf is initiated, the midvein develops acropetally (from the tip towards the base). Subsequently, the lateral veins develop both acropetally into the blade and basipetally into the stem, and finally the intermediate and transverse veins develop in a basipetal direction (Langdale 2005). In most maize inbred lines, the mature leaf has approximately five to seven lateral veins between midrib and margin and a dozen or more intermediate veins between each lateral vein.

The parallel vascular bundles typical of grass leaves can be seen in Figure 7, in which the transverse section of a leaf of *Alloteropsis semialata* is shown. The vascular bundles run all the way from the basis towards the leaf tip and consist of phloem, transporting organic nutrients from the leaves to the rest of the plant and xylem, transporting water from the roots towards the leaves, with the phloem situated abaxially and xylem adaxially in the vascular bundle (Langdale 2005). The stomata that facilitate gas exchange are situated both in the upper and lower epidermis in grasses. In monocot leaves bulliform cells are present. These are situated only at the adaxial side, and are supposed to regulate folding and unfolding of the leaf in response to drought (Li et al. 2010b; Xiang et al. 2012).

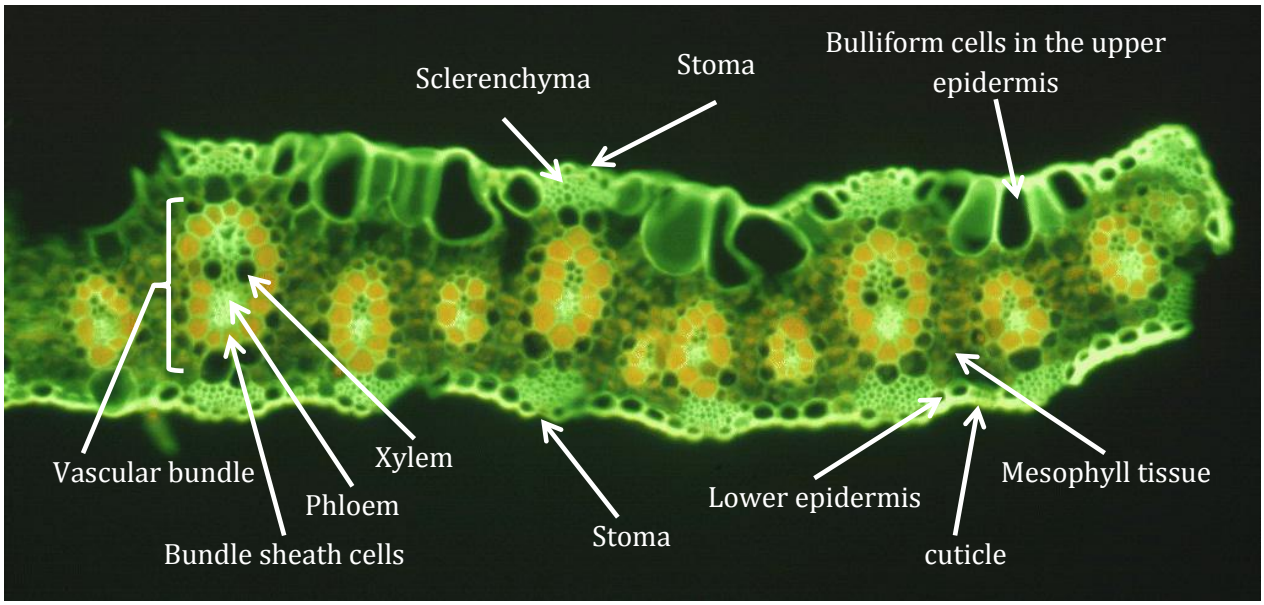


Figure 7. Fluorescence image of a transversal section through a leaf blade of *Alloteropsis semialata* subsp. *semialata* a C4 grass (Watson and Dallwitz 1992).

In monocot leaves not only the veins run in the longitudinal direction, in fact all cells are organized in parallel lines that run from the leaf base towards its tip. The leaves originate from the shoot apical meristem (SAM), which is surrounded by leaf sheaths and thus located inside the pseudostem. The stem cells are located in the upper region of the central zone (CZ) of the meristem (Figure 8), in which cells divide slowly. The progeny produced from the division of the stem cells is used to replenish the stem cells themselves and are also displaced into the peripheral zone (PZ),

where they start to divide more rapidly and lateral organs, such as new leaf primordia, are initiated (Figure 8) (Pautler et al. 2013).

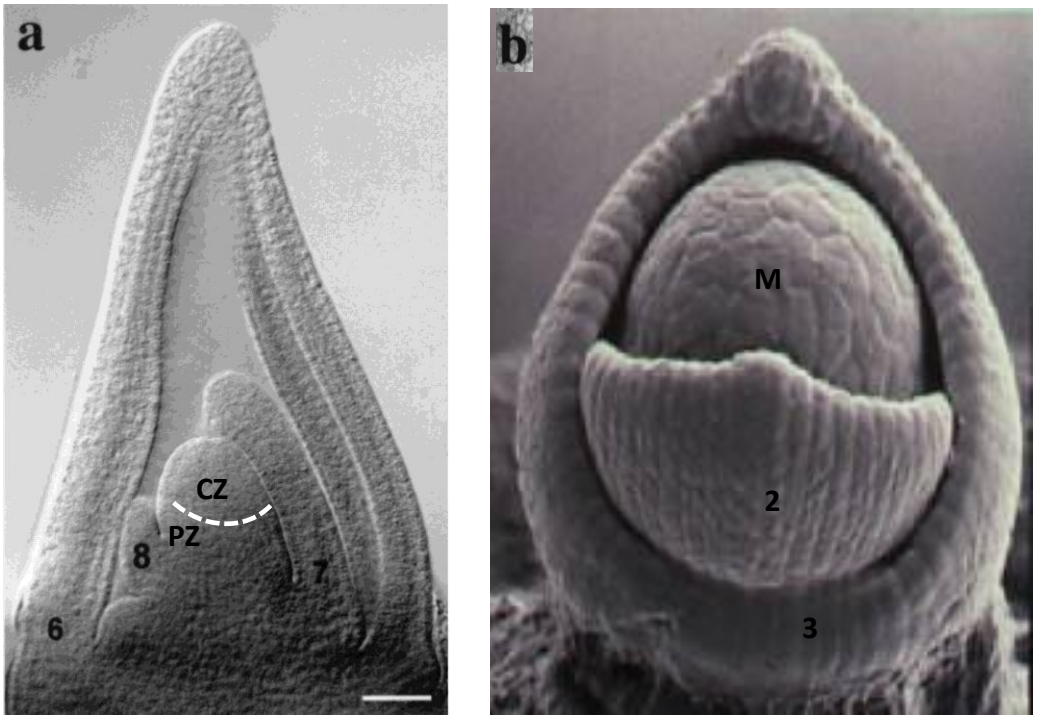


Figure 8. Shoot apices of juvenile maize shoots. (a) Side view of cleared specimen of the shoot apex of an 8-day-old maize seedling with three leaf primordia (leaves 6, 7 and 8). CZ: central zone, PF = peripheral zone (Bar = 100 µm.) Modified from from (Orkwiszewski and Poethig 2000) (b) Top view of a vegetative shoot apex of a maize plant with two leaf primordia and the meristem from http://www.pgec.usda.gov/Hake/ResearchKNOX.html

At least 120 cells are recruited from the SAM to form a leaf primordium (Poethig and Szymkowiak 1995). The cells that constitute the leaf primordium continue to divide and the midvein, lateral and intermediate veins differentiate when the size of the primordium approaches 1 cm in length. At this time, the leaf blade is not yet distinct from the leaf sheath. Shortly after, the ligule and auricle differentiate from the midrib towards the margin, marking the border between sheath and blade (Sylvester et al. 1990). It is hypothesized that the emergence of the tip from the pseudostem triggers the differentiation of the ligular region (Fournier et al. 2005). The next phase is marked by rapid increase in blade length. The sheath and ligule only begin to extend when the blade is nearly fully elongated. The leaf reaches maturity when the sheath stops elongating. In maize, the blade-to-sheath ratio in the mature leaf is around three for juvenile leaves but reaches more than seven for adult leaves (Sylvester 2001). In conclusion, leaf growth in monocots can be summarized chronologically in three phases: (1) a preligular stage, where the leaf primordium is undifferentiated and divides only along its length; (2) differentiation of the veins and ligular region after which the leaf blade grows rapidly and (3) a stage of rapid sheath growth and the cessation of sheath extension marks leaf maturation (Sylvester et al. 1990). The timing of leaf blade, ligule and leaf sheath extension is such that leaf growth is a continuous growth process (Auzanneau et al. 2011). These observations at the organ level can be explained from the dynamics of cell division and elongation. The fact that in the monocot leaf the cells are aligned longitudinally allows

describing the dynamics of leaf growth as a flux of newly born cells out of the basal division zone, where meristematic cells elongate and divide, into the adjacent elongation-only zone, where cells undergo post-mitotic elongation and reach their final length (Kavanová and Lattanzi 2006) (Figure 9). Together these zones are called the growth zone (Kavanová and Lattanzi 2006). This representation of leaf growth in fact simplifies reality by neglecting the differentiation into blade, ligule and sheath and the differential timing of cell elongation in these tissues. However, this model is accurate in the preligular phase of leaf growth, i.e. before the tip emerges from the pseudostem, and can form a good estimation of leaf growth dynamics in later development. Mature cells are situated in the mature zone and furthermore two transition zones in between the previously defined zones are considered (Nelissen et al. 2012). This type of modeling provides the basics for kinematic analysis making it possible to derive, from the spatial profiles of cell length and displacement velocity, rates and durations of cell division and elongation (Silk and Erickson 1979; Rymen et al. 2010; Nelissen et al. 2012).

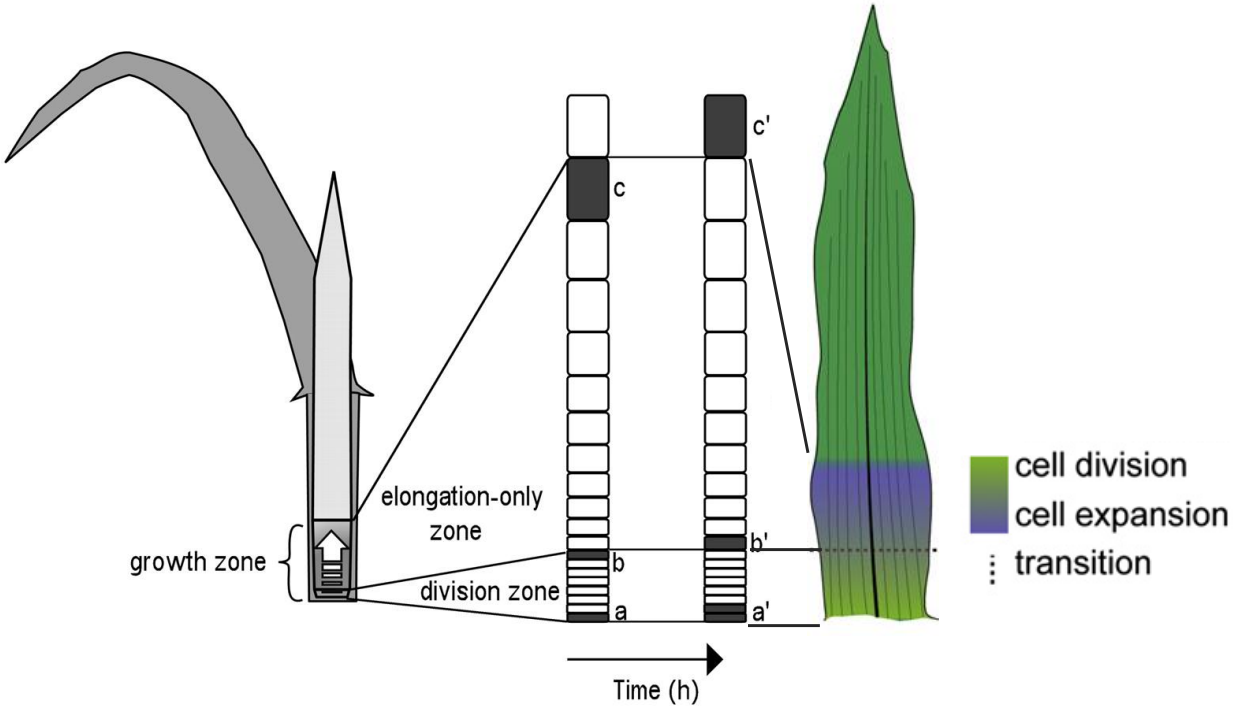


Figure 9. Growth pattern of a grass leaf. Growth is limited to the basal part of the growing leaf, the growth zone, which is enclosed by the sheaths of expanded leaves. Meristematic cells in the division zone elongate and divide simultaneously (a→a′). The progeny of the initial cell at the base of the meristem goes through a certain number of division cycles, thus determining the number of cells per meristematic file. Upon entering the elongation-only zone (b→b′), cells elongate without further divisions until they reach their final length at the distal end of the growth zone (c→c′). Modified from Kavanová and Lattanzi (2006) and Nelissen et al. (2012)

6. The grass stem

The grass stem consists of piled up internodes separated by nodes (Kallenbach 2012). Plant stems play a major role for the vertical distribution of leaf area and thus determine to a large extent the ability to compete for light. Stem dimensions are also closely related to resistance to lodging, which can be an important factor in determining crop yield. For stem development and characteristics, here the maize stem is used as representative for the grasses (Fournier and Andrieu 2000).

In maize the first internode that forms is situated below the coleoptile and elongates rapidly, leading to plant emergence, as described above. Cessation of coleoptile internode extension is triggered by a light signal at emergence (Parvez et al. 1998). The higher internodes originate from the intercalary meristem, which is situated above the individual nodes. As shoots produce new leaf primordia, the initial cell divisions that produce intercalary meristems occur just above the site of leaf initiation (Sharman 1945; Fisher and French 1976). After establishment, intercalary meristems ultimately produce the majority of cells that comprise mature stems (Nemoto et al. 2004). However, intercalary growth would occur only after the transition of the apex to the reproductive phase (Fournier and Andrieu 2000). Before that time, plant growth is visible by the appearance of leaves of which the sheaths form the pseudostem. At that time, the 'true' stem, with piled up internodes and nodes is situated at the bottom of the maize plant (Kallenbach 2012). Stem elongation rates increase after tassel initiation and grow rapidly once the ear is initiated (Siemer et al. 1969; Fournier and Andrieu 2000). Each internode elongates and undergoes differentiation and secondary cell wall thickening in a basipetal fashion, beginning with the first above-ground internode and proceeding acropetally through the stem (Morrison and Kessler 1994) (Figure 10).

Internode growth can be separated in several phases (Fournier and Andrieu 2000). In the first phase, the distribution of cell lengths remains unchanged over the entire internode. Division occurs at a constant cell size, which results in an exponential rate of internode elongation which is homogeneous over the internode. In the second phase, a gradient of cell length develops at the distal end of the internode, as cells start to elongate, with an increase in the rate of internode elongation. Thirdly, after full cell elongation, the most distal cells mature, so that the region with a gradient of cell lengths moves downwards, whereas the region of nonelongated cells, at the base of the internode, remains constant in length. This can be regarded as a temporary `stationary regime' of linear internode elongation, in which the production of new cells that start elongation compensates for the maturation of older cells which cease to elongate. In the fourth phase, as the production of new cells ceases, the region of non-elongated cells at the base of the internode decrease in length, resulting finally in the cessation of elongation. By this developmental pattern, several internodes are elongating at the same time. However, the lowest internodes develop first so that in the maize stem, sequential internodes are found in range of developmental stages; from primordial in the upper stem internodes to fully mature in the lower internodes (Morrison and Kessler 1994). The amount of internodes that elongate simultaneously is coordinated to ensure constant growth rates of the whole stem and might be related to the timing of leaf sheath growth (Morrison and Kessler 1994; Robertson 1994; Fournier and Andrieu 2000).

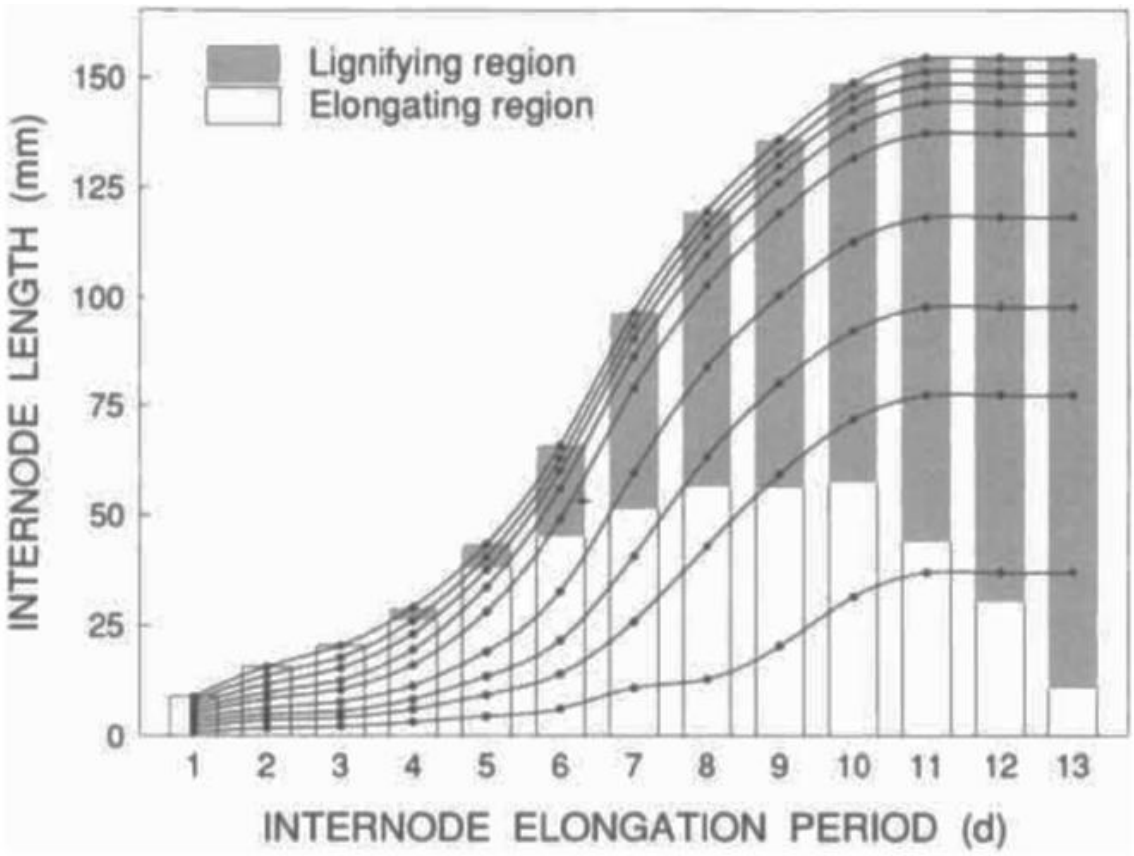


Figure 10. The internode elongates and undergoes differentiation and secondary cell wall thickening in a basipetal fashion. Differentiation is characterized by secondary cell wall thickening and lignification which was monitored in the developing Internode 10 over 13 days of growth. Spline plots indicate the incremental increase in space between paint dots originally placed 1 mm apart in the internode at day 1. Shaded bars indicate where lignin has been deposited in the cell wall, as determined by phloroglucinol-HCl stain. Figure was reproduced from Morrison et al. (1994).

The vascular bundles in the stem consist of phloem, transporting assimilates from the leaves to the rest of the plant and xylem, transporting water from the roots towards the leaves and reproductive organs. In contrast to dicots, in which a single line of vascular bundles is arranged in a ring, forming a cylinder-like structure in the stem, in monocots, the vascular bundles are scattered or disposed in a more complex arrangement (Figure 11). Furthermore, the presence of bundle sheath cells and the lacuna, which is an air space created by the disruption of the protoxylem during fast elongation of the internode, is specific to monocots (Figure 12). The phloem is oriented towards the outside of the stem and xylem towards the inside. The xylem vessels are made up of tracheary elements that are in fact dead, hollow cells with thick secondary walls (see below). In the newly formed stem, the xylem cells in the vascular bundle differentiate and undergo programmed cell death with autolysis of the protoplasm (Fukuda 1997b; Fukuda 1997a; Buckner et al. 1998; Fukuda 2000; Yu et al. 2002). In contrast to xylem, phloem consists of all living cells that are not lignified. The phloem sieve tubes are surrounded by companion cells. The phloem and xylem are surrounded by sclerenchyma cells, which have a thick secondary cell wall and are heavily lignified, constituting strong structural elements of the vascular bundle (Jung and Casler 2006).

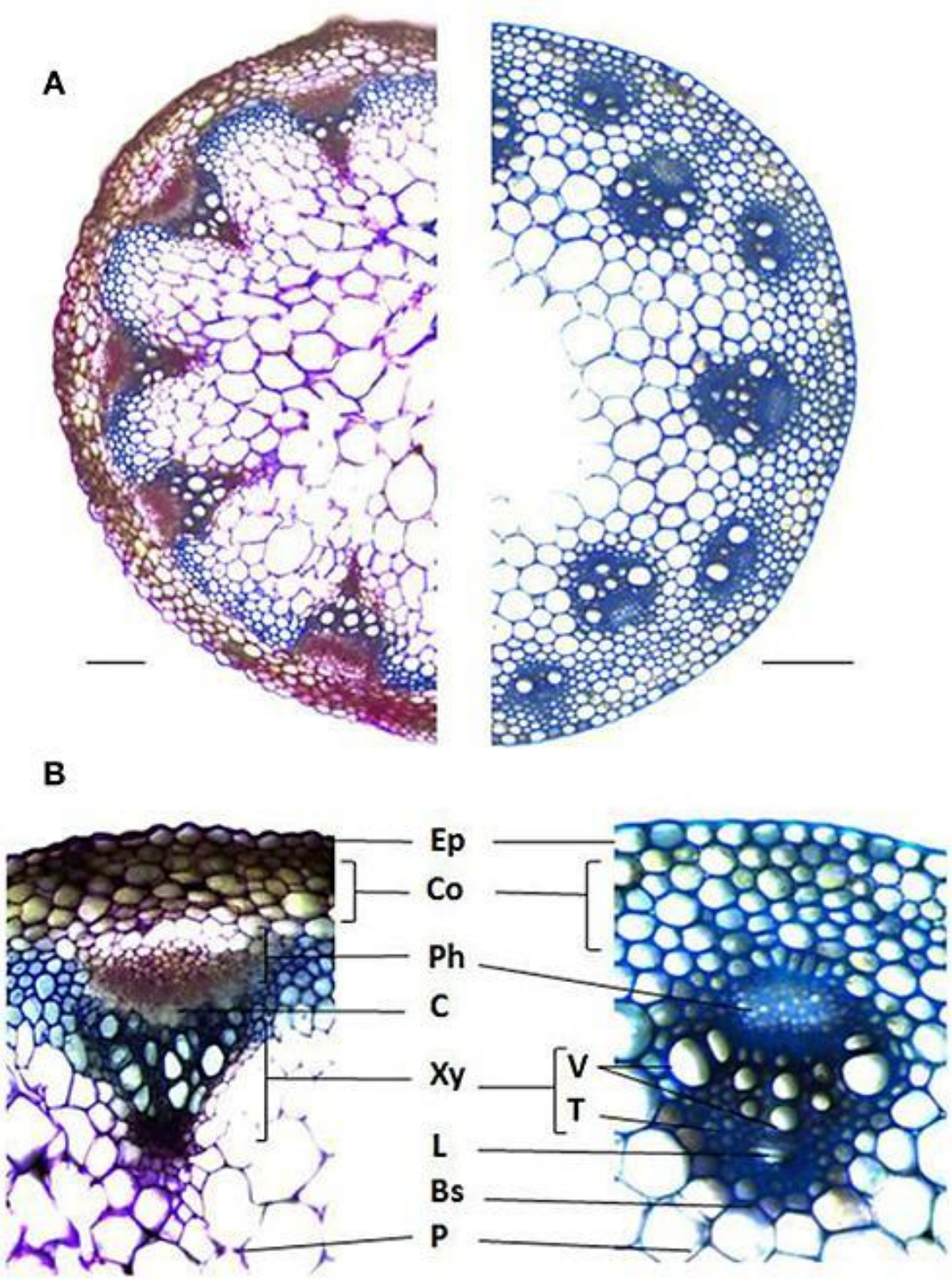


Figure 11. Stem cross-sections illustrating the different cell types and arrangements between dicots and monocots. (A) *Arabidopsis thaliana* (left) and *Brachypodium distachyon* (right) stained with Toluidine blue. (B) Vascular bundle anatomy of *A. thaliana* (left) and *B. distachyon* (right). Ep, Epidermis; Co, Cortex; Ph, Phloem; C, Cambium; Xy, Xylem; V, Vessels; T, Tracheads; L, Lacuna; Bs, Bundle Sheath; P, Pith. Bars = 0.1 mm. (Handakumbura and Hazen 2012)

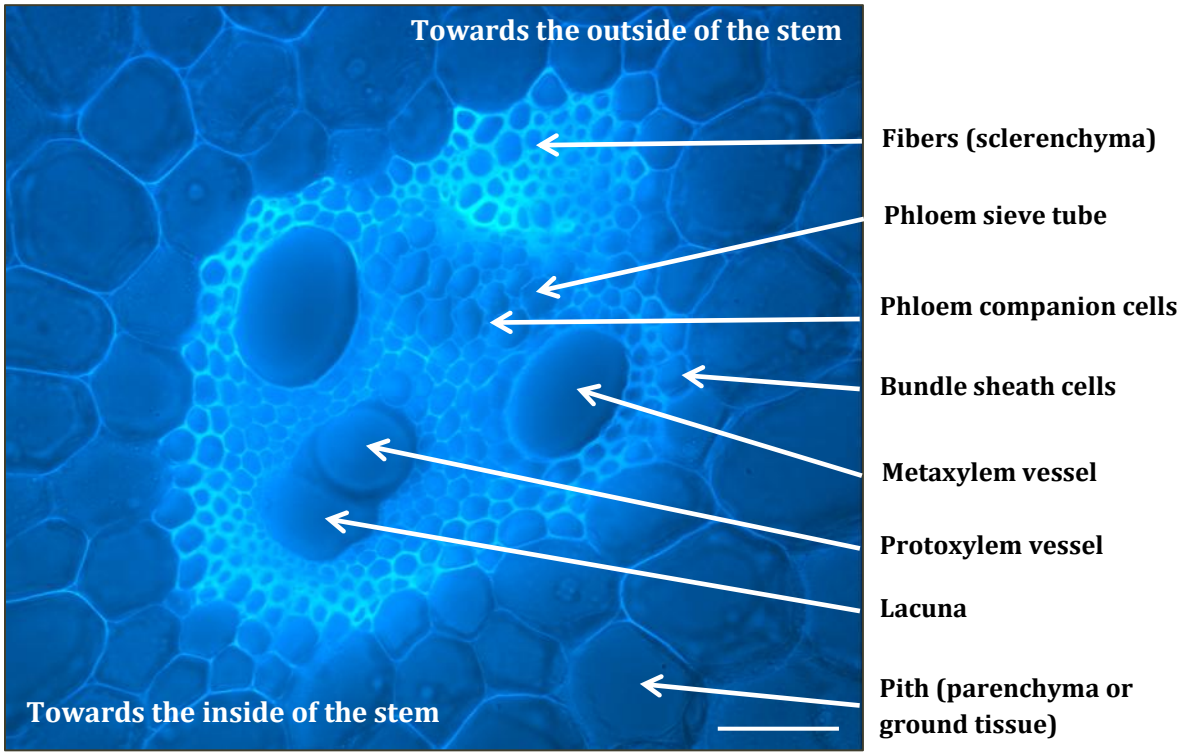


Figure 12. Detail of the maize vascular bundle in the maize ear internode. Bar = 50 µm. The image shows the autofluorescence of lignin, associated with the secondary cell wall and was taken with a Zeiss Axioskop (Carl Zeiss, Oberkochen, Germany) equipped with a Filter Set 02 (488002 – 9901 - 000; excitation: 365 nm; beam splitter: 395 nm, emission ≥ 420 nm).

7. The grass cell wall

In growing cells, the wall is typically a thin, flexible layer (0.1–1 µm) that consists primarily of complex polysaccharides and a small amount of structural proteins. Despite its thinness, the wall forms a strong network that functions like a corset, giving shape to the protoplast within (Cosgrove, 2005). The cell wall of the monocots and thus grasses, differs from that of dicots. Studies of the carbohydrate and aromatic components of the cell walls of a broad spectrum of monocots have revealed that members of the *Poales* which comprises the grasses, their progenitors, and related taxonomic orders have primary cell walls completely different from those of other monocots. The presence of ferulic acid in the primary walls in the *Poales* is one of these features. The cell walls of the *Poales* are composed of cellulose fibers encased in glucuronoarabinoxylans (GAX), high levels of hydroxycinnamates, and very low levels of pectin and structural proteins. In addition, the cell walls of grasses and some related families in the order *Poales* contain significant quantities of mixed linkage glucans (MLG). For that reason, structural models for two types of primary walls were given: the 'Type I' wall of dicots, composed of a cellulose-xyloglucan (XyG) framework embedded in a pectin gel, and the 'Type II' wall, the special wall of the *Poales* (Carpita 1996) (Figure 13). In what follows a detailed description of the cell wall of grasses is provided in order to provide the reader the correct context for the interpretation of the results from the research chapters.

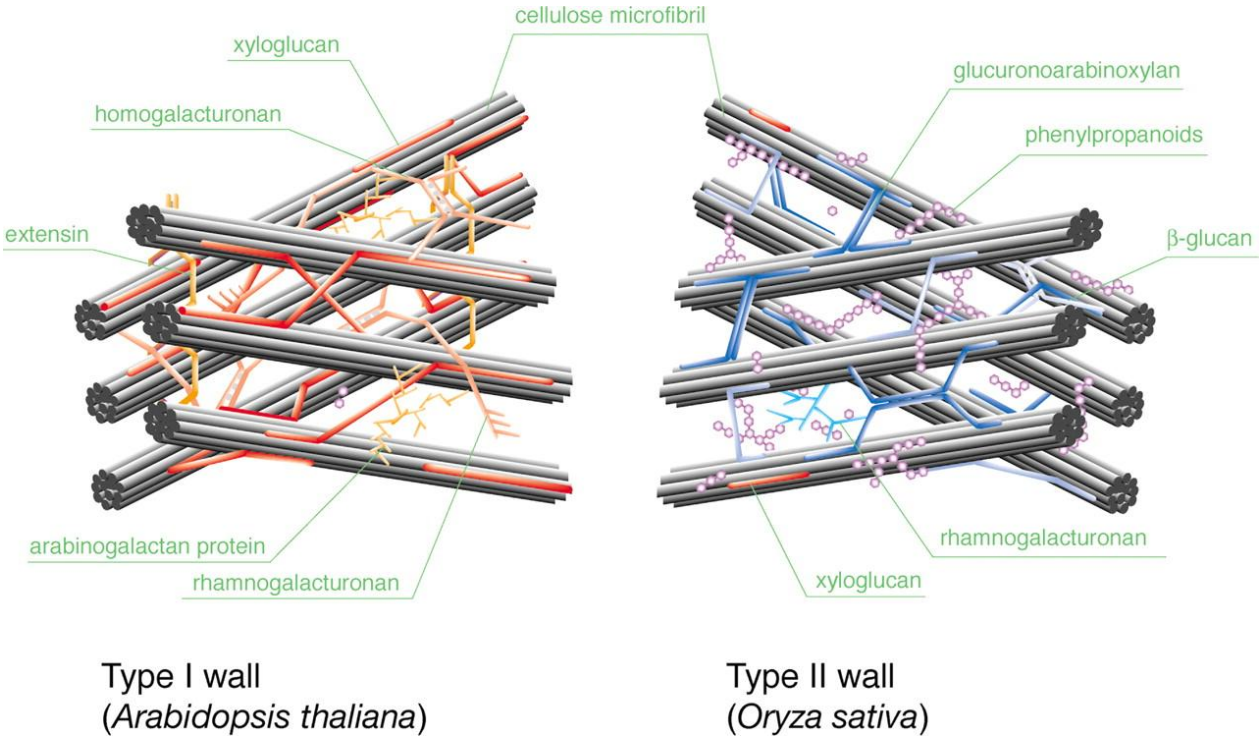


Figure 13. Schematic structural models of type I and type II walls as represented by Arabidopsis and rice cell walls, respectively. From (Yokoyama and Nishitani 2004)**.**

7.1 The primary grass cell wall

7.1.1 Cellulose

Cellulose microfibrils in all flowering plants are composed of about three dozen linear chains of (1->4)-β-linked D-glucose of about 10 to 25 nm in diameter (Figure 14). These microfibrils are synthesized by cellulose synthase (CESA) complexes (Cosgrove, 2005). Multiple microfibrils compose a macrofibril of approximately 500 nm in diameter and these long paracrystalline arrays spool around each cell. Although each chain may be only several thousand units long (approximately 4 μm), they begin and end at different places within a microfibril and are arranged in very long microfibrils whose ends are rarely detected (Raven et al. 1992; Carpita 1996). Despite the fact that cellulose is the most abundant natural biopolymer known, its biosynthesis remains challenging to researchers (Joshi and Mansfield 2007). Cellulose, if accessible for hydrolytic enzymes, provides a huge amount of energy-rich glucose units that could be used for the production of biofuels (see later).

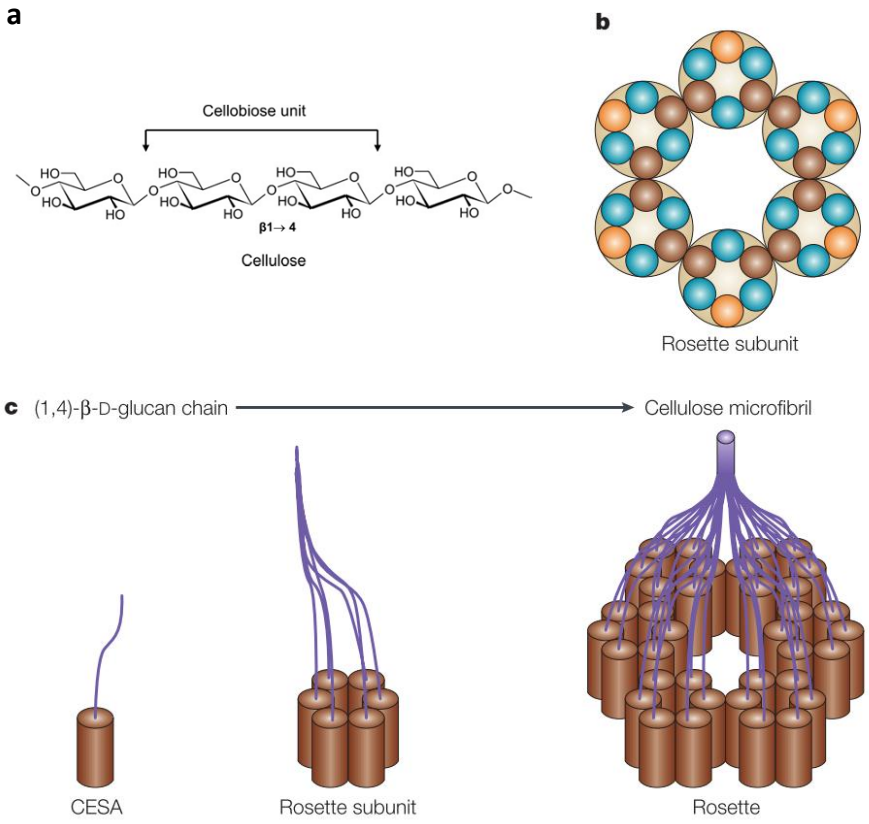


Figure 14. The composition of cellulose and the cellulose-synthesizing machinery of the cell. (a) Simplified representation of a cellulose β-(1-4)-glucan chain, with indication of the repeating cellobiose unit, typically repeated between 200 and 700 times in one cellulose chain. From http://www.namrata.co/category/diet-and-nutrition/theory-notes-diet-and-nutrition. (b) This model of a hexameric particle rosette shows how three different CESA proteins (shown in three different colours: α, orange; β, brown; γ, green) might be organized into rosette subunits and then into a hexameric synthase complex7. CESA assembly into rosette subunits is thought to involve oxidatively reversible disulfide bond formation between cysteines in the N-terminal zinc- finger region of CESA19. (c) A model of how CESA complexes synthesize a cellulose microfibril. Each CESA protein can synthesize a single (1,4)-linked β-D-glucan chain. Cellulose is formed as a crystalline ribbon that is composed of many such glucans. In this model, 36 β-D- glucan chains are formed by a particle rosette, which is composed of a hexamer of CESA hexamers. From (Cosgrove 2005)**.**

7.1.2 Hemicellulose

Glucuronoarabinoxylans (GAX), the major hemicellulose in grass cell walls, are composed of a β 1,4-linked xylose backbone with single arabinose and glucuronic acid side chains primarily attached at the O-3 and O-2 positions, respectively (Figure 15). The highly substituted arabinoxylan is found mainly in vegetative tissues, whereas some grains accumulate large quantities of arabinoxylans that contain little glucuronic acid (Carpita, 1996; Christensen et al., 2010). Although xylans are also present in low amounts in dicot primary cell walls, these xylans contain primarily glucuronic acid and methyl glucuronic acid side-chains attached at the O-2 position of the xylosyl units. GAX in grass primary cell walls takes the place of XyG in dicot primary cell walls and cross-links cellulose microfibrils (Carpita, 1996; Vogel, 2008). The crosslinking of GAX to cellulose fibrils occurs through covalent linkages with ferulate (Molinari et al., 2013).

Mixed linkage glucans (MLGs) (also known as β-glucans) are, like cellulose, unbranched homopolymers of glucose, that are unusual in that they contain both β 1,3- and β 1,4-linkages (Figure 15). MLGs are unique to the cell walls of grasses and a few related families from the order *Poales*. MLGs have been observed in the cell walls of many vegetative cell types and at high concentrations in the endosperm of some grains, where they act as storage carbohydrates. The concentration of MLG in vegetative cells is highly correlated with cell growth and peaks at the same time as cell expansion suggesting that MLG play a role in cell expansion (Vogel 2008).

Only small amounts of xyloglucan (XyG) are found in type II walls, yet it is structurally different from that found in Type I walls, as the xylosyl units appear on different places on the glucan backbone (Carpita, 1996). XyGs are particularly abundant in grass meristematic cells before the onset of enhanced MLG (or β-D-glucan) and GAX synthesis during elongation. XyG biosynthesis is the best understood of any hemicellulose. However, since XyG and (gluco)mannan occur in dicots and are only minor components of grass cell walls, this knowledge is of less relevance to the understanding of the grass cell wall (Vogel 2008).

7.1.3 Pectic Substances

The two fundamental constituents of all flowering plant pectins are polygalacturonic acid (PGA) and rhamnogalacturonan I (RG I) and both are found in grasses, although present in much smaller quantities than in type I walls (Carpita 1996; Christensen et al. 2010; Rancour et al. 2012). To a significant extent, the roles performed by pectins and xyloglucan in the primary type I walls, are replaced in grasses by arabinoxylans and MLGs (Yokoyama and Nishitani, 2004; Christensen et al., 2010).

7.1.4 Aromatic Substances

A major feature of the grasses and their relatives is the enrichment of aromatic substances such as ferulic acid (FA), *p*-coumaric acid (pCA) and sinapic acid in the primary wall. A large portion of the aromatic substances are esters of FA and pCA and can thus be readily released by mild alkaline pretreatment. Mainly FA is found to be esterified to the C5 of the arabinosyl side chains of arabinoxylans (Carpita 1996; Ishii 1997). FA can undergo oxidative dimerization and form different dimer isoforms, and mainly 8-O-4'-, 8-5'-, 8-8'- and 5-5'- coupled diferulate (DiFA) have been identified (Grabber et al. 2004a). In addition to dimers, more complex linkages have been observed (Grabber et al. 2004a). DiFA cross-linking of neighboring arabinoxylan molecules, forming arabinoxylan networks, is commonly thought to play a role in stiffening of the cell wall and in the deceleration of growth, while the function of the pCA esters is still unknown (Grabber et al. 2004a; Yokoyama and Nishitani 2004). FA has also an important role in secondary cell walls, which is described later.

7.1.5 Structural Proteins

Cell wall proteins (CWP) range from structural proteins that are strongly or covalently attached to the polysaccharides, all the way to loosely attached or soluble proteins (Vogel, 2008). The function of structural cell wall proteins (CWP) in the type I wall, which is the crosslinking of carbohydrates to form a non-extensible structure, is largely taken over by hydroxycinnamates in cell walls of grasses (Carpita 1996). A few proteins that are found in the grass cell wall and are of importance to

this work are listed. The extensins in grasses, which are thought to act in cell wall assembly, are called threonine-hydroxyproline-rich glycoproteins (THRGPs) (Cassab 1998). Other cell wall-associated proteins are glycine-rich proteins (GRPs), proline-rich proteins (PRPs) and arabinogalactan proteins (AGPs). Although many of the functions remain elusive, these proteins seem to play important roles in the development of vascular tissues, nodules, and flowers and during wound healing, freezing tolerance, directing planes of growth and development and participating in cell shape (Cassab 1998).

7.1.6 Other Cell-Wall Substances

Several other substances are found in the walls of grasses. Silica is particularly abundant in the walls of grasses, mostly as inclusion bodies in the epidermis, periderm, and other specialized cells of the root, rhizome, and aerial shoots (Carpita 1996). The presence of silica has been associated with grazing tolerance (Briske 1991; Cotterill et al. 2007). Similar to all flowering plants, grasses possess cutin, suberin, and waxes in specialized cells. Cutin and suberin are polymers of fatty acids that occur outside the cell and provide a barrier to water and gas exchange and protection against pathogens (Pollard et al. 2008).

7.2 Structural dynamics during cell elongation

The driving force of cell growth is cell turgor, which is controlled by the cell wall. In order to grow, the cell wall releases its pressure by a mechanism of carefully controlled polymer creep, in which selective loosening and shifting of load-bearing linkages between cellulose microfibrils occurs, allowing the cellulose microfibrils to move apart. Fast elongation can occur without major changes in cell wall composition whereas slower changes, such as the gradual decline in growth that occurs as cells mature and as their walls stiffen, involve substantial changes in wall composition and crosslinking (Cosgrove 2005). This highly controlled process involves several types of enzymes, such as expansins, AGPs and XyG endotransglycosylases.

As cell elongation ceases, both esterified and etherified cinnamic acid constituents accumulate. The cessation of growth is correlated with the appearance of cinnamyl alcohol dehydrogenase (CAD) and specific peroxidases required for cinnamate synthesis and polymerization in the grasses. The accumulation of the phenolic esters and ethers holds the cell in its final shape and provides strength to the cell wall (Carpita 1996).

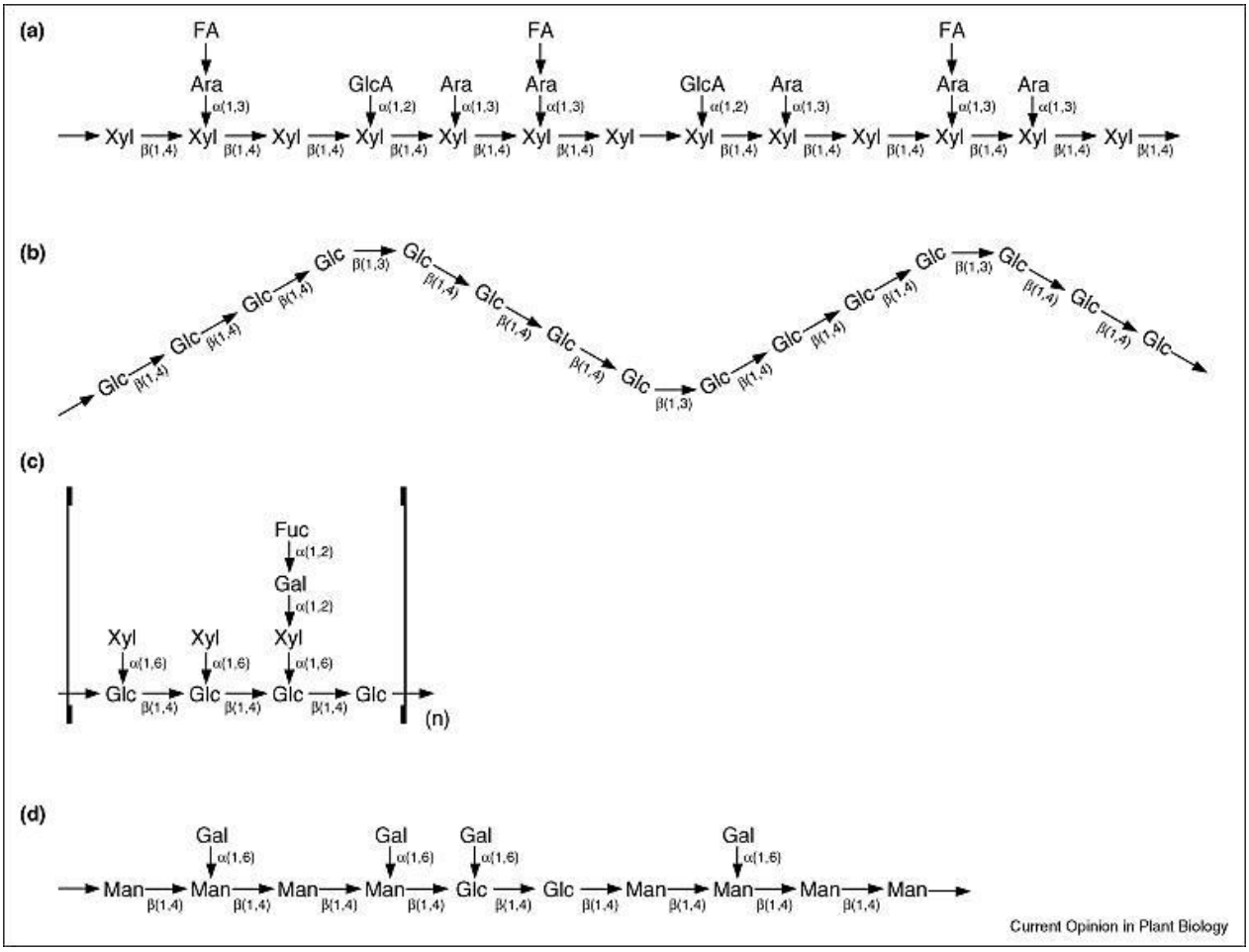


Figure 15. Hemicellulose structures. (a) GAX consists of a b(1,4) linked xylose (Xyl) backbone substituted with arabinose (Ara) (mainly attached at the O-3 position) and less frequently with glucuronic acid (GlcA) (mainly attached at the O-2 position) in a non-repeating fashion. GAX is highly substituted in the Golgi and then some of the Ara and GlcA units are removed in the cell wall. Ferulic acid (FA) is attached to the Ara side chains through various linkages. (b) MLG consists of an unbranched polymer of glucose (Glc) in which variable length stretches of b(1,4) linked Glc units are interrupted with single b(1,3) linked Glc units. The b(1,3) linkages cause the polymer to bend. (c) XyG is composed of a b(1,4) linked Glc backbone substituted in a repeating pattern of four Glc units. The nature of the side chains varies depending on the species. One typical repeat unit containing Xyl, Gal, and fucose (Fuc) is shown. (c) Glucomannans consist of a b(1,4) linked backbone containing both mannose (Man) and Glc. The polymer is variably substituted with galactose (Gal) ranging from not substituted at all to highly substituted. Mannans are similar except they do not contain Glc. From Vogel, 2008

7.3 The secondary grass cell wall

Grass grown for food, feed or biofuel are typically harvested at developmental stages when organs have significant levels of secondary cell wall (Rancour et al. 2012). The secondary cell wall comprises the majority of the total leaf and stem biomass. Its composition is a determinant of biomass quality and therefore of particular interest for saccharification and digestibility improvement (Zeng et al. 2014). Despite its major contribution to plant biomass, not all plant cells have a secondary cell wall. In a newly formed cell, the primary wall is formed first and is continuously modified and fortified during cell elongation (Cosgrove 2005). The secondary cell wall

is deposited inside of the PW and SW thickening occurs only after cells have ceased to elongate and have started differentiation (MacAdam and Grabber 2002). According to Ding et al. (2012) and Zeng et al. (2014), there are two types of secondary walls in mature cells; the parenchyma-type secondary wall (pSW) and sclerenchyma-type secondary wall (sSW). The pSWs are thickened walls in parenchyma, and collenchyma, which are normally found in living cells. The sSWs are secondarily thickened walls in highly differentiated cells, such as tracheary elements and fibers, which are elongated, dead cells. Secondary walls are thus prominent features of xylem, fibers and sclerenchyma, plant tissues that provide support and water transport. The typical grass secondary wall is largely composed of cellulose and GAX, but in contrast to the primary wall is usually highly lignified and consists of a multilayered structure. From outside to inside: a highly lignified compound middle lamellae (CML) containing middle lamellae and primary wall (Engels 1998); a thin S1 layer; a thick, less-lignified middle S2 layer; a thin inner S3 layer; and a warty layer formed by lignin precursors (Agarwal 2006). The cellulose, GAX, and lignin form a complex intertwined network (Figure 16). The GAX found in secondary cell walls has fewer side-chains than the GAX of primary cell walls. This results in a stronger GAX-cellulose interaction (Vogel 2008). In addition, GAX is bound to lignin through FA in a covalent way which tightens the grid of the secondary cell wall (Molinari et al. 2013).

Lignin comprises a substantial portion (~20%) of the grass secondary wall and essentially fills the pores between the polysaccharides. Grass lignin is similar to dicot lignin in that it is primarily composed of guaiacyl (G)(~35–49%) and syringyl (S)(~40–61%) units. However, grass lignin also contains a small but significant percentage (~4–15%) of *ρ*-hydroxyphenyl (H) units that are only found in trace levels in dicot lignin (Grabber et al. 2004a). The mechanism by which the lignin polymer is formed *in muro* is explained in the next section. Nevertheless, lignin is formed after polysaccharide deposition in the secondary wall layers has begun (Iiyama et al. 1994). Lignin deposition is initiated first at cell corners, then in the middle lamella, and proceeds through the primary wall into the secondary wall layers S1, S2 and S3 (Iiyama et al. 1994). The cell corner, being the junction of the CMLs, always has the highest lignin content. The adjacent lignified PW and S1 layers also have relatively high lignin concentrations. The S2 and S3 layers are further away from the lignification initialization sites and contain less lignin. The warty layer next to S3 is composed of highly cross-linked lignin precursors that are formed while the cell is in the final stage of lignification and death. The two types of secondary walls undergo different lignification processes. Taking maize as an example, while there is steady increase in biomass during the plant's vegetative growth phase, the lignin content stays at very low level. At this stage only sSWs are lignified. A dramatic jump in lignin content is observed during the transition from the vegetative to the reproductive growth phase, which is mainly attributed to lignification of pSWs (Zeng et al. 2014) (Figure 17). In the mature plant, the sSWs are always fully lignified on both sides of the CML and warty layers, the pSWs are partially lignified, and there is a lack of the S3 layer and the warty layer (Zeng et al., 2014). At the moment, the biosynthesis and assembly of the monolignols appears to be similar in dicots and grasses (Humphreys and Chapple 2002; Boerjan et al. 2003; Townsley et al. 2013) and is summarized in the next topic. Unlike dicots, grass lignin contains substantial amounts of FA and pCA (MacAdam and Grabber 2002; Jung 2003). FA residues attached to GAX may serve as nucleation sites for lignin formation (Grabber et al. 2002). FA is primarily esterified to arabinosyl residues of arabinoxylan chains, and feruloylated arabinoxylans are later cross-linked to G units of

lignins via ether bonds (Figure 16). The crosslinking of GAX to lignin by FA in the secondary cell wall is considered to affect digestibility (Grabber 2005). Therefore, a decreased feruloylation of GAX in the cell wall might form a good target for bioenergy feedstock improvement (Molinari et al. 2013). pCA is mainly esterified to the γ-position of the side chains of S lignin units (Ralph et al. 1994; Hatfield et al. 2008) and lignified maize cell walls can contain up to 3% pCA (Grabber et al. 2004a).

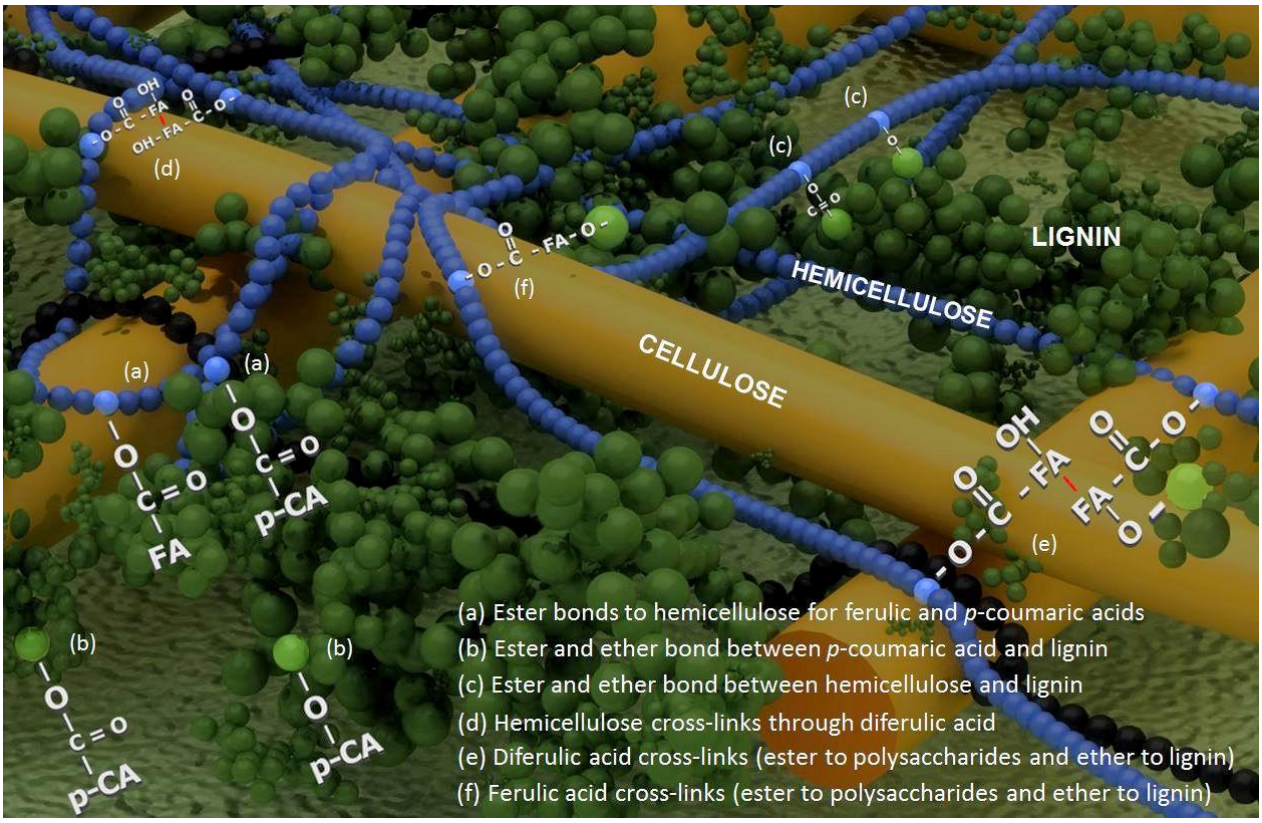


Figure 16. Representation of the secondary cell wall in maize including cellulose, hemicellulose and lignin, with indication of FA and *p*-CA in the lignin polymer and acting as crosslinks in hemicellulose to hemicellulose and hemicellulose to lignin. From Santiago et al. (2013).

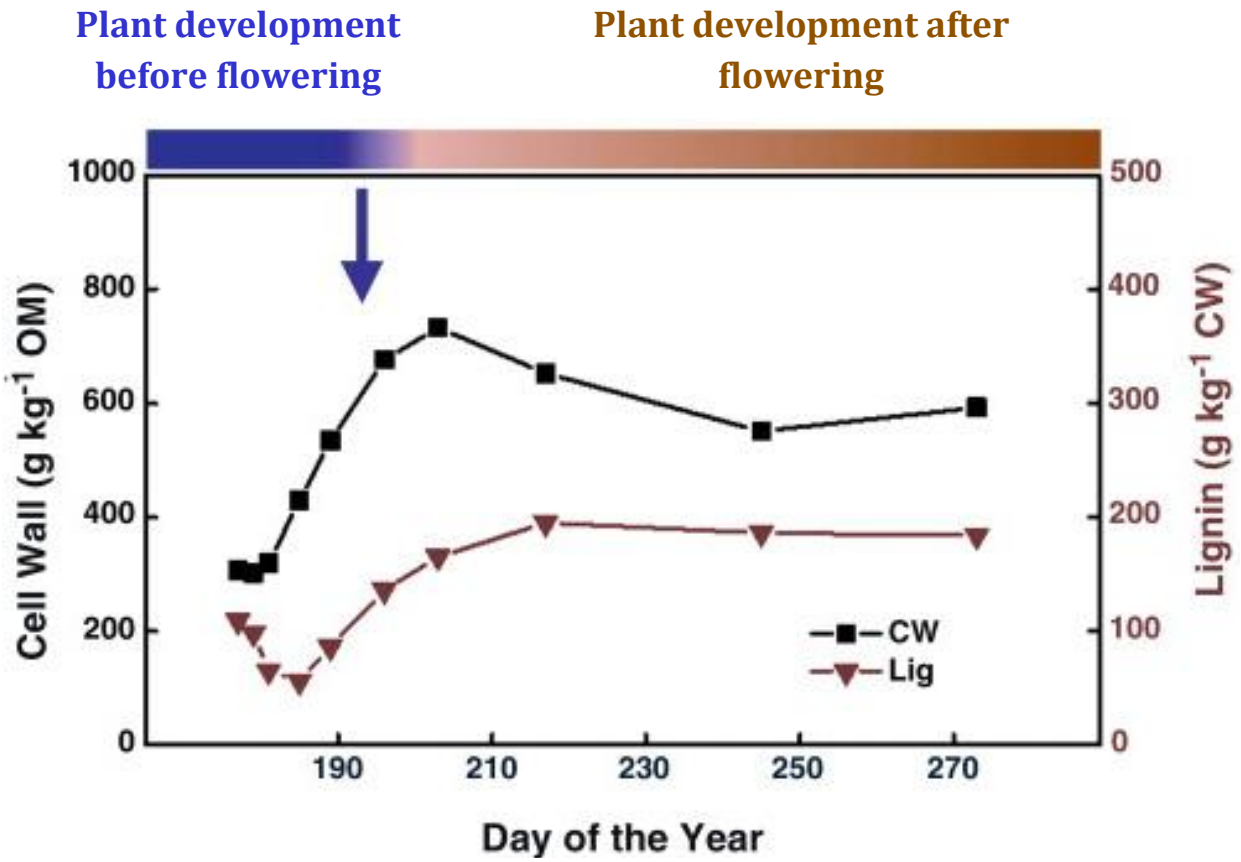


Figure 17. Maize cell wall (CW) biomass and Klason lignin (Lig) content at the plant level during the vegetative and reproductive growth phases of maize, separated by the moment of flowering, as defined above. The blue arrow indicates the approximate date of growth phase transition to plants with visible silks, coinciding with cessation of internode elongation. The data were reproduced from Jung and Casler (2006) in a figure by Zeng et al. (2014). The data was collected from a two-year field trial using three non-related maize hybrids (A632 x A619, A679 x FR481, and Mycogen 2677) of similar relative maturity and were planted on the University of Minnesota St. Paul campus in May 1998 and 1999.

7.4 Lignin biosynthesis

Lignin is synthesized from the oxidative coupling of *p*-hydroxycinnamyl alcohol monomers and related compounds (Boerjan et al. 2003; Ralph et al. 2004; Vanholme et al. 2010b). The main units in the polymer, *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units, are derived from the monolignols *p*-coumaryl, coniferyl and sinapyl alcohol. The monomers differ in the number of methoxyl substituents on the aromatic ring (Sims et al. 2010; Vanholme et al. 2012b). The presence of lignin in the secondary wall negatively influences biomass quality for applications in feed and biofuel industry (Jung et al. 1997; Baucher and Halpin 2003; Boerjan et al. 2003; Boudet et al. 2003; Grabber et al. 2004a; Chen and Dixon 2007; Li et al. 2008; Hisano et al. 2009; Shen et al. 2009; Van Acker et al. 2013; Zeng et al. 2014). Therefore, research efforts are put into developing strategies for either reducing the amount of lignin in the plant without affecting plant fitness, or incorporating novel or alternative structures into the lignin polymer (= lignin engineering) so that the lignin becomes more degradable or better extractable (Vanholme et al. 2012a). For targeted lignin engineering and investigating or anticipating possible effects of lignin perturbation, a good

knowledge is needed of the related pathways and how they are connected (Vogt 2010; Li et al. 2010c; Vanholme et al. 2012a).

7.4.1 The lignin biosynthetic pathway and its regulation

In the following text, the biosynthetic pathway is summarized, based on the publication of Vanholme et al. (2012a) and is of great importance for the work described in chapters 4-5-6-7. The different units of the lignin polymer are produced by the so-called phenylpropanoid pathway (Figure 18). Compounds that are produced by this pathway are called phenylpropanoids and contribute to all aspects of plant responses towards biotic and abiotic stimuli (Dixon et al. 2002; Vogt 2010). For example, they are indicators of plant stress responses upon variation of light or mineral treatment, but are also key mediators in pest resistance responses (La Camera et al. 2004). They promote invasion of new habitats (Bais et al. 2003) and provide the biochemical resources for successful reproduction (Dudareva et al. 2004). Phenylpropanoid-based polymers, like lignin, suberin, or condensed tannins, contribute substantially to the stability and robustness of plants towards mechanical or environmental damage, like wounding or drought (Vogt 2010). The diverse set of phenylpropanoids is derived from a small core set of compounds that is provided by yet another pathway, the shikimate pathway. This seven-step plastid-localized pathway produces the aromatic amino acids phenylalanine, tyrosine and tryptophan. The general phenylpropanoid pathway uses phenylalanine as an entry substrate and is considered 'general' or 'core' until the production of *p*CA in two steps (Dixon et al. 2002), the production of *p*-coumaroyl CoA in three steps (Vogt 2010) or, after seven steps, resulting in feruloyl-CoA (Humphreys and Chapple 2002; Vanholme et al. 2010b; Vanholme et al. 2012a) (Figure 18). In any case, the formation of lignin in angiosperms is, until present, considered conserved between monocots and dicots (Vogel 2008) and is described hereafter.

The phenylpropanoid pathway starts with the deamination of phenylalanine to cinnamate by phenylalanine ammonia-lyase (PAL). Next, the hydroxylation of the aromatic ring leads to *p*CA, a reaction catalyzed by cinnamate 4-hydroxylase (C4H). Alternatively, a PAL isozyme, specific for the grasses, which has also tyrosine ammonia-lyase activity (TAL) might catalyze the deamination of tyrosine directly into *p*CA (Rosler et al. 1997). *p*CA is considered as a central compound in the phenylpropanoid pathway, leading towards the different monolignols but also to various other connected pathways such as flavonoid, coumarin, isoflavonoid, stilbene, aurone, cutin, suberin, proanthocyanidin, lignan, phenylpropene, acylated polyamine and phenylpropanoid ester biosynthesis (Vogt 2010). For the production of monolignols, *p*CA is activated to a thioester by 4-coumarate:CoA ligase (4CL) resulting in *p*-coumaroyl-CoA. The subsequent 3-hydroxylation of *p*-coumaroyl-CoA to caffeoyl-CoA has been revised in the past decade and involves three enzymatic steps, as was demonstrated in dicots (Jouanin and Lapierre 2012). First, *p*-coumaroyl-CoA is transesterified to its quinic or shikimic acid ester derivative by hydroxycinnamoyl-CoA: shikimate/quinate hydroxycinnamoyltransferase (HCT). *p*-Coumaroyl shikimate or quinate is then hydroxylated by *p*-coumarate 3-hydroxylase (C3H) and then transesterified again by HCT to caffeoyl-CoA. However, alternative routes for 3-hydroxylation can be present, as was demonstrated in poplar (Chen et al. 2011). Feruloyl-CoA is formed after methylation of the 3-hydroxyl group by caffeoyl-CoA O-methyltransferase (CCoAOMT).

The monolignol-specific pathway includes four well-studied enzymatic steps that convert feruloyl-CoA into the monolignols coniferyl alcohol and sinapyl alcohol (Humphreys and Chapple 2002; Boerjan et al. 2003). First, feruloyl-CoA is reduced to coniferaldehyde by cinnamoyl-CoA reductase (CCR). Hydroxylation at the 5-position is catalyzed by ferulate 5-hydroxylase (F5H), which is also often called coniferaldehyde 5-hydroxylase (CAld5H) to reflect its preferred substrate, to produce 5-hydroxyconiferaldehyde (Osakabe et al. 1999; Humphreys et al. 1999). The subsequent methylation of the newly formed 5-hydroxyl group is catalyzed by caffeic acid O-methyltransferase (COMT), whose preferred substrate is the aldehyde (Li et al. 2000; Parvathi et al. 2001), to provide sinapaldehyde. Further reduction to their corresponding alcohols, coniferyl alcohol and sinapyl alcohol, is catalyzed by cinnamyl alcohol dehydrogenase (CAD) (Vanholme et al. 2012a).

The formation of lignin in the secondary cell wall is highly dependent on tissue type and is only formed after cells have stopped expanding, as described above. Not surprisingly, a transcriptional network is involved in the coordinated regulation of the biosynthesis of secondary walls (Demura and Ye 2010). Most of the knowledge available has been generated in the dicot Arabidopsis and the extent of conservation between monocots and dicots is unknown. Furthermore, as the cell walls of grasses are so distinct from dicot cell walls, there are likely to be unique aspects in secondary wall regulation (Handakumbura and Hazen 2012). In grasses, most knowledge of these aspects has been generated in maize and rice. The regulatory network acts at different levels with a series of SECONDARY CELL WALL ASSOCIATED NAC (SWN) transcription factors, orthologs of the Arabidopsis SECONDARY WALL-ASSOCIATED NAC-DOMAIN (SWN) and VASCULAR-RELATED NAC DOMAIN (VND) proteins, acting as master switches in a variety of tissues (Demura and Fukuda 2007; Zhong and Ye 2007). These SWNs then regulate a cascade of downstream MYB transcription factors, such as MYB46 for which the maize and rice orthologs were characterized (Zhong and Ye 2012). In grasses, several MYB transcription factors (ZmMYB31, ZmMYB42, ZmMYB2, ZmMYB8, and ZmMYB39) were identified that repressed the lignin biosynthetic *ZmCOMT* gene (Fornalé et al. 2006; Fornalé et al. 2010). Furthermore, the switchgrass PvMYB4 has been identified as a repressor of lignin biosynthetic genes (Shen et al. 2012). At this moment, the identified genes in grass secondary wall regulation are homologs of the known Arabidopsis system (Handakumbura and Hazen 2012). Further elucidation of the transcriptional network, especially in grasses, will provide valuable tools to modify plant biomass production (Demura and Ye 2010).

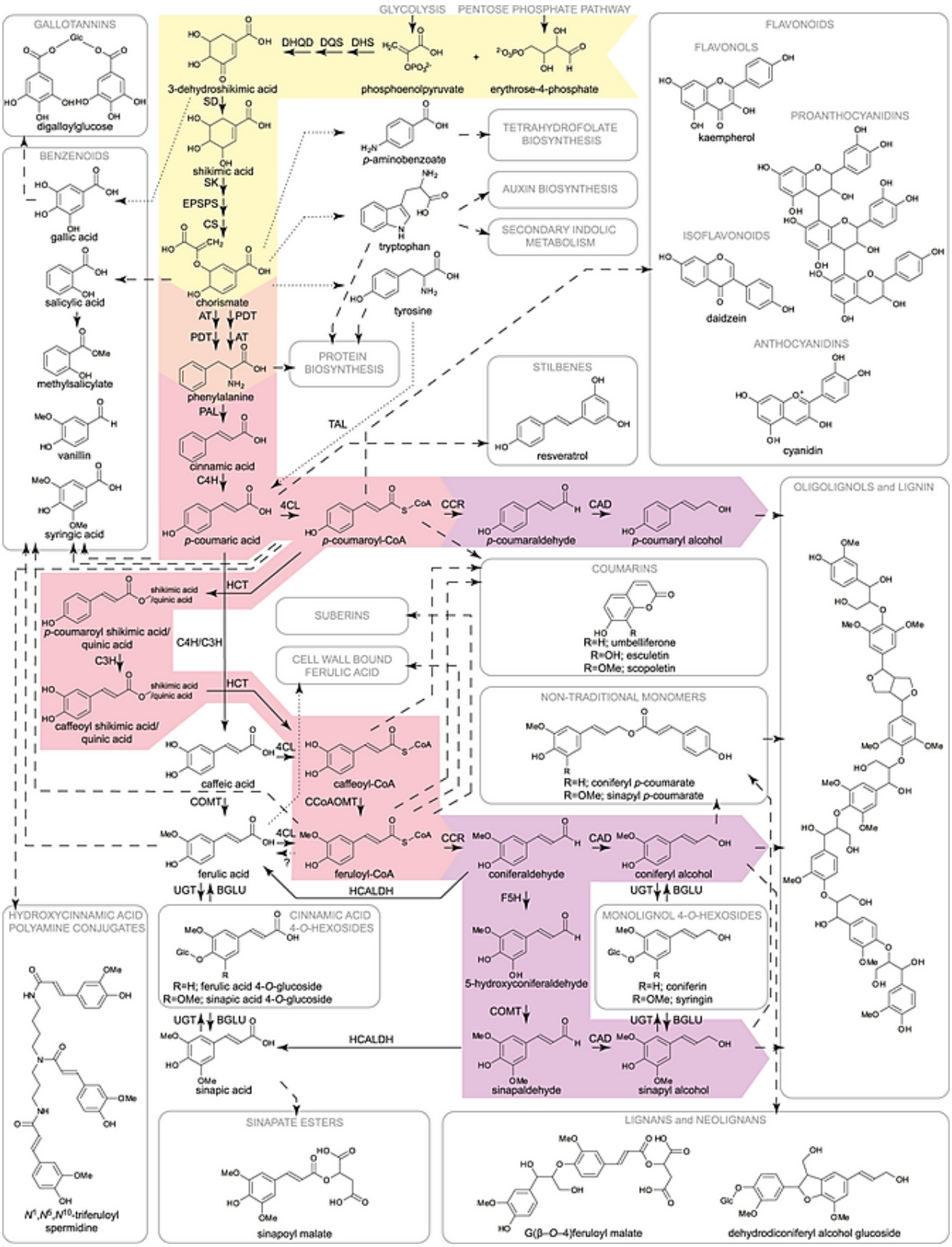

GLYCOLYSIS
PENTOSE PHOSPHATE PATHWAY
GALLOTANNINS
digalloylglucose
BENZENOIDS
gallic acid
salicylic acid
methylsalicylate
vanillin
syringic acid
DHQD
DQS
DHS
3-dehydroshikimic acid
phosphoenolpyruvate
erythrose-4-phosphate
SD
shikimic acid
SK
EPSPS
CS
chorismate
AT
PDT
phenylalanine
PAL
cinnamic acid
C4H
p-coumaric acid
4CL
p-coumaroyl-CoA
CCR
p-coumaraldehyde
CAD
p-coumaryl alcohol
TETRAHYDROFOLATE BIOSYNTHESIS
p-aminobenzoate
AUXIN BIOSYNTHESIS
SECONDARY INDOLIC METABOLISM
tryptophan
tyrosine
PROTEIN BIOSYNTHESIS
TAL
STILBENES
resveratrol
FLAVONOIDS
FLAVONOLS
kaempherol
PROANTHOCYANIDINS
ISOFLAVONOIDS
daidzein
ANTHOCYANIDINS
cyanidin
OLIGOLIGNOLS and LIGNIN
HCT
shikimic acid /quinic acid
p-coumaroyl shikimic acid/ quinic acid
C4H/C3H
C3H
caffeoyl shikimic acid/ quinic acid
SUBERINS
CELL WALL BOUND FERULIC ACID
COUMARINS
R=H; umbelliferone
R=OH; esculetin
R=OMe; scopoletin
NON-TRADITIONAL MONOMERS
R=H; coniferyl p-coumarate
R=OMe; sinapyl p-coumarate
caffeic acid
COMT
caffeoyl-CoA
CCoAOMT
ferulic acid
feruloyl-CoA
HCALDH
coniferaldehyde
coniferyl alcohol
UGT
BGLU
F5H
5-hydroxyconiferaldehyde
MONOLIGNOL 4-O-HEXOSIDES
R=H; coniferin
R=OMe; syringin
sinapic acid
sinapaldehyde
sinapyl alcohol
HYDROXYCINNAMIC ACID POLYAMINE CONJUGATES
N5,N6,N10-triferuloyl spermidine
CINNAMIC ACID 4-O-HEXOSIDES
R=H; ferulic acid 4-O-glucoside
R=OMe; sinapic acid 4-O-glucoside
SINAPATE ESTERS
sinapoyl malate
LIGNANS and NEOLIGNANS
G(β–O–4)feruloyl malate
dehydrodiconiferyl alcohol glucoside

Figure 18. Phenolic metabolism in plants, as was published by Vanholme et al. (2012). The phenolic metabolite classes are given (in gray frames), as well as pathways and metabolic sinks that use phenolic metabolites or shikimate pathway intermediates as substrates. Representative metabolites are given for phenolic classes. Not every phenolic metabolic class shown is present in every plant species. The major route towards the monolignols p-coumaryl, coniferyl and sinapyl alcohol is given in color; the shikimate pathway (yellow), phenylalanine biosynthesis (orange), general phenylpropanoid pathway (pink) and monolignol-specific pathway (purple). Arrows with dashed lines designate knownroutes that involve multiple enzymatic steps; for simplicity, the individual enzymatic steps are not shown. Arrows withdotted lines designate unknown or unauthenticated routes. Arrows with a question mark are routes that have been suggested in the literature. DHS, 3-deoxy-D-arabinoheptulosonate 7-phosphate synthase;DQS,3-dehydroquinate synthase;DHQD,3-dehydroquinate dehydratase; SD, shikimate dehydrogenase; SK, shikimate kinase; EPSPS, 5-enolpyruvylshikimate-3-phosphate synthase; CS, chorismate synthase; AT, amino transferase; TAL, tyrosine ammonia-lyase; PAL, phenylalanine ammonia-lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate: CoA ligase; HCT, hydroxycinnamoyl-CoA: shikimate/quinate hydroxycinnamoyltransferase; C3H, p-coumarate 3-hydroxylase; CCoAOMT, caffeoyl-CoA O-methyltransferase; CCR, cinnamoyl-CoA reductase; F5H, ferulate 5-hydroxylase; COMT, caffeic acid O-methyltransferase; CAD, cinnamyl alcohol dehydrogenase; UGT, UDP-glucosyltransferase; HCALDH, hydroxycinnamaldehyde dehydrogenase; BGLU, b-glucosidase.

7.4.2 The composition of the lignin monomer

After formation in the cytoplasm, the monolignols are transported to the apoplast in a manner that is poorly understood (Bonawitz and Chapple 2010; Zeng et al. 2014). Yet after translocation, lignin is polymerized *in muro* by radical coupling.

Generally, the composition of the lignin polymer is depicted as the radical coupling of sinapyl and coniferyl alcohol as major constitutent and *p*-coumaryl alcohol as minor constituent. However, recent advances in our understanding of lignin composition have shown that the lignin polymer can consist of many more constituents besides the three main lignin units. It has been demonstrated that besides the addition of FA and pCA to the arabinofuranosyl unit of GAX (Molinari et al. 2013; Bartley et al. 2013), other members of the BAHD acyl-coA transferase superfamily can add pCA (and probably also FA) to monolignols. Withers et al. (2012) identified a grass-specific enzyme, pCA monolignol transferase (PMT), that is capable of acylating monolignols with pCA thereby leading to series of monolignol conjugates (demonstrated *in planta* by Petrik et al. (2013). These monolignol-hydroxycinnamate conjugates are primary building blocks for lignin in grasses (but analogously with monolignol acetates and *p*-hydroxybenzoates in other plants) (Hatfield et al. 2008; Ralph 2010). Recent evidence suggests that even the hydroxycinnamic acids themselves (pCA, FA and sinapic acid) can be monomers in lignification in wild-type and transgenic plants, undergoing radical cross-coupling reactions to incorporate into the polymer. Thus, besides the three traditional monomers H, G and S, also pCA, FA, sinapic acid and various cross-coupling products should be considered as true monomers as building blocks of the lignin polymer in grasses (Ralph 2010).

The combinatorial oxidative coupling of the above-stated monomers is under simple chemical control. The units are linked to each other in no set order by 'linkage types' that are defined by the way a monomer couples with another monomer or, more commonly, the way a monomer couples with the growing oligomer, or by the coupling of two oligomer units (Ralph 2010). The two most common monolignols are coniferyl and sinapyl alcohol and these will be used to illustrate the different linkage types, typically found in angiosperm lignin (Figure 19). Following oxidation of the monolignols by peroxidase and/or laccase, the resulting electron-delocalized radical has unpaired

electron density at its 1-, 3-, O–4-, 5-, and 8-positions (Figure 19B); note that much of the lignin literature uses the Greek letter β for the 8-position. As radical coupling at the 8-position is favored, coupling with another monolignol radical affords, after rearomatization, a mixture of dehydrodimers with 8–8-, 8–5-, and 8–O–4-linkages (Figure 19C). Following dimerization, polymerization will continue by the coupling of the 8-position of an incoming monolignol radical to the O–4-position of the dimer's phenolic end. In the case of a G dimer, coupling can also occur, albeit at a lower frequency, to the 5-position. Thus, chain elongation creates 8–5- and 8–O–4-linkages (Morreel et al. 2010b; Morreel et al. 2010a). It is the 8–O–4-coupling that is called a β-ether linkage (Ralph 2010). Besides 8–8-, 8–5-, and 8–O–4-linkages, also 5–5-and 5-O-4-linkages can form, but these occur only between oligomers (Ralph et al., 2008). The plasticity of lignin polymerization permits the incorporation of any phenolic that enters the lignification site, a characteristic that could be explored for creating a lignin polymer hat is more suitable for extraction in biofuel applications (Baucher and Halpin 2003; Grabber et al. 2008; Morreel et al. 2010a; Ralph 2010; Vanholme et al. 2010c).

Analysis of the lignin composition and aspects of its structure can be examined using methods that degrade the lignin polymer in smaller fragments that can be analyzed by analytical methods. The most informative method for lignin structure elucidation is the thioacidolysis method (Lapierre 1993). Thioacidolysis selectively cleaves β-aryl ether units (8-O-4 linkages) in lignin polymers, releasing low-molecular-mass thio-ethylated compounds. For example, the guaiacyl and syringyl monomers (Figure 19A) are routinely used to determine the ratio of such units involved in β-ether units and to infer the composition of lignins from their monomers, coniferyl and sinapyl alcohols (Ralph et al. 2008). However, this method and similar often used methods (e.g. nitrobenzene oxidation, DFRC) only analyze a fraction of the lignin polymer. NMR analysis is perhaps the most informative method for the elucidation of unknown compounds, but it is has low-sensitivity and works best on purified compounds. Because extracted lignins or cell walls are a complex mixture of different components, NMR does not easily discern which units are attached to which, yet the method is very powerful for examining the proportions of the different building blocks and the different bond types (Morreel et al. 2010b). Over the past decade, liquid chromatography (LC) coupled to ion trap (IT) mass spectrometry (MS) has been used to detect and resolve lignin structures such as dilignols, trilignols, and tetralignols of the traditional H, G and S units but also many more and often less abundant units derived from e.g. sinapyl p-hydroxybenzoate (Morreel et al. 2004a), feruloyl tyramine (Dauwe et al. 2007), feruloyl malate (Rohde et al. 2004) or totally new linkage types based on the 5-hydroxyguaiacyl unit (Morreel et al. 2004b). Furthermore, the MS-based fragmentation pattern can be used for the sequencing of lignin oligomers up to hexamers (for which the typical MS spectra are shown in Figure 19), providing valuable information about the nature of the lignin polymer and the type of units and linkages present (Morreel et al. 2010a).

Although much information about phenolic metabolism has been gathered over recent decades, the low proportion of identified metabolites in phenolic profiling studies (Morreel et al. 2010a; Vanholme et al. 2010c) underscores the complexity of these pathways and the need for methods to speed up structural characterization (Vanholme et al. 2012a).

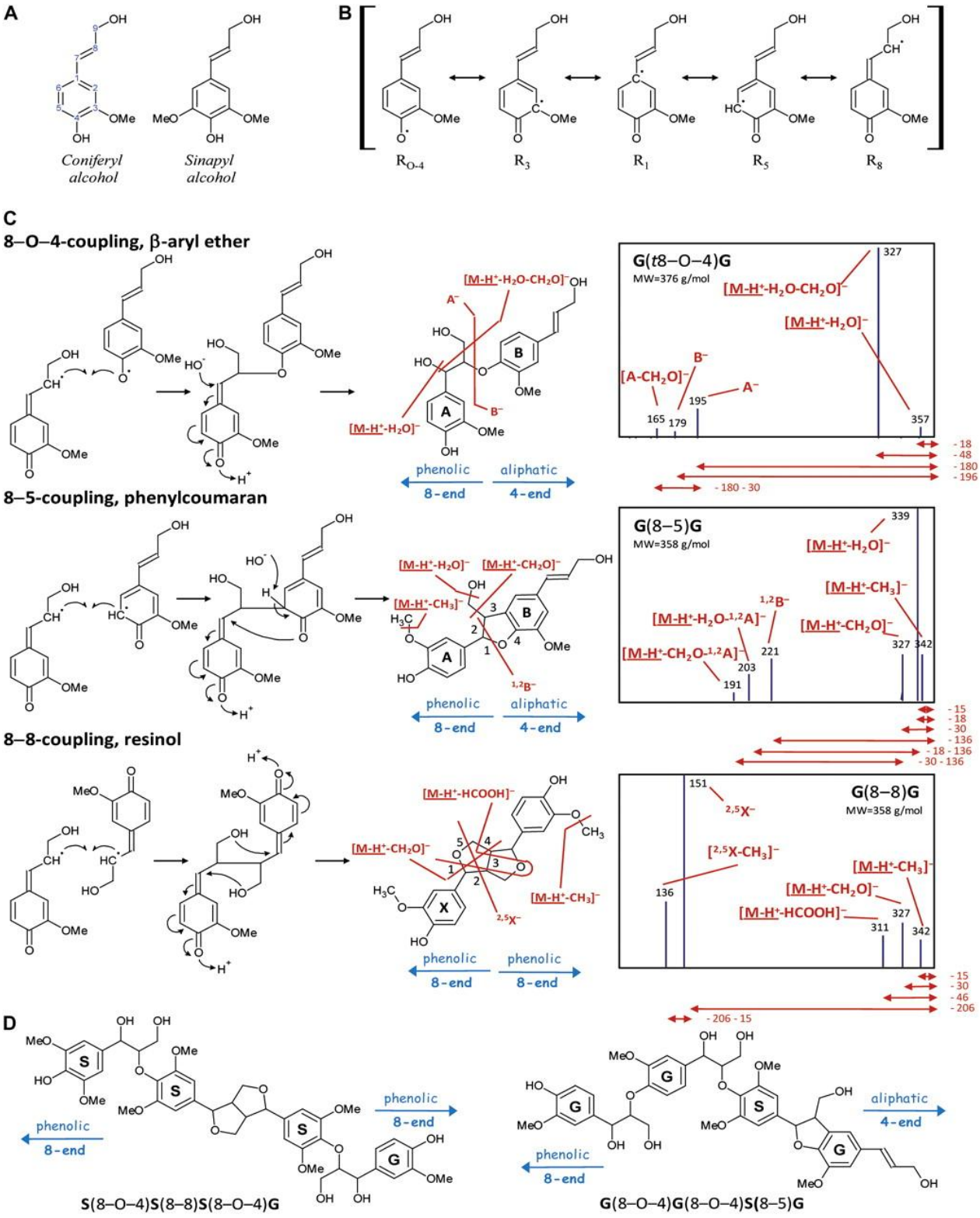


Figure 19. Radical-radical coupling during lignin polymerization as was published by Morreel et al. (2010a). A, Main angiosperm monolignols. B, Delocalized radical following monolignol oxidation. C, Main types of monolignol dimerizations and their MS2 spectra. Below the spectra, the neutral loss for each first product ion is noted in D. For phenylcoumarans and resinols, the pathway II first product ions are specified by superscripts indicating the

bonds that are cleaved. This numbering system has been suggested previously to annotate the first product ions upon CID of flavonoids (Ma et al. 1997; Fabre et al. 2001; Morreel et al. 2006) **and upon collision-induced dissociations (CID) of oligosaccharides** (Domon and Costello 1988)**. Most first product ions are due to charge-driven fragmentations in which the charge center initiates the dissociation, yet the pathway I-associated formaldehyde loss upon CID of β-aryl ethers as well as the pathway I-associated methyl radical loss upon CID of phenylcoumarans and resinols are examples of charge-remote fragmentations (i.e. where the fragmentating center occurs remote from the charge center; Bowie, 1990). D, Examples of tetralignols observed in poplar xylem** (Morreel et al. 2010a)**.**

8. Pleiotropic effects of improving yield and saccharification efficiency

As lignin abundance is negatively correlated with digestibility and saccharification efficiency (Zeng et al. 2014), the main focus for improvement of these traits has been the perturbation of lignin biosynthetic genes. Nevertheless, genes involved in lignin biosynthesis are also functional in other biochemical pathways, as described above. The interconnectivity between lignin biosynthesis and related pathways such as flavonoid, amino acid, benzenoid, suberin, coumarin, stilbene, tetrahydrofolate and auxin biosynthesis is most pronounced in the initial steps of the phenylpropanoid pathway (Vogt 2010; Vanholme et al. 2012b). One could therefore assume that perturbation of *PAL*, *C4H*, *4CL*, *CCR* and *HCT* would cause more severe far-reaching effects than perturbation of later steps in the pathway encoded by *COMT*, *F5H* and *CAD*. However, a study of the systems-wide effects of lignin biosynthesis gene perturbation in Arabidopsis revealed that far-reaching effects on the transcriptome and metabolome level are present for perturbations in both, early and later steps of the phenylpropanoid pathway. For example, based on transcript-based clustering of the different lignin mutants, mutants of the central part of the phenylpropanoid pathway (*c4h-3*, *4cl1-1*, *4cl1-2*, *ccoaomt1-3*, and *ccoaomt1-5*) grouped in a separate transcript-based subcluster, but surprisingly, *pal1* mutants appeared in the same cluster as *f5h1-2* and *comt* mutants (Vanholme et al. 2012b). This indicates that similar responses on the transcript level do not necessarily correlate with the hierarchical position in the pathway. In addition, it remains largely unclear why certain mutants develop a growth defect while others do not. In Arabidopsis, mutants in *C3H* and *HCT* are sublethal. Stem growth rate was slower than the wild type in *c4h-2*, *ccr1-3*, *ccr1-6* and *ccoaomt1-3* which resulted in reduced final stem height for *c4h-2*, *ccr1-3*,and *ccr1-6*. In contrast, mutants in *PAL1*, *PAL2*, *4CL*, *F5H*, *COMT* and *CAD* did not show this reduction or delay in growth despite obvious systems-wide effects on the transcriptomic and metabolic level (Vanholme et al. 2012b). The presence of such a yield penalty might even be dependent on the genetic background and the environmental conditions as was shown for Sorghum *bmr-6* (deficient in CAD) in different genetic backgrounds and grown in different locations (Casler et al. 2003; Pedersen et al. 2005). Furthermore, several different but related biosynthetic pathways can be co-regulated in secondary cell wall biosynthesis (Boudet et al. 2003), therefore genes within one pathway are likely to have pleiotropic effects (Chen 2011). There might also be a cell wall integrity sensing system present in plant cell walls (Hématy et al. 2007; Hématy et al. 2009; Seifert and Blaukopf 2010; Ringli 2010), adding an additional level of regulation to the already complex network of interconnected pathways in secondary metabolism.

Similar effects have been described in maize, where it has been shown that selective breeding for specific agronomics traits can have adverse effects on others. E.g. in the past fifty years, maize

breeding efforts in Western Europe have achieved an impressive improvement in whole plant yield of ~ 4.4 t/ha increase leading to yields of more than 10 t/ha (FAO Statistics Division 2013c). This was accompanied by a substantial decrease in cell wall digestibility, which resulted in a reduced feeding value of elite maize hybrids (Barrière et al. 2006). Selection of alleles for good stalk standability and breakage resistance helped to increase whole plant yield, but also likely eliminated alleles favorable for cell wall digestibility (Chen, 2011). Reversely, the natural brown midrib mutants of maize and sorghum display improved digestibility but perform less in the field (Pedersen et al. 2005).

9. Cellulosic ethanol

Currently, bioethanol is mainly produced by fermentation of sugars derived from starch found in of corn and wheat grains or by fermentation of soluble sugars extracted form sugarcane. These sugars are easily accessible for the degradative enzymes. In contrast, cellulosic ethanol refers to the fermentation of sugars that are derived from the polysaccharides found in plant biomass, mainly the cell walls. This cellulosic ethanol is not so easily produced, as the polysaccharides need to be accessed and degraded to monosaccharides before fermentation can take place (Figure 20). The process involves the allocation or removal of lignin (Zeng et al. 2014). This lignin allocation or removal is performed by a pretreatment step (see below). The complicated combined thermo and biochemical process that is necessary to produce cellulosic ethanol results in high capital costs for the construction of commercial-scale plants. For that reason, the currently installed production capacity is small (100 million liters in US and Canada and 82 million liters in EU; Janssen et al. 2013) and contribution to the total amount of biofuel production is very small. Many production facilities are currently being built and many more are planned, however, as the trend is that construction takes longer than anticipated and many facilities may never be built due to high capital costs, whether second-generation biofuel production will continue to expand in the future is hard to tell (Janssen et al. 2013). The reduction of production costs for cellulosic ethanol aims at improving feedstock properties and improving enzyme and pretreatment efficiencies (Mosier et al. 2005; Eggeman and Elander 2005; Ringli 2010; Limayem and Ricke 2012). Improving lignocellulosic biomass properties include the genetic engineering of plants for reducing the need for pretreatments through lignin modification, for production of cellulases and hemicellulases *in planta* and for increasing plant polysaccharide content and overall biomass (Chapple et al. 2007; Karp and Shield 2008; Gressel 2008; Sticklen 2008; Yuan et al. 2008; Jakob et al. 2009; Abramson et al. 2010; Feltus and Vandenbrink 2012; Jung et al. 2012; Vanholme et al. 2013a).

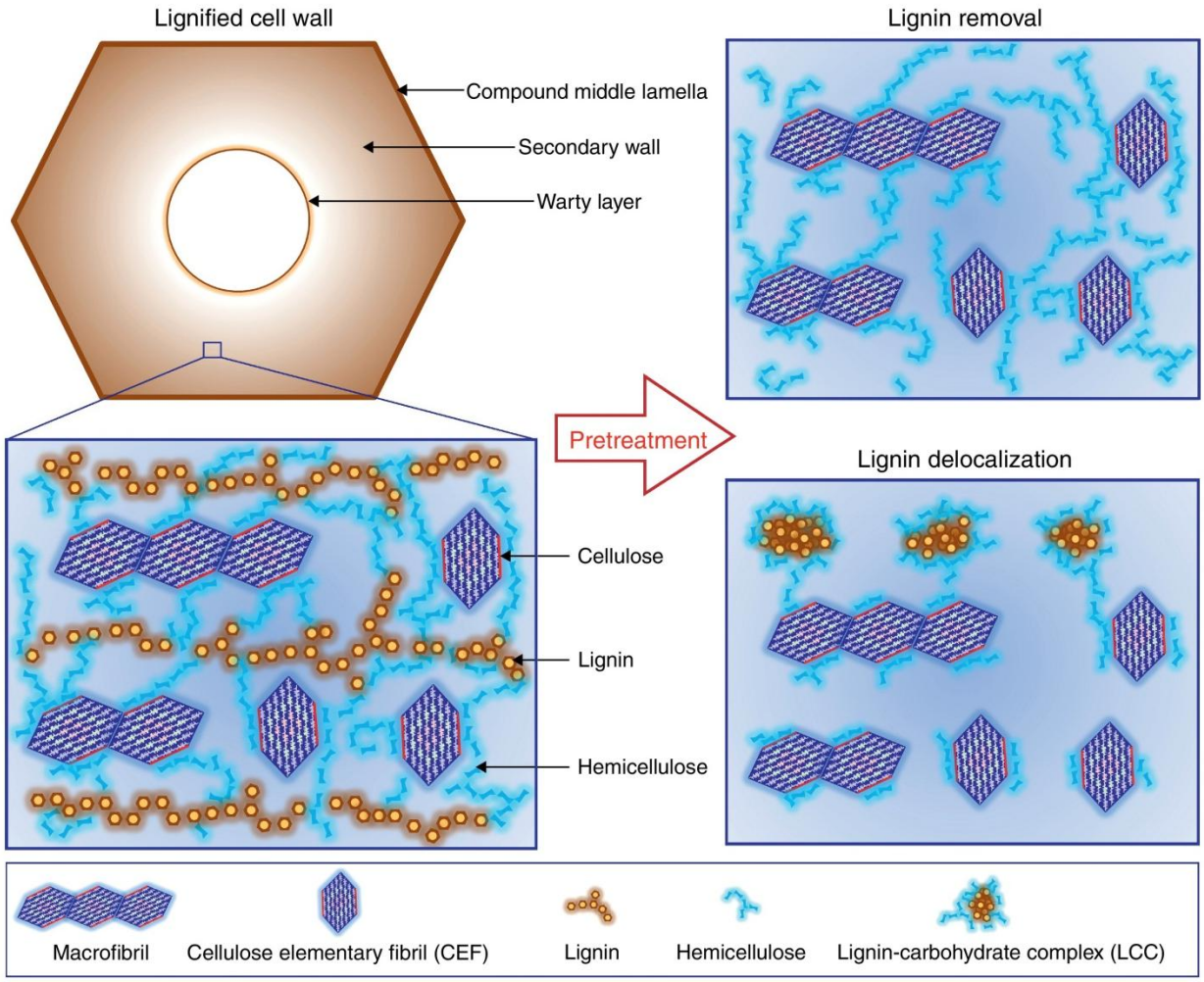


Figure 20. Schematic structure of a lignified cell wall and possible changes by pretreatment. The sSW is enclosed by two condensed lignin layers (i.e. warty layer and CML), whereas the pSW lacks the warty layer. The prerequisite of enzymatic saccharification is the accessibility of enzymes to the polysaccharides that is physically impeded by lignin networks in the SW. Therefore, effective enzyme–cellulose interaction requires either effective removal, such as by lignin bleaching, or delocalization of lignin, such as melting lignin into small LCCs from the active sites on cellulose microfibrils. Reproduced from (Zeng et al. 2014)**.**

9.1 Industrial process of cellulosic ethanol

For the production of second generation bioethanol, the main potential feedstocks are agricultural residues, forest wood, by-products of the wood industry and dedicated energy crops. Several studies have estimated that the availability of biomass resources in Europe do not form a limitation for meeting the future biomass energy targets set by the European union. The goal is a 10% replacement of the transportation fuel by biofuels in 2020, and an estimated demand for 15%-25% biofuels replacement by 2030, given a share of second generation ethanol of 75%-85% in that of biofuels (Ericsson and Nilsson 2006; Fischer et al. 2010a; Fischer et al. 2010b; Gnansounou 2010). However, other issues like the biomass costs and logistics do form possible constraints to the economic potential of second generation bioethanol (Gnansounou 2010).

As commercial-scale cellulosic ethanol production plants are only just starting to operate (Janssen et al. 2013), production processes are likely to still be optimized in the future. An example of such an alternative production process is the one proposed by Kumar et al. (2005) to integrate the feedstock transport to the ethanol production facility and the saccharification process, named simultaneous transport and saccharification. These authors consider that the enzymatic hydrolysis of corn stover can be carried out in pipelines during its transport; the hydrolyzed corn stover could directly enter the ethanol fermentation plant, saving about 0.2 USD cents/L EtOH.

The general production process of cellulosic ethanol is depicted in Figure 21. In short, the lignocellulosic biomass is harvested and transported to a cellulosic ethanol refinery where it is stored. The biomass undergoes first a pretreatment step, which is often physiochemical in nature and conducted at high temperatures. The objective of the pretreatment step is to remove or delocalize lignin before the enzymatic hydrolysis of biomass (Zeng et al. 2014) (depicted in Figure 20). According to Ding et al. (2012), the improvement of the overall efficiency of biomass conversion largely relies on improvement of pretreatment technologies and in particular the effectiveness of lignin modification. It is estimated that pretreatment is the most expensive processing step in cellulosic biomass-to-fermentable sugars conversion (Mosier et al. 2005; Cardona and Sánchez 2007; Limayem and Ricke 2012). Lowering the severity of the pretreatment by lowering the temperature would require a higher enzyme load to achieve a similar fermentable sugar yield. Unfortunately, the enzymes are also expensive, making both the pretreatment and the efficiency of the enzyme complexes a major target for improvement (Himmel et al. 2007; MacLean and Spatari 2009).

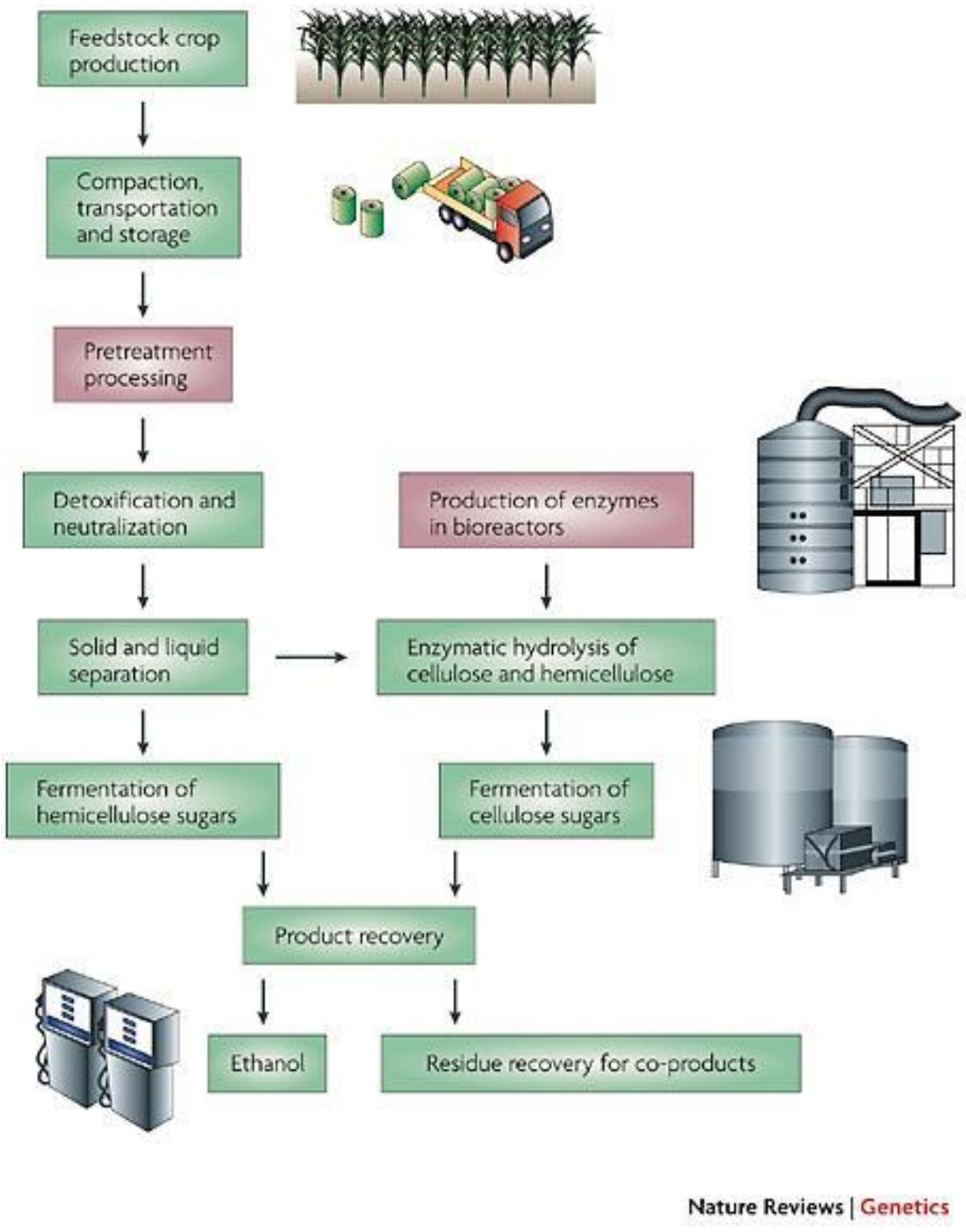


Figure 21. Flow chart showing the steps in the production of cellulosic ethanol from feedstock crops. From Sticklen (2008).

The pretreatment step can be physical (grinding, steam explosion), thermal (elevated temperatures), chemical (alkali or acid) and most commonly, a combination of these. Different types of biomass such as wood chips and vegetative tissues of grasses may require different pretreatment methods (Li et al. 2012; Zeng et al. 2014). Commonly used is the dilute acid pretreatment using H_2SO_4 which is effective at a relatively low cost (Limayem and Ricke 2012). Its main action is the hydrolysis of hemicelluloses, making the cellulose more accessible to enzymatic hydrolysis (Hendriks and Zeeman 2009). The hydrolysis of the glycosyl bonds in hemicelluloses results in the formation of lignin-carbohydrate complexes (LCCs) which are redeposited on the biomass surface as droplets (Figure 20). However, the combination of elevated temperatures and an acidic environment can favor the formation of sugar and lignin degradation compounds such as furfural and HMF and aromatic lignin degradation compounds which affect the microorganism metabolism in the fermentation step (Saha et al. 2005; Alvira et al. 2010). Other commonly used pretreatments, such as ammonia fiber expansion (AFEX) and ammonia recycled percolation (ARP), are based on the use of ammonia, which is an efficient delignification reagent (Kim and Lee 2005). Alternative pretreatment methods are under consideration but all have their advantages and disadvantages so that a clearly improved cost-effective pretreatment method is currently not easily appointed (Eggeman and Elander 2005). Examples of alternative pretreatment technologies, but certainly not all existing ones, that can be used for grass biomass (corn stover, sugar cane bagasse, wheat straw, switchgrass or Miscanthus) are: steam explosion, controlled pH, lime and ionic liquid (Sun and Cheng 2002; Mosier et al. 2005; Eggeman and Elander 2005; Lloyd and Wyman 2005; Wyman et al. 2005; Ballesteros et al. 2006; Lau et al. 2009; Kumar and Wyman 2009; Bals et al. 2010; Li et al. 2010a). All methods are based on either the delocalization of lignin, such as the dilute acid method, or the removal of lignin, such as the AFEX method (Figure 20).

The pretreated biomass undergoes then detoxification, neutralization and separation into its liquid and solid components. The solid components, mostly hemicellulose and cellulose, are then enzymatically hydrolyzed into fermentable sugars (hexose and pentose sugars). Finally, the sugars are separated and fermented into alcohol (Sticklen 2008). The lignin-rich residue is often burned for the production of electricity as is the case for the cellulosic ethanol plant in Northern Italy by Novozymes and beta renewables (http://www.ethanolproducer.com/articles/10332/commercial-scale-cellulosic-ethanol-refinery-opens-in-italy). This sequential configuration employed to obtain cellulosic ethanol is called separate hydrolysis and fermentation (SHF) and is opposed to simultaneous saccharification and fermentation (SSF). The reason why mostly SHF is carried out is that hexose and pentose fermentation is usually separated since microorganisms used for these fermentations have different reaction kinetics and different sensitivities to inhibitors (Cardona and Sánchez 2007).

10. References

Abramson M, Shoseyov O, Shani Z (2010) Plant cell wall reconstruction toward improved lignocellulosic production and processability. Plant Sci 178:61–72. doi: 10.1016/j.plantsci.2009.11.003

Agarwal UP (2006) Raman imaging to investigate ultrastructure and composition of plant cell walls: distribution of lignin and cellulose in black spruce wood (Picea mariana). Planta 224:1141–53. doi: 10.1007/s00425-006-0295-z

Alvira P, Tomás-Pejó E, Ballesteros M, Negro MJ (2010) Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. Bioresour Technol 101:4851–61. doi: 10.1016/j.biortech.2009.11.093

Andersen J, Zein I, Wenzel G (2008) Characterization of phenylpropanoid pathway genes within European maize (Zea mays L.) inbreds. BMC plant … 14:1–14. doi: 10.1186/1471-2229-8-2

Auzanneau J, Huyghe C, Escobar-Gutiérrez AJ, et al. (2011) Association study between the gibberellic acid insensitive gene and leaf length in a Lolium perenne L. synthetic variety. BMC Plant Biol 11:183. doi: 10.1186/1471-2229-11-183

Bais HP, Vepachedu R, Gilroy S, et al. (2003) Allelopathy and exotic plant invasion: from molecules and genes to species interactions. Science 301:1377–80. doi: 10.1126/science.1083245

Ballesteros I, Negro MJ, Oliva JM, et al. (2006) Ethanol production from steam-explosion pretreated wheat straw. Appl Biochem Biotechnol 129-132:496–508.

Bals B, Rogers C, Jin M, et al. (2010) Evaluation of ammonia fibre expansion (AFEX) pretreatment for enzymatic hydrolysis of switchgrass harvested in different seasons and locations. Biotechnol Biofuels 3:1. doi: 10.1186/1754-6834-3-1

Barrière Y, Alber D, Dolstra O, et al. (2006) Past and prospects of forage maize breeding in Europe. II. History, germplasm evolution and correlative agronomic changes. Maydica 51:435–449.

Bartley LE, Peck ML, Kim S-R, et al. (2013) Overexpression of a BAHD acyltransferase, OsAt10, alters rice cell wall hydroxycinnamic acid content and saccharification. Plant Physiol 161:1615–33. doi: 10.1104/pp.112.208694

Baucher M, Halpin C (2003) Lignin: genetic engineering and impact on pulping. Crit. Rev. …

Bevan MW, Garvin DF, Vogel JP (2010) Brachypodium distachyon genomics for sustainable food and fuel production. Curr Opin Biotechnol 21:211–7. doi: 10.1016/j.copbio.2010.03.006

Birch CJ, Carberry PS, Muchow RC, et al. (1990) Development and evaluation of a sorghum model based on CERES-Maize in a semi-arid tropical environment. F Crop Res 24:87–104. doi: 10.1016/0378-4290(90)90023-5

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. Annu Rev Plant Biol 54:519–46. doi:

10.1146/annurev.arplant.54.031902.134938

Bonawitz ND, Chapple C (2010) The genetics of lignin biosynthesis: connecting genotype to phenotype. Annu Rev Genet 44:337–63. doi: 10.1146/annurev-genet-102209-163508

Boudet AM, Kajita S, Grima-Pettenati J, Goffner D (2003) Lignins and lignocellulosics: a better control of synthesis for new and improved uses. Trends Plant Sci 8:576–81. doi: 10.1016/j.tplants.2003.10.001

Briske DD (1991) Developmental morphology and physiology of grasses. In: Heitschmidt RK, Stuth JW (eds) Grazing Manag. An Ecol. Perspect. Timber Press, Portland Oregon, pp 85–108

Brkljacic J, Grotewold E, Scholl R, et al. (2011) Brachypodium as a model for the grasses: today and the future. Plant Physiol 157:3–13. doi: 10.1104/pp.111.179531

Buckner B, Janick-Buckner D, Gray J, Johal GS (1998) Cell-death mechanisms in maize. Trends Plant Sci 3:218–223. doi: 10.1016/S1360-1385(98)01254-0

Cardona CA, Sánchez OJ (2007) Fuel ethanol production: process design trends and integration opportunities. Bioresour Technol 98:2415–57. doi: 10.1016/j.biortech.2007.01.002

Carpita NC (1996) Structure and Biogenesis of the Cell Walls of Grasses. Annu Rev Plant Physiol Plant Mol Biol 47:445–476. doi: 10.1146/annurev.arplant.47.1.445

Carpita NC, McCann MC (2008) Maize and sorghum: genetic resources for bioenergy grasses. Trends Plant Sci 13:415–20. doi: 10.1016/j.tplants.2008.06.002

Casler MD, Pedersen JF, Undersander DJ (2003) Forage Yield and Economic Losses Associated with the Brown-Midrib Trait in Sudangrass. Crop Sci 43:782. doi: 10.2135/cropsci2003.0782

Cassab GI (1998) Plant Cell Wall Proteins. Annu Rev Plant Physiol Plant Mol Biol 49:281–309. doi: 10.1146/annurev.arplant.49.1.281

Chaffey N (2000) Physiological anatomy and function of the membranous grass ligule. New Phytol 146:5–21. doi: 10.1046/j.1469-8137.2000.00618.x

Chapple C, Ladisch M, Meilan R (2007) Loosening lignin's grip on biofuel production. Nat Biotechnol 25:746–8. doi: 10.1038/nbt0707-746

Chen F, Dixon R a (2007) Lignin modification improves fermentable sugar yields for biofuel production. Nat Biotechnol 25:759–61. doi: 10.1038/nbt1316

Chen H-C, Li Q, Shuford CM, et al. (2011) Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis. Proc Natl Acad Sci U S A 108:21253–8. doi: 10.1073/pnas.1116416109

Chen Y (2011) Pleiotropic effects of genes involved in cell wall lignification on agronomic characters.

Christensen U, Alonso-Simon A, Scheller H V, et al. (2010) Characterization of the primary cell walls of seedlings of Brachypodium distachyon--a potential

model plant for temperate grasses. Phytochemistry 71:62–9. doi: 10.1016/j.phytochem.2009.09.019

Clerget B, Dingkuhn M, Gozé E, et al. (2008) Variability of phyllochron, plastochron and rate of increase in height in photoperiod-sensitive sorghum varieties. Ann Bot 101:579–94. doi: 10.1093/aob/mcm327

Colasanti J, Coneva V (2009) Mechanisms of floral induction in grasses: something borrowed, something new. Plant Physiol 149:56–62. doi: 10.1104/pp.108.130500

Corredor DY, Salazar JM, Hohn KL, et al. (2009) Evaluation and characterization of forage Sorghum as feedstock for fermentable sugar production. Appl Biochem Biotechnol 158:164–79. doi: 10.1007/s12010-008-8340-y

Cosgrove D (2005) Growth of the plant cell wall. Nat Rev Mol cell Biol 6:850–861. doi: 10.1038/nrm1746

Cotterill J V, Watkins RW, Brennon CB, Cowan DP (2007) Boosting silica levels in wheat leaves reduces grazing by rabbits. Pest Manag Sci 63:247–53. doi: 10.1002/ps.1302

Craufurd PQ, Bidinger FR (1988) Effect of the Duration of the Vegetative Phase on Shoot Growth, Development and Yield in Pearl Millet (Pennisetum americanum (L.) Leeke). J Exp Bot 39:124–139. doi: 10.1093/jxb/39.1.124

Dauwe R, Morreel K, Goeminne G, et al. (2007) Molecular phenotyping of lignin-modified tobacco reveals associated changes in cell-wall metabolism, primary metabolism, stress metabolism and photorespiration. Plant J 52:263–85. doi: 10.1111/j.1365-313X.2007.03233.x

Demura T, Fukuda H (2007) Transcriptional regulation in wood formation. Trends Plant Sci 12:64–70. doi: 10.1016/j.tplants.2006.12.006

Demura T, Ye Z-H (2010) Regulation of plant biomass production. Curr Opin Plant Biol 13:299–304. doi: 10.1016/j.pbi.2010.03.002

Ding S-Y, Liu Y-S, Zeng Y, et al. (2012) How does plant cell wall nanoscale architecture correlate with enzymatic digestibility? Science 338:1055–60. doi: 10.1126/science.1227491

Dixon R a, Achnine L, Kota P, et al. (2002) The phenylpropanoid pathway and plant defence-a genomics perspective. Mol Plant Pathol 3:371–90. doi: 10.1046/j.1364-3703.2002.00131.x

Domon B, Costello CE (1988) A systematic nomenclature for carbohydrate fragmentations in FAB-MS/MS spectra of glycoconjugates. Glycoconj J 5:397–409. doi: 10.1007/BF01049915

Draper J, Mur LAJ, Jenkins G, et al. (2001) Brachypodium distachyon . A New Model System for Functional Genomics in Grasses 1. 127:1539–1555. doi: 10.1104/pp.010196.1

Dudareva N, Pichersky E, Gershenzon J (2004) Biochemistry of plant volatiles. Plant Physiol 135:1893–902. doi: 10.1104/pp.104.049981

Eckardt NA (2000) Giving Rice the Time of Day: Molecular Identification of a Major Photoperiod Sensitivity Quantitative Trait Locus. PLANT CELL ONLINE

12:2299–2301. doi: 10.1105/tpc.12.12.2299

Eggeman T, Elander RT (2005) Process and economic analysis of pretreatment technologies. Bioresour Technol 96:2019–25. doi: 10.1016/j.biortech.2005.01.017

Engels F (1998) Alfalfa Stem Tissues: Cell-wall Development and Lignification. Ann Bot 82:561–568. doi: 10.1006/anbo.1998.0705

Ericsson K, Nilsson LJ (2006) Assessment of the potential biomass supply in Europe using a resource-focused approach. Biomass and Bioenergy 30:1–15. doi: 10.1016/j.biombioe.2005.09.001

European Academies Science Advisary Council (2012) The current status of biofuels in the European Union, their environmental impacts and future prospects; EASAC policy report 19. 47. doi: ISBN: 978-3-8047-3118-9

Fabre N, Rustan I, de Hoffmann E, Quetin-Leclercq J (2001) Determination of flavone, flavonol, and flavanone aglycones by negative ion liquid chromatography electrospray ion trap mass spectrometry. J Am Soc Mass Spectrom 12:707–15. doi: 10.1016/S1044-0305(01)00226-4

FAO Statistics Division (2013a) FAOSTAT 2011. http://faostat3.fao.org/faostat-gateway/go/to/download/Q/QC/E. Accessed 28 Dec 2013

FAO Statistics Division (2013b) FAOSTAT 2012. http://faostat3.fao.org/faostat-gateway/go/to/download/Q/QC/E. Accessed 7 Feb 2014

FAO Statistics Division (2013c) FAOSTAT 1961-2012. http://faostat3.fao.org/faostat-gateway/go/to/home/E. Accessed 12 Feb 1BC

Feltus FA, Vandenbrink JP (2012) Bioenergy grass feedstock: current options and prospects for trait improvement using emerging genetic, genomic, and systems biology toolkits. Biotechnol Biofuels 5:80. doi: 10.1186/1754-6834-5-80

Fischer G, Prieler S, van Velthuizen H, et al. (2010a) Biofuel production potentials in Europe: Sustainable use of cultivated land and pastures, Part II: Land use scenarios. Biomass and Bioenergy 34:173–187. doi: 10.1016/j.biombioe.2009.07.009

Fischer G, Prieler S, van Velthuizen H, et al. (2010b) Biofuel production potentials in Europe: Sustainable use of cultivated land and pastures. Part I: Land productivity potentials. Biomass and Bioenergy 34:159–172. doi: 10.1016/j.biombioe.2009.07.008

Fisher JB, French JC (1976) The Occurrence of Intercalary and Uninterrupted Meristems in the Internodes of Tropical Monocotyledons. Am. J. Bot. J. Bot. 63:

Fornalé S, Shi X, Chai C, Encina A (2010) ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux. Plant … 633–644. doi: 10.1111/j.1365-313X.2010.04363.x

Fornalé S, Sonbol F-M, Maes T, et al. (2006) Down-regulation of the maize and Arabidopsis thaliana caffeic acid O-methyl-transferase genes by two new

maize R2R3-MYB transcription factors. Plant Mol Biol 62:809–23. doi: 10.1007/s11103-006-9058-2

Fournier C, Andrieu B (2000) Dynamics of the Elongation of Internodes in Maize (Zea mays L.). Effects of Shade Treatment on Elongation Patterns. Ann Bot 86:1127–1134. doi: 10.1006/anbo.2000.1280

Fournier C, Durand JL, Ljutovac S, et al. (2005) A functional-structural model of elongation of the grass leaf and its relationships with the phyllochron. New Phytol 166:881–94. doi: 10.1111/j.1469-8137.2005.01371.x

Fukuda H (1997a) Programmed cell death during vascular system formation. Cell Death Differ 4:684–8. doi: 10.1038/sj.cdd.4400310

Fukuda H (1997b) Tracheary element differentiation. Plant Cell 9:3–7.

Fukuda H (2000) Programmed cell death of tracheary elements as a paradigm in plants. Plant Mol Biol 44:245–53.

Garvin DF, Gu Y-Q, Hasterok R, et al. (2008) Development of Genetic and Genomic Research Resources for , a New Model System for Grass Crop Research. Crop Sci 48:S–69. doi: 10.2135/cropsci2007.06.0332tpg

Gnansounou E (2010) Production and use of lignocellulosic bioethanol in Europe: Current situation and perspectives. Bioresour Technol 101:4842–50. doi: 10.1016/j.biortech.2010.02.002

Grabber JH (2005) How Do Lignin Composition, Structure, and Cross-Linking Affect Degradability? A Review of Cell Wall Model Studies. Crop Sci 45:820. doi: 10.2135/cropsci2004.0191

Grabber JH, Hatfield RD, Lu F, Ralph J (2008) Coniferyl ferulate incorporation into lignin enhances the alkaline delignification and enzymatic degradation of cell walls. Biomacromolecules 9:2510–2516. doi: 10.1021/bm800528f

Grabber JH, Ralph J, Hatfield RD (2002) Model studies of ferulate-coniferyl alcohol cross-product formation in primary maize walls: implications for lignification in grasses. J Agric Food Chem 50:6008–6016. doi: jf0205312 [pii]

Grabber JH, Ralph J, Lapierre C, et al. (2004) Genetic and molecular basis of grass cell-wall degradability. I. Lignin–cell wall matrix interactions. C R Biol 327:455–465. doi: 10.1016/j.crvi.2004.03.004

Grass Phylogeny Working Grp (2001) Phylogeny and subfamilial classification of the grasses (Poaceae). Ann MISSOURI Bot Gard 88:373–457.

Gressel J (2008) Transgenics are imperative for biofuel crops. Plant Sci 174:246–263. doi: 10.1016/j.plantsci.2007.11.009

Hamelinck CN, Hooijdonk G van, Faaij AP (2005) Ethanol from lignocellulosic biomass: techno-economic performance in short-, middle- and long-term. Biomass and Bioenergy 28:384–410. doi: 10.1016/j.biombioe.2004.09.002

Handakumbura PP, Hazen SP (2012) Transcriptional Regulation of Grass Secondary Cell Wall Biosynthesis: Playing Catch-Up with Arabidopsis

thaliana. Front Plant Sci 3:74. doi: 10.3389/fpls.2012.00074

Hatfield R, Ralph J, Grabber JH (2008) A potential role for sinapyl p-coumarate as a radical transfer mechanism in grass lignin formation. Planta 228:919–28. doi: 10.1007/s00425-008-0791-4

Hématy K, Cherk C, Somerville S (2009) Host-pathogen warfare at the plant cell wall. Curr Opin Plant Biol 12:406–13. doi: 10.1016/j.pbi.2009.06.007

Hématy K, Sado P-E, Van Tuinen A, et al. (2007) A receptor-like kinase mediates the response of Arabidopsis cells to the inhibition of cellulose synthesis. Curr Biol 17:922–31. doi: 10.1016/j.cub.2007.05.018

Hendriks a TWM, Zeeman G (2009) Pretreatments to enhance the digestibility of lignocellulosic biomass. Bioresour Technol 100:10–8. doi: 10.1016/j.biortech.2008.05.027

Higgins JA, Bailey PC, Laurie DA (2010) Comparative genomics of flowering time pathways using Brachypodium distachyon as a model for the temperate grasses. PLoS One 5:e10065. doi: 10.1371/journal.pone.0010065

Himmel ME, Ding S-Y, Johnson DK, et al. (2007) Biomass recalcitrance: engineering plants and enzymes for biofuels production. Science 315:804–7. doi: 10.1126/science.1137016

Hisano H, Nandakumar R, Wang Z-Y (2009) Genetic modification of lignin biosynthesis for improved biofuel production. Vitr Cell Dev Biol - Plant 45:306–313. doi: 10.1007/s11627-009-9219-5

Humphreys JM, Chapple C (2002) Rewriting the lignin roadmap. Curr Opin Plant Biol 5:224–9.

Humphreys JM, Hemm MR, Chapple C (1999) New routes for lignin biosynthesis defined by biochemical characterization of recombinant ferulate 5-hydroxylase, a multifunctional cytochrome P450-dependent monooxygenase. Proc Natl Acad Sci U S A 96:10045–50.

Iiyama K, Lam T, Stone BA (1994) Covalent Cross-Links in the Cell Wall. Plant Physiol 104:315–320.

Ishii T (1997) Structure and functions of feruloylated polysaccharides. Plant Sci 127:111–127. doi: 10.1016/S0168-9452(97)00130-1

Jakob K, Zhou F, Paterson AH (2009) Genetic improvement of C4 grasses as cellulosic biofuel feedstocks. Vitr Cell Dev Biol - Plant 45:291–305. doi: 10.1007/s11627-009-9214-x

Janssen R, Turhollow AF, Rutz D, Mergner R (2013) Production facilities for second-generation biofuels in the USA and the EU - current status and future perspectives. Biofuels, Bioprod Biorefining 7:647–665. doi: 10.1002/bbb.1451

Joshi CP, Mansfield SD (2007) The cellulose paradox--simple molecule, complex biosynthesis. Curr Opin Plant Biol 10:220–6. doi: 10.1016/j.pbi.2007.04.013

Jouanin L, Lapierre C (2012) Lignins: Biosynthesis, Biodegradation and Bioengineering, 1st Editio. 440.

Jung HG, Casler MD (2006) Maize Stem Tissues. Crop Sci 46:1793. doi: 10.2135/cropsci2005.02-0085

Jung HG, Mertens DR, Payne a J (1997) Correlation of acid detergent lignin and Klason lignin with digestibility of forage dry matter and neutral detergent fiber. J Dairy Sci 80:1622–8. doi: 10.3168/jds.S0022-0302(97)76093-4

Jung H-JG (2003) Maize stem tissues: ferulate deposition in developing internode cell walls. Phytochemistry 63:543–549. doi: 10.1016/S0031-9422(03)00221-8

Jung H-JG, Samac D a, Sarath G (2012) Modifying crops to increase cell wall digestibility. Plant Sci 185-186:65–77. doi: 10.1016/j.plantsci.2011.10.014

Kallenbach RL (2012) Growth of Pasture Plants, Dairy Grazing Manual M182,. Missouri

Karp A, Shield I (2008) Bioenergy from plants and the sustainable yield challenge. New Phytol 179:15–32. doi: 10.1111/j.1469-8137.2008.02432.x

Kavanová M, Lattanzi F (2006) Phosphorus deficiency decreases cell division and elongation in grass leaves. Plant … 141:766–775. doi: 10.1104/pp.106.079699.766

Kim TH, Lee YY (2005) Pretreatment and fractionation of corn stover by ammonia recycle percolation process. Bioresour Technol 96:2007–13. doi: 10.1016/j.biortech.2005.01.015

Klepper B, Rickman RW, Peterson CM (1982) Quantitative Characterization of Vegetative Development in Small Cereal Grains1. Agron J 74:789. doi: 10.2134/agronj1982.00021962007400050005x

Kumar A, Cameron JB, Flynn PC (2005) Pipeline transport and simultaneous saccharification of corn stover. Bioresour Technol 96:819–29. doi: 10.1016/j.biortech.2004.07.007

Kumar R, Wyman CE (2009) Effect of additives on the digestibility of corn stover solids following pretreatment by leading technologies. Biotechnol Bioeng 102:1544–57. doi: 10.1002/bit.22203

La Camera S, Gouzerh G, Dhondt S, et al. (2004) Metabolic reprogramming in plant innate immunity: the contributions of phenylpropanoid and oxylipin pathways. Immunol Rev 198:267–84.

Langdale J (2005) The then and now of maize leaf development. Maydica 50:459–467.

Lapierre C (1993) Application of new methods for the investigation of lignin structure. Forage Cell Wall Struct. Dig. American Society of Agronomy, Crop Science Society of America, Soil Science Society of America, pp 133–166

Lau MW, Gunawan C, Dale BE (2009) The impacts of pretreatment on the fermentability of pretreated lignocellulosic biomass: a comparative evaluation between ammonia fiber expansion and dilute acid pretreatment. Biotechnol Biofuels 2:30. doi: 10.1186/1754-6834-2-30

Li C, Knierim B, Manisseri C, et al. (2010a) Comparison of dilute acid and ionic liquid pretreatment of switchgrass: Biomass recalcitrance, delignification and enzymatic saccharification.

Bioresour Technol 101:4900–6. doi: 10.1016/j.biortech.2009.10.066

Li L, Popko JL, Umezawa T, Chiang VL (2000) 5-hydroxyconiferyl aldehyde modulates enzymatic methylation for syringyl monolignol formation, a new view of monolignol biosynthesis in angiosperms. J Biol Chem 275:6537–45.

Li L, Shi Z-Y, Li L, et al. (2010b) Overexpression of ACL1 (abaxially curled leaf 1) increased Bulliform cells and induced Abaxial curling of leaf blades in rice. Mol Plant 3:807–17. doi: 10.1093/mp/ssq022

Li M, Foster C, Kelkar S, et al. (2012) Structural characterization of alkaline hydrogen peroxide pretreated grasses exhibiting diverse lignin phenotypes. Biotechnol Biofuels 5:38. doi: 10.1186/1754-6834-5-38

Li X, Bonawitz ND, Weng J-K, Chapple C (2010c) The growth reduction associated with repressed lignin biosynthesis in Arabidopsis thaliana is independent of flavonoids. Plant Cell 22:1620–32. doi: 10.1105/tpc.110.074161

Li X, Weng J-K, Chapple C (2008) Improvement of biomass through lignin modification. Plant J 54:569–81. doi: 10.1111/j.1365-313X.2008.03457.x

Limayem A, Ricke SC (2012) Lignocellulosic biomass for bioethanol production: Current perspectives, potential issues and future prospects. Prog Energy Combust Sci 38:449–467. doi: 10.1016/j.pecs.2012.03.002

Lloyd TA, Wyman CE (2005) Combined sugar yields for dilute sulfuric acid pretreatment of corn stover followed by enzymatic hydrolysis of the remaining solids. Bioresour Technol 96:1967–77. doi: 10.1016/j.biortech.2005.01.011

Ma YL, Li QM, Van den Heuvel H, Claeys M (1997) Characterization of flavone and flavonol aglycones by collision-induced dissociation tandem mass spectrometry. Rapid Commun Mass Spectrom 11:1357–1364. doi: 10.1002/(SICI)1097-0231(199708)11:12<1357::AID-RCM983>3.0.CO;2-9

MacAdam JW, Grabber JH (2002) Relationship of growth cessation with the formation of diferulate cross-links and p-coumaroylated lignins in tall fescue leaf blades. Planta 215:785–93. doi: 10.1007/s00425-002-0812-7

MacLean HL, Spatari S (2009) The contribution of enzymes and process chemicals to the life cycle of ethanol. Environ Res Lett 4:014001. doi: 10.1088/1748-9326/4/1/014001

Matsuoka Y, Vigouroux Y, Goodman MM, et al. (2002) A single domestication for maize shown by multilocus microsatellite genotyping. Proc Natl Acad Sci U S A 99:6080–4. doi: 10.1073/pnas.052125199

McSteen P (2010) Auxin and monocot development. Cold Spring Harb Perspect Biol 2:a001479. doi: 10.1101/cshperspect.a001479

Molinari HBC, Pellny TK, Freeman J, et al. (2013) Grass cell wall feruloylation: distribution of bound ferulate and candidate gene expression in

Brachypodium distachyon. Front Plant Sci 4:50. doi: 10.3389/fpls.2013.00050

Monfreda C, Ramankutty N, Foley JA (2008) Farming the planet: 2. Geographic distribution of crop areas, yields, physiological types, and net primary production in the year 2000. Global Biogeochem Cycles 22:n/a–n/a. doi: 10.1029/2007GB002947

Moore KJ, Moser LE (1995) Quantifying Developmental Morphology of Perennial Grasses. Crop Sci 35:37. doi: 10.2135/cropsci1995.0011183X003500010007x

Morreel K, Dima O, Kim H, et al. (2010a) Mass spectrometry-based sequencing of lignin oligomers. Plant Physiol 153:1464–78. doi: 10.1104/pp.110.156489

Morreel K, Goeminne G, Storme V, et al. (2006) Genetical metabolomics of flavonoid biosynthesis in Populus: a case study. Plant J 47:224–37. doi: 10.1111/j.1365-313X.2006.02786.x

Morreel K, Kim H, Lu F, et al. (2010b) Mass spectrometry-based fragmentation as an identification tool in lignomics. Anal Chem 82:8095–8105. doi: 10.1021/ac100968g

Morreel K, Ralph J, Kim H, et al. (2004a) Profiling of oligolignols reveals monolignol coupling conditions in lignifying poplar xylem. Plant Physiol 136:3537–49. doi: 10.1104/pp.104.049304

Morreel K, Ralph J, Lu F (2004b) Phenolic profiling of caffeic acid O-methyltransferase-deficient poplar reveals novel benzodioxane oligolignols. Plant … 136:4023–4036. doi: 10.1104/pp.104.049312.The

Morrison T, Kessler J (1994) Activity of two lignin biosynthesis enzymes during development of a maize internode. J … 133–139.

Mosier N, Wyman C, Dale B, et al. (2005) Features of promising technologies for pretreatment of lignocellulosic biomass. Bioresour Technol 96:673–86. doi: 10.1016/j.biortech.2004.06.025

Nafziger E (2009) Corn. In: Extension U of I (ed) Illinois Agron. Handb., 24TH 09 ed. University of Illinois, Office of Argicultural, pp 13–26

Neild R, Newman J (1990) Growing season characteristics and requirements in the Corn Belt. Natl. Corn Handb. 40

Nelissen H, Rymen B, Jikumaru Y (2012) A local maximum in gibberellin levels regulates maize leaf growth by spatial control of cell division. Curr Biol 1–5. doi: 10.1016/j.cub.2012.04.065

Nelson T, Dengler N (1997) Leaf Vascular Pattern Formation. Plant Cell 9:1121–1135. doi: 10.1105/tpc.9.7.1121

Nemoto K, Nagano I, Hogetsu T, Miyamoto N (2004) Dynamics of cortical microtubules in developing maize internodes. New Phytol 162:95–103. doi: 10.1046/j.1469-8137.2004.01006.x

O'Keeffe K (2009) Maize growth and development. Maize growth Dev 60.

Orkwiszewski JA, Poethig RS (2000) Phase identity of the maize leaf is determined after leaf initiation. Proc Natl Acad Sci U

S A 97:10631–6. doi: 10.1073/pnas.180301597

Osakabe K, Tsao CC, Li L, et al. (1999) Coniferyl aldehyde 5-hydroxylation and methylation direct syringyl lignin biosynthesis in angiosperms. Proc Natl Acad Sci U S A 96:8955–60.

Parvathi K, Chen F, Guo D, et al. (2001) Substrate preferences of O-methyltransferases in alfalfa suggest new pathways for 3-O-methylation of monolignols. Plant J 25:193–202.

Parvez MM, Wakabayashi K, Hoson T, Kamisaka S (1998) White light-induced sugar distribution controls growth and osmotic properties in the coleoptile and the first leaf in Zea mays seedlings. Physiol Plant 102:1–8. doi: 10.1034/j.1399-3054.1998.1020101.x

Pautler M, Tanaka W, Hirano H-Y, Jackson D (2013) Grass meristems I: shoot apical meristem maintenance, axillary meristem determinacy and the floral transition. Plant Cell Physiol 54:302–12. doi: 10.1093/pcp/pct025

Pedersen JF, Vogel KP, Funnell DL (2005) Impact of Reduced Lignin on Plant Fitness. Crop Sci 45:812. doi: 10.2135/cropsci2004.0155

Petrik DL, Karlen SD, Cass CL, et al. (2013) p-Coumaroyl-CoA:Monolignol Transferase (PMT) acts specifically in the lignin biosynthetic pathway in Brachypodium distachyon. Plant J. doi: 10.1111/tpj.12420

Poethig RS, Szymkowiak EJ (1995) Clonal analysis of leaf development in maize. Maydica (Italy)

Pollard M, Beisson F, Li Y, Ohlrogge JB (2008) Building lipid barriers: biosynthesis of cutin and suberin. Trends Plant Sci 13:236–46. doi: 10.1016/j.tplants.2008.03.003

Ragauskas AJ, Williams CK, Davison BH, et al. (2006) The path forward for biofuels and biomaterials. Science 311:484–9. doi: 10.1126/science.1114736

Ralph J (2010) Hydroxycinnamates in lignification. Phytochem Rev 9:65–83. doi: 10.1007/s11101-009-9141-9

Ralph J, Hatfield RD, Quideau S, et al. (1994) Pathway of p-Coumaric Acid Incorporation into Maize Lignin As Revealed by NMR. J Am Chem Soc 116:9448–9456. doi: 10.1021/ja00100a006

Ralph J, Kim H, Lu F, et al. (2008) Identification of the structure and origin of a thioacidolysis marker compound for ferulic acid incorporation into angiosperm lignins (and an indicator for cinnamoyl CoA reductase deficiency). Plant J 53:368–79. doi: 10.1111/j.1365-313X.2007.03345.x

Ralph J, Lundquist K, Brunow G, et al. (2004) Lignins: Natural polymers from oxidative coupling of 4-hydroxyphenyl-propanoids. Phytochem Rev 3:29–60. doi: 10.1023/B:PHYT.0000047809.65444.a4

Rancour DM, Marita JM, Hatfield RD (2012) Cell wall composition throughout development for the model grass Brachypodium distachyon. Front Plant Sci 3:266. doi: 10.3389/fpls.2012.00266

Raven PH, Evert RF, Eichhorn SE, Evert RF (1992) Biology of Plants, 5th Editio. 791.

Renvoize SA (2002) Grass anatomy. Flora Aust. Vol. 43

Rickman RW, Klepper BL (1995) The Phyllochron: Where Do we Go in the Future? Crop Sci 35:44. doi: 10.2135/cropsci1995.0011183X003500010008x

Ringli C (2010) Monitoring the outside: cell wall-sensing mechanisms. Plant Physiol 153:1445–52. doi: 10.1104/pp.110.154518

Ritchie SW, Hanway JJ, Benson G 0. (1992) How a corn plant develops. Special Report No. 48. Iowa State University of Science and Technology, Ames, Iowa

Rivas-Martinez S, Sánchez-Mata D, Costa. M (1999) North American boreal and western temperate forest vegetation. Itinera Geobot 12:3–331.

Robertson MJ (1994) Relationships between internode elongation, plant height and leaf appearance in maize. F Crop Res 38:135–145. doi: 10.1016/0378-4290(94)90085-X

Rohde A, Morreel K, Ralph J (2004) phenotyping of the pal1 and pal2 mutants of Arabidopsis thaliana reveals far-reaching consequences on phenylpropanoid, amino acid, and carbohydrate metabolism. Plant Cell … 16:2749–2771. doi: 10.1105/tpc.104.023705.phospho-enol-pyruvate

Rosler J, Krekel F, Amrhein N, et al. (1997) Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. Plant Physiol 113:175–179.

Rymen B, Coppens F, Dhondt S, et al. (2010) Kinematic analysis of cell division and expansion. Methods Mol Biol Clift Nj 655:203–227.

Saha BC, Iten LB, Cotta MA, Wu YV (2005) Dilute acid pretreatment, enzymatic saccharification, and fermentation of rice hulls to ethanol. Biotechnol Prog 21:816–22. doi: 10.1021/bp049564n

Santiago R, Barros-Rios J, Malvar R a (2013) Impact of cell wall composition on maize resistance to pests and diseases. Int J Mol Sci 14:6960–80. doi: 10.3390/ijms14046960

Schippmann U (1991) Revision Der Europäischen Arten Der Gattung Brachypodium Palisot de Beauvois (Poaceae).

Schwartz CJ, Doyle MR, Manzaneda AJ, et al. (2010) Natural Variation of Flowering Time and Vernalization Responsiveness in Brachypodium distachyon. BioEnergy Res 3:38–46. doi: 10.1007/s12155-009-9069-3

Seifert GJ, Blaukopf C (2010) Irritable walls: the plant extracellular matrix and signaling. Plant Physiol 153:467–78. doi: 10.1104/pp.110.153940

Sharman BC (1945) Leaf and Bud Initiation in the Gramineae. Bot Gaz 106:269. doi: 10.1086/335298

Shen H, Fu C, Xiao X, et al. (2009) Developmental Control of Lignification in Stems of Lowland Switchgrass Variety Alamo and the Effects on Saccharification Efficiency. BioEnergy Res 2:233–245. doi: 10.1007/s12155-009-9058-6

Shen H, He X, Poovaiah C (2012) Functional characterization of the switchgrass (Panicum virgatum) R2R3-MYB transcription factor PvMYB4 for improvement of lignocellulosic feedstocks. New …. doi: 10.1111/j.1469-8137.2011.03922.x

Siemer EG, Leng ER, Bonnett OT (1969) Timing and Correlation of Major Developmental Events in Maize, Zea mays L.1. Agron J 61:14. doi: 10.2134/agronj1969.00021962006100010005x

Silk WK, Erickson RO (1979) Kinematics of plant growth. J Theor Biol 76:481–501. doi: 10.1016/0022-5193(79)90014-6

Sims REH, Mabee W, Saddler JN, Taylor M (2010) An overview of second generation biofuel technologies. Bioresour Technol 101:1570–80. doi: 10.1016/j.biortech.2009.11.046

Skinner RH, Nelson CJ (1994) Epidermal cell division and the coordination of leaf and tiller development. Ann Bot 74:9–16. doi: 10.1093/aob/74.1.9

Sticklen MB (2008) Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol. Nat Rev Genet 9:433–43. doi: 10.1038/nrg2336

Sun Y, Cheng J (2002) Hydrolysis of lignocellulosic materials for ethanol production: a review. Bioresour Technol 83:1–11.

Sylvester A (2001) Leaf shape and anatomy as indicators of phase change in the grasses: comparison of maize, rice, and bluegrass. Am J … 88:2157–2167.

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

Tanaka W, Pautler M, Jackson D, Hirano H-Y (2013) Grass meristems II: inflorescence architecture, flower development and meristem fate. Plant Cell Physiol 54:313–24. doi: 10.1093/pcp/pct016

Thompson BE, Hake S (2009) Translational biology: from Arabidopsis flowers to grass inflorescence architecture. Plant Physiol 149:38–45. doi: 10.1104/pp.108.129619

Townsley BT, Sinha NR, Kang J (2013) KNOX1 genes regulate lignin deposition and composition in monocots and dicots. Front Plant Sci 4:121. doi: 10.3389/fpls.2013.00121

Valentine J, Clifton-Brown J, Hastings A, et al. (2012) Food vs. fuel: the use of land for lignocellulosic “next generation” energy crops that minimize competition with primary food production. GCB Bioenergy 4:1–19. doi: 10.1111/j.1757-1707.2011.01111.x

Van Acker R, Vanholme R, Storme V, et al. (2013) Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in Arabidopsis thaliana. Biotechnol Biofuels 6:46. doi: 10.1186/1754-6834-6-46

Van der Weijde T, Alvim Kamei CL, Torres AF, et al. (2013) The potential of C4 grasses for cellulosic biofuel production. Front Plant Sci 4:107. doi: 10.3389/fpls.2013.00107

Van Minnebruggen A, Cnops G, Saracutu O, et al. (2013) Processes underlying branching differences in fodder crops. Euphytica 195:301–313. doi: 10.1007/s10681-013-0997-9

Vanholme B, Desmet T, Ronsse F, et al. (2013) Towards a carbon-negative sustainable bio-based economy. Front Plant Sci 4:174. doi: 10.3389/fpls.2013.00174

Vanholme R, Demedts B, Morreel K (2010a) Lignin biosynthesis and structure. Plant … 153:895–905. doi: 10.1104/pp.110.155119

Vanholme R, Morreel K, Darrah C, et al. (2012a) Metabolic engineering of novel lignin in biomass crops. New Phytol 196:978–1000. doi: 10.1111/j.1469-8137.2012.04337.x

Vanholme R, Ralph J, Akiyama T, et al. (2010b) Engineering traditional monolignols out of lignin by concomitant up-regulation of F5H1 and down-regulation of COMT in Arabidopsis. Plant J 64:885–97. doi: 10.1111/j.1365-313X.2010.04353.x

Vanholme R, Storme V, Vanholme B, et al. (2012b) A systems biology view of responses to lignin biosynthesis perturbations in Arabidopsis. Plant Cell 24:3506–29. doi: 10.1105/tpc.112.102574

Verdenal A, Combes D, Escobar-Gutiérrez A (2008) A study of ryegrass architecture as a self-regulated system, using functional–structural plant modelling. Funct Plant Biol 911–924.

Verdenal A, Combes D, Escobar-Gutierrez A (2009) Pseudostem artificial extension with colored tubes led to the modulation of leaf elongation in Tall Fescue (Festuca arundinacea S.). Comp Biochem Physiol Part A Mol Integr Physiol 153:S207. doi: 10.1016/j.cbpa.2009.04.476

Verdenal A, Combes D, Escobar-Gutiérrez AJ (2012) Programmable and Self-Organised Processes in Plant Morphogenesis: The Architectural Development of Ryegrass. In: Doursat R, Sayama H, Michel O (eds) Morphog. Eng. Springer Berlin Heidelberg, Berlin, Heidelberg, pp 501–517

Vogel J (2008) Unique aspects of the grass cell wall. Curr Opin Plant Biol 11:301–7. doi: 10.1016/j.pbi.2008.03.002

Vogel J, Garvin D, Mockler T, The International Brachypodium Initiative (2010) Genome sequencing and analysis of the model grass Brachypodium distachyon. Nature 463:763–8. doi: 10.1038/nature08747

Vogt T (2010) Phenylpropanoid biosynthesis. Mol Plant 3:2–20. doi: 10.1093/mp/ssp106

Watson L, Dallwitz MJ (1992) The families of flowering plants: descriptions, illustrations, identification, and information retrieval. In: Version 19th Oct. 2013. http://delta-intkey.com/angio/www/graminea.htm. Accessed 12 Feb 2014

Williams N (2008) Biofuel debate deepens. Curr Biol 891–892.

Withers S, Lu F, Kim H, et al. (2012) Identification of grass-specific enzyme that acylates monolignols with p-coumarate. J Biol Chem 287:8347–55. doi: 10.1074/jbc.M111.284497

World Commission on Environment and Development (1987) Our Common Future. Oxford University Press, Oxford

Wyman CE, Dale BE, Elander RT, et al. (2005) Coordinated development of leading biomass pretreatment technologies. Bioresour Technol 96:1959–66. doi: 10.1016/j.biortech.2005.01.010

Xiang J-J, Zhang G-H, Qian Q, Xue H-W (2012) Semi-rolled leaf1 encodes a putative glycosylphosphatidylinositol-anchored protein and modulates rice leaf rolling by regulating the formation of bulliform cells. Plant Physiol 159:1488–500. doi: 10.1104/pp.112.199968

Yokoyama R, Nishitani K (2004) Genomic basis for cell-wall diversity in plants. A comparative approach to gene families in rice and Arabidopsis. Plant Cell Physiol 45:1111–21. doi: 10.1093/pcp/pch151

Yu X-H, Perdue TD, Heimer YM, Jones AM (2002) Mitochondrial involvement in tracheary element programmed cell death. Cell Death Differ 9:189–198. doi: 10.1038/sj/cdd/4400940

Yuan JS, Tiller KH, Al-Ahmad H, et al. (2008) Plants to power: bioenergy to fuel the future. Trends Plant Sci 13:421–9. doi: 10.1016/j.tplants.2008.06.001

Zeng Y, Zhao S, Yang S, Ding S-Y (2014) Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. Curr Opin Biotechnol 27:38–45. doi: 10.1016/j.copbio.2013.09.008

Zhong R, Ye Z-H (2007) Regulation of cell wall biosynthesis. Curr Opin Plant Biol 10:564–72. doi: 10.1016/j.pbi.2007.09.001

Zhong R, Ye Z-H (2012) MYB46 and MYB83 bind to the SMRE sites and directly activate a suite of transcription factors and secondary wall biosynthetic genes. Plant Cell Physiol 53:368–80. doi: 10.1093/pcp/pcr185

Chapter 2: General objectives

Chapter 2: General objectives

Switching from a petroleum-based towards a bio-based economy is a necessary step in the direction of a sustainable way of living. The challenges that lie ahead are providing human nutrition, elevating living standards and at the same time stabilizing climate changes by reducing greenhouse gas emissions. Part of the solution is the use of lignocellulosic feedstock to produce so-called second-generation bioethanol that can supply the ever-increasing energy demand in a sustainable way. In contrast to first-generation bioethanol, non-food crops or plant parts that are not used for human consumption are converted into fuel. However, in order to make the production economically attractive it is necessary to improve biomass yields and the conversion efficiency of the lignocellulosic plant material into bioethanol. **The general objective of this work was thus to explore strategies to improve lignocellulosic biomass crops. Two options were explored: improving the quality and improving the quantity of the lignocellulosic feedstock**. **In addition, the interconnection between improved biomass production and the quality of the feedstock was investigated.** The quality of the biomass for bioethanol production is defined here as the amount of fermentable sugars that can be released from the biomass. Accordingly, the biomass quantity or yield is determined by the innate growth potential and the dry matter yield, in particular the lignocellulosic fraction of the plant. In both fields, fundamental insights have been gained using the dicot model plant *Arabidopsis thaliana* (hereafter Arabidopsis). However, high yielding energy crops such as Miscanthus, switchgrass and maize are all monocot species. We thus set out to investigate the transferability of the knowledge available in Arabidopsis to the monocot crop maize. Maize has a huge economic value and has also constituted an important model for biological research for over a century. This translational approach is perhaps easiest achieved by transferring knowledge from a dicot model to a monocot model species first. Therefore, in this PhD also the wild grass species *Brachypodium distachyon* (hereafter Brachypodium), presented as a model species for the grasses, was studied. Mainly due to its small size and compact genome, Brachypodium is considered a model for closely related grass species of high economic value such as wheat and barley. **In this dissertation, we explore strategies for bioenergy crop improvement in maize and investigate the potential of Brachypodium, using available tools and resources, as model plant for bioenergy crop improvement**.

Specifically, the following aspects were investigated:

- To study the potential of a plant to rapidly produce a high amount of biomass, the availability of methods to monitor plant and organ growth closely is a basic requirement. For this type of studies in monocots, the leaf is often used as model system by taking a series of length measurements over the duration of leaf development. This data is commonly used to quantify plant responses to environmental conditions and/or to identify inter-genotype differences. However, the current methods for extracting biologically relevant information from these data either simplify the growth process or require the use of statistical software packages which makes the extraction of relevant data laborious. To facilitate this process, we developed a user-friendly tool that can be used for fast, reliable

and robust analysis of large leaf elongation datasets. We also explored the possibility to use this methodology for fitting cell length measurements along the maize leaf axis and determining the size of the growth zone. This tool is presented in **Chapter 3**, and is used in other chapters to analyze leaf growth data of Brachypodium and maize plants.

- For improvement of the plant's growth potential and bioethanol feedstock properties, we explored the applicability of current knowledge in Arabidopsis on so-called intrinsic yield genes (IYGs) and genes involved in lignin biosynthesis, to Brachypodium. When mutated or overexpressed, IYGs are able to enlarge organ and/or plant size. We tested whether overexpression of one of these IYGs, *GA20ox1*, has similar effects in Brachypodium, and by extension, in grasses. In addition, the effects of mutations in two genes involved in lignin biosynthesis: *cinnamyl alcohol dehydrogenase* (*CAD*) and *4-coumarate:CoA ligase* (*4CL*) were investigated in Brachypodium. These results are summarized in **Chapter 4**. In addition, this chapter contains an overview of available tools and resources for Brachypodium research, and a discussion of the main achievements in Brachypodium research since its introduction as a model species in 2001.

- In contrast to Brachypodium, maize has been a successful genetic model for more than a century and its use in plant biology research has even been boosted by the release of the B73 genome sequence in 2009. In this study, two translational approaches were tested for the improvement of the saccharification efficiency in maize. The first was disruption of *cinnamate-4-hydroxylase* (*C4H*), coding for an enzyme that functions in an early step in the lignin biosynthetic pathway. In Arabidopsis, this results in improved saccharification efficiency. The second was downregulation and disruption of *CAD* of which the enzyme catalyzes the last step in the production of monolignols, the building blocks of lignin. The latter strategy has proven to be effective for improving saccharification efficiency in maize and Arabidopsis. However, as the effect can differ depending on the genetic background in which *CAD* is affected, we tested a transgenic approach for reducing CAD activity in the inbred line B104. In addition, a systems biology approach to interrogate the effects of lignin perturbation, such as disruption of *C4H* and *CAD*, on gene expression levels can provide deeper inside into lignin biosynthesis and the complex metabolic network it is embedded in. Thus, in addition to investigating the effects of the individual perturbations, we also investigated changes in systems-wide gene expression profiles. These results of *C4H* disruption are presented in **Chapter 5** and those of downregulation and disruption of *CAD* in **Chapter 6**.

- **A final objective of this PhD dissertation was to interrogate the connection between biomass quantity and biomass quality of lignocellulosic feedstock.** Numerous literature reports have linked lignin perturbation with improvement of the saccharification efficiency in several species, but in some cases a reduction of the biomass yield was found in plants in which the lignin biosynthesis pathway had been perturbed. Here we investigated whether increases in biomass yields had an effect on the biomass quality. In particular, we investigated the cell wall properties of one GA overproducing maize line that displays

higher growth rates resulting in taller plants. The results of the case study are discussed in **Chapter 7**.

Finally, a general conclusion on bioenergy feedstock improvement using maize and Brachypodium as model systems as well as recommendations on further research is depicted in **Chapter 8**.

Chapter 3: LEAF-E: a tool to analyze grass leaf growth using function fitting

Chapter 3: LEAF-E: a tool to analyze grass leaf growth using function fitting[1]

Wannes Voorend[1,2,3], Peter Lootens[3], Hilde Nelissen[1,2,] Isabel Roldán-Ruiz[3], Dirk Inzé[1,2] and Hilde Muylle[3]*

[1] Department of Plant Systems Biology, VIB, Technologiepark 927, 9052 Gent, Belgium
[2] Department of Plant Biotechnology and Bioinformatics, Ghent University, Technologiepark 927, 9052 Gent, Belgium
[3] Plant Sciences Unit – Growth and Development, Institute for Agricultural and Fisheries Research (ILVO), Caritasstraat 21, 9090 Melle, Belgium

* Correspondence: hilde.muylle@ilvo.vlaanderen.be

Other email addresses:

wannes.voorend@psb.vib-ugent.be, peter.lootens@ilvo.vlaanderen.be; hilde.nelissen@psb.vib-ugent.be; dirk.inze@psb.vib-ugent.be; isabel.roldan-ruiz@ilvo.vlaanderen.be

Authors' contributions:

Wannes Voorend wrote this manuscript and was involved in conception and design of the method and acquisition, analysis and interpretation of the data.

Peter Lootens is responsible for the implementation and optimization of the method as a macro in Microsoft Excel and critical revision of the manuscript.

Hilde Nelissen was involved in providing data, interpretation of data and critical revision of the manuscript.

Isabel Roldán-Ruiz has made substantial contributions to conception and design of the method, interpretation of data and critical revision of the manuscript.

Dirk Inzé was involved in critical revision of the manuscript and has given final approval of the version to be published.

Hilde Muylle has made substantial contributions to conception and design of the method, was involved in interpretation of data and critical revision of the manuscript and has given final approval of the version to be published.

[1] This chapter has been formatted according to the instructions for authors of the journal BMC Plant Methods

1. Abstract

Leaf length measurements that are performed during leaf development form the basis for studying the growth dynamics of the grass leaf. Deriving biologically relevant parameters such as the leaf elongation rate (LER) from these data is essential for detecting inter-genotype differences and/or quantifying plant responses to changing environmental conditions.

Here, we describe a tool that we have given the name LEAF-E, that allows in-depth analysis of the growth behavior of grass leaves in different species. LEAF-E is based on nonlinear regression modeling of leaf length measurements. We demonstrate that the extraction of biologically relevant parameters from these measurements using LEAF-E is both straightforward and user-friendly since the tool is Microsoft Excel-based. The results are stored in tabular form and can easily be analyzed in search of differential responses.

We validate our method and demonstrate its broad application range using published and unpublished data sets of maize, *Miscanthus spp.* and *Brachypodium distachyon*, generated in separate experiments and for different purposes. Moreover, we propose that the method is suited for fitting cell length measurements along the leaf axis, thereby allowing robust determination of the size of the growth zone. The LEAF-E method is available upon request to the second author in the form of an Excel macro.

Keywords: Leaf elongation rate, Beta sigmoid function, Leaf length, Cell length, Growth zone

2. Background

Research in a wide variety of grass species such as maize, rice, wheat, barley, *Lolium*, *Miscanthus, Sorghum* and *Brachypodium* makes use of leaf lengths measured repeatedly at defined time points [1-8]. These data are organized in time series that describe the growth dynamics and are commonly used to quantify plant responses to environmental conditions and/or to identify inter-genotype differences. In grass species, one of the most frequently used parameters derived from repeated leaf length measurements is the leaf elongation rate (LER), providing an integrated estimation of the growth of an individual leaf over a given period of time. In addition, it has been demonstrated that the LER is a major determinant of individual leaf area and of whole plant leaf area [9-14].

Frequently, the LER is calculated as the slope of a linear fit based on two or more consecutive leaf length measurements, thereby assuming that the LER is constant over a longer period during leaf development [1, 3, 9, 10]. When leaf growth is regarded as an exponential or a log-linear relation, a constant relative elongation rate (RER) is assumed. These assumptions limit their utility, as both LER and RER may vary with environmental conditions and developmental stage [15]. As a consequence, in plant growth modeling there is a growing consensus that the traditional approaches that apply linear and exponential models are inadequate [15]. The polynomial model does cope with variations in LER and RER during leaf development. However, for displaying results and extracting data, the use of this model is discouraged since polynomial functions tend to make spurious upward or downward predictions, especially at the extremes of the data [15, 16]. The approach that allows for a more accurate determination of the leaf elongation pattern and biologically relevant parameters is nonlinear regression using a suitable mathematical function [16]. Moreover, the use of such a nonlinear model is the best way to accommodate temporal variation in growth rates [15].

Function fitting has been applied widely in the field of agricultural science to describe and predict plant lengths and weights over time, and under a variety of environmental situations [13, 17-23]. On the contrary, functions suitable for describing the growth pattern of an individual monocot leaf are far less frequently reported (some exceptions are [7, 24, 25]). To our knowledge, only the beta sigmoid function, first used to describe whole plant growth [26], has been successfully applied to model the growth pattern of a single monocot leaf [7, 25]. The beta sigmoid function can be used to fit determinate growth of plants or plant parts, characterized by three distinct phases: an exponential or accelerating growth phase, an approximately linear growth phase, followed by a steadily decelerating growth phase [27]. This growth behavior can be described as 'S-shaped' or 'sigmoid'. Yin and coworkers [26] compared the performance of the beta sigmoid function to that of some other widely used sigmoid functions, such as Gompertz, Weibull and Richards to analyze datasets from maize, pea and wheat and revealed that the beta sigmoid function is unique in dealing with determinate growth. This is due to its high flexibility for describing various asymmetrical sigmoidal patterns, the incorporation of biologically relevant parameters and the prediction of a zero growth rate at both the start and end of a precisely defined growth period [26](Yin et al. 2002).

Here we demonstrate that the beta sigmoid function can be used to model leaf growth of both C3 and C4 grass species. We show how the use of this nonlinear regression method can assist data

analysis and interpretation of experiments in which different genotypes are compared or the response of a single genotype to specific growth conditions is investigated. Additionally, we demonstrate that this method can be used not only to estimate growth parameters, such as the maximal elongation rate, but also to study the temporal evolution of these parameters in a developing leaf. We validate our method using published and unpublished datasets of *Zea mays*, *Brachypodium distachyon* and *Miscanthus spp.* The calculations for data fitting and parameter derivation have been implemented in a Microsoft Excel spreadsheet (LEAF-E), providing a very fast, user-friendly, semi-automated method for the grass leaf phenotyping community.

3. Results and discussion

3.1 Fitting of kinematic individual leaf length measurements using the beta sigmoid function

3.1.1 Evaluation of the goodness of fit

We first investigated to what extent the beta sigmoid function can be used to accurately fit the leaf length measurements in function of thermal time in the three species considered here. Thermal time is a summation of cumulative differences between daily mean temperature and a specified base temperature [41]. Equation 1 was used to fit length measurements of the 4th leaf over thermal time of nine non-transgenic B104 maize plants (dataset 1a, Figure 22). This resulted in R^2 values ranging from 0.9970 to 0.9989 with a mean value of 0.9981. Function fitting of leaf length measurements in *Miscanthus* and *Brachypodium* (datasets 2 and 3, respectively) rendered similar results: an overall mean R^2-value of 0.9931, ranging from 0.9669 to 0.9989 (n = 18) for the two *Miscanthus* species, and an overall mean R^2-value of 0.9932, ranging from 0.9871 to 0.9993 (n=36) for the four *Brachypodium* inbred lines. These results demonstrate that the beta sigmoid function is able to accommodate leaf growth measurements of three grass species with very distinct phenotypic characteristics. Maize and *Miscanthus spp.* both possess a C4 metabolism, however, maize is an annual crop characterized by one stem, whereas *Miscanthus spp.* are rhizomatous perennials that form numerous tillers. *Brachypodium* is a small, annual C3 plant presented as a model for several temperate grain crops such as wheat and barley [28]. Based upon these findings and the results obtained previously in *L. perenne* [7], we can conclude that the beta sigmoid function is probably of broad application for describing leaf growth in both C3 and C4 grass species.

A comparison of the accuracy of the fitting procedure described here and that implemented in standard statistical software packages was out of the scope of this work. However, the correlation coefficients obtained for datasets 1a, 2 and 3 were more than satisfactory (see Table 1, Table 2 and Table 3). One of the main advantages of LEAF-E is its user-friendliness, since it requires no prior knowledge of the mathematics of function fitting or the use of sophisticated calculation tools. Moreover, the output is provided in tabular form summarizing the correlation coefficient, the function parameters and all derived growth parameters for each sample analyzed in one single datasheet.

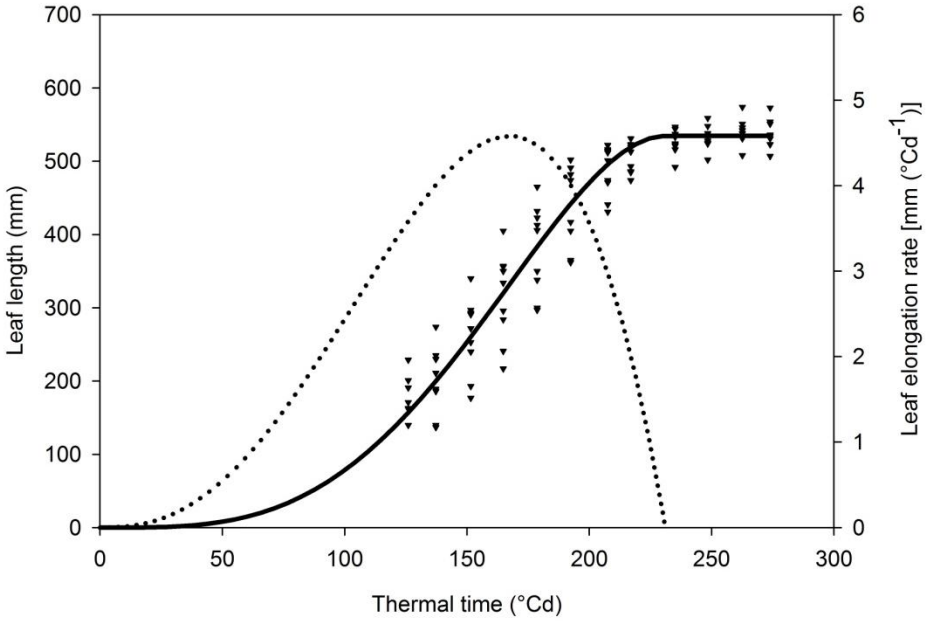


Figure 22. Leaf length and LER of maize B104 non-transgenic plants. Triangles represent length measurements of the 4th leaf of nine non-transgenic maize plants from dataset1a. The measurements of each individual leaf were fitted and biologically relevant growth parameters were extracted with LEAF-E. The S-shaped leaf length curve and bell-shaped LER curve are function plots using the mean values of the function parameters for the nine non-transgenic maize plants.

3.1.2 Deriving biologically relevant function parameters from the data

We used the nine non-transgenic B104 maize plants from dataset 1a to illustrate how leaf growth can be analyzed using LEAF-E (Figure 22, Table 1). The Excel macro that we designed, automatically generates the variables of the fitted function, all additional parameters and a graph showing the original data points, the fitted growth curve and the function variables for each biological replicate. Although the method can be used to fit the measurements of several replicates jointly, using the function to fit data of individual leaves allows the estimation of average and standard deviation values for each growth parameter. This allows for a statistical comparison of the parameters derived from different genotypes or from different treatments, a strategy that is both straightforward and statistically correct (Motulsky et al. 1987). Note that fitting of individual leaves using standard statistical software instead of our Excel-based method would be time-consuming and more prone to errors due to the copying and pasting of the data from text or html outputs.

To illustrate the procedure of extracting biological information from leaf length measurements using LEAF-E, we calculated averages for the function parameters for nine non-transgenic maize plants (Figure 2, Table 1). The final length (*Lm*) of the 4th leaf in the nine non-transgenic maize plants was 535 ± 6 mm on average. This value was attained after 231 ± 5 °Cd (*te*), which, in this experiment, is equivalent to 16.5 days after sowing. The moment at which the LER was maximal, *tm*, was 167 ± 5 °Cd, or 12.0 days after sowing. As illustrated here, the advantage of using the beta sigmoid function in the form of Auzanneau and coworkers [7] is that the function parameters themselves are biologically relevant when assessing monocot leaf growth. This represents a clear

advantage over functions that are based on parameters with no biologically relevant meaning or parameters that are difficult to interpret visually such as in the Weibull equation [26].

3.1.3 Flexibility to extract additional biologically relevant information from the dataset

As described above, the advantage of fitting a continuous function to the data is that for any given thermal time t (°Cd), the leaf length L (mm) can be estimated and vice versa. To illustrate the flexibility of the tool to accommodate specific biological questions, such as partial leaf elongation duration (LED), early leaf growth predictions and comparison of leaf growth independent of final leaf size, several additional parameters are generated.

For example, the nine non-transgenic maize plants of dataset 1a (Table 1) reached 10% of their final length (*t10%*) already at on average 89 ± 4 °Cd or 6.5 days after sowing. For *t100*, the moment at which the 4th leaf reaches 100 mm, we obtained an estimate of 108 ± 4 °Cd. In this experiment this was just before the 4th leaf emerged from the pseudo-stem (approximately 8 days after sowing). Knowing that the average final length (*Lm*) is 535 ± 6 mm, we can state that a considerable share of at least 19% of the final maize leaf length is hidden in the pseudo-stem. Exactly 50% of the final leaf length was attained at on average 153 ± 5 °Cd (*t50%*), which is not significantly different from the moment of the maximal LER, *tm* at 167 ± 5 °Cd. The leaf reached 90% of its final leaf size (*t90%*) at 203 ± 5 °Cd or 14.5 days after sowing. The use of parameters that are independent of final leaf size, such as *t10%*, *t50%* and *t90%*, can however be more meaningful when comparing genotypes that differ inherently in final leaf sizes. Likewise, we calculated a time window between reaching *10%* and *90%* of the final leaf size. The parameter was named [*LED(10%-90%)*] and corresponded to 114 ± 2 °Cd or 8 days in this experiment. Depending on the objective of a particular experiment, several other useful parameters can be extracted from the growth curve. Here, the method is illustrated using non-destructive measurements of leaf length over time. Therefore, no empirical evidence can be provided for the estimation of *t100* and *t10%*. On the other hand, these parameters represent estimations not far from the start of the actual measurements, and their estimations should approximate reality. However, when one wants to investigate early leaf development, i.e. when the leaf is still hidden in the pseudo-stem, we recommend that these parameters are first validated using destructive measurements of leaf elongation. In this way, validated predictions of early leaf growth can form a complementary method for the conventional techniques, such as the pinning or the ink method [29, 30].

Table 1. Effect of *GA20ox1* overexpression on maize leaf elongation, based on the analysis of a segregating population produced by backcrossing (BC) a transgenic plant overexpressing the *Arabidopsis thaliana GIBBERELLIC ACID 20 OXIDASE1 (GA20ox1) gene* to the wild-type line B104. The results are based on the analysis of eleven transgenic and nine non-transgenic BC1 plants. Lm: final leaf length; LERmax: maximal leaf elongation rate; t10%, t50%, t90%, te: time points at which the leaf reaches 10%, 50%, 90% and 100% of the final leaf length, respectively; t100: time point at which the leaf reaches 100 mm; tm: time point at which the leaf reaches LERmax; LEDs: leaf elongation durations between above stated thermal time points

	Growth parameter		AtGA20ox1 OE (mean ± SE)			Control (mean ± SE)			Difference in mean (+)	
	Lm	(mm)	743	±	13	535	±	6	38.9%	***
	LERmax	($mm°C^{-1}d^{-1}$)	6.2	±	0.2	4.6	±	0.1	34.2%	***
Thermal time points	t10%	(°Cd)	97	±	3	89	±	4	9.5%	NS
	t100	(°Cd)	107	±	2	108	±	4	-1.7%	NS
	t50%	(°Cd)	165	±	3	153	±	5	7.6%	*
	tm	(°Cd)	180	±	4	167	±	5	7.7%	*
	t90%	(°Cd)	217	±	4	203	±	5	6.7%	*
	te	(°Cd)	246	±	5	231	±	5	6.3%	*
Leaf elongation durations	LED(100-e)	(°Cd)	139	±	4	123	±	2	13.3%	**
	LED(10%-90%)	(°Cd)	119	±	3	114	±	2	4.5%	NS
	LED(10%-e)	(°Cd)	148	±	4	142	±	2	4.2%	NS
Ratio of initial growth phase over total growth phase	LED(100-m)/LED(100-e)	(-)	0.530	±	0.006	0.480	±	0.007	10.4%	***
	LED(10%-m)/LED(10%-e)	(-)	0.560	±	0.005	0.552	±	0.006	1.4%	NS
	LED(10%-50%)/LED(10%-90%)	(-)	0.568	±	0.002	0.564	±	0.002	0.5%	NS
	LED(10%-m)/LED(10%-90%)	(-)	0.696	±	0.005	0.688	±	0.006	1.1%	NS

+ Statistical significance based on student t-test of non-transgenic plants (n=9) vs GA20ox1 overexpression (n=11), * p<0.05, ** p<0.01, *** p<0.001, NS non-significant. Applied base temperature for thermal time calculation = 10°C, Mean of overall R^2 values = 0.9983 (0.9970-0.9991)

3.1.4 Leaf elongation rate and '(non)steady-state' growth

LEAF-E is particularly appropriate for calculating the LER. The LER can be determined at any thermal time point during leaf development and the maximum of this bell-shaped curve, the result of the first derivative of the beta sigmoid function, is denoted as the maximal LER or *LERmax*.

We estimated a remarkably high LER for the nine non-transgenic maize plants, with a *LERmax* of 4.6 ± 0.1 mm/°Cd, equivalent to an impressive 2.7 mm/h or 64 mm/day in this experiment. Due to this high LER, the 4th leaf completed its growth from embryonic leaf in the seed to its full length in merely 231 ± 5 °Cd or 16.5 days, calculated from the moment of sowing.

In several studies, a steady-state LER is assumed during a given growth period, or in other words, a period of constant LER can be defined [5, 30-32]. Steady-state growth in monocot species such as maize has very often been used as an acceptable simplification of the actual gradual growth process [5, 30-32]. However, it can be argued that this steady-state growth period is short compared to the total leaf growth period [8, 30, 33]. The results shown in Figure 22 support this view. Nonetheless, LEAF-E can be used to determine the duration and the timing at which leaf growth is approximately constant. To illustrate this, we estimated the thermal time window between the points at which LER has a value of 95% of *LERmax*, this is before and after reaching the moment of the maximal LER or *tm*. The same calculation was performed for 90% of *LERmax* and both periods were assayed for the nine non-transgenic plants of dataset 1a. In this experiment the corresponding time windows, which are associated with a relatively stable LER, lasted 34.5 ± 0.6 °Cd or 2.5 days (10.7 until 13.1 days after sowing) and 49.1 ± 0.9 °Cd or 3.5 days (from 10.1 until 13.6 days after sowing) using 95% and 90% of *LERmax*, respectively. The 95% and 90% steady-state windows estimated here comprised thus 15% and 21% of the total leaf growth period (i.e. from sowing until fully expanded 4th leaf), respectively, suggesting that a relatively stable LER is found only for a short time-span (2.5 to 3.5 days) during leaf growth.

3.2 Effect of GA20ox1 overexpression on maize leaf elongation

Comparison of transgenic plants overexpressing the GA biosynthesis gene *GA20ox1* with non-transgenic plants in a previous study, demonstrated that altering GA levels specifically affects the size of the division zone, resulting in proportional changes in leaf and whole plant growth rates [34]. The results obtained for this dataset using LEAF-E are presented in Figure 23A and Table 1 and are discussed below.

Overexpression of *GA20ox1* in maize results into significantly longer leaves (*Lm*) and higher maximum leaf elongation rates (*LERmax*) (39% and 34% respectively, $p<0.001$). This corresponds very well with the 38% increase in LER reported earlier by Nelissen and coworkers [34] and is consistent with an increased growth zone. However, in addition to the calculation of *Lm* and *LERmax*, the LEAF-E method facilitates a more profound analysis of the data. For example, leaves of transgenic plants take slightly, but significantly more time to reach their full length (*te*) than those of non-transgenic plants (Figure 23A and Table 1). As a consequence, also *t50%*, *t90%* and *te*, as well as *tm*, occur significantly later in the transgenics. This means that leaves of transgenic plants take slightly more time to reach 50%, 90% and 100% of their final leaf length. In contrast, the

leaves of the transgenics and non-transgenics attain the size of 100 mm (*t100*, Table 1) at exactly the same thermal time. Therefore, we speculate that the initiation of the growth of the 4th leaf in both groups is synchronous and that the leaf elongation period before *tm* is longer in the leaves of transgenic plants. Attaining *LERmax* coincides with the final cell divisions in the division zone, that start to elongate shortly after exit from the division zone [35]. Accordingly, the timing of *tm* could be linked with cell division activity in the growth zone. The shift in *tm* detected for the *GA20ox1* overexpressing plants, could thus be due to the alteration of the cell division activity in the leaves.

Summarizing, using LEAF-E to reanalyze dataset 1a, we could confirm previous findings. In addition we demonstrate that when LEAF-E is used, the timing of leaf growth can be studied in great detail, thereby facilitating the detection of a shift in attaining *LERmax* in *GA20ox1* overexpressing plants, that could not be quantified using other methods. The observed alterations are consistent with altered cell division activity in the division zone of *GA20ox1* overexpressing leaves [34]. However, to confirm the correlation of the shift in *tm* with altered cell division activity, a kinematic analysis of cells in the division zone, together with a detailed study of the timing of leaf elongation, should be performed in more cases of altered cell division activity.

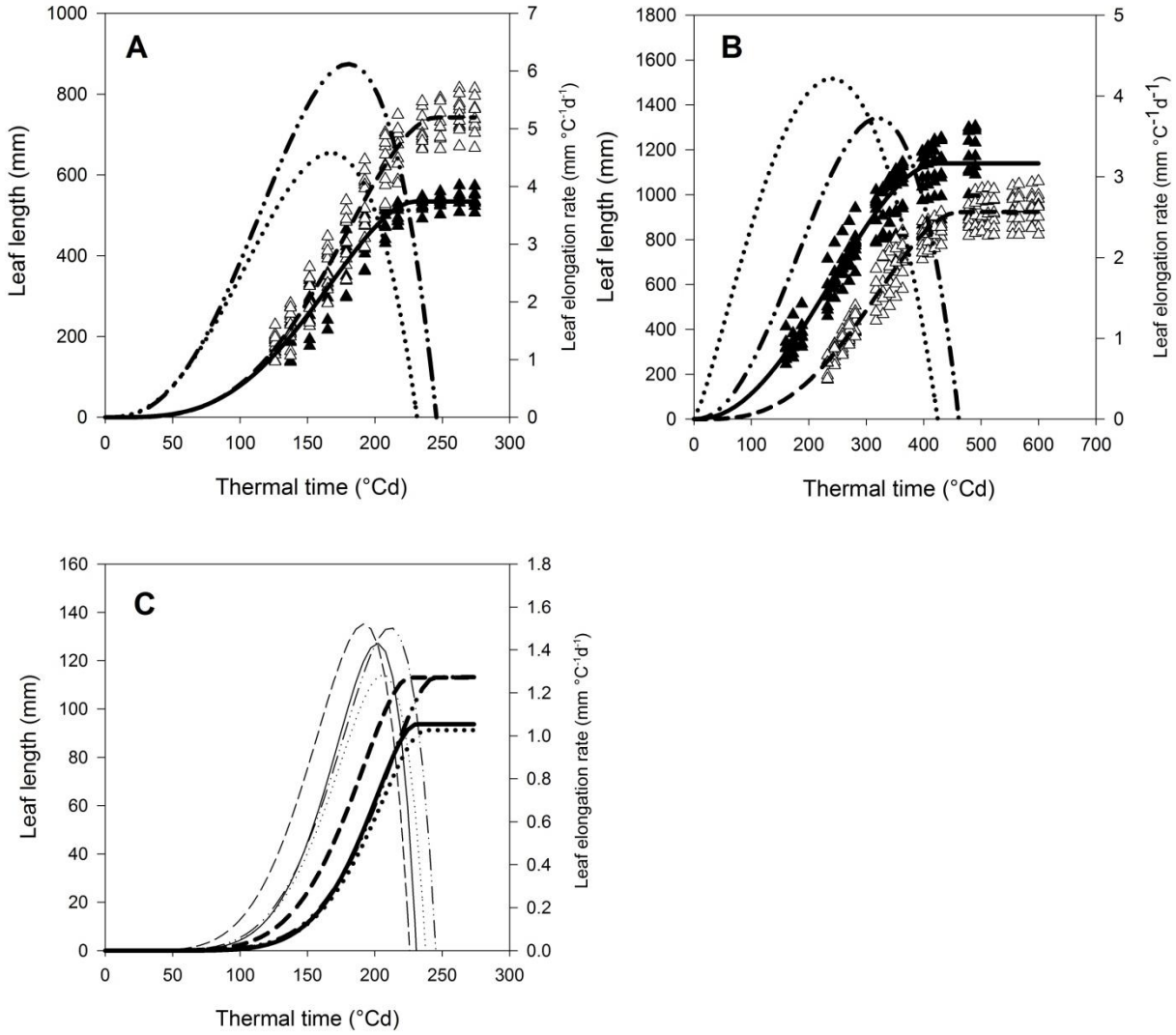


Figure 23. Analysis of the leaf elongation datasets of maize, *Miscanthus* and *Brachypodium* using LEAF-E. (A) Leaf length measurements of transgenic (white triangle) and non-transgenic (black triangle) plants of a segregating population produced by backcrossing a transgenic plant overexpressing the *GA20ox1 gene* to the wild-type line B104 maize, including leaf length (s-shaped) and LER (bell-shaped) function plots for both groups using the mean values of the function parameters. (B) Leaf length measurements of *Miscanthus sinensis* 'Goliath' (black triangle) and *M. x giganteus* (white triangle) plants, including leaf length and LER function plots for both groups using the mean values of the function parameters. (C) Leaf length and LER function plots for *Brachypodium distachyon* inbred lines Bd3-1 (full line), Bd21 (dotted), Bd21-3 (dashed) and Bd2-3 (dash dotted) using the mean values of the function parameters. For the sake of clarity, the individual leaf length measurements are not shown in this case.

3.3 Variation in leaf growth behavior in two *Miscanthus* species

We investigated the leaf growth characteristics of two genotypes of *Miscanthus* belonging to different species with high potential as bio-energy crops, but with contrasting phenotypic characteristics. *M. sinensis 'Goliath'* is characterized by high shoot densities, whereas *M. x giganteus* produces less, but thicker and taller shoots [36, 37].

When comparing leaf growth characteristics of *M. sinensis 'Goliath'* and *M. x giganteus*, we found that *M. sinensis 'Goliath'* had significantly longer leaves than *M. x giganteus* (Table 2, Figure 23B). The leaves of *M. sinensis 'Goliath'* on average grow up to 1140 ±46 mm long, whereas leaves of *M. x giganteus* on average become 923 ± 26 mm in length (Table 2). The longer leaves of *M. sinensis 'Goliath'* plants cannot be explained by significant changes in *LERmax* (differences are apparent but not significant) nor by an extended elongation period (no significant differences for LED values). However, our analysis revealed that leaves of *M. sinensis 'Goliath'* plants display a very strong initial growth compared to *M. x giganteus*. Parameters *t10%*, *t50%* and *tm* are all attained sooner (Table 2). For example, leaves of *M. sinensis 'Goliath'* reached 10% of their final leaf length at 101 ± 6 °Cd, or 61 °Cd sooner than *M. x giganteus* leaves, and 50% of their final leaf length at 230 ± 8 °Cd, or 64 °Cd sooner than *M. x giganteus* plants, at which point the leaves of the last are just emerging from the pseudostem (the spiral arrangement of leaves forming a cylindrical structure from where newly formed leaves emerge). Parameters defining the ratio of the duration of the accelerating phase over total LED show that *M. sinensis 'Goliath'* plants spend relatively less time in the accelerating phase than *M. x giganteus* plants (Table 2). This means that their leaves grow with relatively higher LER before reaching a maximum at *tm*.

The analysis with the LEAF-E method shows that the leaf growth pattern in these two *Miscanthus spp.* is clearly different and that the differences in total leaf length can be attributed to contrasting early leaf growth.

Table 2. Comparison of leaf elongation in two *Miscanthus* genotypes from different species. The results are based on the analysis of nine *M. sinensis 'Goliath'* and eight *M. x giganteus* plants. Lm: final leaf length; LERmax: maximal leaf elongation rate; t10%, t50%, t90%, te: time points at which the leaf reaches 10%, 50%, 90% and 100% of the final leaf length, respectively; tm: time point at which the leaf reaches LERmax; LEDs: leaf elongation durations between above stated thermal time points.

	Growth parameter		*M. Sinensis 'Goliath'* (mean ± SE)			*M. x giganteus* (mean ± SE)			Difference in mean (+)	
	Lm	(mm)	1140	±	46	923	±	26	217	**
	LERmax	($mm°C^{-1}d^{-1}$)	4.2	±	0.2	3.8	±	0.2	0.4	NS
Thermal time points	t10%	(°Cd)	101	±	6	162	±	4	-61	***
	t100	(°Cd)	230	±	8	294	±	8	-64	***
	t50%	(°Cd)	240	±	10	320	±	8	-80	***
	tm	(°Cd)	471	±	13	493	±	16	-21	NS
	t90%	(°Cd)	425	±	10	461	±	18	-37	NS
	te	(°Cd)	329	±	8	295	±	18	35	NS
Leaf elongation Duration	LED(10%-90%)	(°Cd)	370	±	12	330	±	15	40	NS
	LED(10%-e)	(°Cd)	323	±	7	299	±	18	24	NS
Ratio of initial growth phase over total growth phase	LED(10%-m)/LED(10%-e)	(-)	0.43	±	0.01	0.53	±	0.01	-0.10	***
	LED(10%-50%)/LED(10%-90%)	(-)	0.35	±	0.02	0.40	±	0.02	-0.05	NS
	LED(10%-m)/LED(10%-90%)	(-)	0.38	±	0.02	0.48	±	0.02	-0.10	**

+ Statistical significance based on student t-test of M. sinensis 'Goliath' (n=9) vs M. x giganteus (n=8), * p<0.05, ** p<0.01, *** p<0.001, NS non-significant. Applied base temperature for thermal time calculation = 8°C, Mean of overall R^2 values = 0.9931 (0.9669-0.9989)

3.4 Variation in leaf growth behavior in different *Brachypodium distachyon* inbred lines

Fifty *Brachypodium distachyon* inbred lines are currently being used for a study of natural diversity, which is led by 'The International Brachypodium Initiative' [28]. We analyzed the leaf growth behavior of four diploid *Brachypodium* inbred lines that are part of that study: Bd21, Bd21-3, Bd2-3 and Bd3-1 (Figure 23C).

Leaf growth analysis of these four genotypes with LEAF-E revealed distinct leaf growth characteristics. Based upon final leaf length, two groups can be distinguished. Bd21 and Bd3-1 have short leaves and Bd21-3 and Bd2-3 have long leaves (Figure 23C, Table 3). The length of the leaves is determined by both LER and LED. For Bd21, a low LER is probably the underlying factor of the shorter leaves (Table 3). This is in contrast to Bd3-1 which, like Bd21, has short leaves but a *LERmax* that is similar to those of the genotypes with longer leaves (Table 3). Parameters for LED are the smallest for Bd3-1. Thus, the leaf of Bd3-1 is short, most likely due to a short growing period.

Bd21-3 and Bd2-3 both have long leaves, a high *LERmax* and similar LED. Despite these similarities, based upon our analysis, we can conclude that the 3th leaf of Bd21-3 plants starts and finishes its growth significantly earlier in thermal time than that of Bd2-3 (Table 3).

DNA marker analysis showed that the inbred lines Bd21 and Bd21-3 are genetically very closely related, supporting the fact that both lines originate from the same accession [38]. Despite this close genetic relation, quite some variation in leaf growth behavior between these two inbred lines was found here.

Table 3. Comparison of leaf elongation in four *Brachypodium* inbred lines, Bd3-1 (n=7), Bd21 (n=10) plants, Bd21-3 (n=7) plants, Bd2-3 (n=10). Lm: final leaf length; LERmax: maximal leaf elongation rate; t10%, t50%, t90%, te: time points at which the leaf reaches 10%, 50%, 90% and 100% of the final leaf length, respectively; tm: time point at which the leaf reaches LERmax; LEDs: leaf elongation durations between above stated thermal time points.

	Growth parameter		Bd3-1				Bd21				Bd21-3				Bd2-3			
	Lm	(mm)	94	a	±	3	91	a	±	1	113	b	±	3	113	b	±	2
	LERmax	(mm°C^{-1}d^{-1})	1.44	ab	±	0.05	1.29	a	±	0.03	1.52	b	±	0.05	1.51	b	±	0.03
Thermal time points	t10%	(°Cd)	146	a	±	3	146	a	±	2	131	b	±	3	148	a	±	2
	t20	(°Cd)	164	a	±	3	166	a	±	2	145	b	±	3	163	a	±	2
	t50%	(°Cd)	189	a	±	2	193	a	±	2	178	b	±	3	197	a	±	2
	tm	(°Cd)	202	a	±	2	206	a	±	2	192	b	±	3	211	a	±	2
	t90%	(°Cd)	217	ab	±	1	222	bc	±	2	210	a	±	4	229	c	±	2
	te	(°Cd)	230.8	ab	±	0.8	238	bc	±	2	226	a	±	4	245	c	±	2
Leaf elongation durations	LED(20-e)	(°Cd)	67	a	±	3	72	a	±	1	81	b	±	2	82	b	±	2
	LED(10%-90%)	(°Cd)	71	a	±	2	76	ab	±	1	79	b	±	1	81	b	±	1
	LED(10%-e)	(°Cd)	85	a	±	2	92	ab	±	2	95	b	±	2	97	b	±	2
Ratio of initial growth phase over total growth phase	LED(20-m)/LED(20-e)	(-)	0.568	ab	±	0.006	0.559	a	±	0.004	0.579	bc	±	0.003	0.589	c	±	0.003
	LED(10%-m)/LED(10%-e)	(-)	0.660	a	±	0.004	0.654	a	±	0.002	0.642	b	±	0.002	0.651	ab	±	0.002
	LED(10%-50%)/LED(10%-90%)	(-)	0.614	a	±	0.004	0.609	ab	±	0.002	0.604	b	±	0.002	0.608	ab	±	0.001
	LED(10%-m)/LED(10%-90%)	(-)	0.789	a	±	0.004	0.784	ab	±	0.002	0.774	b	±	0.002	0.783	ab	±	0.002

+ Statistical significance indicated with distinct letters based on ANOVA and Scheffé Post hoc test ($p<0.05$) between lines Bd3-1 (n=7), Bd21 (n=10) plants, Bd21-3 (n=7) plants, Bd2-3 (n=10), applied base temperature for thermal time calculation = 11°C, Mean of overall R^2 values = 0.9993 (0.9871-0.9993)

3.5 Fitting of cell length measurements along the leaf axis of maize overexpressing *GA20ox1* using the beta sigmoid function

The cell length profile along the longitudinal axis of an actively growing grass leaf also displays a sigmoid pattern [8]. This sigmoidal profile is determined by the spatial distribution of cells in different stages of differentiation along the leaf axis: a number of dividing cells of small size at the leaf base, a stretch of cells that undergo elongation and thus increase in length when being pushed towards the leaf tip, and finally the tip of the leaf that is made up of cells that have reached their final length. We applied LEAF-E, adapted to use the extended version of the beta sigmoid function (Equation 3), to fit cell length measurements of dataset 1b. We found that fitting was successful and resulted in overall R^2-values ranging from 0.8420 up to 0.8749 for the transgenic and non-transgenic cell length measurements. Knowing that cell lengths can vary considerably, even for adjacent cells of the same cell file in one leaf (Figure 24), these R^2-values are noticeably high. However, a goodness of fit assessment on cell length measurements is, to our knowledge, not reported in literature, making further interpretation of the suitability of the beta sigmoid function for fitting cell length measurements difficult. We investigated the deviation of the data points from the fitted curve and could detect a slight underestimation of cell lengths in the very first 3 mm from the leaf base. This is an artifact inherent to the cell length profile of maize leaves when measured as close to the leaf base as was performed by Nelissen and coworkers [34] and Rymen and coworkers [3]. This is because the cells in the first few millimeter of the division zone are slightly longer than those in the middle of the division zone; this hyperbolic profile at the leaf basis cannot be fitted with the function that we used here, since it assumes an increment in cell lengths. The exploration of mathematical functions that are more appropriate for fitting these initial cell lengths along the longitudinal axis of monocot leaves should be the topic of future research.

The cell length profile along the leaf axis in maize plants overexpressing *GA20ox1* (dataset 1b) was analyzed by Nelissen and coworkers [34] using a polynomial fit approach similar to Rymen and coworkers [3]. To determine the size of the division zone, DAPI staining of actively dividing cells revealed that these cells are on average 20 µm long and that cells that surpass 40 µm proceed to elongation without further divisions [39]. Nelissen and coworkers [34] found that the division zone is 39% larger in *GA20ox1* overexpressing leaves compared to leaves of non-transgenic plants. They calculated the size of the elongation zone, defined as the distance between the end of the division zone and the point at which cells reach 95% of the calculated average mature cell length, being 52% larger in the *GA20ox1* overexpressing leaves compared to leaves of non-transgenic plants. In contrast, mature cell size in these leaves was not affected. Based upon those results they concluded that overexpression of *GA20ox1* results in a larger growth zone that is made up of more cells, without affecting the mature cell size [34].

Using LEAF-E, we determined that the size of the division zone, defined as the stretch of cells near the leaf base by which function fitted cell lengths do not exceed 40 µm, is on average 33% longer ($p<0.01$) in *GA20ox1* overexpressing leaves as compared to non-transgenic leaves (Figure 24). We estimated that the elongation zone, defined as the distance between the end of the division zone and the function parameter *Pe*, is 29% longer ($p<0.01$) in *GA20ox1* overexpressing leaves. Moreover, we can state that the position where maximal cell elongation *Pm* occurs is situated 8 mm (or 38%) further away from the leaf base in *GA20ox1* overexpressing leaves as compared to non-transgenic

plants and that the mature cell length *Lm* (mm) did not differ (p= 0.92). These findings are in accordance with results obtained earlier by Nelissen and coworkers [34] and show that the fitted curve can be very useful for data analysis in future experiments.

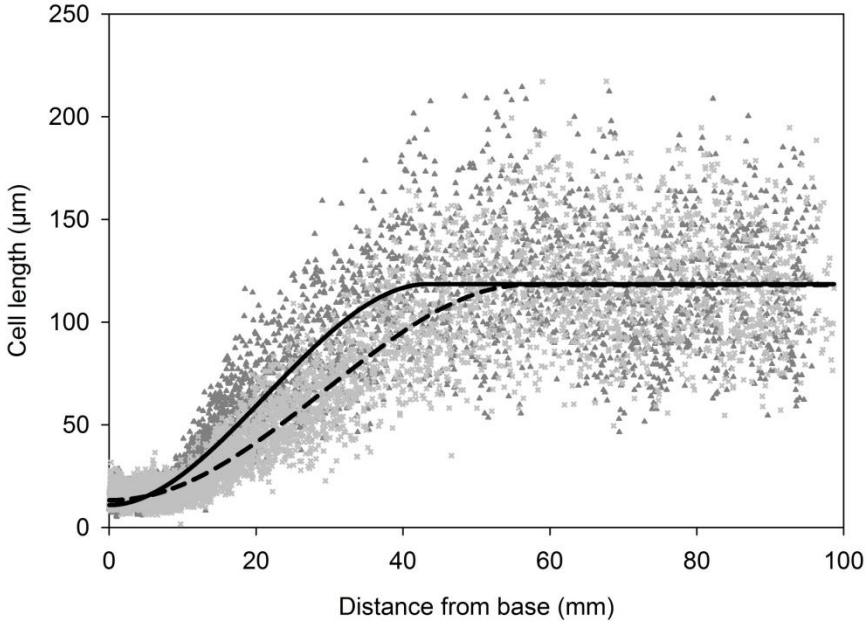


Figure 24. Effect of *GA20ox1* overexpression on the cell length profile of the 4th leaf in maize. The cell length profile along the axis of the 4th leaf is shown for three non-transgenic (triangle, dark gray) and three transgenic (x, light gray) plants. The S-shaped curves are function plots using the mean values of the fitted function parameters for non-transgenic (full) and transgenic (dashed) profiles.

4. Methods

4.1 Datasets used for method validation

4.1.1 Dataset 1a:

The dataset on leaf elongation of the maize B104 inbred line overexpressing the *Arabidopsis thaliana GIBBERELLIC ACID 20 OXIDASE1* (*GA20ox1*) gene, previously described by Nelissen and coworkers [34] was reanalyzed here. A segregating population produced by backcrossing the overexpression line (hemizygous for the transgenic event) to the wild-type B104 inbred line and consisting of 9 non-transgenic and 11 transgenic plants was used.

To determine leaf elongation rates, the length of the 4th leaf of transgenic and non-transgenic plants was measured daily until complete development, as previously described [34](Nelissen et al. 2012). For further details about growth conditions see [34]. The plants were grown in a growth chamber at 24°C. Here we used a base temperature of 10 °C for thermal time (Growth Degree Days, GDD) calculations.

4.1.2 Dataset 1b:

The same segregating population used to generate dataset 1a was previously used by Nelissen and coworkers [34] for the analysis of cell lengths along the leaf axis, based upon methods previously described [39]. In short, the 4th leaf was harvested two days after appearance from the pseudo-stem (stem-like structure composed of concentric rolled or folded blades and sheaths that surround the growing point). At this time point the ligule is only a few mm away from the base of the plant. The length of cell files adjacent to stomatal rows along the proximal-distal axis was measured using a DIC microscope (AxioImager, Zeiss, USA), and image analysis software (AxioVision, Zeiss, USA). The size of the division zone was determined as the distance between the base and the most distally observed mitotic figure in DAPI-stained leaves along the proximal-distal axis, with a fluorescence microscope (AxioImager, Zeiss, USA). Here we reanalyzed the cell length measurements.

4.1.3 Dataset 2:

Eight *Miscanthus* x *giganteus* and nine *M. sinensis* 'Goliath' plants were grown at 20°C (average temperature over the measuring period was 19.1°C) in a greenhouse, in Melle, Belgium, in September 2012 with no supplementary light. Plants were grown from rhizome cuttings in 2-l pots and were hand-watered and not fertilized during the experiment. The rhizomes were excavated during the winter of 2011 and stored in a cold room at 3°C until the start of the experiment. Also in this case, the length of the 4th leaf was measured five times a week (from leaf tip to soil level). The measurements were spread over a time period of approximately four weeks. The calculation of thermal time was based on the average air temperature in the greenhouse taking into account a base temperature of 8°C, based on Farrell and coworkers [40].

4.1.4 Dataset 3:

The *Brachypodium distachyon* inbred lines Bd21, Bd2-3 and Bd3-1 were provided by David F. Garvin from the USDA-ARS (Minnesota, US), and line Bd21-3 was provided by Richard Sibout from INRA-IJPB (Versailles, France). Plants were grown in rootrainers (Haxnicks®, UK) in biological replicates (n=10, 10, 7, 9 respectively) in a greenhouse (average temperature over the measuring period was 21 °C) in Melle, Belgium, August 2012 with no supplementary light. To calculate the thermal time, a base temperature of 10 °C was used. Fertilizer was added with the water supply: conductivity Ec = 1mS/cm; water soluble fertilizer Poly-feed (Haifa, Belgium) (N, P_2O_5, K_20; 20:5:20 + 3 MgO). Measurements were taken from the tip of the 3rd leaf to its basal level on a daily basis, for a period of 10 days.

4.2 A mathematical function for fitting leaf length measurements

For the estimation of leaf growth parameters we used the beta sigmoid function for determinate growth, inspired by the Euler integral, in the form of Equation 1(Yin et al. 2002). This function was used previously by Auzanneau and coworkers [7] and Verdenal and coworkers [25] to model leaf growth after cutting in *Lolium perenne*. The leaf length L (mm) at a given moment in development t (°Cd) is determined by final leaf length Lm (mm) and three particular points in leaf development, expressed as units of thermal time or growing degree days (°Cd). Thermal time is a summation of cumulative differences between daily mean temperature and a specified base temperature [41].

These thermal time points are the moment at which leaf growth starts *t0* (°Cd), the moment of maximal leaf growth rate *tm* (°Cd) and the moment at which leaf growth ceases *te* (°Cd). Estimations of *t0* often result in negative values that are biologically not relevant [7]. Therefore, in the experiments in which seedlings were involved (maize and *Brachypodium*), we assumed that *t0* = 0 was at the moment of sowing. In the case of *Miscanthus*, *t0* = 0 was assumed to be at the moment of potting the rhizomes (no visible leaves at this stage).

$$L = Lm.\ \ 1 + \frac{te-t}{te-tm}\ .\ \frac{t-t0}{te-t0}^{\frac{te-t0}{te-tm}}$$

Equation 1. Beta sigmoid function for fitting leaf length, modified from [7]. Function is applicable for t0 ≤ t ≤ te and t0 ≤ tm < te. For t > te, Equation 1 is reduced to L = Lm

The leaf elongation rate (LER) at any given moment in leaf development *t* (°Cd) can be calculated from the LER function (Equation 2), which is the first derivative of Equation 1. From this equation we determined the maximum leaf elongation rate or *LERmax* (mm/°Cd), as the LER at *tm*.

$$\frac{dL}{dt} = Lm.\ \ \frac{1 + \frac{te-t}{te-tm}\ .\ \frac{t-t0}{te-t0}^{\frac{te-t0}{te-tm}-1}}{te-tm} - \frac{\frac{t-t0}{te-tc}^{\frac{te-t0}{te-tm}}}{te-tm}$$

Equation 2. Leaf elongation rate function, modified from [7].

As Equation 1 is a continuous function, it allows calculating the leaf length *L* (mm) at any given moment in the leaf elongation period *t*, and vice versa (Figure 25). Therefore, in addition to the parameters *Lm*, *t0*, *tm* and *te*, we estimated a set of parameters that can be biologically relevant. For the maize dataset 1a we estimated the time point *t100*, the moment at which the leaf length is 100 mm. The *t100* time point, which is early in development, was chosen since it is close to the moment at which the leaf emerges from the pseudo-stem in maize non-transgenic B104 plants. Furthermore, we estimated *t10%*, *t50%* and *t90%* (°Cd), which are the moments at which the leaf reaches 10%, 50% and 90% of its final length, respectively. For *Brachypodium* (dataset 2), we replaced the *t100* parameter by *t20* (°Cd), the moment at which the leaf reaches 20 mm in length, to accommodate the smaller size of the *Brachypodium* leaf. These extra parameters allow comparing different treatments or inter-genotypic differences in a very detailed fashion.

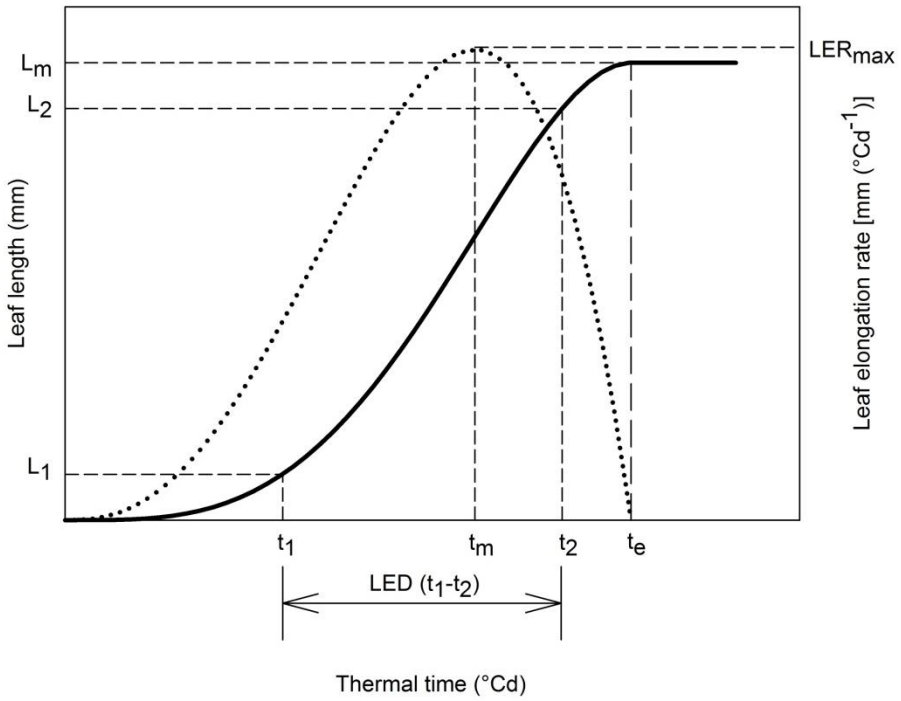


Figure 25. Deriving leaf growth parameters from the fitted leaf length curve (S-shaped) and the LER curve (bell-shaped) using LEAF-E. The leaf length curve is generated by fitting Equation 1 to measurements of a single leaf. Based upon the leaf length curve the final leaf length L_m can be calculated as well as the thermal time *t* needed to reach any given leaf length L, e.g. *te* is the thermal time needed to reach L_m. As a result, LEDs can be calculated between every desirable pair of thermal time points, e.g. LED(t1-t2). The first derivative of Equation 1 renders a bell-shaped LER curve from which maximal leaf elongation rate *LERmax*, occurring at the thermal time point *tm*, can be extracted.

Expressing leaf growth as durations of thermal time, the leaf elongation duration or LED, allows describing leaf growth in a fluctuating environment [31]. On that account, various LEDs were explored. For example, *LED(10%-90%)* defines the leaf growth duration between reaching 10% and 90% of its final length. Parameters *LED(100-e)* and *LED(10%-e)* were defined similarly.

Finally, we also defined ratios between growth durations. For example, the parameter *LED(10%-50%)/LED(10%-90%)* reflects the share of *LED(10%-50%)*, which is the thermal time spent for growing from 10% to 50%, in *LED(10%-90%)*, which is the thermal time necessary to grow from 10% to 90% of the final leaf size. A value of 0.5 would mean that an equal amount of thermal time was spent before and after reaching exactly 50% of the final leaf length, keeping in mind that for this parameter the growth period starts and ends at reaching 10% and 90% of the final leaf size. We explored also *LED(100-m)/LED(100-e)*, *LED(10%-m)/LED(10%-e)* and *LED(10%-m)/LED(10%-90%)*, as possibly biologically relevant parameters in particular experiments.

It should be mentioned that the reliability of the estimates of parameters early in development, before the leaf tip is visible (*t100* for maize and *t10%* for maize, *Miscanthus* and *Brachypodium*), was not tested here. Since non-destructive measurements were taken, no evidence for accurate fitting of these parameters could be provided. However, the estimate for *t100*, the moment at which the 4th leaf in maize B104 plants reaches 100 mm, is close to the actual measurements, and *t10%* and *t100* both are still useful as estimates for early leaf development. For *Brachypodium* plants, the leaf can be

measured already before it reaches 20 mm, and the *t20* parameter is supported by experimental data.

4.3 A mathematical function for fitting cell length measurements

Often, analysis of leaf growth involves cell length measurements along the leaf axis, providing insight in the sizes of dividing, elongating and mature cells, and of the leaf zones encompassing these three cell types [3, 8, 9, 29-31]. Therefore, we analyzed whether the beta sigmoid function could accurately fit cell length measurements along the leaf axis in maize. However, since in this case cells in the division zone have an initial length before proceeding to the elongation phase and eventually toward mature cells, a different equation had to be used as Equations 1 and 2 assume zero values at the beginning of growth. Yin and coworkers [26] describe an extended version of the beta sigmoid function that allows taking into account the initial length of the cell, *Lb* (mm) (Equation 3). In this case the data points do not represent a time series, but a positional-series of cell lengths along the leaf axis, taking the leaf base as position zero. For ease of interpretation, the symbols of the variable *t* and parameters *tm* and *te* from the original equation have been converted into the positions *p* (mm) along the leaf axis, starting from the base towards the tip. The position at which maximal cell elongation occurs is denoted as *Pm* (mm), and the position at which the cells cease to elongate is denoted as *Pe* (mm) (see Equation 3).

$$L = Lb + \ Lm - Lb \ . \ 1 + \frac{Pe-p}{Pe-Pm} \ . \ \frac{p}{Pe}^{\frac{Pe}{Pe-Pm}}$$

Equation 3. Extended version of the beta sigmoid function modified from [26] to fit cell length measurements along the leaf axis.

4.4 LEAF-E: Function fitting using a Microsoft Excel spreadsheet and the SOLVER function

The nonlinear fitting procedure described above was performed using Excel 2010 and the Solver function (32 bit) according to Brown [42]. The automation of the procedure in the form of a macro is innovative. Each row in the datasheet contains the data of one individual leaf (ordered in a time series), the starting values of the parameters of the model, and the formulae to extract the necessary statistical components for the calculation of the least square estimates following Neter and coworkers [43]. First, the macro checks for non-empty rows. When a non-empty row is found, the model is fitted to the data by minimizing the sum of squares of the errors iteratively, and changing the starting values of the parameters at each step. Per row (=leaf), all values described in the previous sections are calculated and stored in tabular form for further statistical analysis. In addition, data is generated and used to automatically produce a graph showing the original data points, the fitted growth curve and the function variables. This enables the evaluation of the correlation coefficient and visual interpretation of the goodness of fit. It also provides an easy way to check for miss fits or errors in the data. Miss fitting can occur when the Solver function fails to minimize the sum of squares of the errors using a particular set of starting values. Accordingly, the procedure can be repeated using more appropriate starting values. This semi-automatic way of working guarantees a fast analysis, avoids errors due to copying and pasting from text files and provides a table in standardized format containing the resulting parameters and derived variables per leaf. The results in the table can easily be analyzed in search of differential responses using

standard statistical software. We gave this Excel tool, which is available upon request, the name LEAF-E.

4.5 Statistical analysis

The statistical analysis of the derived growth parameters comprised a student t-test on datasets containing only two genotypes (dataset 1a, 1b and 2) and an ANOVA followed by post hoc Scheffé tests for datasets containing more than two genotypes (dataset 3). All analyses were carried out in the software package *STATISTICA* version 11 (Statsoft Inc., USA).

5. Conclusions

We provide here a straightforward tool in a simple Microsoft Excel format that does not require prior knowledge of complex mathematics, programming or advanced statistics. The automated the procedure using Excel 2010 and the Solver function (32 bit), described by Brown [42]. The results are stored in tabular form and can easily be analyzed in search of differences due to the applied treatments or to explore inter-genotypic or inter-population differences. Our way of working guarantees a fast, reliable and robust analysis of large datasets, avoids errors due to the copying and pasting from the text files (such as when using SAS or R software) and provides a standardized table with the resulting parameters and derived variables per leaf.

We applied LEAF-E to three datasets containing leaf length measurements of maize, *Miscanthus* and *Brachypodium*. In *Miscanthus* and *Brachypodium*, we have shown that LEAF-E is an appropriate tool for data analysis and that the analyzed species and genotypes display distinct leaf growth characteristics. In maize, the changes in both leaf elongation and cell length profile along the leaf axis as a result of enhanced GA levels, previously demonstrated by Nelissen and coworkers [34], were confirmed. In addition, we demonstrated that using LEAF-E, the timing of leaf growth can be studied in great detail, thereby facilitating the detection of dissimilarities in the timing of leaf growth that could not be quantified using other approaches. Furthermore, analysis with LEAF-E allows for a stable calculation of *LERmax*, which leads to reliable detection of significant changes. We propose this method as an excellent tool for comparing leaf growth behavior in different genotypes or to analyze the response of specific genotypes to a treatment. Moreover, the results presented here demonstrate that the method is probably applicable for most C3 and C4 monocot species. This method is provided in the form of an Excel worksheet to the scientific community.

6. Competing interests

The authors declare that they have no competing interests.

7. Acknowledgements

We thank the Agency for Innovation by Science and Technology in Flanders (IWT) for funding and Simon Fonteyne (ILVO, Belgium) for providing leaf measurements of *Miscanthus spp.* We also like to thank David Garvin and Richard Sibout for providing Brachypodium seed stocks.

8. References

1. Fricke W, McDonald AJS, Mattson-Djos L: Why do leaves and leaf cells of N-limited barley elongate at reduced rates? *Planta* 1997, 202:522-530.
2. Munns R, James RA: Screening methods for salinity tolerance: a case study with tetraploid wheat. *Plant Soil* 2003, 253:201-218.
3. Rymen B, Fiorani F, Kartal F, Vandepoele K, Inzé D, Beemster GTS: Cold nights impair leaf growth and cell cycle progression in maize through transcriptional changes of cell cycle genes. *Plant Physiol* 2007, 143:1429-1438.
4. Clifton-Brown JC, Jones MB: The thermal response of leaf extension rate in genotypes of the C_4-grass *Miscanthus*: an important factor in determining the potential productivity of different genotypes. *J Exp Bot* 1997, 48:1573-1581.
5. Lafarge T, Tardieu F: A model co-ordinating the elongation of all leaves of a sorghum cultivar was applied to both Mediterranean and Sahelian conditions. *J Exp Bot* 2002, 53:715-725.
6. Verelst W, Bertolini E, De Bodt S, Vandepoele K, Demeulenaere M, Pè ME, Inzé D: Molecular and physiological analysis of growth-limiting drought stress in *Brachypodium distachyon* leaves. *Mol Plant* 2013, 6:311-322.
7. Auzanneau J, Huyghe C, Escobar-Gutiérrez AJ, Julier B, Gastal F, Barre P: Association study between the gibberellic acid insensitive gene and leaf length in a *Lolium perenne* L. synthetic variety. *BMC Plant Biol* 2011, 11:183.
8. Parent B, Conejero G, Tardieu F: Spatial and temporal analysis of non-steady elongation of rice leaves. *Plant Cell Environ* 2009, 32:1561-1572.
9. Fiorani F, Beemster GTS, Bultynck L, Lambers H: Can meristematic activity determine variation in leaf size and elongation rate among four *Poa* species? A kinematic study. *Plant Physiol* 2000, 124:845-855.
10. Arredondo JT, Schnyder H: Components of leaf elongation rate and their relationship to specific leaf area in contrasting grasses. *New Phytol* 2003, 158:305-314.
11. Bultynck L, Ter Steege MW, Schortemeyer M, Poot P, Lambers H: From individual leaf elongation to whole shoot leaf area expansion: A comparison of three *Aegilops* and two *Triticum* species. *Ann Bot* 2004, 94:99-108.
12. Reymond M, Muller B, Tardieu F: Dealing with the genotype x environment interaction via a modelling approach: a comparison of QTLs of maize leaf length or width with QTLs of model parameters. *J Exp Bot* 2004, 55:2461-2472.
13. Chenu K, Chapman SC, Hammer GL, McLean G, Ben Haj Salah H, Tardieu F: Short-term responses of leaf growth rate to water deficit scale up to whole-plant and crop levels: an integrated modelling approach in maize. *Plant Cell Environ* 2008, 31:378-391.
14. Dignat G, Welcker C, Sawkins M, Ribaut JM, Tardieu F: The growths of leaves, shoots, roots and reproductive organs partly share their genetic control in maize plants. *Plant Cell Environ* 2013, 36:1105-1119.
15. Paine CET, Marthews TR, Vogt DR, Purves D, Rees M, Hector A, Turnbull LA:

How to fit nonlinear plant growth models and calculate growth rates: an update for ecologists. *Methods Ecol Evol* 2012, 3:245-256.

16. Motulsky HJ, Ransnas LA: Fitting curves to data using nonlinear regression: a practical and nonmathematical review. *FASEB J* 1987, 1:365-374.
17. Evers JB, Vos J, Fournier C, Andrieu B, Chelle M, Struik PC: Towards a generic architectural model of tillering in Gramineae, as exemplified by spring wheat (*Triticum aestivum*). *New Phytol* 2005, 166:801-812.
18. Evers JB, Vos J, Yin X, Romero P, van der Putten PEL, Struik PC: Simulation of wheat growth and development based on organ-level photosynthesis and assimilate allocation. *J Exp Bot* 2010, 61:2203-2216.
19. Karadavut U, Palta Ç, Kökten K, Bakoğlu A: Comparative study on some non-linear growth models describing leaf growth of maize. *Int J Agric Biol* 2010, 12:227-230.
20. Greef JM, Ott H, Wulfes R, Taube F: Growth analysis of dry matter accumulation and N uptake of forage maize cultivars affected by N supply. *J Agric Sci* 1999, 132:31-43.
21. Jame YW, Cutforth HW, Ritchie JT: Temperature response function for leaf appearance rate in wheat and corn. *Can J Plant Sci* 1999, 79:1-10.
22. Hu Y, Camp K-H, Schmidhalter U: Kinetics and spatial distribution of leaf elongation of wheat (*Triticum aestivum* L.) under saline soil conditions. *Int J Plant Sci* 2000, 161:575-582.
23. Guzmán G, Pukkala T, Palahí M, de-Miguel S: Predicting the growth and yield of *Pinus radiata* in Bolivia. *Ann For Sci* 2012, 69:335-343.
24. Stewart DW, Dwyer LM: Appearance time, expansion rate and expansion duration for leaves of field-grown maize (*Zea mays* L.). *Can J Plant Sci* 1994, 74:31-36.
25. Verdenal A, Combes D, Escobar-Gutiérrez AJ: A study of ryegrass architecture as a self-regulated system, using functional-structural plant modelling. *Funct Plant Biol* 2008, 35:911-924.
26. Yin X, Goudriaan J, Lantinga EA, Vos J, Spiertz HJ: A flexible sigmoid function of determinate growth. *Ann Bot* 2003, 91:361-371.
27. Goudriaan J, Van Laar HH: *Modelling potential crop growth processes : textbook with exercises.* Dordrecht; Boston: Kluwer Academic Publishers; 1994.
28. Vogel JP, Garvin DF, Mockler TC, Schmutz J, Rokhsar D, Bevan MW, Barry K, Lucas S, Harmon Smith M, Lail K, et al: Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* 2010, 463:763-768.
29. Ben-Haj-Salah H, Tardieu F: Temperature affects expansion rate of maize leaves without change in spatial distribution of cell length (analysis of the coordination between cell division and cell expansion). *Plant Physiol* 1995, 109:861-870.
30. Muller B, Reymond M, Tardieu F: The elongation rate at the base of a maize leaf shows an invariant pattern during both the steady-state elongation and the establishment of the elongation zone. *J Exp Bot* 2001, 52:1259-1268.

31. Tardieu F, Reymond M, Hamard P, Granier C, Muller B: Spatial distributions of expansion rate, cell division rate and cell size in maize leaves: a synthesis of the effects of soil water status, evaporative demand and temperature. *J Exp Bot* 2000, 51:1505-1514.
32. Bouchabké O, Tardieu F, Simonneau T: Leaf growth and turgor in growing cells of maize (*Zea mays* L.) respond to evaporative demand under moderate irrigation but not in water-saturated soil. *Plant Cell Environ* 2006, 29:1138-1148.
33. Durand J-L, Schäufele R, Gastal F: Grass leaf elongation rate as a function of developmental stage and temperature: morphological analysis and modelling. *Ann Bot* 1999, 83:577-588.
34. Nelissen H, Rymen B, Jikumaru Y, Demuynck K, Van Lijsebettens M, Kamiya Y, Inze D, Beemster GT: A local maximum in gibberellin levels regulates maize leaf growth by spatial control of cell division. *Curr Biol* 2012, 22:1183-1187.
35. Skinner RH, Nelson CJ: Epidermal cell division and the coordination of leaf and tiller development. *Ann Bot* 1994, 74:9-15.
36. Greef JM, Deuter M: Syntaxonomy of *Miscanthus x giganteus* GREEF et DEU. *Angew Bot* 1993, 67:87-90.
37. Hodkinson TR, Renvoize SA, Chase MW: Systematics of *Miscanthus*. *Asp Appl Biol* 1997, 49:189-198.
38. Vogel JP, Tuna M, Budak H, Huo N, Gu YQ, Steinwand MA: Development of SSR markers and analysis of diversity in Turkish populations of *Brachypodium distachyon*. *BMC Plant Biol* 2009, 9:88.
39. Rymen B, Coppens F, Dhondt S, Fiorani F, Beemster GTS: Kinematic analysis of cell division and expansion. In *Plant Developmental Biology (Methods in Molecular Biology, Vol 655).* Edited by Hennig L, Köhler C. New York: Humana Press; 2010: 203-227. [Walker JM (Series Editor)
40. Farrell AD, Clifton-Brown JC, Lewandowski I, Jones MB: Genotypic variation in cold tolerance influences the yield of *Miscanthus*. *Ann Appl Biol* 2006, 149:337-345.
41. Atwell B, Kriedemann P, Turnbull C: *Plants in action: adaptation in nature, performance in cultivation.* Melbourne, Australia: MacMillan Education Australia Pty Ltd; 1999.
42. Brown AM: A step-by-step guide to non-linear regression analysis of experimental data using a Microsoft Excel spreadsheet. *Comput Meth Programs Biomed* 2001, 65:191-200.
43. Neter J, Kutner M, Nachtsheim C, Wasserman W: *Applied Linear Statistical Models.* Fourth edn. Columbus: McGraw-Hill Education; 1996.

Chapter 4: Brachypodium as model for (bioenergy) grasses

Chapter 4: Brachypodium as model for (bioenergy) grasses

Personal contribution to the work:

The transformation of Brachypodium for *GA20ox* overexpression and *BdCAD1* downregulation included cloning of the *BdCAD1* gene, Agrobacterium-mediated transformation of callus derived from immature embryos, *in vitro* cultivation of transformed calli and the selection of transgenic shoots. Transgenic plants were analyzed for leaf growth kinetics, transgene and endogene expression and CAD protein activity. For the practical work, I got assistance from the technical team at ILVO, Plant Unit, Growth and Development group. Furthermore, the construct for GA20ox overexpression was obtained from VIB-PSB.

The screening for TILLING mutants in *Bd4CL1*was performed during a three weeks stay at URGV, Evry, France in collaboration with INRA Versailles, France. The growing, genotyping and screening of these mutants was performed at ILVO with great help from the technical team.

I personally conducted the qRT-PCR expression analysis including data analysis and reporting of the *BdCAD1* TILLING mutant *Bd4179* leading towards co-authorship in:

Bouvier d'Yvoire M, Bouchabke-Coussa O, Voorend W, et al. (2013) Disrupting the *cinnamyl alcohol dehydrogenase 1 gene* (*BdCAD1*) leads to altered lignification and improved saccharification in *Brachypodium distachyon*. Plant J 73:496–508.

1. Summary

Brachypodium has been proposed as a model system for the improvement of food and feed crops as well as for bioenergy grasses. A short lifecycle, small statue, easy growth requirements, high quality genome sequence and close relationship to economically important cereals suggests its suitability as a model plant. In this chapter, the major achievements of Brachypodium from the moment of its introduction as a model plant in 2001 until present have been described. Special attention is paid to biomass yield and cell wall biosynthesis related research. Furthermore, the use of Brachypodium as an alternative model for studying bioenergy related traits such as biomass accumulation and saccharification efficiency is evaluated using the tools available for this new model system. Based on the experience gained during this study, the major hurdle for further implementation of Brachypodium in research is the low throughput of the Agrobacterium-mediated transformation. On the other hand, the alternative TILLING approach has proven to be effective for identifying mutants in lignin biosynthesis with enhanced saccharification efficiency.

2. Introduction

As discussed in Chapter 1, monocotyledonous plant species, particularly grasses, are major resources for food, feed and for the production of biofuels and biobased products (Brown 2003; Fursova et al. 2012; van der Weijde et al. 2013). The four most cultivated grasses worldwide are maize, rice, wheat and sugarcane (FAO Statistics Division 2013b). Promising bioenergy crops that are getting more attention in recent years include perennial grasses such as switchgrass and Miscanthus species (van der Weijde et al. 2013). Most of these grass species are in general physically large, have relatively long life cycles and large, complex genomes, which are inconvenient characteristics for biological research purposes (Opanowicz et al. 2008). Biological research in all these species could therefore strongly benefit from the availability of a suitable grass model system that is small, easy to work with in large numbers, and cheap to maintain (Draper et al. 2001).

2.1 **"Working grass hero"** (Garvin 2007)

The small dicotyledonous model plant *Arabidopsis thaliana* (hereafter 'Arabidopsis') is without doubt the most highly developed and powerful plant model system, and has contributed tremendously to the field of plant biology (Meyerowitz and Somerville 1994; Meinke and Koornneef 1997; Draper et al. 2001; Vogel and Bragg 2009; Flavell 2009; Koornneef and Meinke 2010). Arabidopsis, being a dicotyledonous plant species, is however only distantly related to the grasses (Keller and Feuillet 2000; Figure 26B). The lack of colinearity between the Arabidopsis genome and that of monocots such as rice suggests that a grass model system is a key requirement for the future identification of genes of agronomic interest from cereals and forage grasses (Devos and Gale 2000; Draper et al. 2001; Liu et al. 2001), and by extrapolation also from bioenergy grasses. Moreover, traits such as cell wall composition, plant architecture, grain properties, intercalary meristems and root architecture that differ fundamentally between monocots and dicots are best studied in a grass model (Vogel 2008; Vogel and Bragg 2009; Watt et al. 2009). *Brachypodium distachyon* (hereafter 'Brachypodium') was proposed as a model for the *Poaceae* in 2001 (Draper et al. 2001).

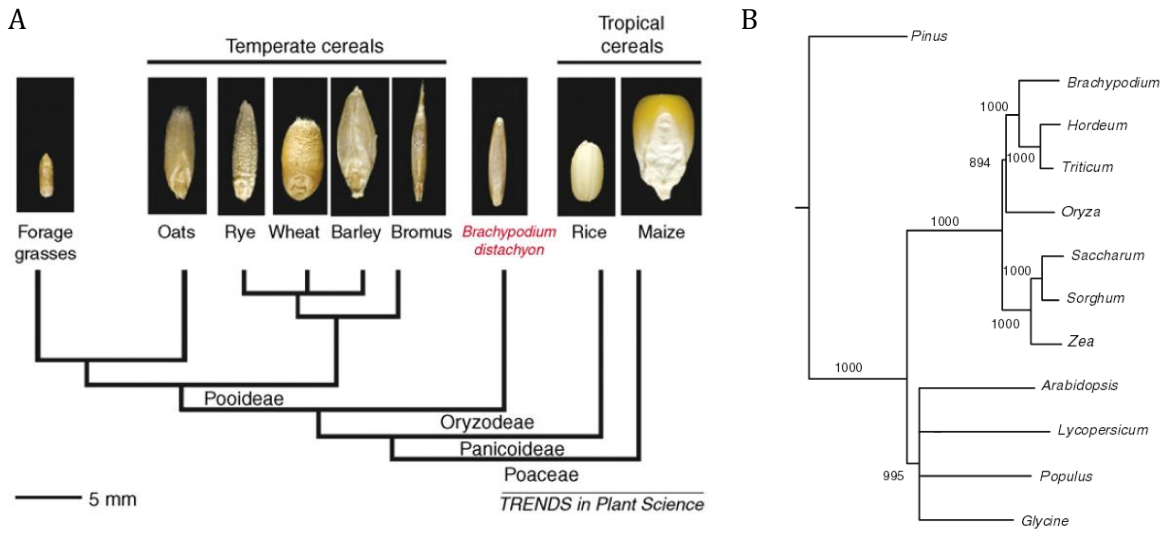


Figure 26. Phylogenetic relationship of Brachypodium to grasses and dicot species. (A) Phylogenetic relationships between Brachypodium and the small grain cereals (from Opanowicz et al. 2008). Examples of the respective grains are shown at the same scale (bar = 5 mm) (B) Rooted phylogenetic tree showing the relationship of Brachypodium to other grasses and shows the more distant relationship to dicot species (from Vogel et al. 2006)

Maize is an important crop, yielding in the USA alone 14 billion bushels or 356 million tonnes of grain (USDA 2013) but can also be considered as an important model organism for fundamental research (Schnable et al. 2009a). Vast collections of mutant stocks, the presence of large heterochromatic chromosomes, extensive nucleotide diversity, and synteny within related grasses, have positioned this species as a centerpiece for genetic, cytogenetic, and genomic research. Maize is the subject of a wide range of biological investigations including plant domestication, genome evolution, developmental physiology, epigenetics, pest resistance, heterosis, quantitative inheritance, and comparative genomics (Strable and Scanlon 2009). With maize being an important agricultural crop and an established model, it would be the model of choice for fundamental research in grasses. However, Brachypodium can provide an alternative model species for the *Poaceae* and has some advantages over maize for certain applications or research areas (Draper et al. 2001; Brkljacic et al. 2011; Catalán et al. 2012). First, Brachypodium possesses one of the smallest monocot and grass genomes (272Mb, Vogel et al. 2010 versus maize: 2.3Gb, Schnable et al. 2009), and comprises mostly single- or low-copy repetitive DNA (Vogel et al. 2010). Second, like Arabidopsis, Brachypodium has a small statue enabling low-cost and high density planting. For comparison, where densities of 1,000 plants/m^2 can easily be achieved in growth chambers or greenhouses for Brachypodium, the same space accommodates only 50 wheat plants, 36 rice plants, four switchgrass plants (Vogel and Bragg 2009) and six maize plants (Eddy and Hahn 2010). Third, the life cycle is short, on average twelve weeks (Opanowicz et al. 2008) but can be sped up to a minimum of a 6 weeks from seed to seed (Garvin et al. 2008; Mur et al. 2011), compared to more than four months for maize (own observations). Moreover, Brachypodium is more closer related to the economically important Triticeae species (*Triticum* spp., *Hordeum* spp.) than maize (Catalán and Shi 1995). The phylogenetic position of Brachypodium is basal to the four grass tribes that collectively encompass the vast majority of domesticated cool season cereal grain, forage, and turf crops (Figure 26; (Kellogg 2001; Bevan et al. 2010). This phylogenetic position gives it great

relevance for the development of a wide range of comparative analyses of gene expression and regulatory mechanisms that can be valuable for its economically valuable relatives (Mur et al. 2011). Fourth, a potential advantage of Brachypodium over maize is the free access to germplasm, which is often restricted in maize due to quarantine regulations and intellectual property concerns (Brkljacic et al., 2011; Jung et al., 2008). Finally, as a member of the *Poaceae*, it possess the type II cell wall (Carpita 1996) typical of all grasses (Vogel 2008) and thus would be a suitable model to study recalcitrance of the cell wall to degradation, an important aspect of bioenergy research.

2.2 International Brachypodium community: Genetic and genomics resources and tools

The success of a model system is highly dependent on its use by a large research community. The major advances that plant biology has made during the last fifty years could be achieved only because the research community agreed on concentrating efforts on a single organism, Arabidopsis. Moreover, the explosive growth of Arabidopsis research over the past 25 years has led to an extensive and effective network that is used to share, stimulate and coordinate research efforts (Meinke and Koornneef 1997; Koornneef and Meinke 2010). With Arabidopsis as an example, the Brachypodium community came up with the International Brachypodium Initiative (IBI), which has led to the sequencing of the genome of the community standard line Bd21 (Vogel et al. 2010). Besides the standard line, six other diploid inbred lines have been resequenced and an additional 50 more are currently being resequenced (http://brachypodium.pw.usda.gov). Public resources are gathered in the Brachypodium.org website. The IBI held its first genomics meeting and workshop at the PAG XIV conference in San Diego, California, in January 2006. The Brachypodium community was brought together in Europe on the First European Brachypodium workshop in Versailles, France (October 2011), and the First International Brachypodium Conference in Modena, Italy (June 2013). Currently, several essential tools to make of Brachypodium a model plant are in place, as summarized on a time scale in Table 4 (see also Vain 2011).

Table 4. Brachypodium resources, tools and major achievements for Brachypodium on a timeline

year	Major achievements or events	Reference
2001	Brachypodium distachyon proposed as model system	(Draper et al. 2001)
2004	cytotaxonomy	(Hasterok et al. 2004)
2005	particle bombardment transformation	(Christiansen et al. 2005)
2006	Agrobacterium-mediated transformation Bd21	(Vogel and Garvin 2006)
	inbred line formation	(Vogel and Garvin 2006)
	BAC libraries and EST sequencing	(Vogel et al. 2006)
	Approval of genome sequencing project	(Brkljacic et al. 2011)
2007	Draft version (4x) of Bd21 genome sequence	(Brkljacic et al. 2011)
2008	T-DNA insertion	(Vain et al. 2008)
	reference genes for qPCR	(Hong et al. 2008)
	saccharification potential and cell wall analysis	(Gomez et al. 2008)
	high efficiency transformation	(Vogel and Hill 2008; Vain et al. 2008; Păcurar et al. 2008)
2009	BAC-based physical map	(Gu et al. 2009)
	"Optimized" Agrobacterium mediated transformation Bd21	(Alves et al. 2009)
	deep sequencing of sRNAs	(Zhang et al. 2009; Wei et al. 2009)
	Illustrated guide to crossing Brachypodium	(Garvin 2009)
2010	Physical, genetic and cytogenic map	(Febrer et al. 2010)
	1000 T-DNA tags	(Thole et al. 2012)
	Genome sequence of Bd21	(Vogel et al. 2010)
	SSR-based linkage map	(Garvin et al. 2010)
	characterization of primary cell walls Bd21	(Christensen et al. 2010)
	VIGS	(Demircan and Akkaya 2009)
	Affymetrix custom Tiling array	http://www.plexdb.org/
	Natural variation in flowering time	(Schwartz et al. 2010)
2011	expression atlas Bradinet	http://aranet.mpimp-golm.mpg.de/bradinet
	First European Brachypodium Workshop (Versailles, France)	https://colloque4.inra.fr/1st_european_brachypodium_workshop
	natural variation in drought response	(Luo et al. 2011)
	grain and endosperm characterization	(Guillon et al. 2011; Opanowicz et al. 2011)
	First report on complementation of a T-DNA mutant	(Vain et al. 2011)
2012	Transient expression transformation system	(Fursova et al. 2012)
	Brachypodium promoters tested in maize	(Coussens et al. 2012)
	Evolution and taxonomic split	(Catalán et al. 2012)
	Protocol for Brachypodium leaf mesophyll protoplasts	(Hong et al. 2012)
2013	first report on successful downregulation of genes	(Trabucco et al. 2013)
	First International Brachypodium Conference (Modena, Italy)	http://www.brachy2013.unimore.it/
	Detailed characterization of cell walls and developmental anatomy of the Brachypodium distachyon stem internode	(Matos et al. 2013)

In what follows, the two tools and resources explored in this thesis, Brachypodium transformation and screening of TILLING (Targeting Induced Local Lesions in Genomes) populations are discussed in more detail.

2.2.1 Transformation of Brachypodium

For a model plant to be successfully introduced, an efficient transformation platform is indispensable. Brachypodium was proposed as model species in 2001, but it was only five years later that an Agrobacterium-mediated transformation protocol was available for the research community (Vogel and Garvin 2006). After that, a protocols with higher efficiency were published for inbred line Bd21-3 and for the standard line Bd21 (Vogel and Hill 2008; Alves et al. 2009, Figure 27). Other protocols for Agrobacterium-mediated transformation of callus derived from immature embryos have been published in the meantime by Lee et al. (2011), Vain et al. (2008), and Păcurar et al. (2008). Protocols for particle bombardment gene transfer (Christiansen et al. 2005) and for transient transformation (Fursova et al. 2012) have been published as well. Despite these efforts, the implementation of an optimal, high-efficiency transformation method is not that straightforward. One of the main difficulties encountered during the transformation of Brachypodium is the production of embryogenic callus that maintains its regenerative capacity which is in turn highly dependent on the quality of the seeds and the stage at which the immature embryo is dissected (Vogel and Hill 2008).

Both basta/bialaphos and Hygromycin resistance are being used with *HYGROMYCIN PHOSPHOTRANSFERASE* (*HPT*) and *PHOSPHINOTHRICIN ACETYL TRANSFERASE* (*BAR*) gene respectively as selectable markers. Anyhow, the presence of a fluorescent signal such as GFP aids greatly in the selection of transformed callus, shoots and plantlets (Vain et al. 2008). Interestingly, an oestradiol-inducible system has recently been used for overexpression of a NAC transcription factor resulting in ectopic secondary cell wall formation in Brachypodium (Valdivia et al. 2013). For Agrobacterium-mediated transformation of Brachypodium, the maize ubiquitin and rice actin promoters are generally used for driving transgene expression and maize ubiquitin and 35S for expression of the selectable marker. The vector pVec8-GFP, with inclusion of the CATI intron in the 5' untranscribed region of the gen of interest or selectable marker was reported with highest transformation efficiencies (20% of embryogenic callus producing at least one transgenic plant, Alves et al. 2009, Figure 27) as compared to previously described protocols (Vogel and Garvin 2006; Vogel and Hill 2008; Vain et al. 2008; Păcurar et al. 2008). The incorporation of an intron sequence in the marker gene was reported to improve transformation efficiencies previously in rice and barley (Cheng et al. 2004). A very efficient way to obtain transgenic Brachypodium plants is to order them at a transformation facility such as the Plant Transformation Facility (PTF) at Iowa State University (http://agron-www.agron.iastate.edu/ptf/service/Brachypodium.aspx). However, at the start of this PhD in 2009-2010, no transgenic seeds yet only 'transformed plantlets' were provided by this facility and these could not be transported overseas, thus not available for European institutes (personal communication with PTF).

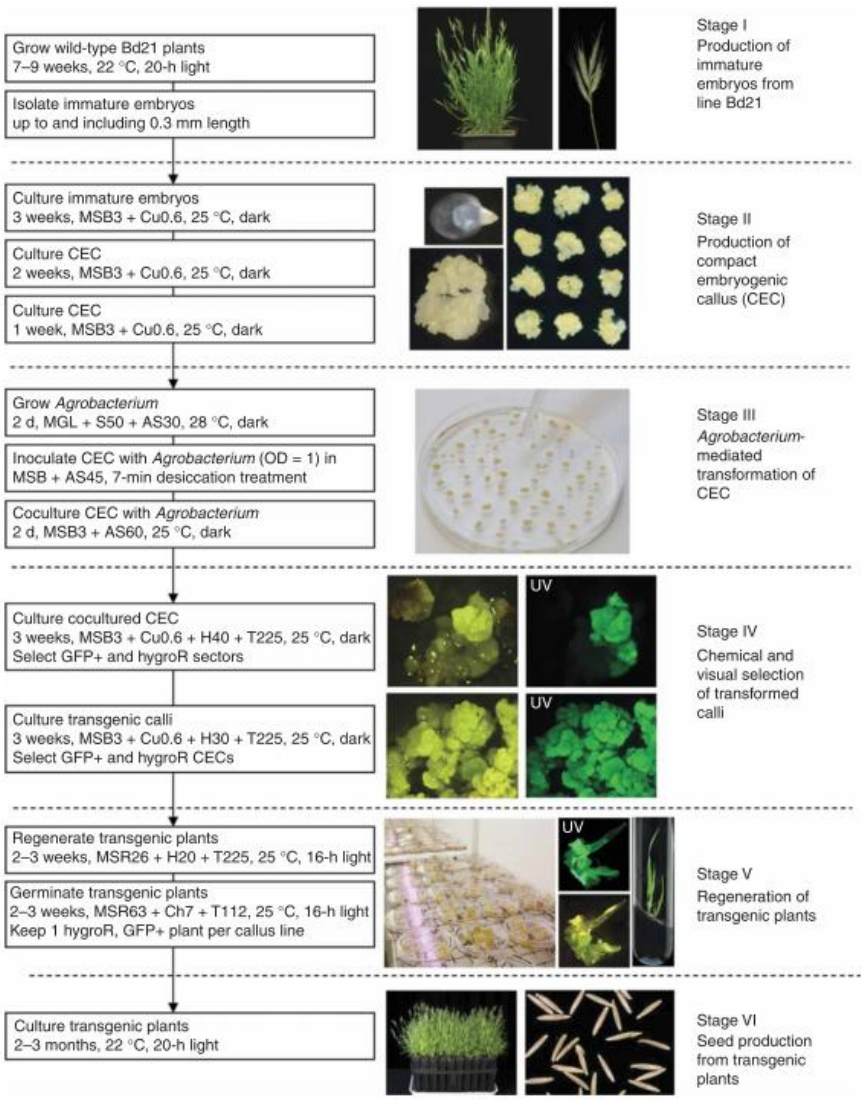


Figure 27. Agrobacterium-mediated transformation protocol for B. distachyon genotype Bd21. HygroR, hygromycin resistant; GFP+, expressing the green fluorescent protein. CEC, compact embryogenic callus. UV, ultraviolet/blue light. (Alves et al. 2009)

2.2.2 TILLING in Brachypodium

To study the function of a gene, the functional analysis of loss-of function mutants is a well-known strategy. For Brachypodium, T-DNA insertion and TILLING collections are available for this reverse genetics approach (Vain 2011; Thole et al. 2012; Dalmais et al. 2013). In the scope of this thesis, the TILLING strategy is further described.

The original concept of TILLING or 'Targeting Induced Local Lesions in Genomes' described the use of mutagens, such as EMS (Ethyl Methanesulfonate) or gamma irradiation to randomly induce genomic changes followed by high-throughput recovery of lesions for reverse genetics applications (McCallum et al. 2000). TILLING can be carried out using different technologies, but the central principle is the same; the detection of rare genetic mutations in pooled DNA samples from large mutant populations in a first round of screening, followed by deconvolution of the pool to identify

the individual carrying the mutation (Wang et al. 2012). To date, only one institute has produced a collection of chemically mutagenized Brachypodium lines that can be screened for forward and/or reverse genetics. The Institut National de Recherche Agronomique – Institut Jean-Pierre Bourgin (INRA-IJPB) (Versailles, France) and the Unité de Recherches en Génétique Végétale (URGV) (Evry, France) produced a sodium azide treated collection of 5530 families called BRACHYTIL. The inbred line Bd21-3 was used for this purpose (Dalmais et al. 2013). Recently, one other group has identified mutants in lignification in a gamma irradiation mutagenized TILLING population of 1773 plants (Lee et al. 2013). However, to my current knowledge, only the BRACHYTIL population allows for a high-throughput screening system that is available for the Brachypodium community. This includes the optimization of nested PCR conditions for tagging a PCR fragment amplified from the sequence of interest with fluorescent labels, the screening of the collection for mutations and providing the seed stocks of plants harboring the identified mutations. Currently, the screening is based on gene-specific PCR, incubation with an endonuclease that preferentially cleaves mismatches in heteroduplexes between wild type and mutant DNA and detection of these mismatches by polyacrylamide gel electrophoresis (Dalmais et al. 2013). However, this system is to be upgraded by using next generation sequencing of the whole collection so that mutations in genes of interest can be mined using software-based approaches rather than wet-lab based (personal communication with Richard Sibout). The BRACHYTIL population has successfully been used for identifying mutants in lignin biosynthesis. Mutations were found in eight genes belonging to five lignin biosynthesis families (Dalmais et al. 2013), as discussed below. The results indicate that TILLING is indeed a good strategy for identifying lignin mutants in Brachypodium.

2.3 Brachypodium research on biomass yield improvement and cell wall analysis

Arabidopsis research has greatly enhanced our knowledge of the fundamental mechanisms of plant growth (Koornneef and Meinke 2010). However, for traits that are fundamentally different between monocots and dicots such as root development, cell wall biosynthesis, tillering and flowering, the use of a closely related monocot model system like Brachypodium might be more appropriate (Opanowicz et al. 2008; Vogel 2008; Watt et al. 2009). Hence, the translation of knowledge obtained in Arabidopsis to Brachypodium would be one step further to close the gap between fundamental research and applications in food, feed and energy grasses.

An example of such a translational approach is the study of the effect of overproduction of Giberellic Acid (GA) on plant and organ growth. The overexpression of the *AtGA20ox1* gene in the dicot model systems Arabidopsis and tobacco resulted in a dramatic increase in biomass production and in plant and organ size (Coles et al. 1999; Biemelt et al. 2004). In maize, overexpression of the *AtGA20ox1* gene causes a 40% increase in leaf length by increasing the number cells in the growth zone (Nelissen et al. 2012). This strategy of increasing biomass production was tested in this thesis.

Furthermore, traits that can directly or indirectly affect biomass and/or seed yield and that have been studied in Brachypodium include root structure (Watt et al. 2009), vernalization and flowering time (Olsen et al. 2006; Opanowicz et al. 2008; Schwartz et al. 2010; Faricelli et al. 2010; Mach 2013; Wu et al. 2013), seed storage proteins (Laudencia-Chingcuanco and Vensel 2008; Charles et al. 2009; Gu et al. 2010), abiotic stress tolerance (Boden et al. 2013) and disease resistance (Parker et al. 2008; Azhaguvel et al. 2008; Cui et al. 2012; Ayliffe et al. 2013). Progress on biomass

improvement may come also from the use of Brachypodium in introducing C4 metabolism into C3 plants (Vain 2011; Weissmann and Brutnell 2012). When browsing through the phenotypic screens of the BRACHYTIL population, plants can be found with “big stems” or “high tiller number”, but no research reports have been published studying these mutants for improved biomass yield. However, more reports in this regard can be expected in the near future.

Detailed analysis of cell walls of mature Brachypodium plants showed high similarity to those of Miscanthus, wheat and barley (Gomez and Bristow 2008; Matos et al. 2013). Also the primary cell wall of Brachypodium displays high similarity to that of wheat and barley (Christensen et al. 2010). Studying cell wall properties of Brachypodium can thus advance crop breeding in bioenergy feedstock (Vain 2011; Matos et al. 2013). However, the presence of a lower degree of substitution on arabinoxylan and a higher percentage of diferulic acid in Brachypodium, shows that probably other species-specific cell wall characteristics are present (Christensen et al. 2010).

Analogous to the naturally occurring *brown midrib* mutants in maize and Sorghum (Halpin et al. 1998a; Sattler et al. 2009), several mutants in the BRACHYTIL collection display a red coloration of the stem. These mutants have been named *brown stem* (*bs*) (Figure 28). Two of these mutants (*Bd4179* and *Bd7591*) have been demonstrated to carry mutations in the *BdCAD1* (BRADI3G06480) gene. They display significantly lower CAD activity and lower lignin content than the wild type, while their development or biomass production remains unaltered. Furthermore, these mutants have higher saccharification efficiency than wild-type plants confirming that CAD constitutes a good target for saccharification improvement in grasses. An analysis of the expression of the different *CAD* gene family members over development in *Bd4179* and control plant was performed in the scope of this thesis. These results were included in the paper and led to co-authorship in Bouvier d’Yvoire et al. (2012).

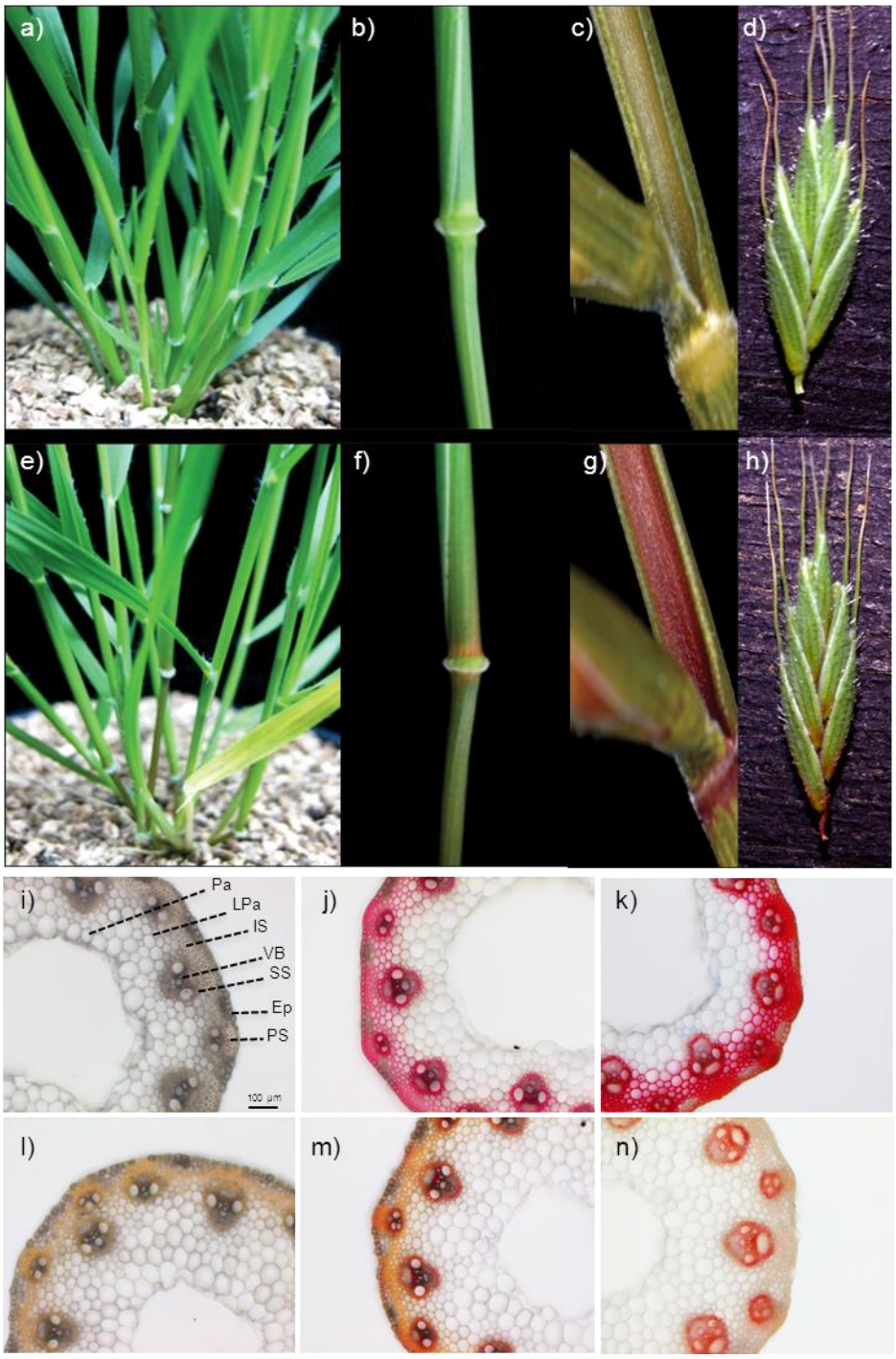


Figure 28. Phenotype of a *brown stem* mutant line (Bd4179) with reddish-brown colored stems compared with the wild-type line (WT). (a–h) Photographs of tillers (a, e), nodes (b, f), spike rachillas (c, g) and lemma (d, h). (i–n) Unstained transverse sections (30 µm) of the internode (i, l) and transverse sections stained using the Wiesner method (j, m) and the Maüle method (k, n). (a–d, i–k) wild-type plant; (e–h, l–n) Bd4179 line. Pa, parenchyma; Lpa, lignified parenchyma; IS, interfascicular sclerenchyma; VB, vascular bundle; Ep, epidermis; PS, polar sclerenchyma; SS, sheath sclerenchyma. From Bouvier d'Yvoire et al. 2013.

The involvement of *BdCAD1* in stem lignification was confirmed by Trabucco et al. (2013) using an amiRNA approach. In addition, a higher stem biomass and a two-fold increase in tiller number were reported in *BdCAD1* downregulated lines (Trabucco et al. 2013) highlighting the interconnection between different pathways. The downregulation of another lignin biosynthetic gene (BRADI3G16530), designated *BdCOMT4* in Trabucco et al. (2013) but identical to *BdCOMT6* in Dalmais et al. (2013), also resulted in reduction of lignin content. The laccase genes function in the polymerization of aromatic compounds such as monolignols, and their involvement in lignification was demonstrated recently (Berthet et al. 2011; Zhao et al. 2013). In Brachypodium, two laccases identified using the BRACHYTIL collection, have been shown to be involved in stem lignification. The respective mutants have significantly altered lignin content and composition (Wang et al. 2013; Dalmais et al. 2013).

Furthermore, Brachypodium has been used as study subject for identifying cell wall biosynthesis mechanisms that are common between monocot and dicots. Analysis of the phylogenetic relationships of CELLULOSE SYNTHASE A (CESA) genes that function in cellulose synthesis have shown that the CESA gene family members playing a key role in secondary cell wall biosynthesis have not expanded since the time of eudicot and monocot divergence 140–150 million years ago (Handakumbura et al. 2013). The downregulation of these genes in Brachypodium using amiRNA resulted in delayed flowering, reduced stature, reduced stem cross-section area and thinner cell walls, features that resemble phenotypes of Arabidopsis, rice and barley CESA mutants (Handakumbura et al. 2013).

Valdivia et al. (2013) presented clear evidence of the NAC transcription factor *SWN* being a master switch of secondary wall synthesis and cell death in Brachypodium, similarly to what has been demonstrated in Arabidopsis. The binding specificity of SWN proteins and therefore the motifs present in their target promoters appear to have been conserved, indicating that *SWN* genes already functioned as master switches of secondary cell-wall synthesis in the last common ancestor of monocots and dicots and that this role has been preserved in both lineages (Valdivia et al. 2013).

Specifically interesting for bioenergy grasses is the relation between cell wall quality and biomass yield. The overexpression of the MYB transcription factor *BdMYB48* which is involved in stem secondary cell wall biosynthesis results in increased aboveground biomass (Handakumbura and Hazen 2012). An increase in stem biomass and tiller production was reported for *BdCAD1* downregulated lines, perhaps related to delayed flowering (Trabucco et al. 2013).

3. Objectives

Here we present an evaluation of Brachypodium as study system for bioenergy-related traits such as biomass accumulation and saccharification efficiency, using available tools and resources. More specifically, we investigated whether overexpression of *GA20ox* in Brachypodium resulted in larger organ size, as previously demonstrated in Arabidopsis and maize. In addition, we studied the effect of RNAi downregulation of *BdCAD1* and mutation of *BdCAD1* and *Bd4CL1* on the saccharification efficiency. On the basis of the results and the experience gained, the benefits and hurdles of using Brachypodium as a model system for bioenergy crop improvement are discussed.

4. Results

4.1 Overexpression of *AtGA20ox1* in Brachypodium

For constitutive overexpression of GA20-oxidase, the *AtGA20OX1* gene (At4g25420) was cloned under control of the maize UBIL promoter in the vector pBbm42GW7 (http://gateway.psb.ugent.be/), the same construct that was successfully used for overexpression of GA20ox in maize (Figure 29). For Brachypodium transformation we used a protocol based on Vogel et al. (2008) and Alves et al. (2009) but as the construct had the BAR gene as marker, we used the herbicide basta as selective agent instead of the antibiotic hygromycin. Selection and regeneration of transgenic plants using basta was shown previously to be effective (Păcurar et al. 2008; Bragg et al. 2012). However, we observed that shoot regeneration on selective medium was completely prevented by the applied basta concentrations (5 mg/l medium, similar to Păcurar et al. 2008). In contrast, in control calli grown on non-selective medium a rapid increase of callus size and shoot regeneration was observed.

This indicates that the presence of the construct in transgenic calli was not able to induce basta resistance. We tested the presence and functionality of the PAT protein in leaves of Brachypodium plants transformed with pXBb7FNFI-UBIL (Supplementary figure 5). These transgenic plants express GFP under the control of the UBIL promoter and light up green under blue light. More importantly, pBbm42GW7 and pXBb7FNFI-UBIL vectors carry the same CaMVp35S-BAR resistance cassette. A basta leaf painting assay (Yao et al. 2006) was conducted for which solutions with various basta concentrations were applied on leaf tips of Brachypodium plants transformed with pXBb7FNFI-UBIL. Unexpectedly, no resistance to the basta herbicide was observed and leaf tips senesced similar to control plants. Next, an immunochromatographic assay was used to detect the PAT protein (AgroStrip, Romer) in leaves of Brachypodium plants transformed with pXBb7FNFI-UBIL. No PAT protein could be detected, explaining why transgenic plants and callus were not resistant to basta application.

In order to identify transgenic plants, the regenerated shoots from the calli that were co-cultivated with Agrobacterium harboring *GA20ox* expression vector but grown on non-selective medium were collected. In total, 10,000 T_0 regenerants were transferred to the greenhouse. Leaf samples were used for DNA extraction and were tested by PCR for the presence of the CaMVp35S promoter. In total, 19 independently transformed T_0 plants were identified.

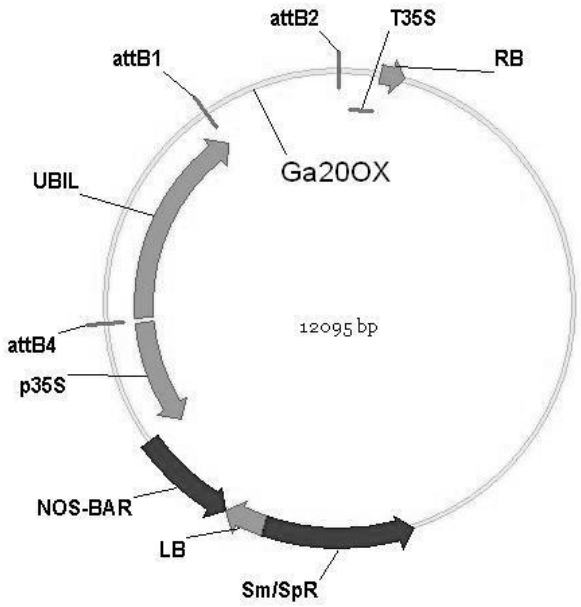


Figure 29. Expression vector used for GA20ox overexpression in Brachypodium. The construct was made by exchanging the ccdB gene in pBbm42GW7 for the maize UBIL promoter and the GA20ox cDNA sequence in a multisite gateway LR reaction. The pBbm42GW7 vector is available at http://gateway.psb.ugent.be/; RB: right border, UBIL: maize ubiquitin promoter, attB1 and attB2: recognition sites for site-specific recombination resulting from a recombination reaction, T35S: CaMV 35S terminator, p35S: CaMV 35S promoter, Bar: basta resistance gene, Tnos: nopaline synthesis terminator, LB: left broder, Sm/SpR: spectitomycin resistance gene

The segregating populations of five independently transformed lines were tested for transgene expression levels and were evaluated phenotypically. The expression level of *AtGA20ox1* varied for individual plants from the segregating population as hemizygous and homozygous plants were not distinguished (Figure 30). *AtGA20ox1* expression was the highest for transgenic plants of lines 874 and 548, intermediate for 471 and 497, and the lowest for 370 plants (Figure 30). As the overexpression of this gene caused longer leaves and increase in plant height in maize and in the dicot models Arabidopsis and tobacco (Coles et al. 1999; Biemelt et al. 2004; Nelissen et al. 2012), here the leaf length was measured over time as a phenotypic screening and analyzed using LEAF-E. No significant differences were detected in *LERmax* or *Lm* when the investigated transgenic plants were compared with the respective NT controls. For illustration, leaf length and leaf elongation rates of plants of line 471 are shown in Figure 31 which were representative for all five investigated lines.

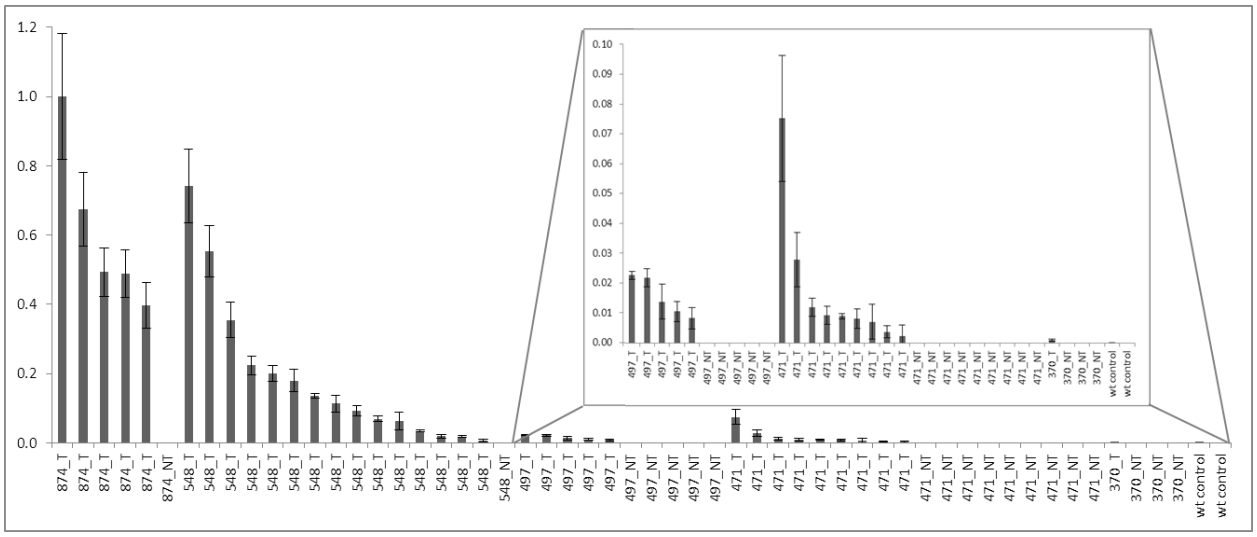


Figure 30. Expression of AtGA20ox in Brachypodium transformants. The different numbers indicate independent transgenic lines and expression values of multiple plants from the segregating population of each line was shown, both transgenic (T; heterozygous and homozygous transgenic) and non-transgenic plants (NT; azygous control) respectively. Error bars represent standard errors over three technical repeats.

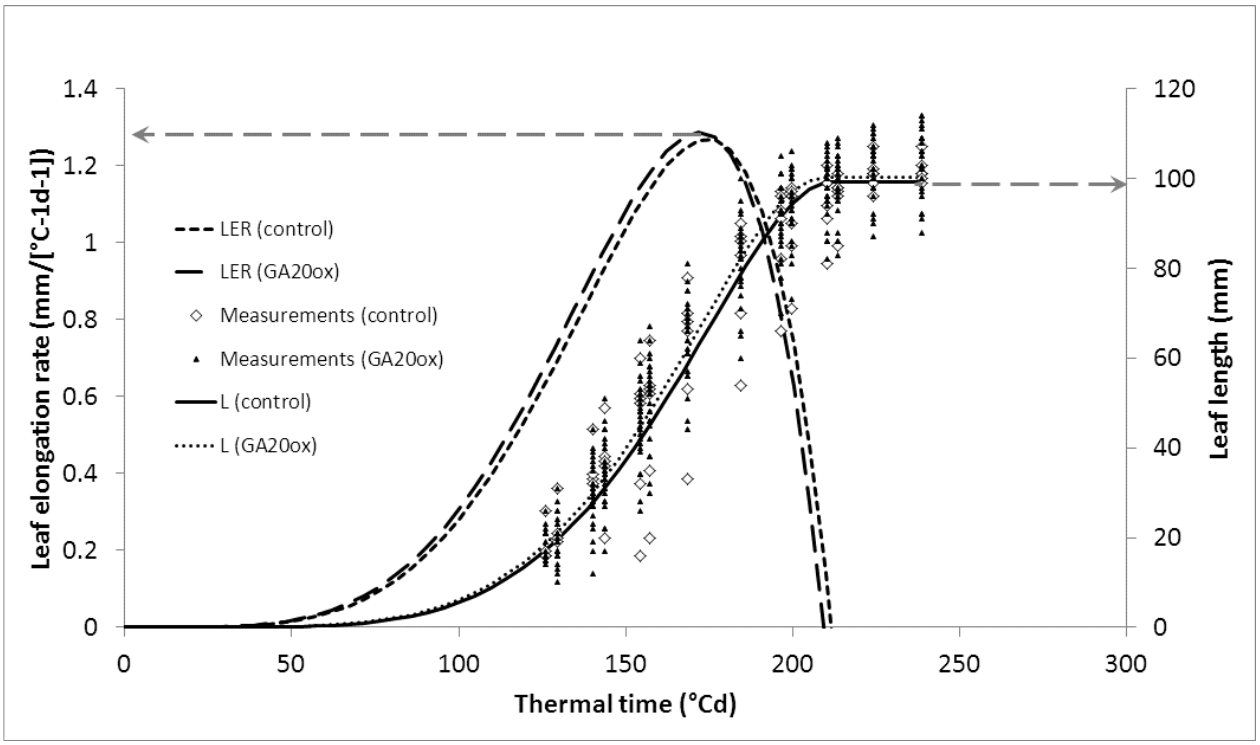


Figure 31. Leaf (leaf#3) growth analysis of the segregating population of the 471 line for *GA20ox* overexpression using LEAF-E. Measurements were taken from 28 transgenic (GA20ox) and 6 control plants. Sigmoid curves show the leaf length curves (right axis) and bell shaped curves show the leaf elongation rate (LER, left axis).

4.2 Targeting the *BdCAD1* gene in Brachypodium

4.2.1 Phylogeny and expression patterns of the CAD family in Brachypodium

In flowering plants, CAD is encoded by a small gene family comprising nine members in Arabidopsis (Raes et al. 2003), twelve members in rice (Zhang et al. 2006) and seven members in maize (Guillaumie et al. 2007). To select a good candidate gene for lignin perturbation in the Brachypodium stem, a phylogenetic analysis and an expression analysis was performed. Based upon protein sequences of CAD family members of Brachypodium, maize, rice and Arabidopsis, a phylogenetic tree constructed using the CLC Genomics workbench (CLC bio, Aarhus, Denmark). Like in maize, the CAD family in Brachypodium consists of seven members. One clade in the phylogenetic tree contained the protein sequences that were identified as the main CAD in stem lignification: ZmCAD2 in maize (Fornalé et al. 2012), OsCAD2 in rice (Hirano et al. 2013) and AtCAD4 and AtCAD5 in Arabidopsis (Kim et al. 2004; Sibout et al. 2005) (Figure 32). Among them is BRADI3G06480 or BdCAD1, now identified as the main CAD in Brachypodium, based on this phylogenetic analysis.

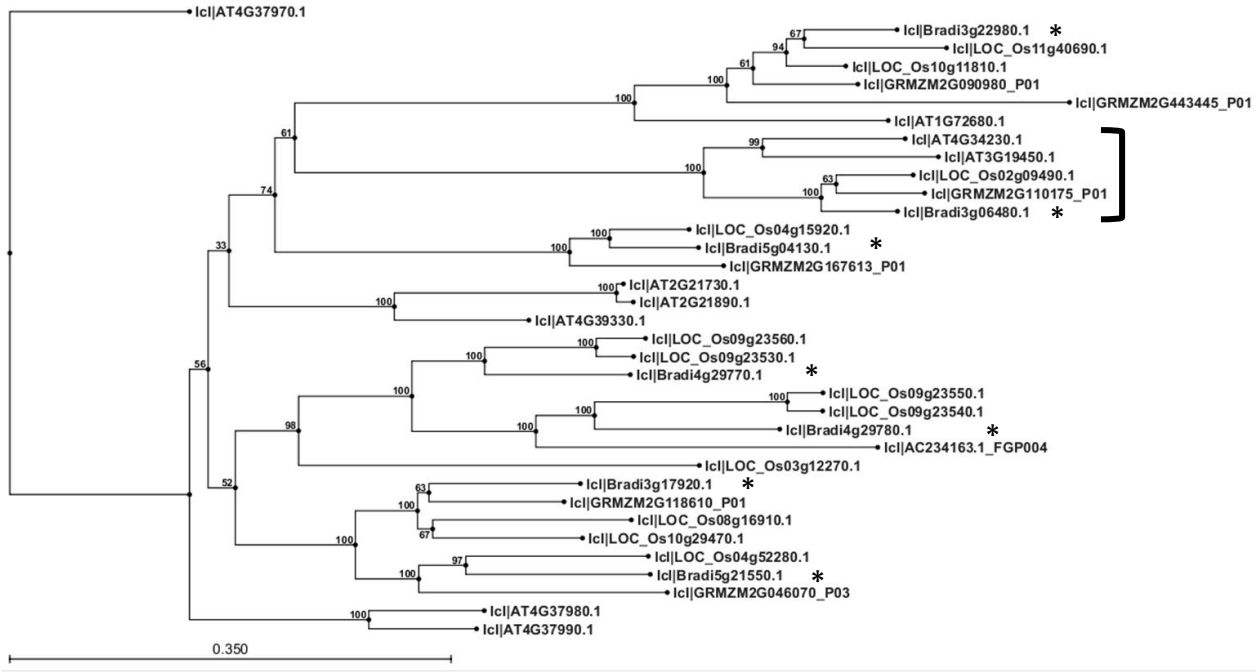


Figure 32. Phylogenetic relationship between CAD family protein sequences of Brachypodium, maize, rice and Arabidopsis. Bracket indicates sequences of the maize, rice and Brachypodium main CAD proteins, ZmCAD2 (Fornalé et al. 2012)**, OsCAD2** (Tobias and Chow 2005; Zhang et al. 2006) **and the main CAD in Brachypodium: BdCAD1. The closest related Arabidopsis CAD members are AtCAD4 and AtCAD5 (encoded by At3G19450 and AT4G34230 respectively), which are the main CAD enzymes in Arabidopsis** (Kim et al. 2004; Sibout et al. 2005)**. The phylogenetic tree was constructed using the CLC Genomics workbench (CLC bio, Aarhus, Denmark).**

The expression of all seven genes in the CAD family of Brachypodium was investigated at four points in stem development (Figure 33). Expression levels were measured at five, seven, nine and twelve weeks after germination corresponding to early vegetative (5w), late vegetative-transition (7w), heading (9w) and fully mature stage (12w). A general trend of rising expression levels from the early vegetative stage (5w) towards late vegetative-transition stage (7w) and then again decreasing expression towards the mature stage (9w to 12w) could be observed for all *CAD* genes except

BdCAD5 and *BdCAD6* (Figure 33). This is consistent with the major onset of lignin and secondary cell wall formation (Matos et al. 2013). The increase in expression from early to late vegetative-transition stage in stems was the strongest for *BdCAD1*. The transcript abundance of *BdCAD1* at these four stages in Brachypodium development was included in Bouvier d'Yvoire et al. (2013). The methodology used, qRT-PCR without dilution series of the cloned genes, does not allow direct comparison of expression levels between the different CAD genes. However, the relative expression levels of different CAD members obtained in our analysis were confirmed in more recent work by Trabucco et al. (2013) using a tiling array. We found that the *BdCAD1* gene was by far most abundantly expressed in stems over development. In comparison with the next most expressed *CAD* gene (*BdCAD4*), *BdCAD1* was expressed 11 fold higher at 5w, 19 fold higher at 7w, 27 fold higher at 9w and 29 fold higher at 12w (Figure 33). Fold changes of *BdCAD1* over the least expressed CAD gene (*BdCAD2*), were 115 fold 212 fold, 155 fold and 170 fold in 5w, 7w, 9w and 12w respectively.

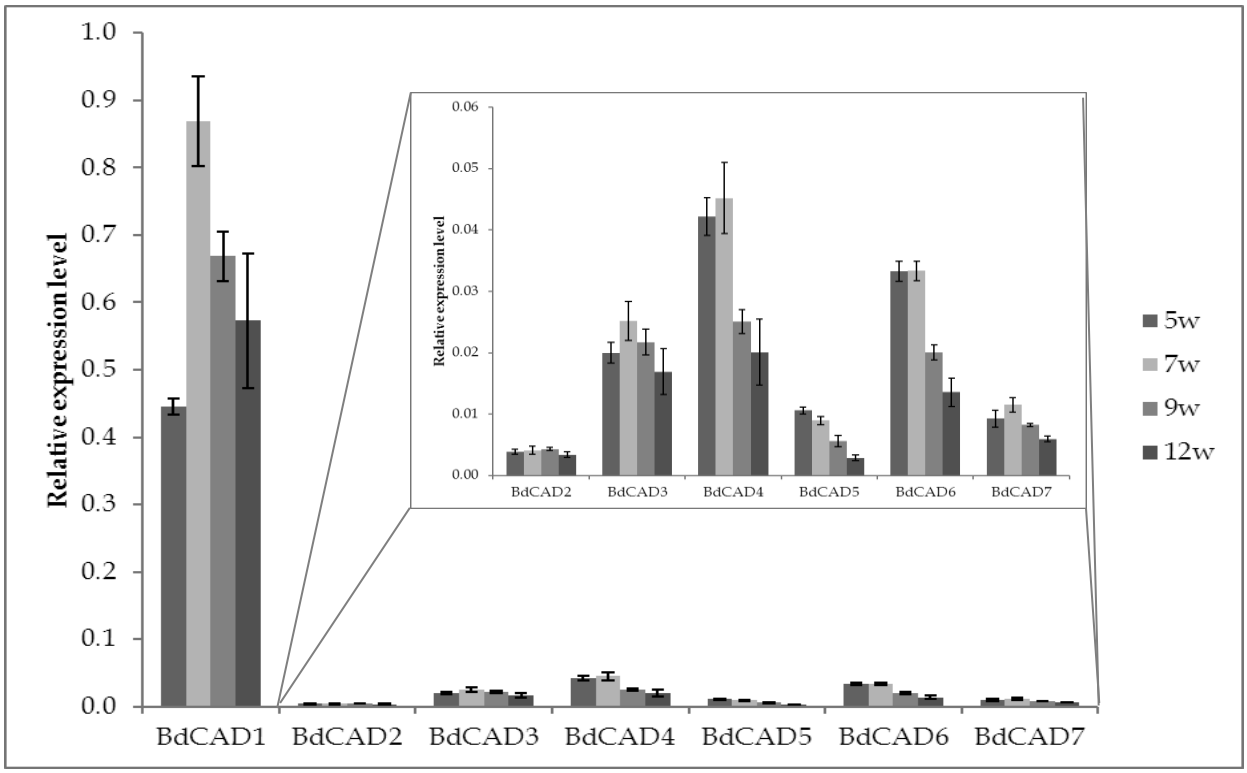


Figure 33. Expression levels of the seven CAD family members Brachypodium stems in four developmental stages: 5, 7, 9 and 12 weeks after germination. These time points in development represent early vegetative stage (5w), late vegetative-transition (7w) heading (9w) and mature stage (12w). Error bars represent standard errors over three biological repeats

4.2.2 Expression analysis of the CAD gene family in Bd4179, a TILLING mutant for BdCAD1

We also investigated eventual differential expression of the different *BdCAD* members in response to *BdCAD1* disruption. Gene expression was monitored in the *Bd4179* mutant, identified by INRA Versailles, in two developmental stages (7w and 9w) and was compared with the segregating wildtype. At 7w, corresponding to late vegetative stage, a two-fold reduction of *BdCAD1* expression was observed in the *Bd4179* mutant compared with the control plants (Figure 34). All other family members either showed a very mild reduction (*BdCAD6*) or showed similar expression levels as control plants. At 9w, plants had just started to flower which is referred to as the major onset of secondary cell wall formation and lignification in the Brachypodium stem (Matos et al. 2013). At this point, there is a clear upregulation in *Bd4179* as compared with control plants of *BdCAD1* (2.7 fold), *BdCAD2* (2.6 fold) and to a lesser extent *BdCAD3* (1.7 fold) transcripts (Figure 34).

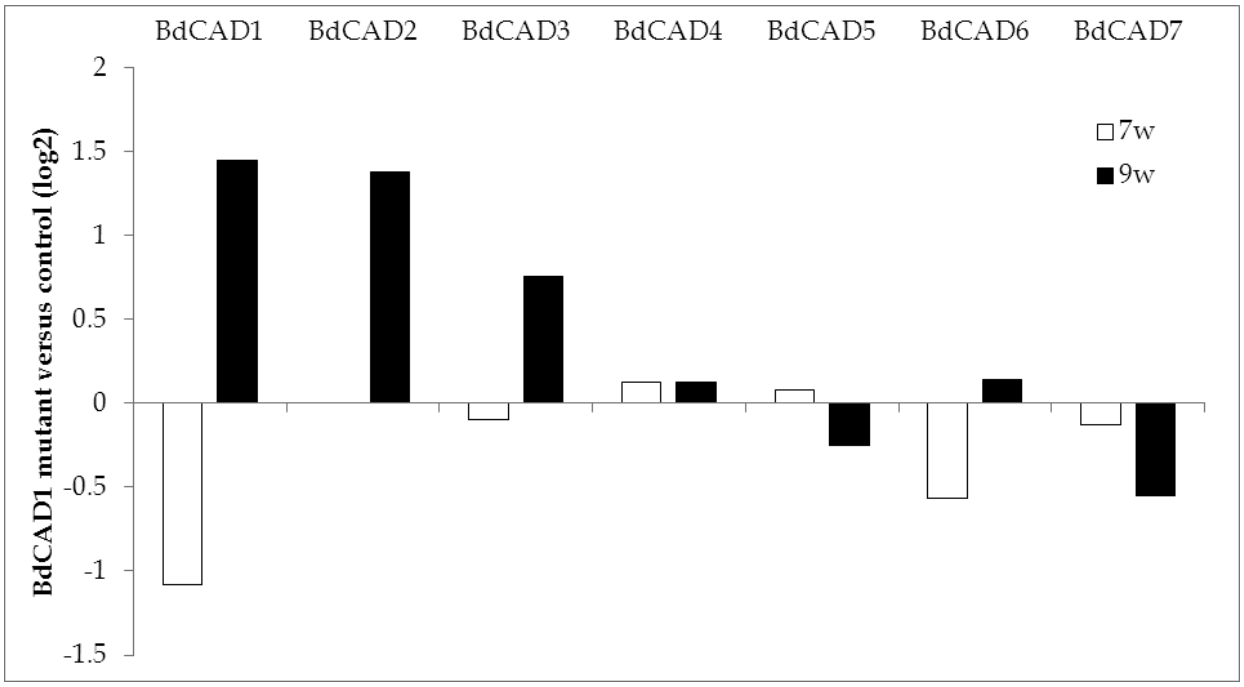


Figure 34. Change (log2 fold) in expression of CAD family members in the *Bd4179* mutant versus control plants in two different stages of development: late vegetative-transition (7w) and heading(9w).

4.2.3 Downregulation of the BdCAD1 gene in Brachypodium using RNAi

Based upon the phylogenetic and expression analysis presented above and the analysis of TILLING mutants in *BdCAD1* by our collaborators at INRA Versailles, the *BdCAD1* gene formed the best target in the CAD gene family for improving saccharification efficiency. Within the frame of this PhD dissertation, we also explored a transgenic approach for improving saccharification efficiency by downregulation of the *BdCAD1* gene using RNAi.

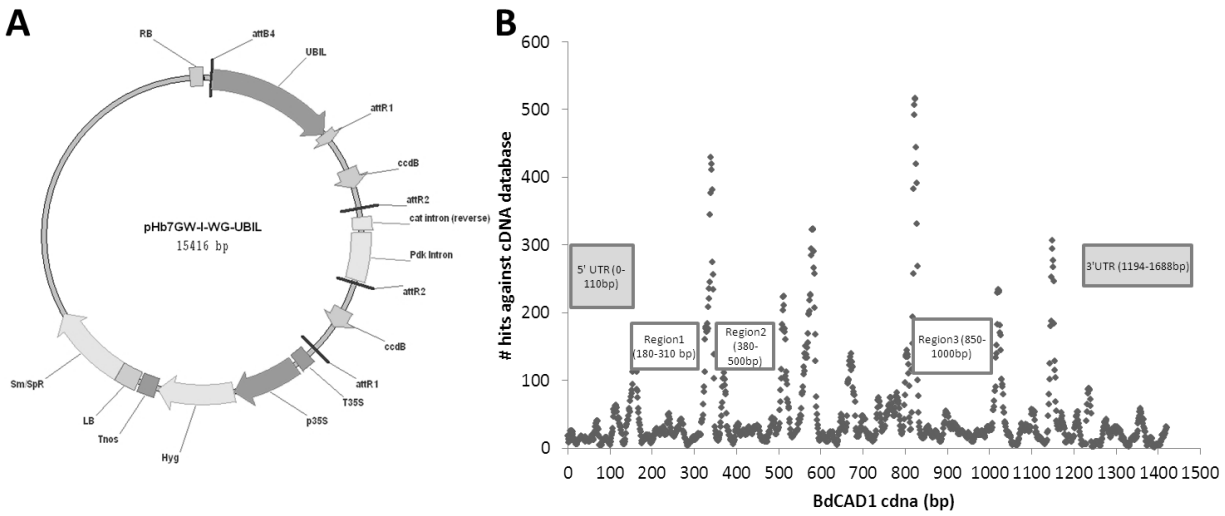


Figure 35. Monocot-specific expression vector (A) and the selection of three regions in the *BdCAD1* cDNA sequence (B) for RNAi-mediated downregulation of the *BdCAD1* gene in Brachypodium. The hairpin vector is available at http://gateway.psb.ugent.be/; RB: right border, UBIL: maize ubiquitin promoter, attR1 abd attR2: recognition sites for site-specific recombination with the target region of the gene of interest, T35S: CaMV 35S terminator, p35S: CaMV 35S promoter, Hyg: Hygromycin resistance gene, Tnos: nopaline synthesis terminator, LB: left broder, Sm/SpR: spectitomycin resistance gene. The graph in B shows the blast output of all possible 21-mers in the *BdCAD1* cDNA against the Brachypodium cDNA database. The number of hits from the blast output is shown and does not indicate necessarily perfect matches. Three regions of 130 bp, 120 bp and 150 bp long were designated to be good targets for an RNAi approach.

For the downregulation of *BdCAD1*, a specific region was chosen in order to reduce the amount of possible off-targets with the RNAi approach. A BLAST search was used to compare all possible 21-mers of the cDNA sequence of *BdCAD1* against the cDNA database of Brachypodium (Figure 35). Based on the number of BLAST hits, three regions with high specificity for BdCAD1 were designated as good targets for hairpin design. Because of the low efficiency that we could achieve, only the RNAi construct targeting region1 was used. A total of 33 independent transgenic lines were moved from 'in vitro' to the greenhouse. A total of seven lines were selected based on seed yield and expected three-to-one ratio of transgenic versus non-transgenic plants in the segregating population indicative for single locus of transgene insertion. Expression levels of the *BdCAD1* for the selected lines are shown in Figure 36A and B respectively. The expression of the hairpin was estimated by qRT-PCR using a set of primers specifically targeting the hairpin which is formed from the transgene upon transcription. The forward primer was designed in the cloned region1 of BdCAD1, and the reverse primer in the intron sequence of the pHb7GW-I-WG-UBIL hairpin vector (Figure 35A). The expression of the hairpin in the investigated lines could be classified from highest to lowest: 22_4 > 4_1 > 30_4 > 33_1 > 8_1 > 13_11 > 37_1 (Figure 36A). More specifically, expression of the hairpin in was more than threefold higher than in 4_1 and more than six-fold higher than in 30_4 and 33_1. The high expression of the hairpin in 22_4 and 30_4 correlated with a modest, but significant downregulation of the target gene *BdCAD1* (18% and 25% respectively, Figure 36B). However, no significant changes in *BdCAD1* gene expression were found for 4_1, 33_1, 8_1, 13_11 and 37_1. A CAD activity assay was used to determine whether the reduction in *BdCAD1* transcript abundance also caused reduced CAD activity, indicative of successful downregulation as was shown by Bouvier d'Yvoire et al. (2012). The CAD activity assay showed a significant reduction for lines 33_1, 22_4 and

8_1 (43%, 35% and 26% respectively) (Figure 37). Among the investigated lines, line 22-4 would be interesting for further investigation based on the expression analysis and CAD activity assay but this will be the topic of future research.

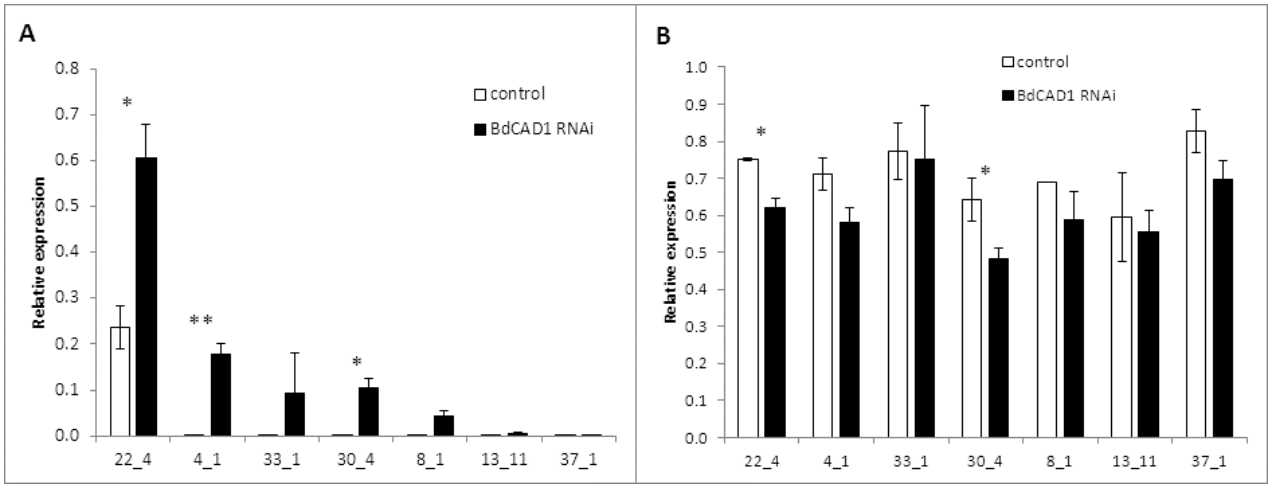


Figure 36. Expression of the hairpin (A) and BdCAD1 expression (B) in transgenic lines of Brachypodium downregulated for BdCAD1. Expression was determined using qRT-PCR in whole-plants samples of plants that just started flowering. Error bars represent standard errors over biological replicates. The number of biological replicates per line can be read between brackets per line as follows "(# biol replicates for control, # boil replicates for BdCAD1 RNAi)": 22_4 (2,8), 4_1 (3,5), 33_1 (4,2), 30_4 (2,6), 8_1 (1,9), 13_11 (2,6) and 37_1 (2,7). *: p<0.05; **: p<0.01

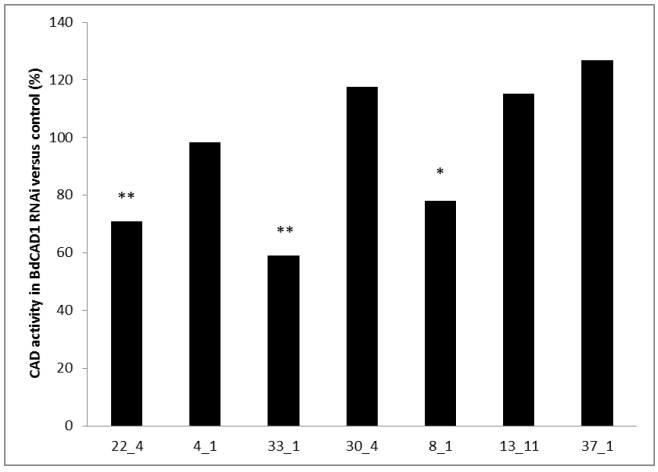


Figure 37. CAD activity assay in seven transgenic lines downregulated for *BdCAD1*. CAD activity in transgenics is expressed as percentage of controls. The same material was used as for the expression analysis in Figure 36. Three biological replicates were used for BdCAD1 RNAi and control for each line. *: p<0.05; **: p<0.01

4.3 TILLING for mutants in *Bd4CL1*

Similar to CAD, also 4CL is encoded by a small gene family in flowering plants comprising four members in Arabidopsis (Raes et al. 2003). In maize, the main 4CL involved in stem lignification is *Zm4CL1*, encoded by the GRMZM2G075333 (refgen v2) gene (Riboulet et al. 2009). A phylogenetic analysis using 4CL protein sequences from Brachypodium, maize, rice, sorghum, Arabidopsis and Medicago showed that the 4CL family in Brachypodium comprises five members (*BRADI3G37300*, *BRADI3G18960*, *BRADI3G05750*, *BRADI1G31320* and *BRADI3G52350*) and indicated *BRADI3G05750* as the orthologous of *Zm4CL1* (Figure 38). These findings were supported by comparative genomics tools such as Plaza 2.5 (Van Bel et al. 2012) and EnsemblPlants (http://plants.ensembl.org/index.html). Exploration of expression profiles in the online database of gene expression data of Brachypodium, Bradinet (http://aranet.mpimp-golm.mpg.de/bradinet), the *BRADI3G05750* gene is indeed highest expressed among the five 4CL family members in growing stems. Moreover, our expression analysis in Brachypodium stems using qRT-PCR confirmed that *BRADI3G05750* is indeed the highest expressed (Supplementary figure 2). Therefore, *BRADI3G05750* can be considered as the best candidate for functional analysis using a TILLING approach for 4CL in Brachypodium. We were the first to name this gene *Bd4CL1*.

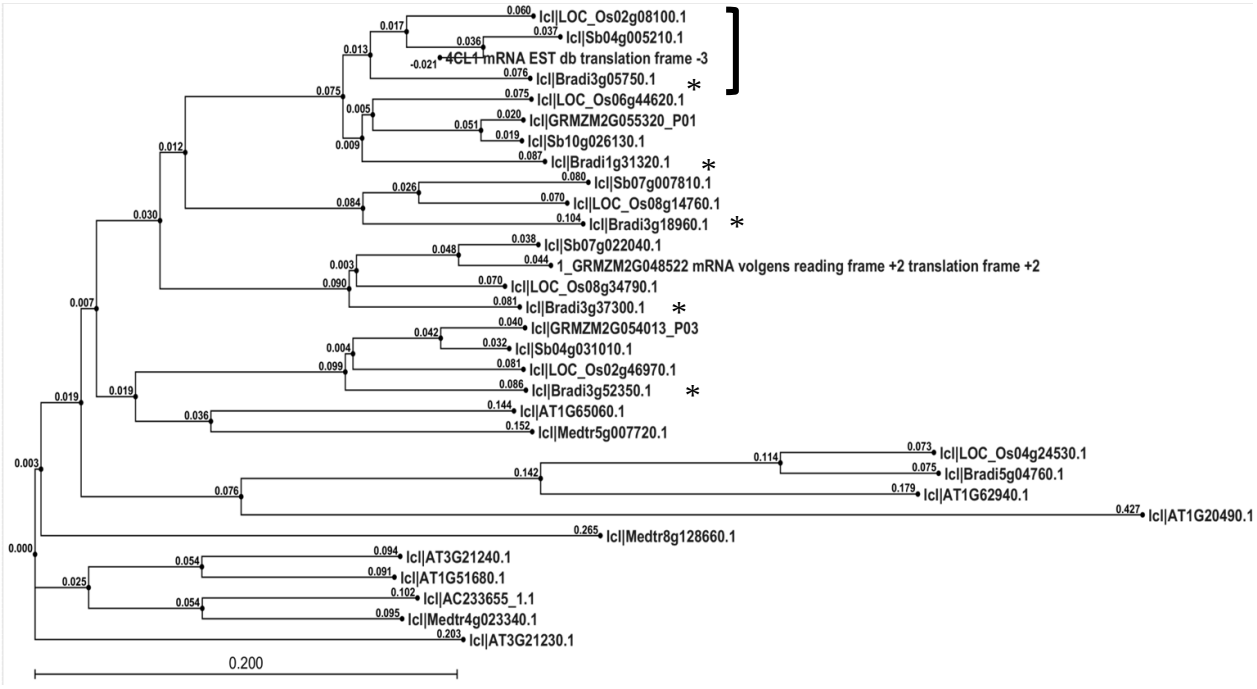


Figure 38. Phylogenetic relationship between 4CL family protein sequences in Brachypodium, maize, rice, Sorghum, Arabidopsis and *Medicago*. Bracket indicates sequences of the rice, maize, sorghum and Brachypodium *4CL1* genes. The closest Arabidopsis ortholog is *At4CL3* (AT1G65060). The phylogenetic tree was constructed using the CLC Genomics workbench (CLC bio, Aarhus, Denmark).

Next, the BRACHYTIL collection was searched for mutations in the *Bd4CL1* gene. A total of 17 mutations were found in a 735 kb region in the first exon of *Bd4CL1* (Table 5). The region included the AMP-binding domain, which is the protein's functional domain (Hu et al. 2010). However, no mutations were found within the 33 bp sequence encoding this domain. Of the 17 mutations found in the neighboring region, 13 were non-synonymous and 4 were synonymous or 'silent' (Table 5,

Figure 39). No premature stop codon mutation was found. To identify the best candidates for loss of function or reduced activity of the 4CL protein, the free software SIFT was used to determine the severity of the expected effect of specific amino acid changes on the protein activity, based upon conservation of the particular amino acid in other plant species (http://sift.jcvi.org/www/SIFT_seq_submit2.html). We could point out six mutations, corresponding to seven lines or plant families, possibly leading towards affected protein function (Table 5). In monolignol biosynthesis, 4CL acts in the general phenylpropanoid pathway, before the split towards the biosynthesis of the different lignin subunits (*p*-hydroxyphenyl, guaiacyl and syringyl units). Therefore, mutations affecting the protein function were expected to result in a reduced total lignin amount, rather than in compositional changes of the lignin. As reduction of the lignin content is highly correlated to improved saccharification efficiency (Van Acker et al. 2013), a saccharification assay was used to phenotype the seven lines mutated at *Bd4CL1* and for which a change in the protein function was predicted. However, no significant effects could be found for any of the lines when compared to control plants (Supplementary figure 3). To confirm this result, the total lignin content was measured on three of these lines (5287, 7413 and 5538). Also in this case no significant changes were found between mutants and control (Supplementary figure 4).

Table 5. Identified mutations in a 735 kb region in the first exon of *Bd4CL1*. The position of the mutations was indicated as well as which nucleotide/amino acid was present in wildtype (in front) and which in the mutant (behind). * homozygous for the mutation; [1] starting from the beginning of the coding sequence; [2] free software tool that predicts whether a certain mutation would affect the protein function or not based on the conservation of the particular nucleotide in the plant kingdom (http://sift.jcvi.org/ww/SIFT_seq_submit2.html)

Mutation	Base position[1]	Protein position	Type of mutation	Plant family	SIFT software prediction[2]
1	G154T	G52W	missense	6141	Affect protein function
2	G164A	G55E	missense	4113	Tolerated
3	G164T	G55V	missense	7413	Affect protein function
4	G164T	G55V	missense	5287	Affect protein function
6	G214A	A72T	missense	6041	Affect protein function
5	C264T	L88L	silent	3684	-
7	C273T	N91N	silent	3586	-
8	C332T	T111I	missense	5538	Affect protein function
9	G386A	G129E	missense	6280	Tolerated
10	G387A	G129G	silent	4551	-
11	C399T	V133V	silent	6227	-
12	G421A	E141K	missense	4056	Tolerated
13	G427A	V143I	missense	4051	Tolerated
14	G431A	R144L	missense	3474	Affect protein function
15	G433A	E145L	missense	6809*	Tolerated
16	G476T	G159V	missense	6523	Affect protein function
17	G505A	E169L	missense	8287	Tolerated

Bradi3g05750.1 Proteic sequence :
MGSVPEESPAAGEETVFRSRLPDIEIPSEQTLQSYCFAKMAEVGSRPCLIDGQTGESYTYSEVESLTRRAAAGLRRMGVGKGDVVMNLLRNCPEFAFSFLG
AARLGAATTTANPFYTPHEIHRQAEAAGAKLVVTEACAVEKVREFAAGKGIPVVTVDGRFDGCVEFEELIGGEEEMDEGEIHPDDVVALPYSSGTTGLPKG
VMLTHRSLITSVAQQVDGENPNLYFSKEDVVLCLLPLFHIYSLNSVLLAGLRAGSAIVIMRKFDIGALVDLVRAHGVTVAPFVPPIVVEIAKSDRVSAADLASIR
MVMSGAAPMGKELQDAFMAKIPNAVLGQGYGMTEAGPVLAMCLAFAKEPFKVKSGSCGTVVRNAELKIVDPDTGASLARNQPGEICIRGEQIMKGYLN
DPESTKNTIDKDGWLHTGDIGFVDDDDEIFIVDRLKEIIKYKGFQVAPAELEALLITHPEIKEAAVVSLKDDLTGEIPVAFVKRIDGSEITEAEIKQFVAKEVVFYK
RIHKVFFTDSIPKSPSGKILRKDLRARLAAGVPSDDTAPRS*

Figure 39. Location of the affected amino acids in the protein sequence of *Bd4CL1*. The AMP binding domain is highlighted in dark green.

5. Discussion

Since its introduction as model plant for temperate grasses and bioenergy crops in 2001, the number of publications related to Brachypodium has increased exponentially (Figure 40), a trajectory that is similar to that of Arabidopsis in the early years (Brkljacic et al. 2011). Mainly, the announcement that the genome would be sequenced by JGI in 2006, stimulated the initiation of Brachypodium-related research projects. In the first ten years of its existence as model plant, papers on Brachypodium had as main topics transformation protocols, establishing community resources, exploiting natural variation for disease resistance and reviews promoting Brachypodium as a model for cereals and bioenergy crops. Thanks to the high quality genome sequence, Brachypodium is used as an efficient tool in comparative genomics. However, only recently, successful overexpression or downregulation of genes of interest were reported (Vain et al. 2011; Handakumbura et al. 2013; Trabucco et al. 2013; Bouvier d'Yvoire et al. 2013; Valdivia et al. 2013). Five years after the first protocol was published for Agrobacterium–mediated transformation of callus derived from immature embryos, still a paper reporting on a transformation protocol and its efficiency was published (Lee et al. 2011). This supports the idea that, despite the documented high transformation efficiency, optimization of a transformation platform in the regular research lab is time-consuming and perhaps not straight-forward.

When this PhD project started, the established Brachypodium transformation platform was inefficient and labor-intensive. Therefore, we could functionally characterize only a limited number of genes using a transgenic approach. Attempts to improve the transformation and regeneration efficiency by addition of $CuSO_4$ to the callus inducing medium (Alves et al. 2009), eliminating the cocultivation without selective agent, altering Hygromycin levels in the selective medium or adding an intron before the HPT gene in the pXHb7FNFI-UBI vector (as in pVEC8-GFP Alves et al., 2009) to reduce the number of escapes did not result in significant efficiency gains. Nevertheless, the transformation platform was used for functional characterization of target genes, yet for fewer genes than initially lined out. Transformations with sufficient number of transgenic events allowing phenotypic characterization were successful for overexpression of *AtGA20ox1* and downregulation of *BdCAD1* using RNAi. Successful overexpression of *AtGA20ox1* was accomplished, however not leading towards a detectable phenotype. This is remarkable and unexpected as overexpression of the same construct in maize triggered such a striking phenotype (see chapter 7 and Nelissen et al., 2012). It is thus not likely that this strategy would not be functional in Brachypodium. Most likely the level of overexpression was insufficient for phenotype detection. At least in Arabidopsis, a very

high overexpression of AtGA20ox1 was necessary in order to obtain the phenotype of enlarged plant and organ size as described by Coles et al. (1999) (personal communication within PSB). Ensuring a higher overexpression e.g. by using stronger promoters or screening more transgenic lines to identify transgenics with a higher level of overexpression could form a solution.

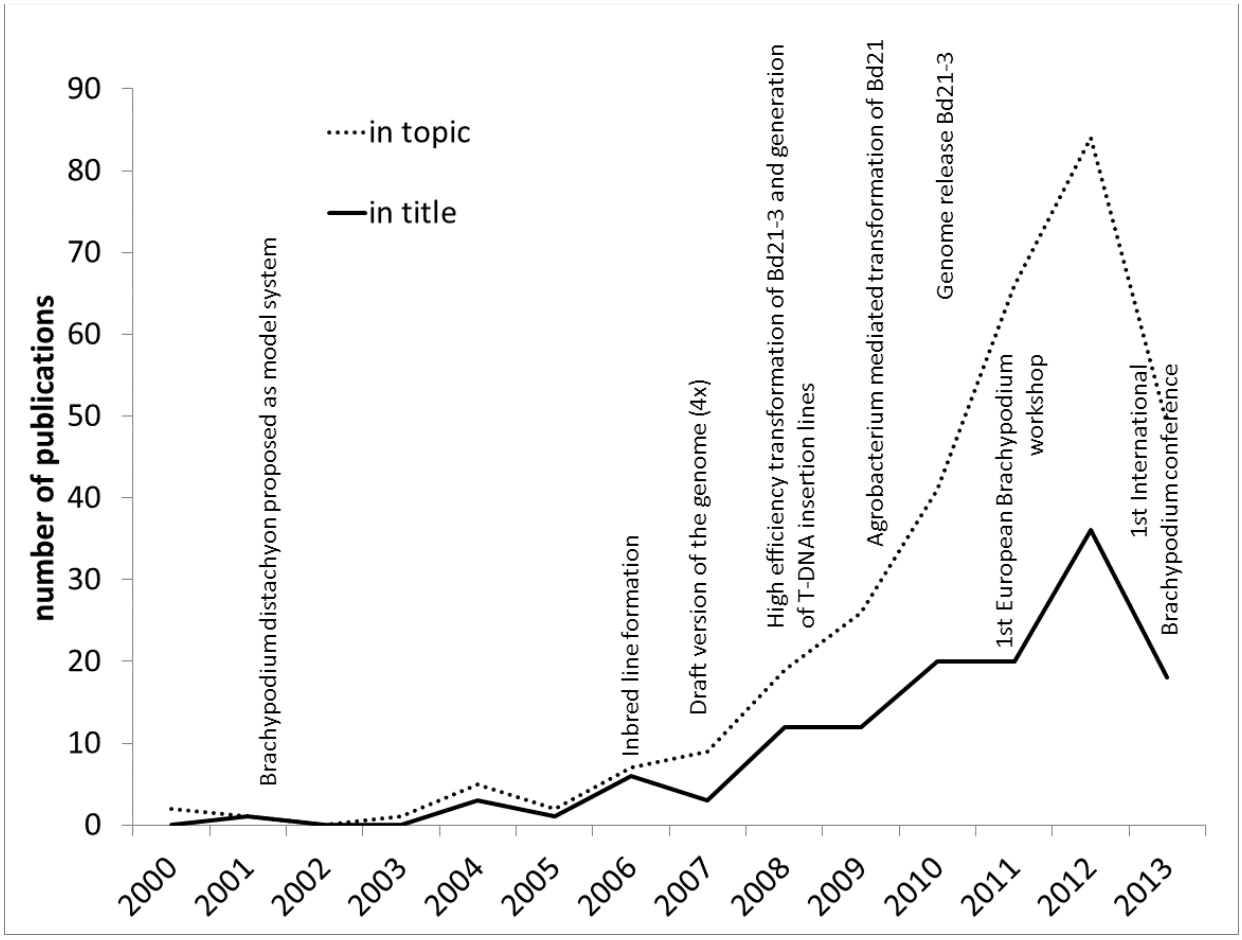


Figure 40. Number of Brachypodium-related publications on 'Web of knowledge' available on October 25, 2013, using the keyword "Brachypodium" in either the topic or the title.

A remarkable fact is the observation that no transformants could be regenerated on selective medium containing basta. The transgenic lines that were retained in this experiment, 19 independent events, were obtained from Agrobacterium cocultivated calli that were cultured on non-selective medium. Yet seven times more calli were cultured on selective medium, supplied with 5 mg/ml basta (as in Păcurar et al. 2008), than on non-selective medium without regenerating any shoot. This indicates that the presence of the construct in transgenic calli was not able to confer basta resistance. Possible reasons could be poor expression of the *BAR* gene, driven by the CaMVp35S promoter or non-functional PAT protein. The functionality and presence of the PAT protein was tested in transgenic Brachypodium plants. No resistance to the basta herbicide was observed using a basta leaf painting assay (Yao et al. 2006) and no PAT protein could be detected using an immunochromatographic assay designed specifically to detect the PAT protein in plant tissues (AgroStrip, Romer). The absence of PAT protein in transgenic plants and calli would explain

the absence of resistance to basta. Yet, selection using basta and the BAR gene as selectable marker has been successfully applied previously (Păcurar et al. 2008). The answer might lie in the design of the resistance cassette. Possibly, the expression of CaMVp35S-BAR that was used in our vector is low in Brachypodium due to non-optimized promoter enhancer sequences and/or absence of intron or other specific sequences in the 5' UTR of the BAR gene to enhance expression in Brachypodium and related cool season cereals.

Successful expression of the hairpin construct targeting the *BdCAD1* gene was accomplished. To my knowledge, this is also the first report that demonstrated the presence and expression level of the hairpin construct using qRT-PCR. For this, a forward primer was designed in the sense oriented cloned region1 of the BdCAD1 gene, inserted in the hairpin vector by recombination, and a reverse primer in the intron sequence, that is used as spacer between the sense and antisense oriented *BdCAD1* region1. Significant downregulation of *BdCAD1* with RNAi could be demonstrated for two out of seven investigated transgenic lines by qRT-PCR. Moreover, CAD activity was significantly reduced in three out of seven investigated lines. However, a correlation between high hairpin expression, *BdCAD1* downregulation and CAD activity reduction was only found for one out of the seven lines: 22_4. This line is a good candidate for phenotypic analysis. Successful downregulation of *BdCAD1* was also accomplished by Trabucco et al. (2013) and resulted in improved saccharification efficiency. The reduction of *BdCAD1* expression levels were stronger than obtained in this work (-55% and 33% for two lines in Trabucco et al., 2013 versus -25% and -18% for the significantly downregulated lines in this work). On the other hand, total CAD activity was only slightly reduced in Trabucco et al. (2013) (-6% and -17%) whereas in this work higher reduction of CAD activity was detected (-43%, -35% and -26%). However, the downregulated lines of Trabucco et al. (2013) showed a brown-reddish coloration of the stem, as was reported for *BdCAD1* TILLING mutants (Bouvier d'Yvoire et al., 2012). The brown-reddish coloration was not observed in our transgenic lines, perhaps resulting from the different construct used for downregulation. Trabucco et al. (2013) used amiRNA instead of a hairpin. Despite similar features such as UBIL promoter driving the transgene and Hygromycin resistance gene driven by the CaMVp35S promoter, the vector backbone is highly different. They used the pOL001 vector (http://brachypodium.pw.usda.gov/files/pOL001_sequence.pdf), which had been used previously in Brachypodium transformation at the Vogel's lab (Vogel and Garvin 2006; Vogel and Hill 2008; Handakumbura et al. 2013). Besides an intron in the HPT gene and a CaMVp35S promoter sequence that was not identical to the one used in our work, additional differences could probably positively influence transgene expression.

An alternative strategy for cell wall research in Brachypodium is the generation and identification of lignin mutants using a TILLING approach (Dalmais et al., 2013). We were successful in finding a total of 17 mutations in a 735 kb region of the first exon of *Bd4CL1*, by screening the BRACHYTIL population containing 5530 sodium aide (NaN_3) mutated families. The data of the identified mutations was incorporated into a publication describing the BRACHYTIL platform in Dalmais et al. (2013). However, no lignin-related phenotype could be detected. There may be three possible explanations for not finding a phenotype: (1) the mutations did not affect the protein function severely enough to cause a reduction of the product of the enzymatic reaction, (2) the background mutations, one in every 400bp on average (Dalmais et al., 2013), affect normal plant growth to the

extent that small differences in lignification and saccharification efficiency due to *4CL* perturbation became undetectable and perhaps the most logical explanation, (3) the functional redundancy of *4CL* genes in the *4CL* family. Nevertheless, the BRACHYTIL population is a value reverse and forward genetics tool (Bouvier d'Yvoire et al., 2012; Dalmais et al., 2013). The characterization of *brown stem* (*bs*) mutants in the BRACHYTIL population confirmed that the brown coloration in the stems was caused by mutations in *BdCAD1*. Within the scope of this thesis, an expression analysis in the *bs_4179* mutant revealed that a feedback mechanism is present, thereby upregulating the expression of *BdCAD1*, *BdCAD2* and *BdCAD3*, a phenomenon that was also observed in Arabidopsis *cad* mutants (Sibout et al. 2003; Sibout et al. 2005). It was suggested that *BdCAD1* and some of its paralogs try to compensate for the CAD activity deficiency at the time of flowering when the plant needs to reinforce its stems to carry flowers and seeds upright. The upregulation of the affected gene and its paralogs was also observed in the bmr2, encoding 4CL, mutant in Sorghum (Saballos et al. 2012). The authors claim to have direct evidence for an auto-regulatory mechanism for 4CL genes that either involves enhanced gene expression or an increased half-life of the transcript. The same mechanism might also play for the CAD family.

Based on the experience gained in this work, Brachypodium can be used as a model for studying cell wall biosynthesis provided that considerable time and energy is put in the optimization and set up of an efficient transformation platform. Supporting this view is the fact that after the introduction of Brachypodium as a model (Draper et al. 2001), a decade has passed before the first publication of successful overexpression or downregulation of genes of interest (Vain et al. 2011). Nevertheless, its small size, the high quality of its genome sequence, and its close phylogenetic relationship to the major cereal and energy crops maize, wheat, Sorghum, Miscanthus and switchgrass (Vogel et al. 2010) are great assets of Brachypodium as model system.

Yet, Brachypodium possesses C3 metabolism (Brutnell et al. 2010), while maize, sorghum and the high yielding and emerging biomass crops Miscanthus and switchgrass are C4 species (Brkljacic et al. 2011). *Setaria viridis*, another small, annual grass being developed as a C4 model, could be particularly useful in this regard (Doust et al. 2009; Brutnell et al. 2010; Brkljacic et al. 2011). Instead, maize itself can serve as an excellent model system provided the availability of an efficient transformation facility for maize, which is the case at the Plant Systems Biology department. Therefore, during this PhD, a switch from Brachypodium as model to maize as model and crop species was made.

6. Conclusion

In this PhD research, Brachypodium was evaluated as a model for bioenergy research. The high quality of its genome sequence and small plant size makes it suitable for integration as model in research labs that are equipped for Arabidopsis research. A TILLING approach has proven to be effective for identifying mutants in lignin biosynthesis with enhanced saccharification efficiency. However, based on the experience gained during this study, the major hurdle for further implementation of Brachypodium in research might be the low throughput of the Agrobacterium-mediated transformation. For the purpose of this PhD research, a switch to maize as crop model for bioenergy feedstock improvement was made, making use of the transformation facility for maize at the Plant Systems Biology department.

7. Material and methods

7.1 Brachypodium transformation (based on Vogel and Hill 2008; Alves et al. 2009)

Brachypodium plants were grown in a growth chamber with 16h light/8h dark rhythm with controlled humidity (70%) under cool white fluorescent bulb illumination. Spikelets were harvested approximately two weeks after emergence and seeds were sterilized by 5 min submergence in a 4.375% v/v hypochlorite solution with 0.1% of Triton X-100 (Sigma-Aldrich, St. Louis, MO, USA) and washed three times with sterile water. Immature embryos were dissected using a binocular and placed on callus induction medium (CIM). CIM medium was made by adding 4.43 g LS-salt (Labconsult, Schaarbeek, Belgium), 30g sucrose (Labconsult, Schaarbeek, Belgium) and 2 g of Phytagel (Sigma-Aldrich, St. Louis, MO, USA) to 1 L milliQ water and adjusting pH to 5.8 using KOH. After autoclaving, 2.5 mg/l 2,4-D(Sigma-Aldrich, St. Louis, MO, USA), prepared in 95% ethanol and filter sterilized and 0.6 mg/l $CuSO_4$ (company, country) from filter sterilized stock solution in H_2O, was added before pouring plates. Dissected embryos on CIM were incubated four weeks in dark (28°C). Embryogenic callus was transferred to fresh medium on the second and third week. After incubation, yellow, well-structured callus was collected in a falcon tube and cocultivated for 5 min while softly shaking with two days old transformed Agrobacterium cultures suspended in LS medium [4.43 g LS-salt (Labconsult, Schaarbeek, Belgium) and 30 g sucrose (Labconsult, Schaarbeek, Belgium) in 1L milliQ water and pH adjusted to 5.8 using KOH, autoclaved after which 2.5 mg/l 2,4-D (Sigma-Aldrich, St. Louis, MO, USA), prepared in 95% ethanol and filter sterilized, was added] supplied with 200 µM acetosyringone (Sigma-Aldrich, St. Louis, MO, USA). The LS suspension was removed from the calli by pipetting and were then dried by placing onto a whatman filter paper. The calli on filter paper were then incubated in dark (22°C) for three. After cocultivation, calli were then placed onto CIM medium supplemented with 250 µl/ml Ticarcillin disodium/potassium clavulanate (Duchefa Biochemie, Haarlem, the Netherlands) and 40mg/l hygromycin B (Duchefa Biochemie, Haarlem, the Netherlands) or 5 mg/l phosphinotricin (ppt) (Duchefa Biochemie, Haarlem, the Netherlands) and incubated in dark at 28°C. After two weeks, calli were placed on fresh medium and incubated in light, 25°C, 12 h light/12 h dark, 80-110 $\mu molm^{-2}s^{-1}$. After another two weeks, calli were placed on REG medium [4.43 g LS-salt (Labconsult, Schaarbeek, Belgium), 30 g maltose (Labconsult, Schaarbeek, Belgium) and 2g Phytagel (Sigma-Aldrich, St. Louis, MO, USA) in 1L milliQ water and pH adjusted to 5.8 using KOH, autoclaved after which 0.2 mg/l kinetin (6-furfurylaminopurine; Sigma-Aldrich, St. Louis, MO, USA), prepared in 1 M NaOH and filter sterilized, was added] supplemented with 250 µl/ml Ticarcillin disodium/potassium clavulanate (Duchefa Biochemie, Haarlem, the Netherlands) and 40mg/l hygromycin B (Duchefa Biochemie, Haarlem, the Netherlands) or 5 mg/l ppt (Duchefa Biochemie, Haarlem, the Netherlands) and incubated in light at 25°C, 12 h light/12 h dark, 80-110 $\mu molm^{-2}s^{-1}$. Every two weeks calli were placed on fresh medium. Appearing shoots were transferred to MS medium [4.43 g MS+vitamins (Labconsult, Schaarbeek, Belgium), 30g sucrose (Labconsult, Schaarbeek, Belgium) and 2 g of Phytogel (Sigma-Aldrich, St. Louis, MO, USA) to 1 L milliQ water and pH adjusted to 5.7 using KOH] supplemented with 250 µl/ml Ticarcillin disodium/potassium clavulanate (Duchefa Biochemie, Haarlem, the Netherlands) and 40mg/l hygromycin B (Duchefa Biochemie, Haarlem, the Netherlands) or 5mg/l ppt (Duchefa Biochemie, Haarlem, the Netherlands)] for rooting. These

plantlets were transferred to new medium every two weeks until enough roots and leaves were present for transfer to soil.

7.2 Plant material

7.2.1 Plants overexpressing AtGA20OX1

For constitutive overexpression of GA20-oxidase, the *AtGA20OX1* gene (At4g25420) was cloned behind the maize UBI1 promoter in the vector pBbm42GW7 (http://gateway.psb.ugent.be/) and introduced into Brachypodium as described above. No selection pressure was applied since the addition of ppt interfered with shoot regeneration. In total, 10000 regenerated T0 plants were collected in the greenhouse, grown in rootrainers (Haxnicks®, UK) at 21 °C and supplementary light was added to ensure 16h light/8h dark rhythm using high-pressure sodium vapour lamps. Fertilizer was added with the water supply: conductivity Ec = 1mS/cm; water soluble fertilizer Poly-feed (Haifa, Belgium) (N, P2O5, K20; 20:5:20 + 3 MgO). Leaf samples were used for DNA extraction according to Edwards et al. (1991) and were tested by PCR for the presence of the CaMVp35S promoter using p35S-for2:TGT TAG ATC CTC GAT CTG AAT TTT TG and p35S-rev2:CCA CAG ATG GTT AGA GAG GCC TAC. Nineteen independent transgenic T_0 plants were detected. For expression analysis and phenotypic screen segregating populations were generated by selfing of the transgenic plants (Brachypodium self-pollinates by default). Three-to-one p35S positive/p35S negative ratio in the offspring was used as test that these plants were hemizygous for *AtGA20ox1*. Seeds were first stratified at 4°C on moist soil for 3 days and placed at 21°C, and 12h light for synchronized germination. Rooted plantlets were transferred to soil in rootrainers in the greenhouse at 24°C, in 16h light/8h dark rhythm using high-pressure sodium vapour lamps. Fertilizer was added with the water supply: conductivity Ec = 1mS/cm; water soluble fertilizer Poly-feed (Haifa, Belgium) (N, P2O5, K20; 20:5:20 + 3 MgO). For expression analysis and phenotyping leaf growth, transgenic and non-transgenic control plants of each of the five segregating populations were compared

7.2.2 Plants downregulated for BdCAD1

For downregulation of the *BdCAD1* gene, a 130 bp region (180-310 bp in CDS sequence) of the *BdCAD1* (BRADI3G06480) was cloned into the gateway vector pHb7GW-I-WG-UBIL (http://gateway.psb.ugent.be/). Brachypodium plants were transformed as described above using Hygromycin as selection agent. Regenerated shoots were screened for the presence of the CaMVp35S promoter using p35S-for2:TGT TAG ATC CTC GAT CTG AAT TTT TG and p35S-rev2:CCA CAG ATG GTT AGA GAG GCC TAC primers and for the presence of the cloned region using Agi51 agri56 primer recognition sites in the gateway vector. DNA was extracted from leaf samples using NucleoSpin Plant II kit (Macherey-Nagel, Düren, Germany). For all qRT-PCR and protein activity assays, tissues of transgenic and non-transgenic control plants of each segregating population were compared. For gene expression and CAD activity analysis, whole plant samples were used when plants started to flower.

7.2.3 Plants mutated in BdCAD1

For plant material, growth conditions, chemical mutagenesis, DNA extraction and mutation detection of the plants in the BRACHYTILL collection I refer to Dalmais et al. (2013). The plant

material of the TILLING mutant *Bd4179* in *BdCAD1* that was used for the expression analysis of the CAD family was grown at INRA Versailles. For details I refer to Bouvier d'Yvoire et al. (2013).

7.2.4 Plants mutated in Bd4CL1

The seeds of *Bd4CL1* were sent by Richard Sibout from INRA Versailles and grown at the facilities of ILVO using the standard conditions as described above. For genotyping, DNA was extracted using NucleoSpin Plant II kit (Macherey-Nagel, Düren, Germany) from leaf samples. The genotyping of the mutants was performed using custom-made KASP assays that targeted the mutated locus (formerly KBioscience now LGC, Teddington, UK) and run on Roche Lightcycler480 system using the manufacturers guidelines. For phenotypic analysis, homozygous mutated plants were compared to azygous control plants of the segregating population grown from a seed stock obtained by selfing the plant hemizygous for the mutation.

7.3 RNA extraction and qRT-PCR

RNA was extracted using the RNeasy kit (Qiagen, Valencia, CA) and a DNase treatment was performed using DNA-free™ (Ambion, Life technologies, Carlsbad, California, U.S.). Extracted RNA was quantified using a Nanodrop® ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and diluted so that a total of 400 ng RNA was used for cDNA synthesis using the First strand cDNA synthesis kit (Thermo Scientific, Thermo Fisher Scientific, Waltham, MA, USA). Samples were run on a Roche Lightcycler480 (LC480) system in 384 well plates using the SYBR Green Kit from Roche. The cDNA was diluted 10 times and used to run in technical duplicates on the LC480 with following protocol: 1 activation cycle of 10 min at 95°C; 45 amplification cycles of 10 s at 95°C, 10 s at 60 °C and 10 s at 72 °C; 1 melting curve cycle measuring from 65 to 95 °C. Fluorescence values were exported from the LC480 program whereupon Ct values, normalization factors and primer efficiencies were calculated based upon Ramakers et al. (2003). Normalization was performed using Brachypodium actin and UBI10 expression using primers as in Supplementary Table 3.

For the different plant materials, the following genes were analyzed:

- In plants overexpressing the *AtGA20ox1* gene, *AtGA20ox1* expression was monitored using qGA20OX1F1 (CATCAACGTTCTCGAGCTTGATGTTC) and qGA20OX1R1 (GCGGCTCGTGTATTCATGAGCG) primers.
- The qRT-PCR analysis of wildtype Brachypodium CAD and 4CL families was performed on RNA extracted from whole stems and was performed using primer sequences in Supplementary Table 1 and Supplementary Table 2 respectively.
- Expression of the hairpin and *BdCAD1* gene in the *BdCAD1* RNAi lines was monitored using primer sequence in Supplementary Table 3
- RNA extractions from the TILLING mutant *Bd4179* in *BdCAD1* were performed at INRA Versailles and for details I refer to Bouvier d'Yvoire et al. (2013).

7.4 Phenotyping leaf growth

For phenotyping the leaf growth of Brachypodium plants overexpressing the *AtGA20ox* gene, 40 seeds per transgenic line were sown and grown as described above. Measurements were taken from the tip of the 3rd leaf to its basal level on a daily basis, for a period of 10 days and analyzed using LEAF-E as described in Chapter 3. For analysis, transgenic and non-transgenic control plants from five segregating populations were compared. To calculate the thermal time, a base temperature of 10 °C was used.

7.5 CAD activity assay

CAD activity assay was performed according to Fornalé et al. (2011) with minor modifications. 100 mg of ground stem material was used for protein extraction and 20 µg of total protein extract was used as loading for the enzymatic assay. Substrate conversion was measured after 20 minutes.

7.6 Saccharification analysis

Aliquots of 20 mg of dry stem material were used. The biomass was pretreated with 1 ml of 1M HCl at 80°C for 2h, while shaking (850 rpm). The supernatant was removed and the pellet containing pretreated material was washed three times with water to obtain a neutral pH. Subsequently, the material was incubated in 1 ml 70% (v/v) ethanol overnight at 55°C. The remaining biomass was washed three times with 1 ml 70% (v/v) ethanol, once with 1 ml acetone, dried under vacuum for 45 min and weighed. The pretreated ethanol-extracted residue was dissolved in 1 ml acetic acid buffer solution (pH 4.8) and incubated at 50°C. Accelerase® 1500 (Genencor, Denmark) enzyme mix was first desalted over an Econo-Pac 10DG-column (Bio-Rad, Hercules, CA, USA), stacked with Bio-gel® P-6DG gel (Bio-rad) according to the manufacturer's guidelines. The activity of the enzyme mix was measured with a filter paper assay (Xiao et al., 2004). To each sample, dissolved in acetic acid buffer (pH 4.8), the enzyme mix with an activity of 0.04 filter paper units was added. After a short spinning to remove droplets from the lid of the reaction tubes, 20 µl aliquots of the supernatant were taken after 0h, 4h, 7h, 24h and 48h incubation at 50°C and 10 fold diluted with acetic acid buffer (pH4.8). The concentration of glucose in these diluted samples was measured indirectly with a spectrophotometric color reaction (glucose oxidase-peroxidase; GOD-POD) A 100 ml aliquot of the reaction mix from this color reaction contained 50 mg ABTS, 44.83 mg GOD (Sigma-Aldrich, St. Louis, MO, USA) and 173 µl of 4% (w/v) POD (Roche Diagnostics, Brussels, Belgium) in acetic acid buffer (pH 4.5). To measure the concentration of glucose, 50 µl of the diluted samples was added to 150 µl GOD_POD solution and incubated for 30 min at 37 °C. The absorbance was measured spectrophotometrically at a wavelength of 405 nm. The concentration in the original sample was calculated with a standard curve based on known D-glucose concentrations (company, country). Glucose release was then expressed per unit dry weight..

7.7 Acetyl bromide lignin analysis

Aliquots of 5 mg ground stem material were subjected to a sequential extraction to obtain a purified CWR. The extractions were done in 2-ml vials, each time for 30 min, at near boiling temperatures for water (98°C), ethanol (76°C), chloroform (59°C) and acetone (54°C). The remaining CWR was dried under vacuum. Lignin was quantified according to a modified version of the acetyl bromide method (Dence, 1992), optimized for small amounts of plant tissue. The dried CWR was dissolved in 0.1 freshly made 25% acetyl bromide in glacial acetic acid and 4 µl 60% perchloric acid. The solution

was incubated for 30 min at 70°C while shaking (850 rpm). After incubation, the slurry was centrifuged at 14000 rpm for 15 min. To the supernatant, 0.2 ml of 2M sodium hydroxide and 0.5 ml glacial acetic acid was added. The pellet was washed with 0.5 ml glacial acetic acid. The supernatant and the washing phase were combined and the final volume was adjusted to 2 ml with glacial acetic acid. After 20 min at room temperature, the absorbance at 280 nm was measured with a nanodrop® ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). The lignin concentrations were calculated by means of the Bouguer-Lambert-Beer law: $A = \varepsilon \times l \times c$ (A = absorbance, ε = extinction coefficient, l = path length, c = concentration), with $\varepsilon = 17.164$ L g-1 cm-1 (Fukushima and Hatfield 2004) and l = 0.1 cm.

8. References

Alves SC, Worland B, Thole V, et al. (2009) A protocol for Agrobacterium-mediated transformation of Brachypodium distachyon community standard line Bd21. Nat Protoc 4:638–49. doi: 10.1038/nprot.2009.30

Ayliffe M, Singh D, Park R, et al. (2013) Infection of Brachypodium distachyon with selected grass rust pathogens. Mol Plant Microbe Interact 26:946–57. doi: 10.1094/MPMI-01-13-0017-R

Azhaguvel P, Li W, Rudd JC, et al. (2008) Aphid feeding response and microsatellite-based genetic diversity among diploid Brachypodium distachyon (L.) Beauv accessions. Plant Genet Resour 7:72. doi: 10.1017/S1479262108994235

Berthet S, Demont-Caulet N, Pollet B, et al. (2011) Disruption of LACCASE4 and 17 results in tissue-specific alterations to lignification of Arabidopsis thaliana stems. Plant Cell 23:1124–1137.

Bevan MW, Garvin DF, Vogel JP (2010) Brachypodium distachyon genomics for sustainable food and fuel production. Curr Opin Biotechnol 21:211–7. doi: 10.1016/j.copbio.2010.03.006

Biemelt S, Tschiersch H, Sonnewald U (2004) Impact of altered gibberellin metabolism on biomass accumulation, lignin biosynthesis, and photosynthesis in transgenic tobacco plants. Plant Physiol 135:254–265. doi: 10.1104/pp.103.036988.254

Boden SA, Kavanová M, Finnegan EJ, Wigge PA (2013) Thermal stress effects on grain yield in Brachypodium distachyon occur via H2A.Z-nucleosomes. Genome Biol 14:R65. doi: 10.1186/gb-2013-14-6-r65

Bouvier d'Yvoire M, Bouchabke-Coussa O, Voorend W, et al. (2013) Disrupting the cinnamyl alcohol dehydrogenase 1 gene (BdCAD1) leads to altered lignification and improved saccharification in Brachypodium distachyon. Plant J 73:496–508. doi: 10.1111/tpj.12053

Bragg JN, Wu J, Gordon SP, et al. (2012) Generation and characterization of the Western Regional Research Center Brachypodium T-DNA insertional mutant collection. PLoS One 7:e41916. doi: 10.1371/journal.pone.0041916

Brkljacic J, Grotewold E, Scholl R, et al. (2011) Brachypodium as a model for the grasses: today and the future. Plant Physiol 157:3–13. doi: 10.1104/pp.111.179531

Brown RC (2003) Chapter 3: The Biorenewable Resource Base. Biorenewable Resour. Eng. New Prod. from Agric. Iowa state press, Wiley-Blackwell, p 286

Brutnell TP, Wang L, Swartwood K, et al. (2010) Setaria viridis: a model for C4 photosynthesis. Plant Cell 22:2537–44. doi: 10.1105/tpc.110.075309

Carpita NC (1996) Structure and Biogenesis of the Cell Walls of Grasses. Annu Rev Plant Physiol Plant Mol Biol 47:445–476. doi: 10.1146/annurev.arplant.47.1.445

Catalán P, Müller J, Hasterok R, et al. (2012) Evolution and taxonomic split of the model grass Brachypodium distachyon.

Ann Bot 109:385–405. doi: 10.1093/aob/mcr294

Catalán P, Shi Y (1995) Molecular phylogeny of the grass genus Brachypodium P. Beauv. based on RFLP and RAPD analysis. Bot J … 263–280.

Charles M, Tang H, Belcram H, et al. (2009) Sixty million years in evolution of soft grain trait in grasses: emergence of the softness locus in the common ancestor of Pooideae and Ehrhartoideae, after their divergence from Panicoideae. Mol Biol Evol 26:1651–61. doi: 10.1093/molbev/msp076

Cheng M, Lowe BA, Spencer TM, et al. (2004) Factors influencing Agrobacterium-mediated transformation of monocotyledonous species. Vitr Cell Dev Biol - Plant 40:31–45. doi: 10.1079/IVP2003501

Christensen U, Alonso-Simon A, Scheller H V, et al. (2010) Characterization of the primary cell walls of seedlings of Brachypodium distachyon--a potential model plant for temperate grasses. Phytochemistry 71:62–9. doi: 10.1016/j.phytochem.2009.09.019

Christiansen P, Andersen CH, Didion T, et al. (2005) A rapid and efficient transformation protocol for the grass Brachypodium distachyon. Plant Cell Rep 23:751–8. doi: 10.1007/s00299-004-0889-5

Coles JP, Phillips a L, Croker SJ, et al. (1999) Modification of gibberellin production and plant development in Arabidopsis by sense and antisense expression of gibberellin 20-oxidase genes. Plant J 17:547–56.

Coussens G, Aesaert S, Verelst W, et al. (2012) Brachypodium distachyon promoters as efficient building blocks for transgenic research in maize. J Exp Bot 63:4263–73. doi: 10.1093/jxb/ers113

Cui Y, Lee MY, Huo N, et al. (2012) Fine mapping of the Bsr1 barley stripe mosaic virus resistance gene in the model grass Brachypodium distachyon. PLoS One 7:e38333. doi: 10.1371/journal.pone.0038333

Dalmais M, Antelme S, Ho-Yue-Kuang S, et al. (2013) A TILLING Platform for Functional Genomics in Brachypodium distachyon. PLoS One 8:e65503. doi: 10.1371/journal.pone.0065503

Demircan T, Akkaya MS (2009) Virus induced gene silencing in Brachypodium distachyon, a model organism for cereals. Plant Cell, Tissue Organ Cult 100:91–96. doi: 10.1007/s11240-009-9623-x

Devos KM, Gale MD (2000) Genome relationships: the grass model in current research. Plant Cell 12:637–646.

Doust AN, Kellogg EA, Devos KM, Bennetzen JL (2009) Foxtail Millet: A Sequence-Driven Grass Model System. PLANT Physiol 149:137–141. doi: 10.1104/pp.108.129627

Draper J, Mur LAJ, Jenkins G, et al. (2001) Brachypodium distachyon . A New Model System for Functional Genomics in Grasses 1. 127:1539–1555. doi: 10.1104/pp.010196.1

Eddy R, Hahn D (2010) Optimizing Greenhouse Corn Production: What Is the Best Lighting and Plant Density?

Edwards K, Johnstone C, Thompson C (1991) A simple and rapid method for the preparation of plant genomic DNA for PCR analysis. Nucleic Acids Res 19:1349.

FAO Statistics Division (2013) FAOSTAT 2012. http://faostat3.fao.org/faostat-gateway/go/to/download/Q/QC/E. Accessed 7 Feb 2014

Faricelli ME, Valárik M, Dubcovsky J (2010) Control of flowering time and spike development in cereals: the earliness per se Eps-1 region in wheat, rice, and Brachypodium. Funct Integr Genomics 10:293–306. doi: 10.1007/s10142-009-0146-7

Febrer M, Goicoechea JL, Wright J, et al. (2010) An integrated physical, genetic and cytogenetic map of Brachypodium distachyon, a model system for grass research. PLoS One 5:e13461. doi: 10.1371/journal.pone.0013461

Flavell R (2009) Role of model plant species. Methods Mol Biol 513:1–18. doi: 10.1007/978-1-59745-427-8_1

Fornalé S, Capellades M, Encina A, et al. (2012) Altered lignin biosynthesis improves cellulosic bioethanol production in transgenic maize plants down-regulated for cinnamyl alcohol dehydrogenase. Mol Plant 5:817–30. doi: 10.1093/mp/ssr097

Fukushima RS, Hatfield RD (2004) Comparison of the acetyl bromide spectrophotometric method with other analytical lignin methods for determining lignin concentration in forage samples. J Agric Food Chem 52:3713–20. doi: 10.1021/jf035497l

Fursova O, Pogorelko G, Zabotina O a (2012) An efficient method for transient gene expression in monocots applied to modify the Brachypodium distachyon cell wall. Ann Bot 110:47–56. doi: 10.1093/aob/mcs103

Garvin D (2007) Brachypodium: a new monocot model plant system emerges. J Sci Food Agric 1179:1177–1179. doi: 10.1002/jsfa

Garvin DF (2009) Illustrated Guide to Crossing Brachypodium.

Garvin DF, Gu Y-Q, Hasterok R, et al. (2008) Development of Genetic and Genomic Research Resources for , a New Model System for Grass Crop Research. Crop Sci 48:S–69. doi: 10.2135/cropsci2007.06.0332tpg

Garvin DF, McKenzie N, Vogel JP, et al. (2010) An SSR-based genetic linkage map of the model grass Brachypodium distachyon. Genome 53:1–13. doi: 10.1139/g09-079

Gomez L, Bristow J (2008) Analysis of saccharification in Brachypodium distachyon stems under mild conditions of hydrolysis. Biotechnol … 12:1–12. doi: 10.1186/1754-6834-1-15

Gomez LD, Bristow JK, Statham ER, McQueen-Mason SJ (2008) Analysis of saccharification in Brachypodium distachyon stems under mild conditions of hydrolysis. Biotechnol Biofuels 1:15. doi: 10.1186/1754-6834-1-15

Gu YQ, Ma Y, Huo N, et al. (2009) A BAC-based physical map of Brachypodium distachyon and its comparative analysis with rice and wheat. BMC Genomics 10:496. doi: 10.1186/1471-2164-10-496

Gu YQ, Wanjugi H, Coleman-Derr D, et al. (2010) Conserved globulin gene across eight grass genomes identify fundamental units of the loci encoding seed storage proteins. Funct Integr Genomics 10:111–22. doi: 10.1007/s10142-009-0135-x

Guillaumie S, San-Clemente H, Deswarte C, et al. (2007) MAIZEWALL. Database and developmental gene expression profiling of cell wall biosynthesis and assembly in maize. Plant Physiol 143:339–363.

Guillon F, Bouchet B, Jamme F, et al. (2011) Brachypodium distachyon grain: characterization of endosperm cell walls. J Exp Bot 62:1001–15. doi: 10.1093/jxb/erq332

Halpin C, Holt K, Chojecki J (1998) Brown-midrib maize (bm1)–a mutation affecting the cinnamyl alcohol dehydrogenase gene. Plant …

Handakumbura PP, Hazen SP (2012) Transcriptional Regulation of Grass Secondary Cell Wall Biosynthesis: Playing Catch-Up with Arabidopsis thaliana. Front Plant Sci 3:74. doi: 10.3389/fpls.2012.00074

Handakumbura PP, Matos D a, Osmont KS, et al. (2013) Perturbation of Brachypodium distachyon CELLULOSE SYNTHASE A4 or 7 results in abnormal cell walls. BMC Plant Biol 13:131. doi: 10.1186/1471-2229-13-131

Hasterok R, Draper J, Jenkins G (2004) Laying the cytotaxonomic foundations of a new model grass, Brachypodium distachyon (L.) Beauv. Chromosome Res 12:397–403. doi: 10.1023/B:CHRO.0000034130.35983.99

Hong S, Seo P, Yang M (2008) Exploring valid reference genes for gene expression studies in Brachypodium distachyon by real-time PCR. BMC Plant Biol 11:1–11. doi: 10.1186/1471-2229-8-112

Hong S-Y, Seo PJ, Cho S-H, Park C-M (2012) Preparation of leaf mesophyll protoplasts for transient gene expression in Brachypodium distachyon. J Plant Biol 55:390–397. doi: 10.1007/s12374-012-0159-y

Hu Y, Gai Y, Yin L, et al. (2010) Crystal structures of a Populus tomentosa 4-coumarate:CoA ligase shed light on its enzymatic mechanisms. Plant Cell 22:3093–104. doi: 10.1105/tpc.109.072652

Keller B, Feuillet C (2000) Colinearity and gene density in grass genomes. Trends Plant Sci 5:246–51.

Kellogg E (2001) Evolutionary history of the grasses. Plant Physiol 125:1198–1205.

Kim S-J, Kim M-R, Bedgar DL, et al. (2004) Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in Arabidopsis. Proc Natl Acad Sci U S A 101:1455–1460.

Koornneef M, Meinke D (2010) The development of Arabidopsis as a model plant. Plant J 61:909–21. doi: 10.1111/j.1365-313X.2009.04086.x

Laudencia-Chingcuanco DL, Vensel WH (2008) Globulins are the main seed storage proteins in Brachypodium distachyon. Theor Appl Genet 117:555–63. doi: 10.1007/s00122-008-0799-y

Lee M, Jeon W, Kim D, Bold O (2011) Agrobacterium-mediated transformation of Brachypodium distachyon inbred line Bd21 with two binary vectors containing hygromycin resistance and GUS reporter. J Crop Sci … 2011:233–238.

Lee MB, Kim DY, Hong MJ, et al. (2013) Identification of gamma irradiated Brachypodium mutants with altered genes responsible for lignin biosynthesis. Genes Genomics 36:65–76. doi: 10.1007/s13258-013-0142-0

Liu H, Sachidanandam R, Stein L (2001) Comparative genomics between rice and Arabidopsis shows scant collinearity in gene order. Genome Res 2020–2026. doi: 10.1101/gr.194501.1

Luo N, Liu J, Yu X, Jiang Y (2011) Natural variation of drought response in Brachypodium distachyon. Physiol Plant 141:19–29. doi: 10.1111/j.1399-3054.2010.01413.x

Mach J (2013) Small RNAs and the Big Decisions: MicroRNA Regulation of Photoperiodic Flowering in Brachypodium distachyon. Plant Cell 25:4283. doi: 10.1105/tpc.113.251112

Matos D a, Whitney IP, Harrington MJ, Hazen SP (2013) Cell Walls and the Developmental Anatomy of the Brachypodium distachyon Stem Internode. PLoS One 8:e80640. doi: 10.1371/journal.pone.0080640

McCallum CM, Comai L, Greene EA, Henikoff S (2000) Targeting induced local lesions IN genomes (TILLING) for plant functional genomics. Plant Physiol 123:439–42.

Meinke D, Koornneef M (1997) Community standards for Arabidopsis genetics. Plant J 12:247–253.

Meyerowitz EM, Somerville CR (1994) Arabidopsis. Cold Spring Harbor Laboratory Press,, Plainview , N.Y. :

Mur L a J, Allainguillaume J, Catalán P, et al. (2011) Exploiting the Brachypodium Tool Box in cereal and grass research. New Phytol 191:334–47. doi: 10.1111/j.1469-8137.2011.03748.x

Nelissen H, Rymen B, Jikumaru Y (2012) A local maximum in gibberellin levels regulates maize leaf growth by spatial control of cell division. Curr Biol 1–5. doi: 10.1016/j.cub.2012.04.065

Olsen P, Lenk I, Jensen CS, et al. (2006) Analysis of two heterologous flowering genes in Brachypodium distachyon demonstrates its potential as a grass model plant. Plant Sci 170:1020–1025. doi: 10.1016/j.plantsci.2006.01.012

Opanowicz M, Hands P, Betts D, et al. (2011) Endosperm development in Brachypodium distachyon. J Exp Bot 62:735–48. doi: 10.1093/jxb/erq309

Opanowicz M, Vain P, Draper J, et al. (2008) Brachypodium distachyon: making hay with a wild grass. Trends Plant Sci 13:172–7. doi: 10.1016/j.tplants.2008.01.007

Păcurar DI, Thordal-Christensen H, Nielsen KK, Lenk I (2008) A high-throughput Agrobacterium-mediated transformation system for the grass model species Brachypodium distachyon L. Transgenic Res 17:965–75. doi: 10.1007/s11248-007-9159-y

Parker D, Beckmann M, Enot DP, et al. (2008) Rice blast infection of Brachypodium distachyon as a model system to study dynamic host/pathogen interactions. Nat Protoc 3:435–45. doi: 10.1038/nprot.2007.499

Raes J, Rohde A, Christensen JH, et al. (2003) Genome-wide characterization of the lignification toolbox in Arabidopsis. Plant … 133:1051–1071. doi: 10.1104/pp.103.026484.role

Riboulet C, Guillaumie S, Méchin V (2009) Kinetics of phenylpropanoid gene expression in maize growing internodes: relationships with cell wall deposition. Crop … 211–223. doi: 10.2135/cropsci2008.03.0130

Saballos A, Sattler SE, Sanchez E, et al. (2012) Brown midrib2 (Bmr2) encodes the major 4-coumarate:coenzyme A ligase involved in lignin biosynthesis in sorghum (Sorghum bicolor (L.) Moench). Plant J 70:818–30. doi: 10.1111/j.1365-313X.2012.04933.x

Sattler SE, Saathoff AJ, Haas EJ, et al. (2009) A nonsense mutation in a cinnamyl alcohol dehydrogenase gene is responsible for the Sorghum brown midrib6 phenotype. Plant Physiol 150:584–95. doi: 10.1104/pp.109.136408

Schnable P, Ware D, Fulton R, Stein J (2009) The B73 maize genome: complexity, diversity, and dynamics. Science (80-). doi: 10.1126/science.1178534

Schwartz CJ, Doyle MR, Manzaneda AJ, et al. (2010) Natural Variation of Flowering Time and Vernalization Responsiveness in Brachypodium distachyon. BioEnergy Res 3:38–46. doi: 10.1007/s12155-009-9069-3

Sibout R, Eudes A, Mouille G, et al. (2005) CINNAMYL ALCOHOL DEHYDROGENASE-C and -D are the primary genes involved in lignin biosynthesis in the floral stem of Arabidopsis. Plant Cell 17:2059–76. doi: 10.1105/tpc.105.030767

Sibout R, Eudes A, Pollet B, et al. (2003) Expression pattern of two paralogs encoding cinnamyl alcohol dehydrogenases in Arabidopsis. Isolation and characterization of the corresponding mutants. Plant Physiol 132:848–60. doi: 10.1104/pp.103.021048

Strable J, Scanlon MJ (2009) Maize (Zea mays): a model organism for basic and applied research in plant biology. Cold Spring Harb Protoc 2009:pdb.emo132. doi: 10.1101/pdb.emo132

Thole V, Peraldi A, Worland B, et al. (2012) T-DNA mutagenesis in Brachypodium distachyon. J Exp Bot 63:567–76. doi: 10.1093/jxb/err333

Tobias CMC, Chow EKE (2005) Structure of the cinnamyl-alcohol dehydrogenase gene family in rice and promoter activity of a member associated with lignification. Planta 220:678–88. doi: 10.1007/s00425-004-1385-4

Trabucco GM, Matos D a, Lee SJ, et al. (2013) Functional characterization of cinnamyl alcohol dehydrogenase and caffeic acid O-methyltransferase in Brachypodium distachyon. BMC Biotechnol 13:61. doi: 10.1186/1472-6750-13-61

USDA (2013) USDA Crop prdocuction report released on November 8 2013.

Vain P (2011) Brachypodium as a model system for grass research. J Cereal Sci 54:1–7. doi: 10.1016/j.jcs.2011.04.002

Vain P, Thole V, Worland B, et al. (2011) A T-DNA mutation in the RNA helicase eIF4A confers a dose-dependent dwarfing phenotype in Brachypodium distachyon. Plant J 66:929–40. doi: 10.1111/j.1365-313X.2011.04555.x

Vain P, Worland B, Thole V, et al. (2008) Agrobacterium-mediated transformation of the temperate grass Brachypodium distachyon (genotype Bd21) for T-DNA insertional mutagenesis. Plant Biotechnol J 6:236–45. doi: 10.1111/j.1467-7652.2007.00308.x

Valdivia ER, Herrera MT, Gianzo C, et al. (2013) Regulation of secondary wall synthesis and cell death by NAC transcription factors in the monocot Brachypodium distachyon. J Exp Bot 64:1333–43. doi: 10.1093/jxb/ers394

Van Acker R, Vanholme R, Storme V, et al. (2013) Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in Arabidopsis thaliana. Biotechnol Biofuels 6:46. doi: 10.1186/1754-6834-6-46

Van Bel M, Proost S, Wischnitzki E, et al. (2012) Dissecting plant genomes with the PLAZA comparative genomics platform. Plant Physiol 158:590–600. doi: 10.1104/pp.111.189514

Van der Weijde T, Alvim Kamei CL, Torres AF, et al. (2013) The potential of C4 grasses for cellulosic biofuel production. Front Plant Sci 4:107. doi: 10.3389/fpls.2013.00107

Vogel J (2008) Unique aspects of the grass cell wall. Curr Opin Plant Biol 11:301–7. doi: 10.1016/j.pbi.2008.03.002

Vogel J, Bragg J (2009) Genetics and Genomics of the Triticeae. 427–449. doi: 10.1007/978-0-387-77489-3

Vogel J, Garvin D (2006) Agrobacterium-mediated transformation and inbred line development in the model grass Brachypodium distachyon. Plant Cell, Tissue … 199–211. doi: 10.1007/s11240-005-9023-9

Vogel J, Garvin D, Mockler T, The International Brachypodium Initiative (2010) Genome sequencing and analysis of the model grass Brachypodium distachyon. Nature 463:763–8. doi: 10.1038/nature08747

Vogel J, Gu Y, Twigg P (2006) EST sequencing and phylogenetic analysis of the model grass Brachypodium distachyon. Theor Appl … 186–195. doi: 10.1007/s00122-006-0285-3

Vogel J, Hill T (2008) High-efficiency Agrobacterium-mediated transformation of Brachypodium distachyon inbred line Bd21-3. Plant Cell Rep 471–478. doi: 10.1007/s00299-007-0472-y

Wang TL, Uauy C, Robson F, Till B (2012) TILLING in extremis. Plant Biotechnol J 10:761–72. doi: 10.1111/j.1467-7652.2012.00708.x

Wang Y, Antelme S, Dalmais M, et al. (2013) The identification of laccases involved in lignin formation in Brachypodium distachyon. 1st Int. Brachypodium Conf. Abstr. B.

Watt M, Schneebeli K, Dong P, Wilson IW (2009) The shoot and root growth of Brachypodium and its potential as a model for wheat and other cereal crops. Funct Plant Biol 36:960. doi: 10.1071/FP09214

Wei B, Cai T, Zhang R, et al. (2009) Novel microRNAs uncovered by deep sequencing of small RNA transcriptomes in bread wheat (Triticum aestivum L.) and Brachypodium distachyon (L.) Beauv. Funct Integr Genomics 9:499–511. doi: 10.1007/s10142-009-0128-9

Weissmann S, Brutnell TP (2012) Engineering C4 photosynthetic regulatory networks. Curr Opin Biotechnol 23:298–304. doi: 10.1016/j.copbio.2011.12.018

Wu L, Liu D, Wu J, et al. (2013) Regulation of FLOWERING LOCUS T by a MicroRNA in Brachypodium distachyon. Plant Cell 25:4363–77. doi: 10.1105/tpc.113.118620

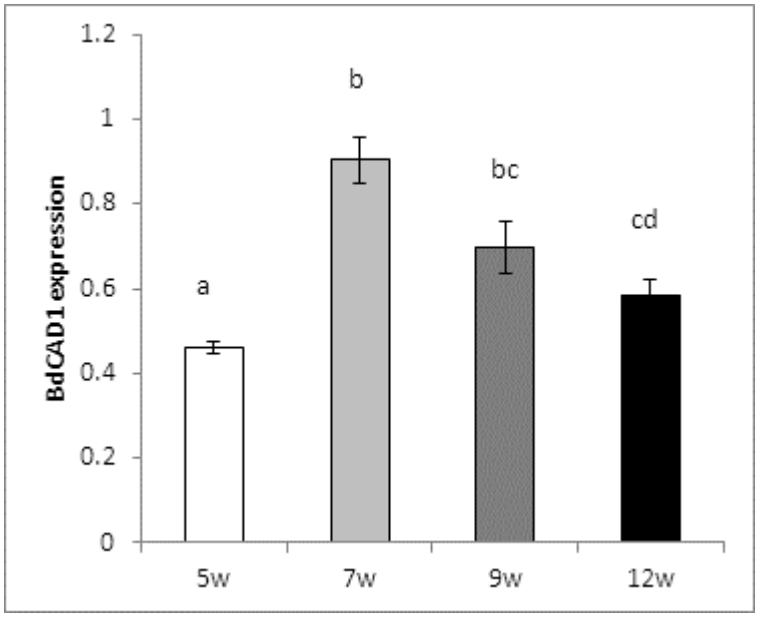
Yao Q, Cong L, Chang JL, et al. (2006) Low copy number gene transfer and stable expression in a commercial wheat cultivar via particle bombardment. J Exp Bot 57:3737–46. doi: 10.1093/jxb/erl145

Zhang J, Xu Y, Huan Q, Chong K (2009) Deep sequencing of Brachypodium small RNAs at the global genome level identifies microRNAs involved in cold stress response. BMC Genomics 10:449. doi: 10.1186/1471-2164-10-449

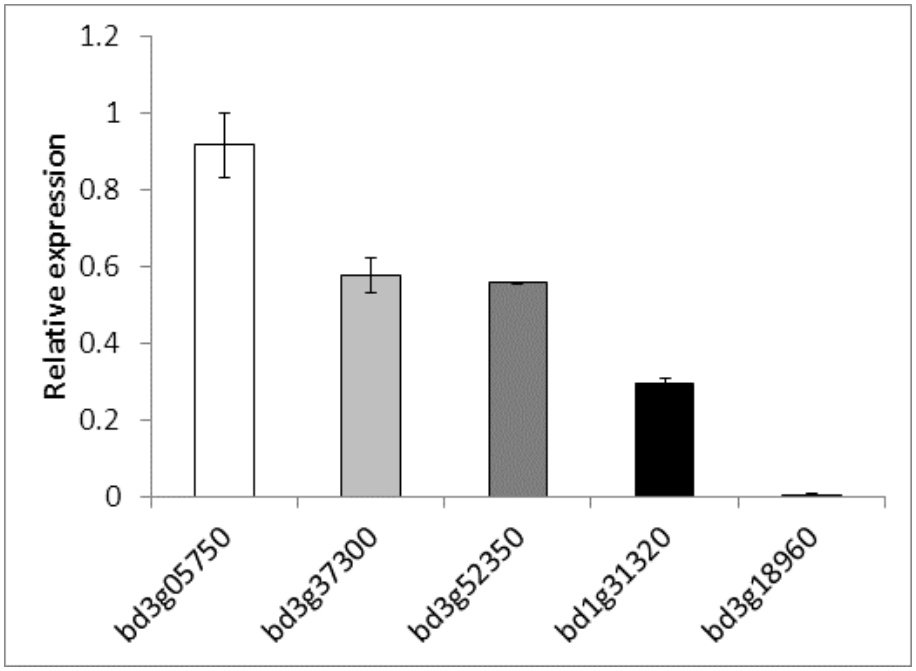
Zhang K, Qian Q, Huang Z, et al. (2006) GOLD HULL AND INTERNODE2 encodes a primarily multifunctional cinnamyl-alcohol dehydrogenase in rice. Plant … 140:972–983. doi: 10.1104/pp.105.073007.cies

Zhao Q, Nakashima J, Chen F, et al. (2013) Laccase is necessary and nonredundant with peroxidase for lignin polymerization during vascular development in Arabidopsis. Plant Cell 25:3976–87. doi: 10.1105/tpc.113.117770

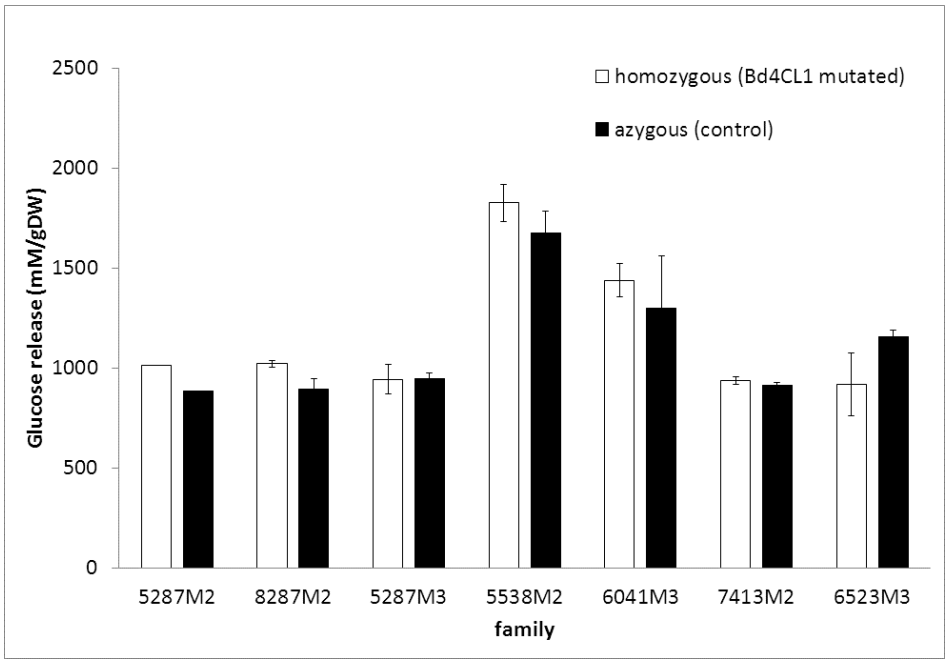
9. Supplementary figures and tables



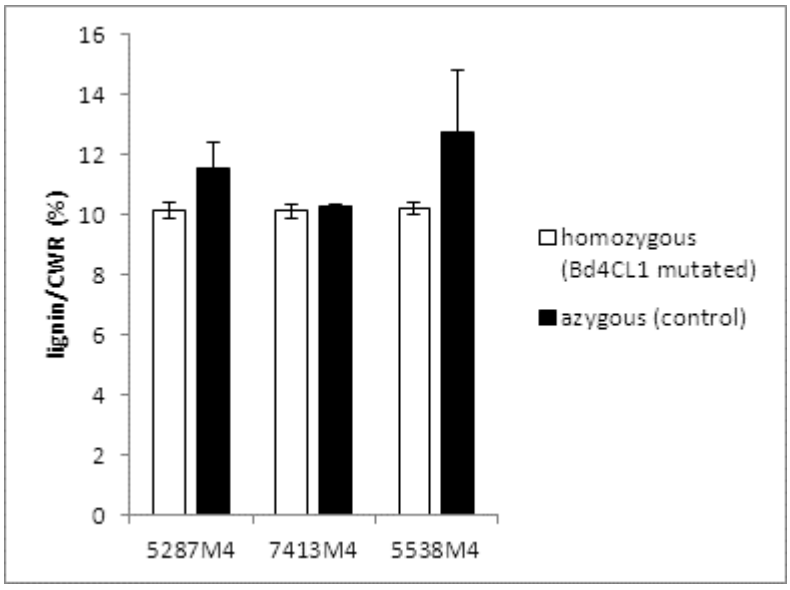
Supplementary figure 1. *BdCAD1* expression in various stages of development. The stages correspond to early vegetative (5w), late vegetative (7w), transition to flowering (9w) and reproductive (12w). Error bars represent standard errors of three biological replicates. Categories a, b, c and d are based on ANOVA with Scheffé post hoc test ($p<0.05$).



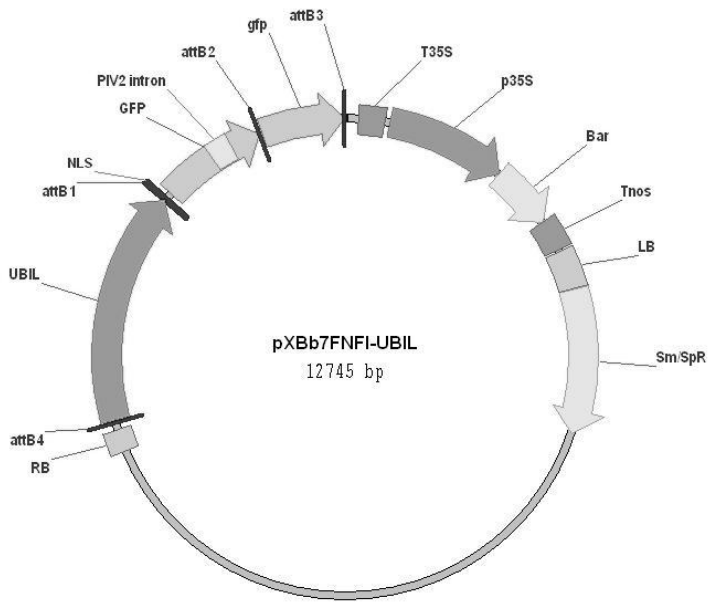
Supplementary figure 2. Relative expression values for 4CL family members in stems of a Brachypodium plant. Error bars represent standard errors over two technical replicates. *Bd3g05750* or *Bd4CL1* is highest expressed in the stem.



Supplementary figure 3. Saccharification assay on Brachypodium plants carrying mutations in Bd4CL1. Glucose release upon enzymatic hydrolysis is shown for plants homozygous and azygous for the mutation in a segregating population of mutated lines 5287, 8287, 5538, 6041, 7413 and 6523. Plants from either the M2 or M3 generation were used. Error bars indicate standard errors over biological replicates indicated as "line (#homozygous mutant replicates, #azygous control replicates)": 5287M2 (1,1), 8287M2 (2,2), 5287M3 (3,3), 5538M2 (2,2), 6041M3 (4,4), 7413M2 (2,2) and 6523M3 (4,4).



Supplementary figure 4. Lignin content in Brachypodium plants carrying mutations in *Bd4CL1*. Plants in M4 generation were used. Error bars indicate standard errors over biological replicates indicated as "line (#homozygous mutant replicates, #azygous control replicates)": 5287M4 (4,6), 7413M4 (6,6) and 5538M4 (3,4).



Supplementary figure 5. pBHb7FNFI-UBIL Expression vector used for GFP expression in Brachypodium. The vector is available at http://gateway.psb.ugent.be/; RB: right border, UBIL: maize ubiquitin promoter,NLS: nuclear localization signal, GFP: green fluorescent protein, PIV2 intron: intron sequence inducing splicing , attB1, attB2, attB3 and attB4: recognition sites for site-specific recombination resulting from a recombination reaction, T35S: CaMV 35S terminator, p35S: CaMV 35S promoter, Bar: basta resistance gene, Tnos: nopaline synthesis terminator, LB: left broder, Sm/SpR: spectitomycin resistance gene

Supplementary Table 1. Primers for qRT-PCR for Brachypodium CAD genes

Gene name	Gene code	Primer sequence FW (5'->3')	Tm (°c)	Primer sequence (5'->3')	Tm (°c)	Amplicon length (bp)
BdCAD1	BRADI3G06480	GGACGGAGCTAGGTAGAGTGC	60	AAGGCCCAGTTTACATCGAA	59	182
BdCAD3	BRADI3G22980	ATGTTAGCAGCTCGCACCTT	60	AATCGTTTGGCACGGTAGAT	59	182
BdCAD4	BRADI4G29770	GACGTCGAGGTCGTCAAGAT	60	TCACGTGGCAATATGAAAGC	59	166
BdCAD6	BRADI5G04130	CTGCTGTGCTCTGGCTACAC	59	TGCAATTGAAGTACAGGCAAC	58	161
BdCAD7	BRADI5G21550	CTGGCGTCAATAAAGAGTGG	58	CAAATGGTGCCTATAGCGAAG	59	243
BdCAD2	BRADI3G17920	GCCCTCGTTGCTAAGAACA	59	TACCTAACGTCGGCCTTGAC	60	173
BdCAD5	BRADI4G29780	CCTCTCGGCCATTTGTATGT	59	ATCATCCACGGCAGTGTTTT	60	152
Actin	BRADI4G41850	GTGAGTATGATGAGTCTGGTCCAG	60	TACGAGTCTAGGAGGTACACAG	53	206
UBI10	BRADI1G32860	CACGCTTGAAGTTGAGTCATC	58	CCATGGACAGGCACTTACTGG	62	191
SamDC	BRADI5G14640	CGGCAAGCTTGCTAATCTGCTGGAAT	62	CAGAGCAACAATAGCCTGGCTGGC	63	164

Supplementary Table 2. Primers for qRT-PCR for Brachypodium 4CL genes

name	sequence (5'->3')
4CL Bd3g52350 cDNA FW1	GATCTTCCGGTCCAAGCTC
4CL Bd3g52350 cDNA RV1	GAGCTCAAATCGTGCCTCTC
4CL Bd3g37300 cDNA FW1	ATTCCATTCCACAGCTCCAC
4CL Bd3g37300 cDNA RV1	GATGTCCTGCAGCTCCTTG
4CL Bd3g18960 cDNA FW1	GATCATCTTCCGGTCGAAAC
4CL Bd3g18960 cDNA RV1	ACCTTGAGCTCCCCGTTG
4CL Bd3g05750 cDNA FW1	GGTCTGTCCCGGAGGAGT
4CL Bd3g05750 cDNA RV1	ACCCGGATTTGACCTTGAAC
4CL Bd1g31320 cDNA FW1	GTCGCGGCTGAGGAGTTC
4CL Bd1g31320 cDNA RV1	GTGAGCATGACGCCCTTG

Supplementary Table 3. Primers for qRT-PCR for Brachypodium BdCAD1 RNAi lines

name	sequence (5'->3')
WV_BRADICAD2_endo_qPCR_FW	GCTGGACTACGTCATCGACA
WV_BRADICAD2_endo_qPCR_RV	TTTCTCGACGCAGAACTTGA
WV_BRADICAD2_RNAi1_CATintron_qPCR_FW	CCTGAAGATGTGCTGGTGAA
WV_BRADICAD2_RNAi1_CATintron_qPCR_RV	CTGCACAAGCTTTGGATCCTC
WV_BRADICAD2_RNAi1_pdkintron_qPCR_FW	TACCGAATTCCTCGAGACCA
WV_BRADICAD2_RNAi1_pdkintron_qPCR_RV	CCTGAAGATGTGCTGGTGAA

Chapter 5: A systems-wide approach to investigate lignification and perturbation of *CINNAMATE 4-HYDROXYLASE* in maize

Chapter 5: A systems-wide approach to investigate lignification and perturbation of *CINAMATE 4-HYDROXYLASE* in maize

Chapter 5: A systems-wide approach to investigate lignification and perturbation of *CINNAMATE 4-HYDROXYLASE* in maize

This chapter was written in preparation of a research paper that will report on the perturbation of the *ZmC4H1* gene in maize, the first report on *C4H* perturbation in monocots. My personal contribution to this work was performing the cell wall analysis of stem material (lignin amount, lignin composition and saccharification efficiency), GO and Pageman analysis of the transcriptome data and statistical analysis of the NIRS data. I was also in charge of plant growing and harvesting and data generation to build the CSPP network. Metabolite extraction from the *zmc4h1* mutant and control samples for UPLC-MS analysis was also one of my tasks. Finally I interpreted all the datasets and wrote the manuscript

Chapter 5: A systems-wide approach to investigate lignification and perturbation of *CINAMATE 4-HYDROXYLASE* in maize

1. Abstract

Bioenergy grasses, such as maize, provide a major source of lignocellulosic biomass that can be converted into second generation bioethanol. Currently, the major goal in developing a cost-effective cellulosic ethanol industry is the improvement of the saccharification efficiency of bioenergy crops. As it is the presence of lignin in the cell wall that hinders the enzymatic breakdown of cellulose into fermentable sugars, lignin engineering is effective in improving saccharification efficiency. A genetic basis of lignin formation in maize has been presented previously, but candidate genes for the early steps in the lignin biosynthetic pathway have yet to be experimentally confirmed.

Here we present a mutant in the lignin biosynthesis gene *ZmC4H1*, the first report of *C4H* perturbation in monocots. Despite possible functional redundancy in the maize *C4H* gene family, cell wall analysis of *zmc4h1* plants at silage stage showed a decrease in the amount of ADL lignin and a compensatory increase in the amount of hemicellulose as compared to control plants. As a result, the saccharification efficiency of *zmc4h1* plants was improved compared to control plants, using acid as pretreatment step. In addition, the systems-wide effects of *ZmC4H1* perturbation on the transcriptome and the metabolome level were studied. Unfortunately, the results from this section of the chapter had to be removed due to inconsistencies of the genotyping results of the samples. However, the methodology that was followed for these transcriptome and metabolome analyses was retained.

We conclude that *C4H* perturbation forms a promising strategy to improve saccharification efficiency in maize and, by extrapolation, in grasses.

2. Introduction

2.1 Maize as energy crop

Bioenergy grasses are defined as members of the grass family (*Poaceae*) that employ C4 metabolism and are capable of producing high biomass yield in the form of lignocellulose, fermentable juice, or fermentable grain (Vermerris 2011; Feltus and Vandenbrink 2012; van der Weijde et al. 2013). *Zea mays* (maize) is next to *Saccharum* spp. (sugarcane), *Sorghum bicolor* (sorghum), *Miscanthus* spp. (Miscanthus), and *Panicum virgatum* (switchgrass) one of the five major bioenergy grasses, given their proven utility as feedstock and their academic and industrial interest (Feltus and Vandenbrink 2012). Currently maize and sugarcane provide the bulk of the so-called first-generation bioethanol that is produced from the starch in the grain (maize) or the sugar extract from the stem (sugarcane). However, it would be advantageous to also use the lignocellulosic fraction (i.e. stems, leaves and cobs) for the production of second generation bioethanol because the fuel yield per hectare would substantially increase and direct competition between food and fuel would decrease (Valentine et al. 2012). For the production of second-generation bioethanol, the lignocellulosic biomass undergoes a biochemical process to convert polysaccharides into fermentable products through enzymatic hydrolysis (saccharification) (Carroll and Somerville 2009; Sims et al. 2010; Carriquiry et al. 2011). Maize, and also sorghum, can be used as dual purpose crops with the grain used for food or feed and the vegetative biomass for the production of second-generation (-cellulosic) ethanol (Carpita and McCann 2008; Bennetzen and Hake 2009; Olsen and Wendel 2013). Likewise, after sugar extraction from sugarcane, the lignocellulosic mass or bagasse could be used for fermentation to cellulosic ethanol. Miscanthus and switchgrass are high yielding, fast growing perennial crops with low input requirements that can be grown with the sole purpose to provide lignocellulosic feedstock. However, maize has some advantages over other bioenergy grasses: (i) the maize (as well as sorghum) genome has been sequenced, enabling genomics-based research and efficient exploitation of natural variation for crop improvement (Carpita and McCann 2008; Jakob et al. 2009; Schnable et al. 2009b; Feuillet et al. 2011); (ii) maize has a long history as model organism, with many useful genetic and genomic resources (Vermerris 2011); (iii) maize and sugarcane are currently the predominant bioenergy grasses, based on the volume of biofuels produced and their well-established production chains can supply the biorefinery with large amounts of agricultural residues (Waclawovsky et al. 2010; van der Weijde et al. 2013). Maize, as the largest crop worldwide in terms of production (FAO Statistics Division 2013a) is therefore expected to play an essential role in the development and large-scale commercialization of cellulosic fuels (Stewart 2007; Penning et al. 2009; van der Weijde et al. 2013). This will probably be accomplished by cultivation as a dual-purpose crop, with optimal grain yield for food and a high stem biomass production with optimal saccharification efficiency for biofuel production.

Currently, the major goal in developing a cost-effective second generation bioethanol industry is the improvement of the saccharification efficiency of bioenergy crops (Carroll and Somerville 2009). The bioenergy grass feedstock traits that underlie conversion efficiency and thus the targets for genetic engineering are related to cellulase inhibition, cellulose accessibility, crystallinity index and enzyme adsorption (Feltus and Vandenbrink 2012). In fact, it is the presence of lignin that negatively influences cellulose accessibility and thus saccharification efficiency (Zeng et al. 2014). Therefore, it is of great importance to investigate the genetic basis of lignin biosynthesis in

bioenergy crops such as maize to define targets for genetic engineering to modify the lignin content and/or composition (Gressel 2008), and to exploit genetic variation. This can greatly benefit from knowledge and methods developed for forage quality, as it has been demonstrated that forage quality and composition data may be used to predict cellulosic ethanol yield in maize (Andersen et al. 2008; Lorenz et al. 2009), Sorghum (Han et al. 2013) and switchgrass (Sarath et al. 2011; Vogel et al. 2013). Suitable methodologies include e.g. *in vitro* digestibility, where digestibility is determined by dry matter disappearance after a 48-hour incubation with rumen fluid (Stern et al. 1997) or enzyme cocktails (Aufrère 1982). Other forage quality data are based on the widely used sequential detergent system developed by Van Soest et al. (1991), as the neutral detergent fiber (NDF) is an approximation of the total amount of cell wall and the acid detergent lignin (ADL) is an approximation of the amount of lignin. If an *in vitro* digestibility test is performed on the NDF fraction, it estimates the digestibility and availability of the cell wall carbohydrates providing information on the conversion efficiency of plant biomass into bioethanol (Lorenz et al. 2009). These methods are well-established and routinely used, and in many cases prediction models based on infrared reflectance spectroscopy (NIRS) are available, in particular for maize. Thus for each wet-lab method, a specific prediction equation is generated that is valid for a specific NIRS instrument and a specific set of samples. Robust estimation of a specific quality parameter is attained by inclusion of plant materials obtained in a wide variety of environmental conditions and derived from different populations in the comparison, e.g. the maize stem harvested in different locations over many successive years using many different accessions (Liu et al. 2008). In this way, Lorenz et al. (2009) constructed a NIRS regression model including NDF and its ruminal digestibility that explained 95% of the variation in ethanol yield in twelve maize varieties.

Strategies for engineering lignin in monocots have been mainly based on results obtained in dicotyledonous model species (Shen et al. 2013). A genetic basis of lignin formation in maize has been presented (Guillaumie et al. 2007; Riboulet et al. 2009), but the corresponding enzymatic functions encoded by many of the gene candidates have yet to be verified biochemically or genetically. Genetic approaches for functional characterization in maize have been reported previously for a *caffeoyl-CoA O-methyltransferase* (*CCoAOMT*) (Li et al. 2013), *cinnamoyl-CoA reductase* (*CCR*) (Tamasloukht et al. 2011), *caffeic acid O-methyltransferase* (*COMT*) (Vignols et al. 1995; Morrow et al. 1997; Piquemal et al. 2002) and *cinnamyl alcohol dehydrogenase* (*CAD*) (Halpin et al. 1998b; Fornalé et al. 2012). However, candidate genes for the early steps in the lignin biosynthetic pathway have yet to be experimentally confirmed.

2.2 Phenotypic consequences of *cinnamate 4-hydroxylase* perturbation in dicots

In dicots, cinnamate 4-hydroxylase (C4H) acts early in the phenylpropanoid pathway by converting cinnamate into 4-hydroxycinnamate or *p*-coumarate, thereby catalyzing an important step in the general phenylpropanoid pathway downstream of phenylalanine ammonia lyase (PAL) and upstream of 4-coumarate ligase (4CL) (Vanholme et al. 2012a and Figure 41). The membrane-associated C4H cytochrome P450 is thought to anchor a complex consisting of PAL, and possibly other phenylpropanoid pathway enzymes, to the endoplasmic reticulum (Wagner and Hrazdina 1984; Winkel-Shirley 1999; Ro et al. 2001; Achnine and Blancaflor 2004; Chen et al. 2011). In dicot species like Tobacco, Arabidopsis and alfalfa, there is only one *C4H* gene. For that reason and because C4H catalyzes a biochemical reaction essential for plant growth, a full knockout is lethal

(Schilmiller et al. 2009). Arabidopsis mutants with severely reduced *C4H* activity have a delay in development, a reduction in inflorescence stem height, inflorescence stem weight and lignin quantity, an increase in matrix polysaccharides and syringyl/guaiacyl (S/G) ratio and a greatly improved saccharification yield both with (4-fold increase) and without acid pretreatment (three-fold) (Van Acker et al. 2013). Downregulation of *C4H* has been reported in tobacco (Sewalt et al. 1997; Cook et al. 2012) and alfalfa (Chen and Dixon 2007), also resulting in reductions in lignin content but in contrast to Arabidopsis here a decrease in S/G ratio was reported as a result of a higher reduction in S units as compared to the reduction in G units. In the tobacco transgenic line with sense *C4H* suppression, sugar release from stem material was increased ~two-fold compared to wildtype (Cook et al. 2012). The saccharification yield increase was more modest in alfalfa plants downregulated for *C4H*, 40% with and without acid pretreatment (Chen and Dixon 2007). Thus, *C4H* appears to be a good candidate to improve saccharification efficiency in maize, albeit potentially with a reduced biomass yield and delayed development.

2.3 *Cinnamate 4-hydroxylase* genes in maize

In contrast to Arabidopsis, tobacco and alfalfa, three *C4H* genes are present in the maize genome. Having multiple *C4H* members reduces the chance to affect lignin when perturbing the function of one of the paralogs, but it will also reduce the chance on having a lethal phenotype, as was the case for the Arabidopsis full knock-out. The genes [redacted] and [redacted] (hereafter *ZmC4H1* and *ZmC4H2*) originate from a recent duplication event, specific for maize (Plaza 2.5, Van Bel et al. 2012). The coding sequences of these two genes are very similar (94%, Supplementary figure 6). The third *C4H* gene is [redacted] (hereafter *ZmC4H3*), with 79% identity to *ZmC4H1* and *ZmC4H2* coding sequences. In the expression database of maize genes involved in cell wall biosynthesis MAIZEWALL (Guillaumie et al. 2007, unfortunately not updated since then), only two C4H sequences are present, one most similar to *ZmC4H1* and one most similar to *ZmC4H3*. Probably *ZmC4H1* and *ZmC4H2* could not be distinguished due to their high homology. According to a gene expression study using the MAIZEWALL array, *ZmC4H1* and *ZmC4H3* are expressed in all organs investigated (roots, leaves and young stems) and in IN1 (basal internode) and IN6 (internode below the ear) at the silking stage (Guillaumie et al. 2007). *ZmC4H1* was by far the predominantly expressed gene in IN6. *ZmC4H3*, was expressed predominantly in young plants. In another study using the MAIZEWALL array to investigate the kinetics of phenylpropanoid gene expression in the growing IN6, *ZmC4H1* had a maximum expression around tasseling (male flowering) followed by a rapid decrease at later developmental stages, while *ZmC4H3* had again a maximum and increasing expression after silking (female flowering) (Riboulet et al. 2009). Thus, the expression of *ZmC4H1* has a higher and earlier maximum than that of *ZmC4H3*. Based on these expression studies, both *ZmC4H1* and *ZmC4H3* might play a role in maize stem lignification, but the higher maximal expression of *ZmC4H1* points this gene as the main target for lignin biosynthesis perturbation. However, as kinetics of *ZmC4H2* expression in the maize internode is unknown, the possibility of functional redundancy cannot be excluded.

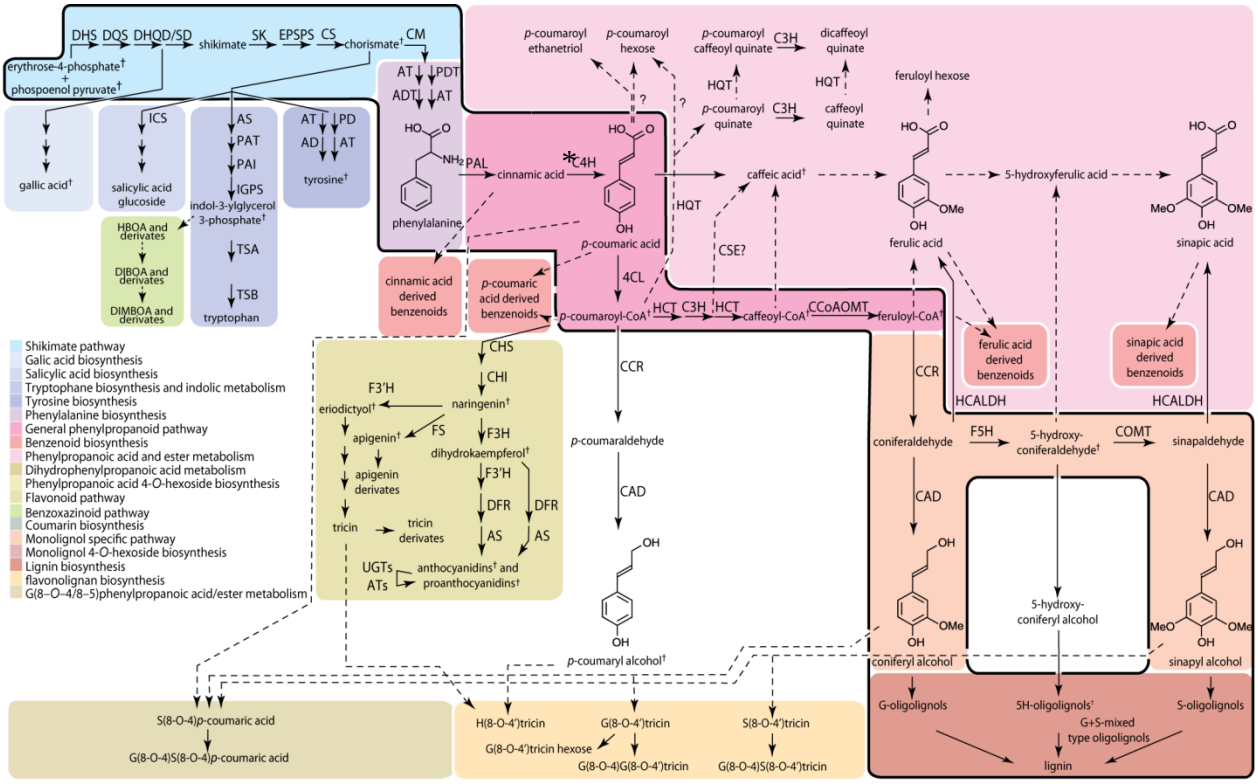


Figure 41. Metabolic map of phenolic metabolism in the internode of maize based on the metabolic map of the Arabidopsis stem (Vanholme et al. 2012b) **and adapted for maize using maize literature on flavonoid biosynthesis** (Sharma et al. 2012)**, benzoxazinoids** (Jonczyk et al. 2008) **and maize phenolic profiling (this study). The main pathways involved in lignin biosynthesis are framed with a bold black border. For the C4H enzyme, highlighted with a red asterisk, a mutant was analyzed in this study. Only the expression of the general phenylpropanoid, monolignol-specific and flavonoid biosynthetic pathway genes were determined via microarray. The relative abundance of the metabolites was determined via UPLC-MS. Metabolites indicated with a cross were not detected or not identified via the targeted UPLC-MS approach. This pathway representation was used to map expression and metabolite data of the *zmc4h1* mutant and control samples. AD, arogenate dehydrogenase; ADT, arogenate dehydratase; AS, anthranilate synthase; AS, anthocyanidin synthase; AT, amino transferase; ATs, acyltransferases; C3H, *p*-coumarate 3-hydroxylase; C4H, cinnamate 4-hydroxylase; CAD, cinnamyl alcohol dehydrogenase; CCoAOMT, caffeoyl-CoA O-methyltransferase; CCR, cinnamoyl-CoA reductase; 4CL, 4-coumarate:CoA ligase; CHI, chalcone isomerase; CHS, chalcone synthase; COMT, caffeic acid O-methyltransferase; CM, chorismate mutase; CS, chorismate synthase; CSE, caffeoyl shikimate esterase; DFR, dihydroflavonol 4-reductase; DHS, 3-deoxy-D-arabino-heptulosonate 7-phosphate synthase; DHQD/ SD, 3-dehydroquinate dehydratase/shikimate dehydrogenase; DQS, 3-dehydroquinate synthase; EPSPS, 5-enolpyruvylshikimate-3-phosphate synthase; F3H, naringenin 3-dioxygenase; F3'H, flavonoid 3-hydroxylase; F5H, ferulate 5-hydroxylase; FLS, flavonol synthase; HCALDH, hydroxy-cinnamaldehyde dehydrogenase; HCT, hydroxycinnamoyl-CoA shikimate/quinate hydroxycinnamoyl transferase; HQT, hydroxycinnamoyl-CoA: quinate hydroxycinnamoyltransferase; ICS, isochorismate synthase; IGPS, indole-3-glycerol phosphate synthase; PAI, phosphoribosylanthranilate isomerase; PAL, phenylalanine ammonia lyase; PAT, phosphoribosylanthranilate transferase; PD, prephenate dehydrogenase; SK, shikimate kinase; TSA, Trp synthase a-subunit; TSB, Trp synthase b-subunit; UGTs, UDP-glucosyltransferases. For nomenclature of aromatic molecules, see** Error! Reference source not found. **and Morreel et al. (2010b and 2010a).**

2.4 Systems biology of lignification

The perturbation of lignin biosynthetic genes has been proven to be a good strategy to improve lignocellulosic breakdown in many plant species (Shi et al. 2006; Chen and Dixon 2007; Leplé et al. 2007; Fu et al. 2011; Mansfield et al. 2012; Jung et al. 2012; Papa et al. 2012; Van Acker et al. 2013; Dalmais et al. 2013). However, most of these studies have focused on the effects of genetic perturbation on biochemical properties of the cell wall. If transcript and metabolite profiling methods were used, they did not allow distillation of transcriptome and metabolome-wide conclusions. Hence, knowledge of how plants cope with altered lignin levels or composition remains fragmented (Vanholme et al. 2008; Vanholme et al. 2010b; Vanholme et al. 2010a; Vanholme et al. 2012b; Vanholme et al. 2012a). This knowledge is crucial to understand and avoid the documented negative impact of lignin reduction on plant yield and agronomic properties. Severe reductions in lignin content can result in dwarfing due to the collapse of xylem vessels which affects water transport in the plant. However, also mild reductions in lignin content can provoke a biomass reduction or growth delay of which the cause is not always known (Pedersen et al. 2005; Voelker et al. 2010). On the other hand, better insights could also explain why reduced lignin content is sometimes correlated with higher yield, as is the case for fast-growing poplar and eucalyptus trees (Novaes et al. 2010). To get a better understanding of metabolic fluxes in and between pathways, a systems biology approach can be taken, i.e. the study of the consequences of pathway perturbations, followed by computational analysis of the data (Ideker et al. 2001; Ideker et al. 2006). This approach preferably involves the combination of different types of data and requires knowledge of genes, regulatory sequences, transcripts, proteins and metabolites (Shi et al. 2010). In that way, a predictive model can be built for complex biological systems such as secondary metabolite and cell wall biosynthesis. Attempts to elucidate systems-wide responses of lignin perturbation have been undertaken in Arabidopsis (Vanholme et al. 2010c; Vanholme et al. 2012b), poplar (Shi et al. 2010) and to a lesser extent in maize (Shi et al. 2006). Systems-wide studies also allow to identify new candidate genes, that co-express with already known lignin biosynthetic genes, as was demonstrated by Vanholme et al. (2012b and 2013). The current advances in transcriptomics and metabolomics (Morreel et al. 2010b; Morreel et al. 2010a; Mochida and Shinozaki 2011) allow this type of approach and are likely to expand the present knowledge of plant biology greatly.

3. Objectives

Results from lignin perturbation studies in the dicot species Arabidopsis, tobacco and alfalfa have indicated that the downregulation of *C4H* expression greatly enhances saccharification efficiency. To our knowledge, no *C4H* perturbation or downregulation has been reported in any monocot species. In this work, we examined the effect of *C4H* disruption in a transposon insertion maize mutant on cell wall characteristics and saccharification efficiency in the context of improving lignocellulosic feedstock for bio-ethanol production. In addition, the systems-wide effects of *ZmC4H* perturbation on the transcriptome and the metabolome level is studied to get further insight in lignin biosynthesis in maize, to pinpoint new candidates involved in stem lignification. The systems-biology approach also allowed to get a deeper insight into the regulation of the phenylpropanoid pathway and the metabolic fluxes between the phenylpropanoid and other metabolic pathways.

4. Results

4.1 A transposon insertion mutant for *ZmC4H1*

Here we present a mutant in *ZmC4H1* (hereafter *zmc4h1*) which was isolated by transposon tagging with a Mutator element by the Biogemma (France). In this chapter, we investigated the phenotype of field grown *zmc4h1* plants by comparison with the corresponding control plants. The *zmc4h1* and control plants resulted from three generations of self-pollinations after five generations of backcrossing (BC5S3) the original mutant plant with an elite line of the Limagrain company. To clarify the origin of the mutant, selected by transposon tagging, a scheme is provided in addendum to this chapter.

Here, I would also like to inform the reviewers that during the last weeks, doubts have been raised about the identity of the plant material that was used for transcript and phenolic profiling, but not for the cell wall analysis. The identity of mutant and wildtype samples could not be confirmed lately by a genotyping assay performed on the plant material. Nevertheless, based upon quality control of the datasets, such as clustering analysis of the transcriptome and metabolome results, we believe that what is described in this part of the work is most likely reliable. Currently, researchers at both Biogemma and PSB are working to solve this issue but conclusive evidence could not be delivered before the deadline of this PhD dissertation. Therefore, we advise that care should be taken while interpreting the obtained results for transcript and metabolic profiling in this chapter.

4.2 Cell wall related characteristics of field-grown *zmc4h1* mutant and control plants using near infrared reflectance spectroscopy

Oven-dried, milled stem material from field grown *zmc4h1* and control plants at silage stage was subjected to NIRS analysis and standard forage quality parameters were determined. The value of NIRS-derived parameters putatively relevant for describing the phenotype of *ZmC4H1* perturbation are presented in Table 6. The full set of parameters estimated by NIRS is listed in Supplementary Table 4 and additional information on the methods used can be found in the addendum to this chapter.

NIRS results revealed a modified cell wall content and composition, and an improved digestibility in the *zmc4h1* mutant compared to control plants (Table 6). The total dry matter digestibility and the digestibility of the cell wall fraction (IVNDFD) were significantly higher in the mutant. This can probably be explained by the higher cell wall content of *zmc4h1* stems (NDF, + 4%) and the lower (11%) acid detergent lignin (ADL; an approximation of the amount of lignin in the cell wall). This reduction in ADL is reflected in a general reduction (~20%) of *p*-hydroxybenzaldehyde, vanillin and syringaldehyde, related to the abundance of the lignin units *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) respectively, based on extraction using the nitrobenzene method. The cellulose content, calculated as ADF-ADL, was unaltered but the hemicellulose fraction (NDF-ADF) was significantly ($p<0.01$) increased by 10% in the *zmc4h1* mutant. The amount of *p*-coumaric acid (*p*CA) in the cell wall (*p*CA is the product of the C4H enzymatic step) was reduced by 13% in the mutant, confirming functional disruption of the *C4H* gene. The *p*CA that was estimated here is ester-linked *p*CA which can be found linked to the lignin polymer (Ralph et al. 1994; Lam et al. 2001; Xu et al. 2005; Harris and DeBolt 2010). The amount of ester-linked FA which is attached to glucoarabinoxylan (GAX), the

most abundant form of hemicellulose in grasses, was not significantly altered. Also the ether-linked FA, which can be incorporated into the lignin polymer by oxidative coupling was not altered according to NIRS estimations. In contrast, both 5-5- and 8-O-4-linked diferulic acid (diFA), which play a major role in the crosslinking of glucoarabinoxylan (GAX) chains (Ishii 1997; Hatfield et al. 1999; MacAdam and Grabber 2002; Jung 2003; Harris and Trethewey 2010; Molinari et al. 2013), were 21% and 15% higher respectively in the *zmc4h1* mutant.

Table 6. NIRS estimation of biomass quality parameters in *zmc4h1* mutant and control plants. NDF: neutral detergent fiber, ADF: acid detergent fiber, ADL: acid detergent lignin, IVDMD: *in vitro* dry matter digestibility, IVNDFD: *in vitro* NDF digestibility. NIRS parameters were estimated according to Barrière et al. (2008). p-values were calculated based upon student t-test in six biological replicates for each group.

NIRS Parameter	units	control	*zmc4h1*	fold change (%)	p-value
NDF	% of Dry matter	54.46	56.78	4.3	**0.004**
ADF	% of Dry matter	32.80	32.93	0.4	0.844
ADL	% of Dry matter	4.01	3.58	-10.7	**0.007**
ester Ferulic Acid	% of Dry matter	6.63	6.69	1.0	0.392
ether Ferulic Acid	% of Dry matter	8.47	8.49	0.1	0.862
5-5 diFerulic Acid	% of Dry matter	0.16	0.20	20.9	**0.022**
8-O-4 diFerulic Acid	% of Dry matter	0.30	0.34	15.0	**0.042**
ester *p*-Coumaric Acid	% of Dry matter	14.76	12.83	-13.1	**0.006**
p-Hydroxybenzaldehyde	% of Dry matter	1.66	1.36	-17.7	**0.005**
Vanillin	% of Dry matter	8.43	7.27	-13.7	**0.004**
Syringaldehyde	% of Dry matter	8.27	6.26	-24.3	**0.001**
IVDMD	% of Dry matter	50.96	53.97	5.9	**0.008**
IVNDFD	% of Dry matter	27.97	36.54	30.6	**0.000009**

4.3 Investigation of the saccharification efficiency

The same dried and milled stem material from field grown *zmc4h1* and control plants was used for wet-lab analysis. A saccharification assay without and with acid pretreatment was performed and the lignin content and composition were determined. In contrast to the predicted 11% decrease in ADL using NIRS (Table 6), no significant changes were detected when the lignin content of mutant and control were compared using the acetyl bromide method (Figure 42). On the other hand, analysis of the lignin composition by thioacidolysis confirmed a reduction in the abundance of H and G units, both per lignin (-65% and -21% respectively) as per dry matter (70% and 32% respectively) amount in the mutant (Figure 43A and B). The abundance of S units remained unaltered. As a result, the S/G ratio increased from 1.17 in the control to 1.54 (32%,) in the *zmc4h1* mutant. Consistent with the NIRS prediction of improved digestibility, the saccharification efficiency of stem biomass pretreated with 1M HCl was higher in *zmc4h1* than in control plants. An increase in glucose release of 26% and 23% was observed after nine and 24 hours of hydrolysis, respectively, at 50°C (Figure 44). Without the acid pretreatment, we could not detect a significant increase in glucose release.

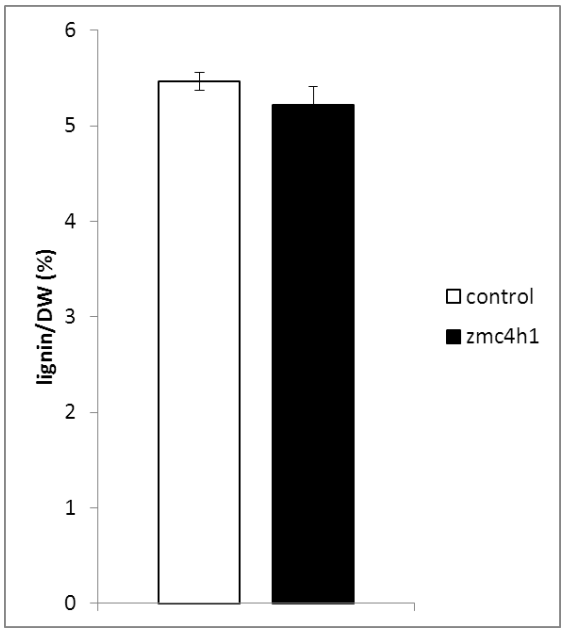


Figure 42. Lignin content in mature stems per dry weight (DW) of *zmc4h1* mutant and control plants determined by acetyl bromide. Error bars represent standard errors over four *zmc4h1* mutant and six control biological replicates.

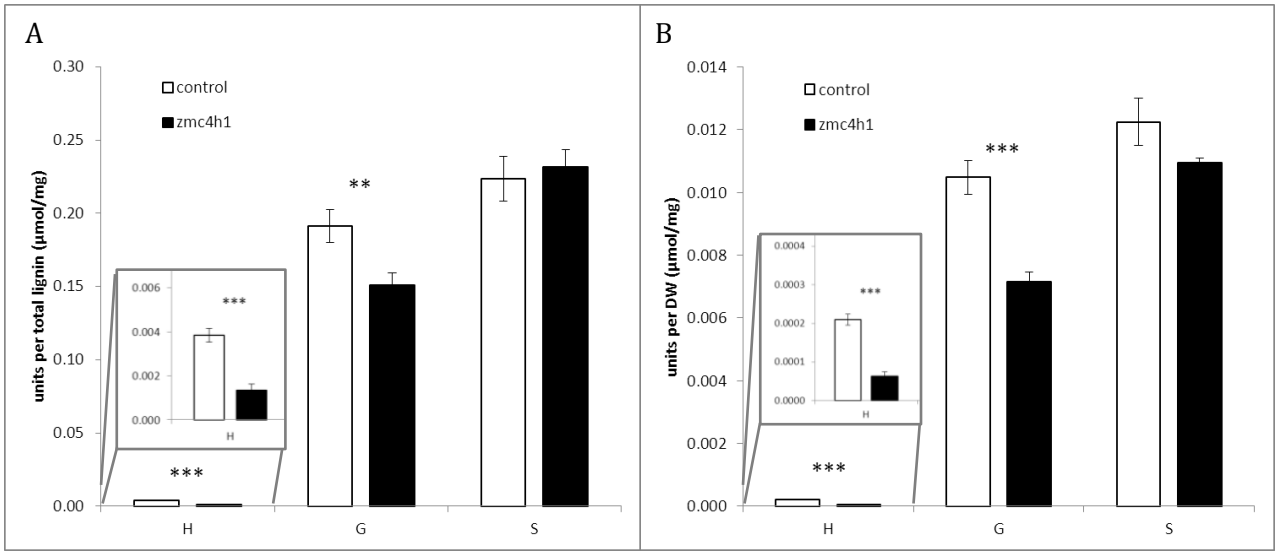


Figure 43. Lignin composition in the stems of control and *zmc4h1* mutant plants determined by thioacidolysis, expressed as abundance of *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) units per lignin amount (A) and per dry weight (B). Error bars represent standard errors of four *zmc4h1* mutant and six control biological replicates. **: p<0.01, ***: p<0.001.

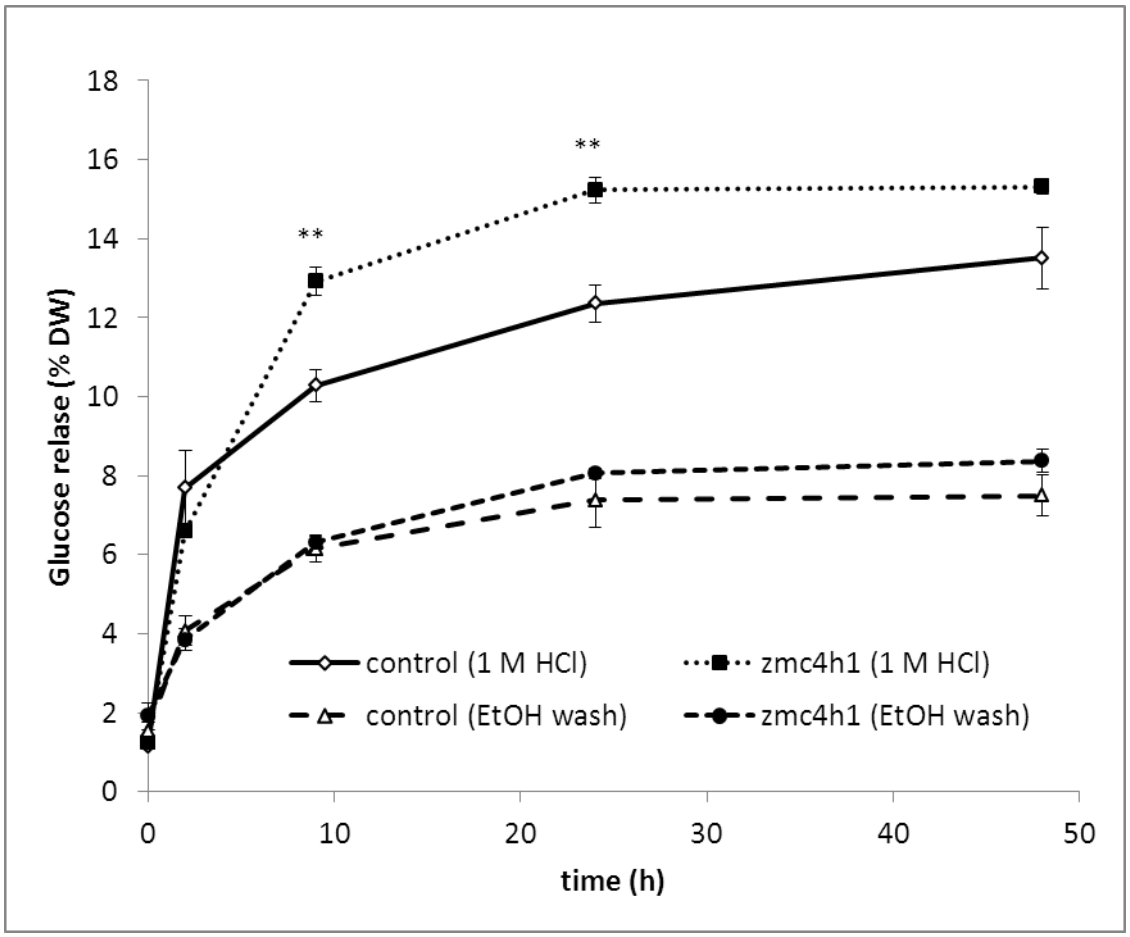


Figure 44. Saccharification efficiency in *zmc4h1* mutant and control stems pretreated with ethanol only (EtOH wash) and ethanol and acid (1 M HCl) expressed as glucose release per dry weight (DW). Error bars represent standard errors over four *zmc4h1* mutant and six control biological replicates. **: p<0.01.

4.4 Identification of genes involved in maize internode lignification

Gene expression analysis at four developmental stage of control plants was used to identify the genes that function in lignification of the maize internode. This will benefit the systems-biology based study of the effects of *ZmC4H1* perturbation, in particular the effect on expression of genes within the same pathway, which will be discussed in the next sections.

Gene expression was analyzed in the ear internode (internode bearing the ear) of control maize plants using a custom Nimblegen microarray containing 37,670 probes. The different developmental stages were V10 (ten visible leaf collars), S (silking), S+7d (seven days after silking) and S+14d (fourteen days after silking). Candidate genes involved in the phenylpropanoid pathway in maize were identified based on orthology relationships with known lignin biosynthesis genes from Arabidopsis using the online comparative genomics tool Plaza v2.5 (Van Bel et al. 2012). These include all phenylpropanoid genes, namely *PAL*, *C4H*, *4CL*, *HCT*, *C3H*, *CCoAOMT*, *CCR*, *F5H*, *COMT* and *CAD*. Based on previously described studies (Guillaumie et al. 2007; Riboulet et al. 2009) and orthologous gene searches, a total of 45 maize orthologs were found. In total, ten *PAL*, three *C4H*,

five *4CL*, two *C3H*, two *HCT*, six *CCoAOMT*, seven *CCR*, two *F5H*, one *COMT* and seven *CAD* genes were identified in the maize genome (Supplementary Table 5).

The identification of phenylpropanoid genes that are involved in lignification was based on two assumptions: 1) gene expression must be high in the internodes and 2) the expression must increase from vegetative V10 stage to the reproductive S stage in agreement with the onset of internode lignification (Northcote 1989; Morrison and Kessler 1994; Matos et al. 2013) and preferably decrease slightly thereafter. A total of six PAL, one *C4H*, one *4CL*, one *C3H*, two *HCT*, three *CCoAOMT*, one *CCR*, one *F5H*, one *COMT* and one *CAD* met these criteria (Table 7).

Table 7. Phenylpropanoid genes selected for high and increasing expression in development in maize control internodes. V10 (ten visible leaf collars), S (silking), S+7d (seven days after silking) and S+14d (fourteen days after silking). Increasing and decreasing expression was calculated by subtracting the mean expression per gene over development from the mean normalized expression value per gene in each stage and by color coding all subsequent values for higher (red) and lower (blue) expression than average. The expression level was labeled according to maximal expression: "very high" >=15, 15>"high">=13, 13>"moderate">=11, 11>"low".

			mean normalized expression (log2)					(normalized expression)-(mean per gene) (log2)			
protein function	gene name	probe name	V10	S	S+7d	S+14d	expression level	V10	S	S+7d	S+14d
PAL	Hidden		14.13	15.07	15.07	14.80	very high	-0.63	0.30	0.30	0.03
PAL			14.47	14.71	15.01	14.40	very high	-0.17	0.06	0.37	-0.25
PAL			13.96	14.95	14.98	14.94	high	-0.75	0.24	0.27	0.23
PAL			13.37	14.81	14.66	14.41	high	-0.94	0.50	0.34	0.10
PAL			13.41	14.16	14.28	14.07	high	-0.57	0.18	0.30	0.09
PAL			12.45	13.07	13.38	13.34	high	-0.61	0.01	0.32	0.28
C4H			12.99	13.24	13.33	13.43	high	-0.26	0.00	0.08	0.18
4CL			12.45	13.04	12.87	11.58	high	-0.03	0.56	0.39	-0.91
HCT			11.52	12.80	12.46	12.35	moderate	-0.76	0.51	0.18	0.07
HCT			11.64	12.54	12.32	12.40	moderate	-0.58	0.31	0.09	0.18
C3H			13.72	13.77	13.83	13.77	high	-0.05	0.00	0.05	0.00
CCoAOMT			12.89	14.54	14.57	13.90	high	-1.08	0.56	0.60	-0.08
CCoAOMT			12.58	13.69	13.58	13.62	high	-0.79	0.32	0.22	0.25
CCoAOMT			12.58	13.62	13.28	13.10	high	-0.56	0.48	0.13	-0.05
CCR			12.22	13.46	13.17	12.84	high	-0.70	0.54	0.25	-0.08
F5H			9.79	11.10	11.12	11.63	moderate	-1.12	0.19	0.21	0.72
COMT			13.32	14.51	14.29	14.15	high	-0.75	0.44	0.22	0.09
CAD			12.66	14.16	13.98	13.50	high	-0.91	0.58	0.41	-0.07

For *PAL*, six out of ten genes showed high and increasing expression in the developing maize internode ([redacted], [redacted], [redacted], [redacted], [redacted] and [redacted]), suggesting a partially redundant role for individual PAL enzymes (Table 7). The three *C4H* genes followed highly similar expression profiles over internode development but showed different levels of expression ranging from low to moderate to high (Supplementary Table 5). These relative expression levels of all three *C4H* genes in the maize internode were not reported previously. Based on the highest expression level in maize internodes, [redacted] or *ZmC4H1* is most likely the main *C4H* gene involved in internode lignification, explaining why this gene is the best candidate for downregulation or perturbation strategies for enhancing saccharification efficiency. Out of the five candidates for *4CL*, [redacted] showed the highest expression and increase in expression from V10 to S stage suggesting its importance for

internode lignification. Three other moderately expressed *4CL* genes, [redacted], [redacted] and [redacted] might also be involved in lignification of which the latter two have a tendency to increase in expression towards the most mature stage. Based on the expression level, [redacted] was selected to be the best candidate for internode lignification. For *C3H*, [redacted] was much higher expressed than [redacted] and was selected accordingly. The two *HCT* genes showed almost identical gene expression. Therefore, both might be as important for lignification. Among the *CCoAOMT* family members, all genes (except [redacted] and [redacted]) displayed a developmental gradient in gene expression, suggesting that also for CCoAOMT, the individual enzymes might be redundant in function. Three *CCoAOMT* genes were selected based on their high expression levels ([redacted], [redacted] and [redacted]). A total of seven *CCR* genes was investigated for their role in maize internode lignification. Among the *CCR* family members, the highest expression was found for [redacted] and [redacted]. Yet, only the expression of [redacted] followed the desired profile of increasing expression from V10 towards S stage and was therefore selected as putatively involved in lignification. The two *F5H* genes followed highly similar profiles but expression of [redacted] was slightly higher than [redacted]. The higher expressed [redacted] was retained as representative but both *F5H* genes might have a role in lignification. The maize unique *COMT* gene showed a high expression and increased from V10 towards S stage, confirming its established role in lignification of the maize stem. In the *CAD* family, three genes show high to moderate expression levels ([redacted], [redacted] and [redacted]). Whereas expression of [redacted] remained constant over development, expression of [redacted] declined from V10 to S suggesting different roles than acting in lignification. The main *CAD* member involved in stem lignification however is the well characterized *bm1* gene or [redacted], which showed high expression and a developmental gradient of increasing expression from V10 towards S stage.

4.5 Lignin biosynthetic gene expression is altered in the *zmc4h1* mutant

This part of the chapter was removed from the final version of the thesis because of inconsistencies in the genotyping results of the samples, obtained from Biogemma, that were used for transcriptome and metabolome analysis.

In this section the expression of genes with a putative role in lignification were compared in mutant and control samples.

4.6 Systems-wide effects of ZmC4H1 perturbation on gene expression levels in internodes

This part of the chapter was removed from the final version of the thesis because of inconsistencies in the genotyping results of the samples, obtained from Biogemma, that were used for transcriptome and metabolome analysis.

C4H acts early in the phenylpropanoid pathway, just downstream of PAL and upstream of 4CL, and catalyzes the production of *p*-coumaric acid, an important intermediate of the general phenylpropanoid pathway (Figure 41). Thus, perturbation of *ZmC4H1* possibly affects processes that are downstream of this enzymatic step. These include monolignol and thus lignin biosynthesis, but also processes related to but beyond the phenylpropanoid pathway such as dihydrophenylpropanoic acid metabolism, phenylpropanoic acid 4-*O*-hexoside biosynthesis, flavonoid biosynthesis, suburin biosynthesis and coumarin biosynthesis (Vanholme et al. 2012a).

In this section the systems-wide responses on gene expression levels were investigated using a GO enrichment analysis using the genes that displayed differential expression between mutant and control samples. For the GO enrichment analysis, the web-based Gene Ontology Enrichment Analysis Software Toolkit (GOEAST, Zheng and Wang 2008) was used. Additionally, Pageman and MapMan (Thimm et al. 2004; Usadel et al. 2005; Usadel et al. 2009) were used for analysis of the systems-wide effects, providing an additional level of analysis beyond the functional enrichment of typical GO categories (Kakumanu et al. 2012).

4.7 Differential expression of genes involved in cell wall and secondary metabolism

This part of the chapter was removed from the final version of the thesis because of inconsistencies in the genotyping results of the samples, obtained from Biogemma, that were used for transcriptome and metabolome analysis.

The GOEAST GO enrichment and Pageman overrepresentation analyses do not provide information at the level of individual genes. Therefore, the functional classes “cell wall”, “secondary metabolism” and “miscellaneous” in MapMan were screened to identify highly upregulated or downregulated genes by selecting for a log fold change >1 or <-1.

4.8 Phenolic profiling reveals metabolic shift in *zmc4h1* internodes

This part of the chapter was removed from the final version of the thesis because of inconsistencies in the genotyping results of the samples, obtained from Biogemma, that were used for transcriptome and metabolome analysis.

Intermediates and products of the phenylpropanoid and monolignol biosynthetic pathways were investigated by UPLC-MS. Using a targeted identification process, a total of 53 compounds present in the maize internode were tentatively structurally identified based on their retention time, exact mass and/or mass spectral fragmentation pattern (MS^2) (Morreel et al. 2010a). The abundance of these compounds in mutant and control samples was investigated and placed onto a map for secondary metabolites in maize to visualize the metabolic shift in *zmc4h1* plants.

5. Discussion

The perturbation of lignin biosynthesis is a major strategy for improving saccharification efficiency (Chen and Dixon 2007; Grabber et al. 2008; Simmons et al. 2008; Gressel 2008; Sticklen 2008; Vanholme et al. 2008; Grabber et al. 2009; Grabber et al. 2010; Vanholme et al. 2010a; Van Acker et al. 2013). The C4H enzyme catalyzes one of the first steps in the phenylpropanoid pathway which, among others, provides the building blocks for lignin biosynthesis in the cell wall. To the author’s knowledge, this is the first report on *C4H* perturbation in a monocot species.

5.1 Multiple copies of *C4H* in the maize genome

In contrast to Arabidopsis, which has only one *C4H* gene, maize has three *C4H* genes with two genes that arose from a recent duplication event and have highly similar coding sequences. We identified and studied the effect of perturbation of one of these genes, *ZmC4H1*, by means of transposon tagging. In Arabidopsis, the full knockout of *AtC4H* is lethal and severe growth defects are even apparent in a weak allele (Schilmiller et al. 2009). In maize, *zmc4h1* had only a moderate delay in development that becomes apparent at silking stage. Under field conditions, *zmc4h1* plants started silking one week later than the corresponding control plants (see addendum to this chapter). A reduction in biomass yield cannot be ruled out since measurements of biomass production were not performed. The fact that the developmental delay in maize *zmc4h1* plants is modest compared to the strong phenotypical abnormalities reported in Arabidopsis *c4h* mutants is probably because of partial functional redundancy in maize. The presence of multiple copies of *C4H* in the genome suggests selection for maintenance of additional copies in the genome of longer living grasses and woody species (Hamberger et al. 2007; Kumar et al. 2013), and apparently also in maize, with a tropical origin. Another explanation could be that in grasses PAL also has TAL activity (Barrière et al. 2007; Rösler et al. 1997). TAL activity catalyzes the tyrosine deamination directly into *p*CA, thereby bypassing the function of C4H. Nevertheless, *zmc4h1* plants show a lignin phenotype, despite possible functional redundancy.

5.2 Plasticity of the cell wall composition and relation to saccharification efficiency

In accordance with previous studies (Andersen et al. 2008; Lorenz et al. 2009; Sarath et al. 2011; Vogel et al. 2013; Han et al. 2013), the predicted enhanced *in vitro* digestibility and NDF and reduced ADL resulted in enhanced saccharification efficiency. In contrast, the reduction in ADL could not be confirmed by biochemical analysis using the acetyl bromide method. The discrepancy between acetyl bromide and ADL, two different methods to determine lignin content, has been observed before (Fukushima and Hatfield 2004). ADL typically underestimates actual lignin levels in grasses, yielding lignin concentrations that are only one half to one quarter of acetyl bromide and Klason methods (Hatfield and Fukushima 2005), with Klason being yet another frequently used lignin quantification method, whereas the acetyl bromide method may overestimate lignin (Voelker et al. 2010). The *zmc4h1* mutant displays a 10% reduction of ADL while the total cell wall fraction is even slightly increased. The reduction in lignin content is compensated by a comparable increase in hemicellulose, but not in cellulose content. In addition, the increased abundance of diFA, which functions in the crosslinking of hemicellulose chains, could also be linked to increased hemicellulose content. A detailed analysis of hemicellulose composition (using the method of Foster et al. 2010) would provide further insight. Notably, the compensation for reduced lignin has been observed previously and was attributed to increased amounts of cellulose in poplar trees (Hu et al. 1999; Jouanin et al. 2000) and to matrix polysaccharides in Arabidopsis lignin mutants (Van Acker et al. 2013). The increased abundance of ferulates in *zmc4h1* plants is surprising because ferulate biosynthesis is downstream of the C4H step in the phenylpropanoid pathway. However, the same phenomenon was observed in Arabidopsis *c4h* mutants which hints that there is an alternative route to ferulate biosynthesis (Vanholme et al. 2010b) in both monocots and dicots.

The maize *zmc4h1* mutant showed a 20% increase in saccharification efficiency with respect to the control using HCl as pretreatment step, but no significant increase with only ethanol wash. The

function of the acid pretreatment is mainly to hydrolyze the hemicellulose, which is higher in the mutant. The removal of the higher amount of hemicellulose might lead to more available space for the hydrolytic enzymes to access the cellulose. The higher hemicellulose content in *zmc4h1* plants can be beneficial for commercial bioethanol production as pentose sugars, the hydrolysis product of hemicellulose are nowadays also fermented into ethanol (Cardona et al. 2010).

A second possible reason for the higher saccharification efficiency after acid pretreatment may be the lignin composition. The lignin in the cell walls of the *zmc4h1* mutant is less rich in H and G units (65% and 21 % respectively) than of the control, as determined by thioacidolysis. As a consequence, the S/G ratio is significantly higher (32%) in *zmc4h1* mutant plants. Similar observations were made in the Arabidopsis *c4h* mutant (Van Acker et al. 2013). Yet, debate remains on the role of the different lignin subunits in determining saccharification efficiency (Grabber and Ralph 1997; Feltus and Vandenbrink 2012). Grabber et al. (1997) suggested that the altered S/G ratios reflect other compositional changes in the lignin that influence saccharification efficiency. Most interestingly, a saccharification model that was built using data from Arabidopsis lignin mutants (Van Acker et al. 2013) showed that a high S/G ratio has a negative effect on saccharification carried out without pretreatment but had a positive effect on saccharification efficiency with acid pretreatment. These findings suggest that cell walls with a high S/G ratio form a matrix in which the hemicelluloses render the cellulose less accessible by cellulases (Van Acker et al. 2013). For the maize *zmc4h1* mutant, this would mean that the higher S/G ratio would mask the positive effect of the reduction of lignin content on saccharification without pretreatment but enhance glucose release when pretreated with acid. Our study in the monocot species maize supports the findings of Van Acker et al. (2013) and might indicate that compensation for lignin reduction by an increase in hemicellulose content is not restricted to dicots.

5.3 Lignin biosynthetic gene expression follows a developmental gradient in the maize internode

This part of the chapter was removed from the final version of the thesis because of inconsistencies in the genotyping results of the samples, obtained from Biogemma, that were used for transcriptome and metabolome analysis.

5.4 *ZmC4H1* perturbation leads to feedback on lignin biosynthetic gene expression

This part of the chapter was removed from the final version of the thesis because of inconsistencies in the genotyping results of the samples, obtained from Biogemma, that were used for transcriptome and metabolome analysis.

5.5 Differential response for different members of the laccase gene family in maize upon *ZmC4H1* perturbation

This part of the chapter was removed from the final version of the thesis because of inconsistencies in the genotyping results of the samples, obtained from Biogemma, that were used for transcriptome and metabolome analysis.

5.6 The common stress-response

This part of the chapter was removed from the final version of the thesis because of inconsistencies in the genotyping results of the samples, obtained from Biogemma, that were used for transcriptome and metabolome analysis.

5.7 Perturbation of *ZmC4H1* leads to metabolic shift in phenolic metabolism

This part of the chapter was removed from the final version of the thesis because of inconsistencies in the genotyping results of the samples, obtained from Biogemma, that were used for transcriptome and metabolome analysis.

6. Conclusion

This study has shown that, despite possible functional redundancy in the maize *C4H* gene family, perturbation of *ZmC4H1* results in an altered cell wall composition with enhanced saccharification efficiency using HCl as pretreatment. *C4H* perturbation is thus a promising strategy to improve saccharification efficiency in maize and, by extrapolation, in grasses.

Unfortunately, the results of the transcriptome and metabolome analysis had to be removed from this chapter because of inconsistencies in the genotyping results of the samples, obtained from Biogemma. However, the methodology that was followed for the investigation of the systems-wide effects can be used for future experiments involving maize lignin mutants. We believe that with the additional identification of phenolic compounds and the careful interpretation of integrated transcriptome and metabolome data of the maize *zmc4h* mutant, this will improve the current understanding of interconnections between pathways in secondary cell wall formation and secondary metabolism.

7. References

Achnine L, Blancaflor E (2004) Colocalization of L-phenylalanine ammonia-lyase and cinnamate 4-hydroxylase for metabolic channeling in phenylpropanoid biosynthesis. Plant Cell … 16:3098–3109. doi: 10.1105/tpc.104.024406.pathways

Andersen J, Zein I, Wenzel G (2008) Characterization of phenylpropanoid pathway genes within European maize (Zea mays L.) inbreds. BMC plant … 14:1–14. doi: 10.1186/1471-2229-8-2

Aufrère J (1982) Etude de la prévision de la digestibilité des fourrages par une méthode enzymatique. Ann. Zootech

Barrière Y, Riboulet C (2007) Genetics and genomics of lignification in grass cell walls based on maize as model species. … Genomes Genomics

Barrière Y, Thomas J, Denoue D (2008) QTL mapping for lignin content, lignin monomeric composition, p-hydroxycinnamate content, and cell wall digestibility in the maize recombinant inbred line progeny F838×F286. Plant Sci 175:585–595. doi: 10.1016/j.plantsci.2008.06.009

Bennetzen J, Hake S (2009) Handbook of Maize: Genetics and Genomics.

Cardona C a, Quintero J a, Paz IC (2010) Production of bioethanol from sugarcane bagasse: Status and perspectives. Bioresour Technol 101:4754–66. doi: 10.1016/j.biortech.2009.10.097

Carpita NC, McCann MC (2008) Maize and sorghum: genetic resources for bioenergy grasses. Trends Plant Sci 13:415–20. doi: 10.1016/j.tplants.2008.06.002

Carriquiry MA, Du X, Timilsina GR (2011) Second generation biofuels: Economics and policies. Energy Policy 39:4222–4234. doi: 10.1016/j.enpol.2011.04.036

Carroll A, Somerville C (2009) Cellulosic biofuels. Annu Rev Plant Biol 60:165–82. doi: 10.1146/annurev.arplant.043008.092125

Chen F, Dixon R a (2007) Lignin modification improves fermentable sugar yields for biofuel production. Nat Biotechnol 25:759–61. doi: 10.1038/nbt1316

Chen H-C, Li Q, Shuford CM, et al. (2011) Membrane protein complexes catalyze both 4- and 3-hydroxylation of cinnamic acid derivatives in monolignol biosynthesis. Proc Natl Acad Sci U S A 108:21253–8. doi: 10.1073/pnas.1116416109

Cook CM, Daudi A, Millar DJ, et al. (2012) Transcriptional changes related to secondary wall formation in xylem of transgenic lines of tobacco altered for lignin or xylan content which show improved saccharification. Phytochemistry 74:79–89. doi: 10.1016/j.phytochem.2011.10.009

Dalmais M, Antelme S, Ho-Yue-Kuang S, et al. (2013) A TILLING Platform for Functional Genomics in Brachypodium distachyon. PLoS One 8:e65503. doi: 10.1371/journal.pone.0065503

FAO Statistics Division (2013) FAOSTAT 2011. http://faostat3.fao.org/faostat-

gateway/go/to/download/Q/QC/E. Accessed 28 Dec 2013

Feltus FA, Vandenbrink JP (2012) Bioenergy grass feedstock: current options and prospects for trait improvement using emerging genetic, genomic, and systems biology toolkits. Biotechnol Biofuels 5:80. doi: 10.1186/1754-6834-5-80

Feuillet C, Leach JE, Rogers J, et al. (2011) Crop genome sequencing: lessons and rationales. Trends Plant Sci 16:77–88. doi: 10.1016/j.tplants.2010.10.005

Fornalé S, Capellades M, Encina A, et al. (2012) Altered lignin biosynthesis improves cellulosic bioethanol production in transgenic maize plants down-regulated for cinnamyl alcohol dehydrogenase. Mol Plant 5:817–30. doi: 10.1093/mp/ssr097

Foster CE, Martin TM, Pauly M (2010) Comprehensive compositional analysis of plant cell walls (lignocellulosic biomass) part II: carbohydrates. J. Vis. Exp.

Fu C, Xiao X, Xi Y, et al. (2011) Downregulation of Cinnamyl Alcohol Dehydrogenase (CAD) Leads to Improved Saccharification Efficiency in Switchgrass. BioEnergy Res 4:153–164. doi: 10.1007/s12155-010-9109-z

Fukushima RS, Hatfield RD (2004) Comparison of the acetyl bromide spectrophotometric method with other analytical lignin methods for determining lignin concentration in forage samples. J Agric Food Chem 52:3713–20. doi: 10.1021/jf035497l

Grabber J, Ralph J (1997) p-Hydroxyphenyl, guaiacyl, and syringyl lignins have similar inhibitory effects on wall degradability. J Agric … 8561:2530–2532.

Grabber JH, Hatfield RD, Lu F, Ralph J (2008) Coniferyl ferulate incorporation into lignin enhances the alkaline delignification and enzymatic degradation of cell walls. Biomacromolecules 9:2510–2516. doi: 10.1021/bm800528f

Grabber JH, Mertens DR, Kim H, et al. (2009) Cell wall fermentation kinetics are impacted more by lignin content and ferulate cross-linking than by lignin composition. J Sci Food Agric 89:122–129. doi: 10.1002/jsfa.3418

Grabber JH, Schatz PF, Kim H, et al. (2010) Identifying new lignin bioengineering targets: 1. Monolignol-substitute impacts on lignin formation and cell wall fermentability. BMC Plant Biol 10:114.

Gressel J (2008) Transgenics are imperative for biofuel crops. Plant Sci 174:246–263. doi: 10.1016/j.plantsci.2007.11.009

Guillaumie S, San-Clemente H, Deswarte C, et al. (2007) MAIZEWALL. Database and developmental gene expression profiling of cell wall biosynthesis and assembly in maize. Plant Physiol 143:339–363.

Halpin C, Holt K, Chojecki J, et al. (1998) Brown-midrib maize (bm1)–a mutation affecting the cinnamyl alcohol dehydrogenase gene. Plant … 14:545–53. doi: 10.1046/j.1365-313X.1998.00153.x

Hamberger B, Ellis M, Friedmann M, et al. (2007) Genome-wide analyses of phenylpropanoid-related genes in Populus trichocarpa, Arabidopsis thaliana, and Oryza sativa: the Populus

lignin toolbox and conservation and diversification of angiosperm gene familiesThis article is one of a selection of papers p. Can J Bot 85:1182–1201. doi: 10.1139/B07-098

Han K-J, Pitman WD, Kim M, et al. (2013) Ethanol production potential of sweet sorghum assessed using forage fiber analysis procedures. GCB Bioenergy 5:358–366. doi: 10.1111/j.1757-1707.2012.01203.x

Harris D, DeBolt S (2010) Synthesis, regulation and utilization of lignocellulosic biomass. Plant Biotechnol J 8:244–62. doi: 10.1111/j.1467-7652.2009.00481.x

Harris PJ, Trethewey J a. K (2010) The distribution of ester-linked ferulic acid in the cell walls of angiosperms. Phytochem Rev 9:19–33. doi: 10.1007/s11101-009-9146-4

Hatfield R, Fukushima RS (2005) Can Lignin Be Accurately Measured? Crop Sci 45:832. doi: 10.2135/cropsci2004.0238

Hatfield RD, Ralph J, Grabber JH (1999) Cell wall cross-linking by ferulates and diferulates in grasses. J Sci Food Agric 79:403–407. doi: 10.1002/(SICI)1097-0010(19990301)79:3<403::AID-JSFA263>3.0.CO;2-0

Hu WJ, Harding S a, Lung J, et al. (1999) Repression of lignin biosynthesis promotes cellulose accumulation and growth in transgenic trees. Nat Biotechnol 17:808–12. doi: 10.1038/11758

Ideker T, Galitski T, Hood L (2001) A new approach to decoding life: systems biology. Annu Rev Genomics Hum Genet 2:343–372. doi: 10.1146/annurev.genom.2.1.343

Ideker T, Winslow LR, Lauffenburger DA (2006) Bioengineering and systems biology. Ann Biomed Eng 34:1226–1233. doi: 10.1007/s10439-006-9119-3

Ishii T (1997) Structure and functions of feruloylated polysaccharides. Plant Sci 127:111–127. doi: 10.1016/S0168-9452(97)00130-1

Jakob K, Zhou F, Paterson AH (2009) Genetic improvement of C4 grasses as cellulosic biofuel feedstocks. Vitr Cell Dev Biol - Plant 45:291–305. doi: 10.1007/s11627-009-9214-x

Jonczyk R, Schmidt H, Osterrieder A, et al. (2008) Elucidation of the final reactions of DIMBOA-glucoside biosynthesis in maize: characterization of Bx6 and Bx7. Plant Physiol 146:1053–63. doi: 10.1104/pp.107.111237

Jouanin L, Goujon T, de Nadaï V, et al. (2000) Lignification in transgenic poplars with extremely reduced caffeic acid O-methyltransferase activity. Plant Physiol 123:1363–74.

Jung H-JG (2003) Maize stem tissues: ferulate deposition in developing internode cell walls. Phytochemistry 63:543–549. doi: 10.1016/S0031-9422(03)00221-8

Jung H-JG, Samac D a, Sarath G (2012) Modifying crops to increase cell wall digestibility. Plant Sci 185-186:65–77. doi: 10.1016/j.plantsci.2011.10.014

Kakumanu A, Ambavaram MMR, Klumas C, et al. (2012) Effects of drought on gene expression in maize reproductive and leaf meristem tissue revealed by RNA-

Seq. Plant Physiol 160:846–67. doi: 10.1104/pp.112.200444

Kumar S, Omer S, Patel K, Khan BM (2013) Cinnamate 4-Hydroxylase (C4H) genes from Leucaena leucocephala: a pulp yielding leguminous tree. Mol Biol Rep 40:1265–74. doi: 10.1007/s11033-012-2169-8

Lam TB, Kadoya K, Iiyama K (2001) Bonding of hydroxycinnamic acids to lignin: ferulic and p-coumaric acids are predominantly linked at the benzyl position of lignin, not the beta-position, in grass cell walls. Phytochemistry 57:987–992. doi: 10.1016/S0031-9422(01)00052-8

Leplé J-C, Dauwe R, Morreel K, et al. (2007) Downregulation of cinnamoyl-coenzyme A reductase in poplar: multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. Plant Cell 19:3669–91. doi: 10.1105/tpc.107.054148

Li X, Chen W, Zhao Y, et al. (2013) Downregulation of caffeoyl-CoA O-methyltransferase (CCoAOMT) by RNA interference leads to reduced lignin production in maize straw. Genet Mol Biol 36:540–6. doi: 10.1590/S1415-47572013005000039

Liu X, Han L, Yang Z, Xu C (2008) Prediction of silage digestibility by near infrared refl ectance spectroscopy. 631–639.

Lorenz AJ, Anex RP, Isci A, et al. (2009) Forage quality and composition measurements as predictors of ethanol yield from maize (Zea mays L.) stover. Biotechnol Biofuels 2:5. doi: 10.1186/1754-6834-2-5

MacAdam JW, Grabber JH (2002) Relationship of growth cessation with the formation of diferulate cross-links and p-coumaroylated lignins in tall fescue leaf blades. Planta 215:785–93. doi: 10.1007/s00425-002-0812-7

Mansfield SD, Kang K-Y, Chapple C (2012) Designed for deconstruction--poplar trees altered in cell wall lignification improve the efficacy of bioethanol production. New Phytol 194:91–101. doi: 10.1111/j.1469-8137.2011.04031.x

Matos D a, Whitney IP, Harrington MJ, Hazen SP (2013) Cell Walls and the Developmental Anatomy of the Brachypodium distachyon Stem Internode. PLoS One 8:e80640. doi: 10.1371/journal.pone.0080640

Mochida K, Shinozaki K (2011) Advances in omics and bioinformatics tools for systems analyses of plant functions. Plant Cell Physiol 52:2017–38. doi: 10.1093/pcp/pcr153

Molinari HBC, Pellny TK, Freeman J, et al. (2013) Grass cell wall feruloylation: distribution of bound ferulate and candidate gene expression in Brachypodium distachyon. Front Plant Sci 4:50. doi: 10.3389/fpls.2013.00050

Morreel K, Dima O, Kim H, et al. (2010a) Mass spectrometry-based sequencing of lignin oligomers. Plant Physiol 153:1464–78. doi: 10.1104/pp.110.156489

Morreel K, Kim H, Lu F, et al. (2010b) Mass spectrometry-based fragmentation as an identification tool in lignomics. Anal Chem 82:8095–8105. doi: 10.1021/ac100968g

Morrison T, Kessler J (1994) Activity of two lignin biosynthesis enzymes during development of a maize internode. J … 133–139.

Morrow S, Mascia P, Self K, Altschuler M (1997) Molecular characterization of a brown midrib3 deletion mutation in maize. Mol Breed 351–357.

Northcote D (1989) Control of plant cell wall biogenesis: an overview. Plant cell wall Polym. Biog. …

Novaes E, Kirst M, Chiang V, et al. (2010) Lignin and biomass: a negative correlation for wood formation and lignin content in trees. Plant Physiol 154:555–61. doi: 10.1104/pp.110.161281

Olsen KM, Wendel JF (2013) A bountiful harvest: genomic insights into crop domestication phenotypes. Annu Rev Plant Biol 64:47–70. doi: 10.1146/annurev-arplant-050312-120048

Papa G, Varanasi P, Sun L, et al. (2012) Exploring the effect of different plant lignin content and composition on ionic liquid pretreatment efficiency and enzymatic saccharification of Eucalyptus globulus L. mutants. Bioresour Technol 117:352–9. doi: 10.1016/j.biortech.2012.04.065

Pedersen JF, Vogel KP, Funnell DL (2005) Impact of Reduced Lignin on Plant Fitness. Crop Sci 45:812. doi: 10.2135/cropsci2004.0155

Penning BW, Hunter CT, Tayengwa R, et al. (2009) Genetic resources for maize cell wall biology. Plant Physiol 151:1703–28. doi: 10.1104/pp.109.136804

Piquemal J, Chamayou S, Nadaud I, et al. (2002) Down-regulation of caffeic acid O-methyltransferase in maize revisited using a transgenic approach. Plant … 130:1675–1685. doi: 10.1104/pp.012237.rides

Ralph J, Hatfield RD, Quideau S, et al. (1994) Pathway of p-Coumaric Acid Incorporation into Maize Lignin As Revealed by NMR. J Am Chem Soc 116:9448–9456. doi: 10.1021/ja00100a006

Riboulet C, Guillaumie S, Méchin V (2009) Kinetics of phenylpropanoid gene expression in maize growing internodes: relationships with cell wall deposition. Crop … 211–223. doi: 10.2135/cropsci2008.03.0130

Ro DK, Mah N, Ellis BE, Douglas CJ (2001) Functional Characterization and Subcellular Localization of Poplar (populus trichocarpa x populus deltoides) cinnamate 4-hydroxylase. Plant Physiol 126:317–329.

Rosler J, Krekel F, Amrhein N, et al. (1997) Maize phenylalanine ammonia-lyase has tyrosine ammonia-lyase activity. Plant Physiol 113:175–179.

Sarath G, Dien B, Saathoff AJ, et al. (2011) Ethanol yields and cell wall properties in divergently bred switchgrass genotypes. Bioresour Technol 102:9579–85. doi: 10.1016/j.biortech.2011.07.086

Schilmiller AL, Stout J, Weng J-K, et al. (2009) Mutations in the cinnamate 4-hydroxylase gene impact metabolism, growth and development in Arabidopsis. Plant J 60:771–782. doi: 10.1111/j.1365-313X.2009.03996.x

Schnable PS, Ware D, Fulton RS, et al. (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326:1112–5. doi: 10.1126/science.1178534

Sewalt V, Ni W, Blount JW, et al. (1997) Reduced Lignin Content and Altered Lignin Composition in Transgenic Tobacco Down-Regulated in Expression of L-Phenylalanine Ammonia-Lyase or Cinnamate 4-Hydroxylase. Plant Physiol 115:41–50.

Sharma M, Chai C, Morohashi K, et al. (2012) Expression of flavonoid 3'-hydroxylase is controlled by P1, the regulator of 3-deoxyflavonoid biosynthesis in maize. BMC Plant Biol 12:196. doi: 10.1186/1471-2229-12-196

Shen H, Mazarei M, Hisano H, et al. (2013) A genomics approach to deciphering lignin biosynthesis in switchgrass. Plant Cell 25:4342–61. doi: 10.1105/tpc.113.118828

Shi C, Koch G, Ouzunova M, et al. (2006) Comparison of maize brown-midrib isogenic lines by cellular UV-microspectrophotometry and comparative transcript profiling. Plant Mol Biol 62:697–714. doi: 10.1007/s11103-006-9049-3

Shi R, Sun Y-H, Li Q, et al. (2010) Towards a systems approach for lignin biosynthesis in Populus trichocarpa: transcript abundance and specificity of the monolignol biosynthetic genes. Plant Cell Physiol 51:144–63. doi: 10.1093/pcp/pcp175

Simmons B a, Loque D, Blanch HW (2008) Next-generation biomass feedstocks for biofuel production. Genome Biol 9:242. doi: 10.1186/gb-2008-9-12-242

Sims REH, Mabee W, Saddler JN, Taylor M (2010) An overview of second generation biofuel technologies. Bioresour Technol 101:1570–80. doi: 10.1016/j.biortech.2009.11.046

Stern M, Bach A, Calsamiglia S (1997) Alternative techniques for measuring nutrient digestion in ruminants. J Anim Sci 2256–2276.

Stewart C (2007) Biofuels and biocontainment. Nat Biotechnol 25:2006–2008.

Sticklen MB (2008) Plant genetic engineering for biofuel production: towards affordable cellulosic ethanol. Nat Rev Genet 9:433–43. doi: 10.1038/nrg2336

Tamasloukht B, Wong Quai Lam MS-J, Martinez Y, et al. (2011) Characterization of a cinnamoyl-CoA reductase 1 (CCR1) mutant in maize: effects on lignification, fibre development, and global gene expression. J Exp Bot 62:3837–48. doi: 10.1093/jxb/err077

Thimm O, Bläsing O, Gibon Y, et al. (2004) MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. Plant J 37:914–939. doi: 10.1111/j.1365-313X.2004.02016.x

Usadel B, Nagel A, Thimm O, et al. (2005) Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of coresponding genes, and comparison with known responses. Plant … 138:1195–1204. doi: 10.1104/pp.105.060459.et

Usadel B, Poree F, Nagel A, et al. (2009) A guide to using MapMan to visualize and compare Omics data in plants: a case study in the crop species, Maize. Plant Cell Environ 32:1211–29. doi: 10.1111/j.1365-3040.2009.01978.x

Valentine J, Clifton-Brown J, Hastings A, et al. (2012) Food vs. fuel: the use of land for lignocellulosic "next generation" energy crops that minimize competition with primary food production. GCB Bioenergy 4:1–19. doi: 10.1111/j.1757-1707.2011.01111.x

Van Acker R, Vanholme R, Storme V, et al. (2013) Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in Arabidopsis thaliana. Biotechnol Biofuels 6:46. doi: 10.1186/1754-6834-6-46

Van Bel M, Proost S, Wischnitzki E, et al. (2012) Dissecting plant genomes with the PLAZA comparative genomics platform. Plant Physiol 158:590–600. doi: 10.1104/pp.111.189514

Van der Weijde T, Alvim Kamei CL, Torres AF, et al. (2013) The potential of C4 grasses for cellulosic biofuel production. Front Plant Sci 4:107. doi: 10.3389/fpls.2013.00107

Van Soest PJ, Robertson JB, Lewis B a (1991) Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. J Dairy Sci 74:3583–97. doi: 10.3168/jds.S0022-0302(91)78551-2

Vanholme R, Cesarino I, Rataj K, et al. (2013) Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in Arabidopsis. Science 341:1103–6. doi: 10.1126/science.1241602

Vanholme R, Demedts B, Morreel K (2010b) Lignin biosynthesis and structure. Plant … 153:895–905. doi: 10.1104/pp.110.155119

Vanholme R, Morreel K, Darrah C, et al. (2012a) Metabolic engineering of novel lignin in biomass crops. New Phytol 196:978–1000. doi: 10.1111/j.1469-8137.2012.04337.x

Vanholme R, Morreel K, Ralph J, Boerjan W (2008) Lignin engineering. Curr Opin Plant Biol 11:278–85. doi: 10.1016/j.pbi.2008.03.005

Vanholme R, Ralph J, Akiyama T, et al. (2010c) Engineering traditional monolignols out of lignin by concomitant up-regulation of F5H1 and down-regulation of COMT in Arabidopsis. Plant J 64:885–97. doi: 10.1111/j.1365-313X.2010.04353.x

Vanholme R, Storme V, Vanholme B, et al. (2012b) A systems biology view of responses to lignin biosynthesis perturbations in Arabidopsis. Plant Cell 24:3506–29. doi: 10.1105/tpc.112.102574

Vanholme R, Van Acker R, Boerjan W (2010a) Potential of Arabidopsis systems biology to advance the biofuel field. Trends Biotechnol 28:543–7. doi: 10.1016/j.tibtech.2010.07.008

Vermerris W (2011) Survey of genomics approaches to improve bioenergy traits in maize, sorghum and sugarcane. J Integr Plant Biol 53:105–19. doi: 10.1111/j.1744-7909.2010.01020.x

Vignols F, Rigau J, Torres M (1995) The brown midrib3 (bm3) mutation in maize occurs in the gene encoding caffeic acid O-methyltransferase. Plant Cell … 7:407–416.

Voelker SL, Lachenbruch B, Meinzer FC, et al. (2010) Antisense down-regulation of 4CL expression alters lignification, tree growth, and saccharification potential of field-grown poplar. Plant Physiol 154:874–86. doi: 10.1104/pp.110.159269

Vogel KP, Mitchell RB, Sarath G, et al. (2013) Switchgrass Biomass Composition Altered by Six Generations of Divergent Breeding for Digestibility. Crop Sci 53:853. doi: 10.2135/cropsci2012.09.0542

Waclawovsky AJ, Sato PM, Lembke CG, et al. (2010) Sugarcane for bioenergy production: an assessment of yield and regulation of sucrose content. Plant Biotechnol J 8:263–76. doi: 10.1111/j.1467-7652.2009.00491.x

Wagner GJ, Hrazdina G (1984) Endoplasmic reticulum as a site of phenylpropanoid and flavonoid metabolism in hippeastrum. Plant Physiol 74:901–6.

Winkel-Shirley B (1999) Evidence for enzyme complexes in the phenylpropanoid and flavonoid pathways. Physiol Plant 142–149.

Xu F, Sun R-C, Sun J-X, et al. (2005) Determination of cell wall ferulic and p-coumaric acids in sugarcane bagasse. Anal Chim Acta 552:207–217. doi: 10.1016/j.aca.2005.07.037

Zeng Y, Zhao S, Yang S, Ding S-Y (2014) Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. Curr Opin Biotechnol 27:38–45. doi: 10.1016/j.copbio.2013.09.008

Zheng Q, Wang X-J (2008) GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis. Nucleic Acids Res 36:W358–63. doi: 10.1093/nar/gkn276

8. Supplemental Figures and Tables

Supplementary Table 4. Full list of estimated NIRS parameters.

NIRS Parameter	units	control	*zmc4h1*	fold change (%)	p-value
Acid detergent Fiber (ADF)	% of Dry matter	32.80	32.93	0.4	0.844
Neutral Detergent Fiber (NDF)	% of Dry matter	54.46	56.78	4.3	**0.004**
acid detergent lignin (ADL)	% of Dry matter	4.01	3.58	-10.7	**0.007**
Raw cellulose rate (CB)	% of Dry matter	28.04	27.95	-0.3	0.876
Organic matter content (MO)	% of Dry matter	0.93	0.92	-1.2	**0.047**
analytic dry matter (MSA)	gram	94.32	94.09	-0.3	0.256
% Dry matter digestibility Aufrère enzymatic method (CASEAUF)	% of Dry matter	50.96	53.97	5.9	**0.008**
% Organic matter digestibility Aufrère enzymatic method (CASEMOAUF)	% of Dry matter	48.24	51.37	6.5	**0.011**
% Dry matter digestibility (DMSAUF)	% of Dry matter	58.15	59.71	2.7	**0.020**
Organic matter digestibility (DMO)	% of Dry matter	62.29	64.72	3.9	**0.010**
Digestible part of the cell wal (FDP)	% of Dry matter	27.97	36.54	30.6	**0.000**
Soluble sugars rate (SSR)	% of Dry matter	15.42	13.86	-10.1	**0.014**
In Vitro Digestibility of the "Non-Starch, soluble Carbohydrates part" (DINAG)	% of Dry matter	42.02	46.57	10.8	**0.002**
Dry matter protein rate (MPT)	% of Dry matter	7.34	8.22	12.0	0.174
energy value for milk production /kg of dry matter (UFLMS)	French Milk Feed Unit for Energy	0.71	0.75	4.8	**0.016**
energy value for meat production /kg of dry matter (UFVMS)	French Meat Feed Unit for Energy	0.59	0.64	7.3	**0.010**
Organic matter protein rate (MATO)	% of Organic matter	78.63	89.16	13.4	0.153
ester ferulic acid rate (FE25)	% of Dry matter	6.63	6.69	1.0	0.392
ether ferulic acid rate (FE170)	% of Dry matter	8.47	8.49	0.1	0.862
5-5 diFerulic acid rate (DIF55)	% of Dry matter	0.16	0.20	20.9	**0.022**
8-O-4 diFerulic acid rate (DIF8O4)	% of Dry matter	0.30	0.34	15.0	**0.042**
ester p-coumaric acid rate (CO25)	% of Dry matter	14.76	12.83	-13.1	**0.006**
p-Hydroxybenzaldehyde. H sub-units rate (H4N)	% of Dry matter	1.66	1.36	-17.7	**0.005**
Syringaldehyde. S sub-units rate (SAN)	% of Dry matter	8.27	6.26	-24.3	**0.001**
Vanillin. G sub-units rate (VAN)	% of Dry matter	8.43	7.27	-13.7	**0.004**
quality parameter I		1.12	1.05	-6.0	0.664
quality parameter 2		1.59	1.70	7.4	0.435

Supplementary Table 5. Expression levels and representation of change in expression over development of putative phenylpropanoid genes involved in lignification of maize wild type internodes. Increasing and decreasing expression was illustrated by subtracting the mean expression per gene over development from the mean normalized expression value per gene in each stage and color code all subsequent values for higher (red) and lower (blue) expression than average. The expression level was labeled according to maximal expression: "very high" >=15, 15>"high">=13, 13>"moderate">=11, 11>"low". V10 (ten visible leaf collars), S (silking), S+7d (seven days after silking) and S+14d (fourteen days after silking).

			mean normalized expression (log2)					[(normalized expression)-(mean per gene)] (log2)			
protein name	gene name	probe name	V10	S	S+7d	S+14d	expression level	V10	S	S+7d	S+14d
PAL			14.13	15.07	15.07	14.80	very high	-0.63	0.30	0.30	0.03
PAL			14.47	14.71	15.01	14.40	very high	-0.17	0.06	0.37	-0.25
PAL			13.41	14.16	14.28	14.07	high	-0.57	0.18	0.30	0.09
PAL			14.48	14.59	14.60	14.58	high	-0.08	0.03	0.04	0.02
PAL			13.96	14.95	14.98	14.94	high	-0.75	0.24	0.27	0.23
PAL			13.37	14.81	14.66	14.41	high	-0.94	0.50	0.34	0.10
PAL			12.45	13.07	13.38	13.34	high	-0.61	0.01	0.32	0.28
PAL			10.58	10.87	10.87	10.92	low	-0.23	0.06	0.06	0.11
PAL			10.10	10.25	10.09	9.93	low	0.01	0.16	0.00	-0.17
PAL			9.43	9.33	9.56	9.37	low	0.01	-0.09	0.14	-0.05
PAL			9.70	9.68	10.17	9.67	low	-0.11	-0.12	0.36	-0.13
C4H			12.99	13.24	13.33	13.43	high	-0.26	0.00	0.08	0.18
C4H			11.58	11.57	11.89	11.80	moderate	-0.13	-0.14	0.18	0.09
C4H			9.76	9.90	9.98	9.88	low	-0.12	0.02	0.10	0.00
4CL			12.45	13.04	12.87	11.58	high	-0.03	0.56	0.39	-0.91
4CL			11.32	11.12	11.18	11.50	moderate	0.04	-0.16	-0.10	0.22
4CL			12.16	12.85	12.59	11.55	moderate	-0.13	0.56	0.30	-0.74
4CL			11.73	10.54	10.48	10.38	moderate	0.94	-0.24	-0.30	-0.40
4CL			11.47	11.64	11.51	11.58	moderate	-0.08	0.09	-0.04	0.03
4CL			9.12	9.32	9.23	9.26	low	-0.11	0.09	0.00	0.03
4CL			9.42	9.48	9.43	9.46	low	-0.03	0.03	-0.02	0.02
4CL			10.87	9.91	9.88	9.72	low	0.78	-0.19	-0.21	-0.38
HCT			10.87	12.38	12.24	11.93	moderate	-0.98	0.53	0.38	0.08
HCT			11.52	12.80	12.46	12.35	moderate	-0.76	0.51	0.18	0.07
HCT			11.64	12.54	12.32	12.40	moderate	-0.58	0.31	0.09	0.18
C3H			13.72	13.77	13.83	13.77	high	-0.05	0.00	0.05	0.00
C3H			10.95	10.04	10.03	9.73	low	0.76	-0.15	-0.16	-0.46
CCoAOMT			12.89	14.54	14.57	13.90	high	-1.08	0.56	0.60	-0.08
CCoAOMT			12.82	14.27	14.28	13.65	high	-0.93	0.51	0.52	-0.10
CCoAOMT			12.63	14.32	14.07	13.64	high	-1.03	0.65	0.40	-0.02
CCoAOMT			12.58	13.62	13.28	13.10	high	-0.56	0.48	0.13	-0.05
CCoAOMT			12.19	13.38	13.43	13.26	high	-0.87	0.32	0.36	0.19
CCoAOMT	Hidden		12.10	13.44	13.32	13.25	high	-0.93	0.41	0.30	0.22
CCoAOMT			12.58	13.69	13.58	13.62	high	-0.79	0.32	0.22	0.25
CCoAOMT			10.50	10.88	10.91	11.21	moderate	-0.38	0.01	0.04	0.33
CCoAOMT			11.44	12.96	13.00	12.60	moderate	-1.06	0.46	0.50	0.10
CCoAOMT			9.92	11.21	11.12	10.95	moderate	-0.88	0.41	0.32	0.15
CCoAOMT			9.32	9.43	9.37	9.46	low	-0.07	0.04	-0.03	0.07
CCoAOMT			10.02	10.00	10.06	10.02	low	0.00	-0.02	0.04	-0.01
CCR			13.45	12.06	12.23	12.20	high	0.97	-0.42	-0.26	-0.29
CCR			14.38	13.21	13.25	13.36	high	0.83	-0.34	-0.30	-0.19
CCR			12.22	13.46	13.17	12.84	high	-0.70	0.54	0.25	-0.08
CCR			11.19	10.24	10.15	10.27	moderate	0.73	-0.22	-0.31	-0.19
CCR			12.18	10.52	10.67	10.55	moderate	1.20	-0.46	-0.31	-0.43
CCR			11.36	11.92	12.01	11.96	moderate	-0.45	0.11	0.19	0.14
CCR			12.64	12.59	12.33	12.38	moderate	0.16	0.10	-0.15	-0.11
CCR			10.35	9.42	9.24	9.41	low	0.74	-0.19	-0.36	-0.19
CCR			9.15	10.26	10.29	9.80	low	-0.73	0.39	0.42	-0.07
CCR			9.68	10.35	10.18	9.99	low	-0.37	0.30	0.13	-0.06
CCR			9.25	9.31	9.11	9.21	low	0.03	0.09	-0.11	-0.01
CCR			9.53	9.60	9.74	9.76	low	-0.13	-0.06	0.08	0.10
CCR			10.11	9.91	9.64	9.47	low	0.33	0.12	-0.14	-0.31
F5H			9.79	11.10	11.12	11.63	moderate	-1.12	0.19	0.21	0.72
F5H			9.61	10.51	10.71	10.98	low	-0.84	0.06	0.26	0.53
COMT			13.32	14.51	14.29	14.15	high	-0.75	0.44	0.22	0.09
COMT			12.64	13.73	13.28	13.35	high	-0.61	0.48	0.03	0.11
CAD			13.73	12.77	12.65	12.42	high	0.84	-0.12	-0.25	-0.47
CAD			13.63	12.43	12.37	11.87	high	1.05	-0.15	-0.20	-0.70
CAD			12.66	14.16	13.98	13.50	high	-0.91	0.58	0.41	-0.07
CAD			12.55	12.54	12.41	12.68	moderate	0.00	0.00	-0.14	0.14
CAD			10.38	10.48	10.59	10.48	low	-0.10	-0.01	0.11	0.00
CAD			9.98	9.75	9.60	9.67	low	0.23	0.00	-0.15	-0.08
CAD			9.42	10.60	10.29	10.00	low	-0.66	0.53	0.21	-0.08
CAD			9.59	9.67	9.72	9.64	low	-0.06	0.01	0.06	-0.01
CAD			9.45	9.87	9.76	9.70	low	-0.25	0.17	0.07	0.01
CAD			9.94	10.15	10.02	10.32	low	-0.17	0.04	-0.09	0.21

Hidden

	1	2	3
1		3	17
2	93,77		16
3	78,71	79,14	

Supplementary figure 6. Pairwise comparison of maize C4H coding sequences. Below the diagonal the percentage identity and above the number of gaps in the alignment are shown. Analysis was performed using CLC genomics workbench (CLC Bio, Denmark). [redacted], [redacted] and [redacted] were named transcripts of *ZmC4H2*, *ZmC4H1* and *ZmC4H3* respectively in the text.

Addendum to Chapter 5 and 6: Material and methods of maize lignin mutants

Addendum to Chapter 5 and 6: Material and methods of maize lignin mutants

Addendum to chapter 5 and 6: Material and methods of maize lignin mutants

1. Plant material

The maize *ZmC4H1*, *ZmCAD2*, and *ZmCCR1* transposon insertion mutants originate from a population that was built by crossing four Limagrain hybrid lines as female with two Mutator lines (MuTaylor and Mu Schnable) as male. A crossing scheme to clarify the subsequent steps is provided in Figure 45. Selfing of the G1 generation resulted in a G2 generation, in which lines with insertions in *ZmC4H1*, *ZmCAD2*, and *ZmCCR1* were identified. These were then crossed with a 'Limagrain elite line' (of which the name falls under IP rights by Biogemma) and additionally five times backcrossed with the 'Limagrain elite line' as recurrent parent. Individuals that were homozygous and azygous for the transposon insertion were then selfed two times to ensure enough biological replicates for the field plot. Given the number of backcrosses, the mutant and control plants thus share 98.4 % of their genetic background. Summer nursery was conducted at Clermont-ferrand, France and winter nursery in Graneros, Chile.

Plant material for expression profiling, metabolite profiling and cell wall characterization was taken on a field trial conducted in 2009 at Clermont-ferrand, France. The field trial was designed in five blocks, four for each developmental stage in expression and metabolite profiling and one for the cell wall characterization at silage stage (Figure 46). For each line, genotypes homozygous for presence and absence of the transposon insertion were sown, resulting in six objects. The first four blocks contained six plots, consisting of 20-25 plants, per object and the last block 2 plots per object. In total, thus 26 plots were present per object/genotype, making 156 plots in total. The ear internode (internode below the cob) was harvested in four developmental stages: V10 (ten visible leaf collars stage), S (silking stage), S+7d (seven days after silking stage) and S+14d (14 days after silking stage). The harvest dates are depicted in Table 8. One biological sample contained multiple ear internodes that were pooled from a random set of plants within one block; ten individual ear internodes fore V10 stage and five the other stages. Per genotype and per developmental stage, six biological replicates were taken. These samples were kept frozen at all times, ground using a tissuelyzer and aliquots were taken for transcriptome (4 of the six replicates) and metabolome analysis (all six replicates). Whole plants without the cob were harvested at silage stage (corresponding to kernel milk stage), three individuals per genotype and this in two replicates, which makes six biological replicates in total per genotype. Oven-dried and ground material was used for cell wall, saccharification efficiency and near infrared reflectance spectroscopy (NIRS) forage quality analysis.

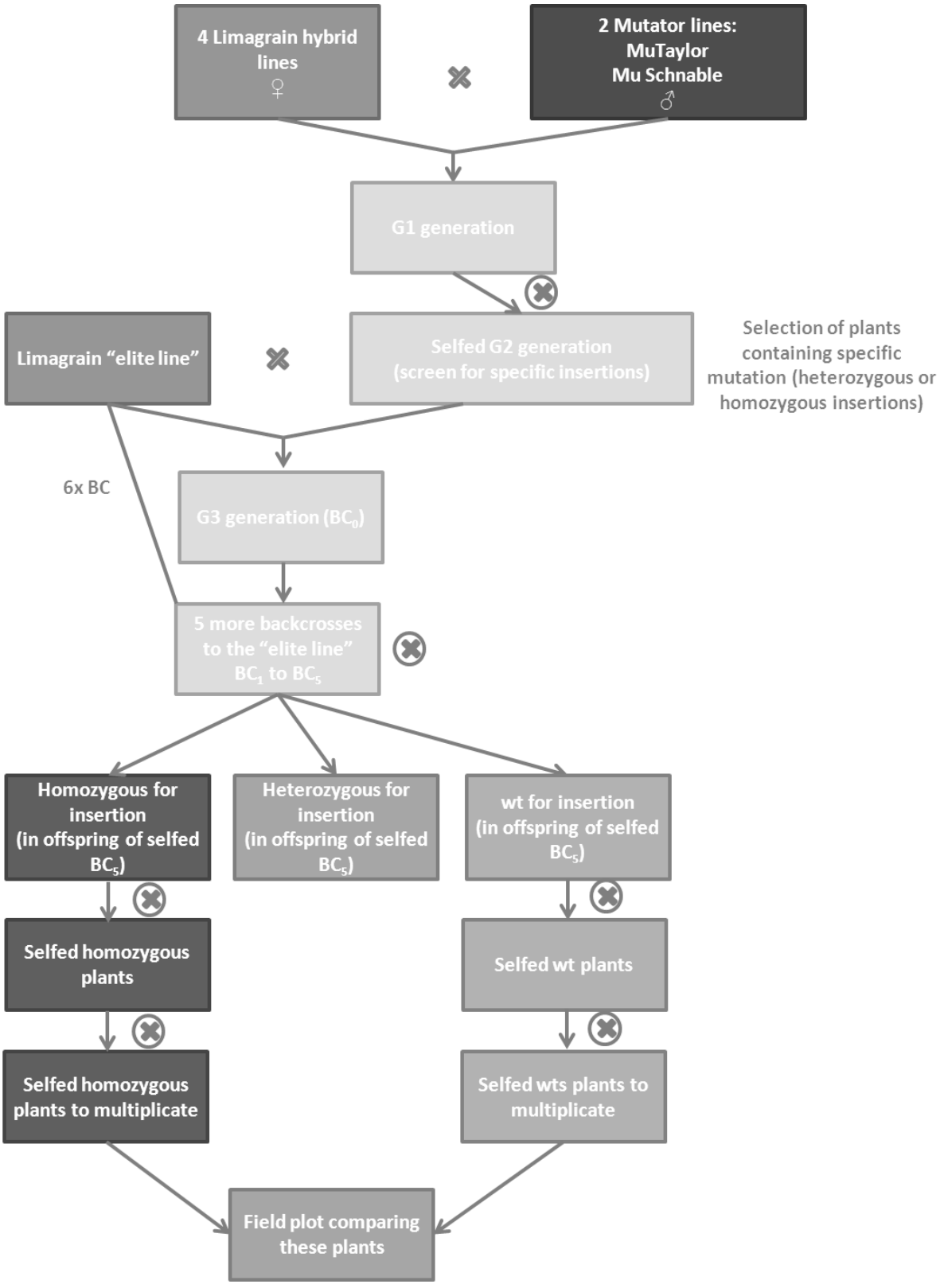


Figure 45. Scheme depicting the origin of the transposon insertional mutants for lignin biosynthetic genes selected by the Biogemma company

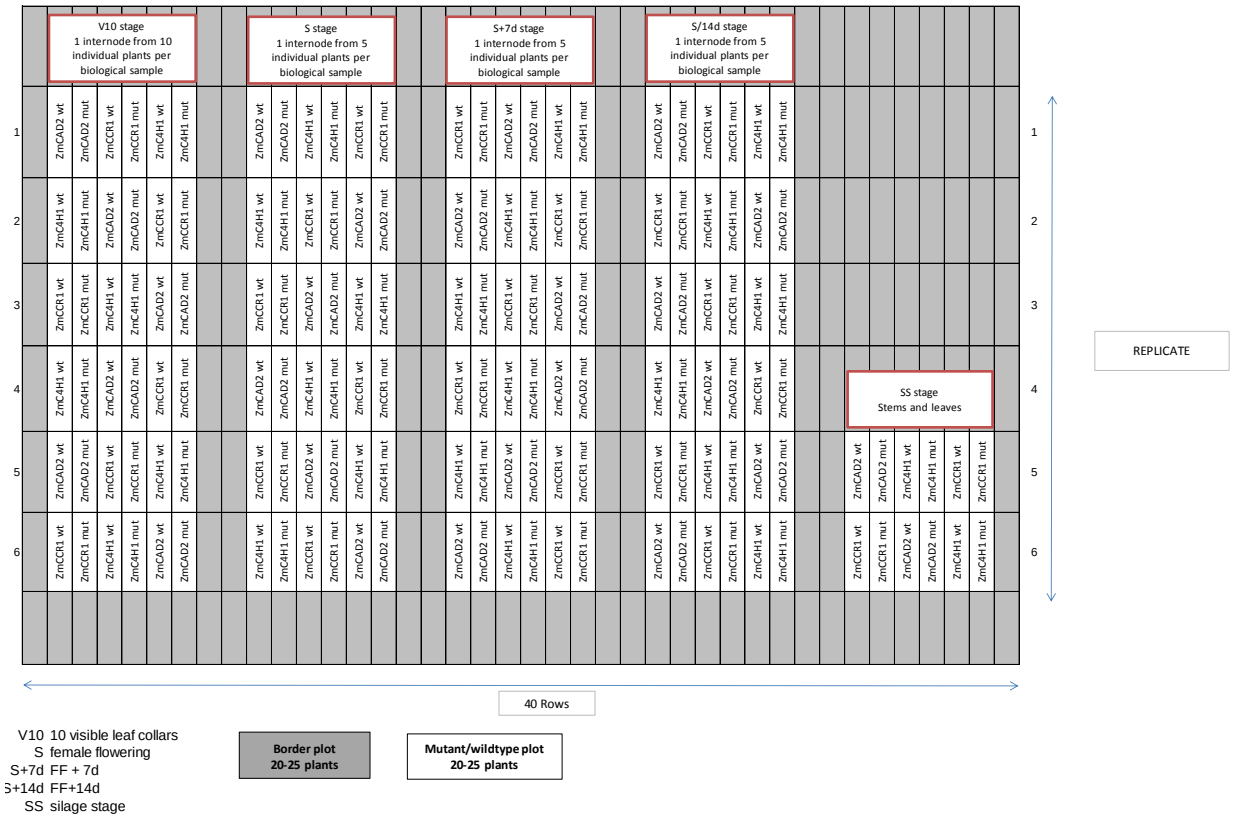


Figure 46. Field plot design for harvesting the ear internode in four developmental stages and whole stems and leaves at silage stage as communicated by Biogemma.

Table 8. Harvest dates and number of biological replicates of the ear internodes at the four developmental stages: V10 (ten visible leaf collars stage), S (silking stage), S+7d (seven days after silking stage) and S+14d (14 days after silking stage) as communicated by Biogemma.

line	V10	S	S+7d	S+14d
	July 7	August 2	August 9	August 16
zmc4h1	6	6	6	6
	July 7	July 26	August 2	August 9
WT-control zmc4h1	6	6	5	6
	July 7	August 9	August 16	August 23
Zmccr1	6	6	6	6
	July 7	August 9	August 16	August 23
WT-control zmccr1	6	6	6	6
	July 7	August 9	August 16	August 23
zmcad2	6	6	6	6
	July 7	August 9	August 16	August 23
WT-control zmcad2	6	6	6	6

2. Genotyping of the lignin mutants by Biogemma

The mutants were identified by screening the G2 generation (Figure 45) for transposon insertion sites using the OmuA primer, designed for binding in the terminal inverted repeat (TIR) of the transposon, to sequence the flanking sequences (Table 9). In this way, transposon insertion mutants were identified in *ZmC4H1* (in refgenv1 and v2 as corresponding B73 gene), *ZmCAD2* (in refgenv1 and in refgenv2 as corresponding B73 gene) and *ZmCCR1* (in refgenv1 and v2 as corresponding B73 gene). For *ZmC4H1*, the insertion occured between position 2279 and 2280 in the second intron (Figure 47). The insertion in *ZmCAD2* occured in the fourth exon between position 3499 and 3500 on the refgenv1 model and between position 3497 and 3498 on the refgenv2 model (Figure 47). For *ZmCCR1*, the insertion occured in the first intron, which was an indel polymorphic region. This made it difficult to determine the exact location, but it was determined as situated between position 545 and 555 (Figure 47). For the follow up of the mutations in the introgressions into the 'Limagrain elite line', the molecular screening was conducted with The OmuA primer binding the TIR of the transposon and two primers in the gene of interest, one upstream and one downstream of the insertion (Table 9).

Table 9. Primer sequences for the molecular screening of the transposon insertion mutations in *ZmC4H1*, *ZmCAD2* and *ZmCCR1*

Gene	OmuA	forward (upstream)	reverse (downstream)
ZmCAD2	Hidden		
ZmC4H1			
ZmCCR1			

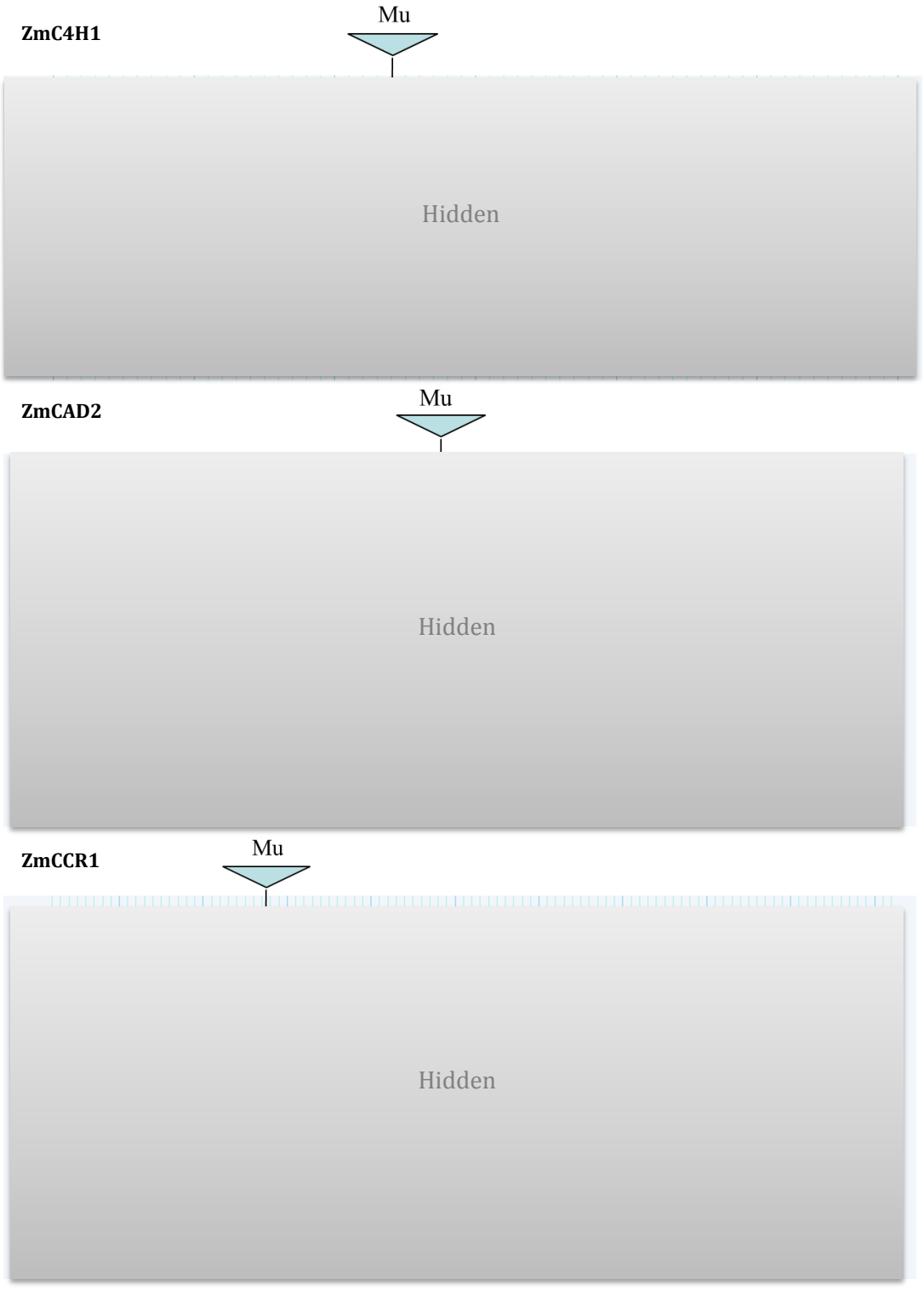


Figure 47. Positions of the Mu transposon insertion sites in *ZmC4H1*, *ZmCAD2* and *ZmCCR1* genes as determined by sequencing (communicated by Biogemma). The insertion site was marked on a genome view using a genome browser such as maizeGDB.org

3. NIRS estimation by Limagrain

After drying the plant samples in a ventilated oven (65°C), the dry samples were ground with a hammer mill to pass through a 1 mm screen. Soluble carbohydrates (SSR) (Lila 1977), Neutral Detergent Fiber (NDF), Acid Detergent Fiber (ADF) and Acid Detergent Lignin (ADL) contents were estimated according to Goering & Soest, (1970). Raw cellulose rate (CB) was determined according to the Weende method (Henneberg and Stohman 1860). The percentage of dry matter (CASEAUF) and organic matter digestibility (CASEMOAUF), the digestible part of the cell wall (FDP) and the organic matter content were all determined based on the method of Aufrère (1982). In parallel, dry matter (DMSAUF) and organic matter digestibility (DMO) were also determined according to the methods of Dardenne et al. (1993) and Barrière et al. (2004). The *in vitro* digestibility of the "non-starch, soluble carbohydrates part" (DINAG) was determined according to the method of Argillier et al. (1995). *p*-Hydroxycinnamic acid contents were measured after treating NDF fractions with NaOH according to the double procedure previously described by Morrison et al. (1993) and used by Méchin and Argillier (2000). This procedure involves a mild alkaline treatment allowing the release of esterified FA (esterFA) and *p*-coumaric acid (*p*CA), and a severe alkaline treatment allowing the release of etherified FA (etherFA) and *p*CA. Because most *p*CA is esterified (Ralph et al. 1994; Hatfield et al. 1999), only esterified *p*CA content was investigated. The concentration of etherFA was calculated as the difference between FA amounts released by the severe and mild alkaline treatments, as all etherFA is also involved in esterified linkages. Besides esterFA and etherFA contents, two FA dimers were also reported as the content in the 5-5 and 8-O-4 FA dimers. The latter is the predominant FA dimer out of the six shown in maize cell wall (Lindsay and Fry 2008). Oxidation of cell wall residues with alkaline nitrobenzene was performed according to a method adapted from Roadhouse and MacDougall (1956) according to Higuchi et al. (1967). During alkaline nitrobenzene oxidation, H, G, and S monomers are oxidized into p-hydroxybenzaldehyde (pHb), vanillin (Va) and syringaldehyde (Sg), respectively, with disruption of Ca and Cb linkages. A part of FA is also oxidized into Va, and the amount of G units of lignins is therefore partly overestimated. Extracted pHb, Va, and Sg were analyzed by HPLC.

All these traits were estimated in the biological samples using robust specific near infrared reflectance spectroscopy (NIRS) using a NIRS system 6500 spectrophotometer (Foss A/S, Hillerød, Denmark), with wavelengths spaced every 4 nm from 400 to 2500 nm. NIRS calibrations, developed at INRA Lusignan, France and CRA Gembloux, Belgium for samples of plants without ear were used. Accurate r^2 values were obtained for all traits investigated in plants without ears based on nearly 1300 laboratory analyses for global traits (such as NDF, ADF, ADL) and moderate but reliable r^2 values were obtained for *p*-CA, ferulate, pHb, Va, and Sg (based on nearly 1300 laboratory analyses) and diFA estimates (using 500 laboratory analyses) (Riboulet et al. 2008; Barrière et al. 2008).

4. Lignin analysis

Aliquots of 5 mg ground stem material were subjected to a sequential extraction to obtain a purified CWR. The extractions were done in 2-ml vials, each time for 30 min, at near boiling temperatures for water (98°C), ethanol (76°C), chloroform (59°C) and acetone (54°C). the remaining CWR was dried under vacuum. Lignin was quantified according to a modified version of the acetyl bromide method

(Dence 1992), optimized for small amounts of plant tissue. The dried CWR was dissolved in 0.1 freshly made 25% acetyl bromide in glacial acetic acid and 4 µl 60% perchloric acid. The solution was incubated for 30 min at 70°C while shaking (850 rpm). After incubation, the slurry was centrifuged at 14000 rpm for 15 min. To the supernatant, 0.2 ml of 2M sodium hydroxide and 0.5 ml glacial acetic acid was added. The pellet was washed with 0.5 ml glacial acetic acid. The supernatant and the washing phase were combined and the final volume was adjusted to 2 ml with glacial acetic acid. After 20 min at room temperature, the absorbance at 280 nm was measured with a nanodrop® ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). The lignin concentrations were calculated by means of Bouguer-Lambert-Beer law: $A = \varepsilon \times l \times c$, with $\varepsilon = 20.48$ L g^{-1} cm^{-1} (Fukushima and Hatfield 2004) and $l = 0.1$ cm.

The lignin composition was investigated with thioacidolysis as previously described (Robinson et al., 2009). The monomers involved in β-O-4 ether bonds, released upon thioacidolysis, were detected with gas chromatography (GC) as their trimethylsilyl (TMS) ether derivates on a Hewlett-Packard HP 6890 Series system (Agilent, Santa Clara, SA, USA) coupled with a HP-5973 mass-selective detector. The GC conditions were as described (Robinson and Mansfield 2009). The quantitative evaluation was carried out based on the specific prominent ions for each compound. A summary of the specific ions for each specific compound can be found in Table 10. Response factors for H, G and S units were taken from Yue et al. (2012).

Table 10. List of specific prominent ions used to extract the ion-specific chromatograms and quantify the different lignin units, released during thioacidolysis.. Target ions and qualifiers are m/z values

Compound	Target ion	qualifier 1	qualifier 2	qualifier 3
H	239	205	179	-
G	269	235	209	418
S	299	265	239	448
β-O-4 FA-I	338	323	308	249
β-O-4 FA-II	339	192	207	385
G aldehydes	293	354	-	-
S aldehydes	323	384	-	-

5. Saccharification assay

Aliquots of 10 mg of dry stem material was used. The biomass was either pretreated with 1 ml of 1M HCl at 80°C for 2h, while shaking (850 rpm) or not pretreated. The extract was removed and the pretreated material was washed three times with water to obtain a neutral pH. Subsequently, the material was incubated with 1 ml 70% (v/v) ethanol overnight at 55°C. The remaining biomass was washed three times with 1 ml 70% (v/v) ethanol, once with 1 ml acetone, and dried under vacuum

for 45 min and weighed. The pretreated ethanol-extracted residue was dissolved in 1 ml acetic acid buffer solution (pH 4.8) and incubated at 50°C. The enzyme mix added to the dissolved material contained cellulase from *Trichoderma reesei* ATCC 26921 and β-glucosidase (Sigma-Aldrich, St. Louis, MO, USA) in a 5:3 ratio. Both enzymes were first desalted over an Econo-Pac 10DG column (Bio-Rad, Hercules, CA, USA), stacked with Bio-gel® P-6DG gel (Bio-Rad) according to the manufacturer's guidelines. The desalted *β*-glucosidase was 350-fold diluted prior to mixing with desalted cellulase. The enzyme mix was further diluted 10-fold and the activity of the diluted enzyme mix was measured with a filter paper assay (Xiao et al., 2004). To each biological sample, dissolved in acetic acid buffer (pH 4.8), the enzyme mix with an activity of 0.002 filter paper units was added. After a short spinning to remove droplets from the lid of the reaction tubes, 20 µl aliquots of the supernatant were taken after 0h, 4h, 7h, 24h and 48h incubation at 50°C and 10 fold diluted with acetic acid buffer (pH4.8). The concentration of glucose in these diluted samples was measured indirectly with a spectrophotometric color reaction (glucose oxidase-peroxidase; GOD-POD) A 100 ml aliquot of the reaction mix from this color reaction contained 50 mg ABTS (Roche Diagnostics, Brussels, Belgium), 44.83 mg GOD (Sigma-Aldrich, St. Louis, MO, USA) and 173 µl of 4% (w/v) POD (Roche Diagnostics, Brussels, Belgium) in acetic acid buffer (pH 4.5). To measure the concentration of glucose, 50 µl of the diluted samples was added to 150 µl GOD_POD solution and incubated for 30 min at 37 °C. The absorbance was measured spectrophotometrically at a wavelength of 405 nm. The concentration in the original sample was calculated with a standard curve based on known D-glucose (Sigma-Aldrich) concentrations.

6. Expression analysis of ear internodes

RNA extraction was performed on 144 ear internode samples in total: six biological repeats in four developmental stages (V10 stage, siliking stage, seven days after silking stage and 14 days after silking stage) of 6 lines: the control and mutated lines for *Zmc4h1*, *Zmcad2* and *ZmCCR1*. From the six biological repeats, four were selected based on RNA quality and quantity. Expression data was generated with a custom made Nimblegen 12-plex microarray containing 37670 probes, including probes for mitochondrial and chloroplast DNA and miRNA encoding genes. The microarray was designed on the basis of Maize genome v1 (Refgenv1, 4a53, B73 line) to be as Full Genome as possible. 95% of the 36392 genes in the Filtered Gene Set (http://www.maizegdb.org/cgi-bin/termrefs.cgi?id=2366450) were covered. Normalization was done via RMA (Robust Multi-Array Analysis), implemented in NimbleScan (Nimblegen software for Nimblegen array analysis). RMA consists of three steps: an optional background adjustment, quantile normalization and finally signal summarization via MedianPolish algorithm. The contribution of outlier probes is reduced in the reported gene expression level, which has been demonstrated to improve the sensitivity and reproducibility of microarray results. Quality control of raw and normalized data indicated good quality of the data and very good correlation between biological replicates (coefficient of determination R^2: 0.95 -> 0.99).

7. Statistical analysis

To identify differential expression in the dataset of six genotypes in four developmental stages, a two-way ANOVA was carried out using the statistical software package SAS 9.2 (SAS Institute Inc., 2008, Cary, North Carolina) according to the model: Log2(expr)=line+stage+line*stage with model

specifications REML and KW df. All F-statistics were extracted for the interaction term and multiple testing using the false discovery rate (FDR). For 19008 probes the interaction term was significant (FDR_p<0.05). The model for the remaining probes was Log2(expr)=line+stage. A total of 2385 probes showed a significant line effect (FDR_p<0.05), 9895 probes showed a significant stage effect (FDR_p<0.05). Post-hoc tests were performed to identify significant differences in gene expression between different stages within control or mutant samples and between control and mutant in each stage.

Student t-tests using SPSS Statistics 22 (IBM, USA) were performed for NIRS, acetyl bromide lignin, lignin composition and saccharification efficiency.

8. Use of MapMan to identify pathways affected by *zmc4h1* perturbation

The MapMan tool facilitates the classification of transcripts (as well as other biological entities) into hierarchical categories (known as bins) in a manner that alleviates the redundancy present in other commonly used ontologies (Usadel et al., 2009); therefore, the tool provides an additional level of analysis beyond the functional enrichment of typical Gene Ontology categories (Kakamanu et al., 2012). Using this schema, the user may view a metabolic pathway or process of interest annotated by groups of participatory entities (Maize transcripts, in this case), where each entity within a given group is represented by a discrete signal visualized using intensity of color (Thimm et al., 2004). Additionally, the tool includes a variety of relevant statistical packages (Usadel et al., 2005) and the ability to visually filter data based on user-defined statistical cutoffs. With the maize annotation used, the differentially expressed genes per sampling stage were classified into bins and could thus be located on the metabolic map in MapMan. The number of mapped probes per total number of differentially expressed probes added to MapMan is 6509 out of 6272, 8239 out of 7952, 6472 out of 6239 and 6898 out of 6657 for V10, S, S+7d and S+14d stages respectively for *zmc4h1* mutants and 925 out of 895, 5940 out of 5728, 6264 out of 5998 and 3532 out of 3404 for V10, S, S+7d and S+14d stages respectively for *zmcad2* mutants. Some of the data points may be mapped multiple times to different bins.

Within MapMap, Pageman was used to perform an overrepresentation analysis. This is a classical test to test for each class, when given the number of objects chosen, the total number of objects, and the class size, one could expect the number of objects from this class by chance. There are several different ways to investigate this; commonly the hypergeometric distribution is used. The differentially expressed transcripts in the *c4h* mutant in four developmental stages were used as input. The differentially expressed transcripts in each bin exceeding the value of 1 (log2 scale) were tested for overrepresentation with two hypergeometric tests (one for upregulated and one for downregulated transcripts) and resulting p-values were corrected for multiple testing using the false discovery rate (FDR) (Benjamini and Yekutieli 2001). In this way, a significant over- or underrepresentation in each bin can be visualized and provide a general idea of altered expression levels (Usadel et al. 2005).

9. GO enrichment analysis

Gene Ontology Enrichment Analysis Software Toolkit (GOEAST), a web-based software toolkit was used for fast identification of underlining biological relevance of the microarray results. GOEAST

discovers statistically significantly enriched GO terms among the given gene list, and provides thorough, unbiased and visible results (Zheng and Wang 2008). Since a custom array was used in this study, the GO annotation file had to be uploaded to the GOEAST web-tool (http://omicslab.genetics.ac.cn/GOEAST/php/customized_microarray.php). The GO annotation file was composed by combining two approaches to link the probe ID with a particular GO category. The ArrayIDer tool (http://www.agbase.msstate.edu/cgi-bin/tools/index.cgi) was used to identify GO categories associated with a subset of probe IDs that are also present on the GeneChip Maize Genome Array chip (BT-codes). A different subset of probes carrying filtered gene set (FGS) codes were associated with GO categories through the FGS parent gene IDs on the maizesequence.org website (http://ftp.maizesequence.org/current/functional-annotations/). The combined list provided GO terms associated with 32955 probes, which is 87% of all probes in the custom array (37669). For the remainder of the probes, mainly a set of MZ codes (5008 probes), also present on a 46 K chip from the University of Arizona, no GO terms were associated. Only the GO categories for "biological process" were considered. The probes of differentially expressed transcripts in the *zmc4h1* mutant per developmental stage were used for GO enrichment analysis. The GOEAST tool generates graphical and text results. Only the text results were used since using the graphical results for data interpretation proved to be inefficient due to the high number of enriched terms. The text results include GOID (Identifiers (GOID) used in Gene Ontology Project), Ontology (To which ontology category does the GOID belong to, namely "biological process", "cellular component" or "molecular function"), Term (the GOID term definition), Level, (the level of this GO term, defined by the longest path connecting back to the root of the GO hierarchical tree), q, (count of probes/probesets/genes associated with the listed GOID (directly or indirectly) in the dataset.), k (total number of probes/probesets/genes in your dataset), m (count of probes/probesets/genes associated with the listed GOID (directly or indirectly) on the chozen microarray platform, t (total number of probes/probesets/genes on the chozen microarray), probes (probesets/targets/genes), IDs of probes/probesets/targets/genes belong to, log_odds_ratio (Logarithm, base 2, of the odds ratio of the enrichment of the GOID. The larger this number is, the stronger the enrichment the GOID among probes in your dataset) and p (p-value of the significance for the enrichment in the dataset of the listed GOID, multiple-test adjusted false discovery rate (FDR) (Zheng and Wang 2008).

10. Metabolic profiling of ear internodes

For all 6 genotypes, the six biological replicates were used for metabolic profiling. 250 µg of ground material in liquid nitrogen was extracted with 250 µl methanol (HPLC grade) at 70°C for 10 min while shaking. After centrifugation, methanol was evaporated and samples were suspended in 300 µl water/cyclohexane (2:1) for extraction. A 15 µl sample of the aqueous phase was subjected to LC-MS analysis using Acquity ultra-performance liquid chromatography system (Waters) connected to a Synapt HDMS quadrupole time-of-flight mass spectrometer (Micromass). Chromatographic separation was performed on a Waters Acquity BEH C18 column (2.1 mm × 150 mm, 1.7 µm) with a gradient elution, with the mobile phase composed of water containing 1% acetonitrile and 0.1% formic acid (A) and acetonitrile containing 1% water and 0.1% formic acid (B). During the gradient elution, a flow rate of 350 µL min−1 was applied, with initialization at time 0 min, 5% B, 30 min, 50% B, and 33 min, 100% B. The mass spectrometry parameters were used as described by Grunewald et al. (2012). The identity of phenolic compounds was confirmed by accurate mass

measurements and fragmentation patterns using Masslynx software (Waters, Milford MA, USA). For relative quantification of each identified compound, chromatograms were integrated and aligned with Transomics software software (Waters, Milford MA, USA). The abundance of the identified compounds was calculated as the peak area.

11. References

Argillier O, Barrière Y, Hebert Y (1995) Genetic variation and selection criterion for digestibility traits of forage maize. Euphytica 82:175–184. doi: 10.1007/BF00027064

Aufrère J (1982) Etude de la prévision de la digestibilité des fourrages par une méthode enzymatique. Ann. Zootech

Barrière Y, Emile JC, Traineau R, et al. (2004) Genetic variation for organic matter and cell wall digestibility in silage maize. Lessons from a 34-year long experiment with sheep in digestibility crates. Maydica 49:115–126.

Barrière Y, Thomas J, Denoue D (2008) QTL mapping for lignin content, lignin monomeric composition, p-hydroxycinnamate content, and cell wall digestibility in the maize recombinant inbred line progeny F838×F286. Plant Sci 175:585–595. doi: 10.1016/j.plantsci.2008.06.009

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Ann Stat 29:1165–1188. doi: 10.2307/2674075

Dardenne P, Andrieu J, Barrière Y, et al. (1993) Composition and nutritive value of whole maize plants fed fresh to sheep. II. Prediction of the in vivo organic matter digestibility. Ann Zootech 42:251–270. doi: 10.1051/animres:19930302

Dence CW (1992) The Determination of Lignin. In: Stephen YL, Dence CW (eds) Methods Lignin Chem. Springer Berlin Heidelberg, pp 33–61

Fukushima RS, Hatfield RD (2004) Comparison of the acetyl bromide spectrophotometric method with other analytical lignin methods for determining lignin concentration in forage samples. J Agric Food Chem 52:3713–20. doi: 10.1021/jf035497l

Goering HK, Soest PJ Van (1970) Forage Fiber Analyses (apparatus, Reagents, Procedures, and Some Applications).

Grunewald W, De Smet I, Lewis DR, et al. (2012) Transcription factor WRKY23 assists auxin distribution patterns during Arabidopsis root development through local control on flavonol biosynthesis. Proc Natl Acad Sci U S A 109:1554–9. doi: 10.1073/pnas.1121134109

Hatfield RD, Ralph J, Grabber JH (1999) Cell wall cross-linking by ferulates and diferulates in grasses. J Sci Food Agric 79:403–407. doi: 10.1002/(SICI)1097-0010(19990301)79:3<403::AID-JSFA263>3.0.CO;2-0

Henneberg W, Stohman F (1860) Beitrage zur Begrundung einer rationellen Futterung der Weiderkäur. Heft 1.

Higuchi T, Ito Y, Kawamura I (1967) p-hydroxyphenylpropane component of

grass lignin and role of tyrosine-ammonia lyase in its formation. Phytochemistry 6:875–881. doi: 10.1016/S0031-9422(00)86035-5

Lila M (1977) Influence des modalités de séchage sur la mesure de la teneur des fourrages en éléments azotés et glucidiques. Conséquences lors des récoltes d'essais au champ. Ann Amélior Plantes 27:465–475.

Lindsay SE, Fry SC (2008) Control of diferulate formation in dicotyledonous and gramineous cell-suspension cultures. Planta 227:439–52. doi: 10.1007/s00425-007-0630-z

Méchin V, Argillier O (2000) Relationship of cell wall composition to in vitro cell wall digestibility of maize inbred line stems. J. … 580:

Morrison WH, Akin DE, Himmelsbach DS, Gamble GR (1993) Investigation of the ester- and ether-linked phenolic constituents of cell wall types of normal and brown midrib pearl millet using chemical isolation, microspectrophotometry and13C NMR spectroscopy. J Sci Food Agric 63:329–337. doi: 10.1002/jsfa.2740630311

Ralph J, Hatfield RD, Quideau S, et al. (1994) Pathway of p-Coumaric Acid Incorporation into Maize Lignin As Revealed by NMR. J Am Chem Soc 116:9448–9456. doi: 10.1021/ja00100a006

Riboulet C, Lefevre B, Denoue D, Barriere Y (2008) Genetic variation in maize cell wall for lignin content, lignin structure, p-hydroxycinnamic acid content, and digestibility in set of 19 lines at silage harvest maturity. Maydica 53:11–19.

Roadhouse EF, MacDougall D (1956) A study of the nature of plant lignin by means of alkaline nitrobenzene oxidation.

Robinson AR, Mansfield SD (2009) Rapid analysis of poplar lignin monomer composition by a streamlined thioacidolysis procedure and near-infrared reflectance-based prediction modeling. Plant J 58:706–14. doi: 10.1111/j.1365-313X.2009.03808.x

Usadel B, Nagel A, Thimm O, et al. (2005) Extension of the visualization tool MapMan to allow statistical analysis of arrays, display of coresponding genes, and comparison with known responses. Plant … 138:1195–1204. doi: 10.1104/pp.105.060459.et

Yue F, Lu F, Sun R-C, Ralph J (2012) Syntheses of lignin-derived thioacidolysis monomers and their uses as quantitation standards. J Agric Food Chem 60:922–8. doi: 10.1021/jf204481x

Zheng Q, Wang X-J (2008) GOEAST: a web-based software toolkit for Gene Ontology enrichment analysis. Nucleic Acids Res 36:W358–63. doi: 10.1093/nar/gkn276

Chapter 6: Targeting the *Cinnamyl Alcohol Dehydrogenase 2* gene in maize for improved saccharification efficiency

Chapter 6: Targeting the *Cinnamyl Alcohol Dehydrogenase 2* gene in maize for improved saccharification efficiency

Chapter 6: Targeting the *Cinnamyl Alcohol Dehydrogenase 2* gene in maize for improved saccharification efficiency

My personal contribution to this work can be summarized as follows. For the transgenic approach, I cloned the target sequences of the *ZmCAD2* gene into an expression vector, screened all the transgenic lines for CAD activity and carried out the following analyses for three selected transgenic lines: cell wall characterization, analysis of saccharification efficiency, expression and enzyme activity analysis, leaf growth analysis and biomass accumulation measurements. For the practical work, I received great assistance from technicians from ILVO, Plant Unit, Growth and Development group. The transformation of maize was performed by the VIB-PSB Chromatin and Growth Control group.

For the *zmcad2* mutant, I performed the cell wall analysis of the stem material (lignin amount, lignin composition and saccharification efficiency), the statistical analysis of the NIRS data and the mining of the transcriptome data, received from Biogemma (France). Furthermore, I performed UPLC-MS metabolite analyses of the *zmcad2* mutant and control samples. In addition, the experimental data and interpretation and discussion of the results were written down in this chapter in preparation for a paper.

Chapter 6: Targeting the *Cinnamyl Alcohol Dehydrogenase 2* gene in maize for improved saccharification efficiency

1. Abstract

CAD activity reduction is a solid strategy for improving saccharification efficiency in maize. We evaluated a transposon insertion mutant in *ZmCAD2* in the background of an elite breeding line, proprietary of Biogemma (France). These plants had reduced ADL lignin, improved *in vitro* digestibility and improved saccharification efficiency compared to the control, consistent with previous reports of *ZmCAD2* perturbation in maize. An evaluation of the systems-wide effects of this perturbation by transcriptome and metabolome profiling revealed the presence of a general stress response, besides altered gene expression and phenolic compound abundance related to lignin biosynthesis.

Alternatively, we targeted the *ZmCAD2* gene for downregulation by means of RNAi in the elite maize inbred line B104. Having reduced CAD activity as a dominant trait is ideal for gene stacking for lignocellulosic feedstock improvement. We were successful in generating 35 transgenic lines of which three selected lines had reduced CAD activity and a reduction in *ZmCAD2* transcript, depending on the tissue and developmental stage. However, no obvious lignin-related phenotype could be detected in these three selected lines, most likely due to insufficient overexpression of the hairpin using the pUBIL promoter.

2. Introduction

In the phenylpropanoid pathway three cinnamyl alcohols are produced that constitute the main building blocks for lignin synthesis. The last step in this pathway is catalyzed by cinnamyl alcohol dehydrogenase (CAD) which converts *p*-coumaraldehyde into *p*-coumaryl alcohol, coniferaldehyde into coniferyl alcohol and sinapyl aldehyde into sinapyl alcohol in an NADP-dependent manner (Mansell et al. 1974; Morrison and Kessler 1994)(see Figure 41 in Chapter 5). The CAD enzyme was one of the first enzymes of the lignin biosynthetic pathway studied and was isolated from a variety of plant species including bryophytes, pteridophytes, gymnosperms and angiosperms (Mansell et al. 1974). The genes encoding CAD are found in multi-gene families but not all are important for lignin biosynthesis. In Arabidopsis, the *CAD* family consists of nine members of which two are responsible for structural lignin biosynthesis (Raes et al. 2003; Sibout et al. 2005). Other *CADs* have expression patterns excluding a role in lignification and can have preferred substrates other than for lignin biosynthesis (Kim et al. 2004; Eudes et al. 2006). Some of them are involved in response to pathogen attack and responsiveness to stress (Kiedrowski et al. 1992; Lauter 1996; Quirino et al. 1999; Cheong et al. 2002), potentially by acting in the production of soluble phenolics and defense lignin (Eudes et al. 2006). In rice, the *CAD* family consists of 12 genes with *OsCAD2*, better known as *goldhullandinternode2* (*GH2*), as the main *CAD* gene involved in lignification (Zhang et al. 2006). In maize, the CAD family consists of seven members with *ZmCAD2* as the main *CAD* gene in stem lignification (Guillaumie et al. 2007).

Considering the importance of CAD for lignin formation in plant cell walls, mutants or transgenic plants with reduced CAD activity have been generated and thoroughly investigated in many dicot species such as Arabidopsis (Sibout et al. 2005), tobacco (Halpin and Knight 1994; Vailhé and Andrée 1998; Chabannes and Barakate 2001), alfalfa (Baucher et al. 1999), poplar (Baucher et al.

1996; Ralph et al. 2001) and eucalyptus (Valério et al. 2003) and monocot species such as maize (Halpin et al. 1998b; Fornalé et al. 2012), Sorghum (Sattler et al. 2009), switchgrass (Saathoff et al. 2011; Fu et al. 2011), rice (Zhang et al., 2006), tall fescue (Chen et al., 2007) and Brachypodium (Trabucco et al. 2013; Bouvier d'Yvoire et al. 2013). In general, CAD perturbation did not result in great reductions of lignin quantity but caused incorporation of cinnamyl aldehydes into the lignin polymer instead of cinnamyl alcohols with an altered lignin composition and ultrastructure as result (Ralph et al. 2001). This leads to an increase in the frequency of free phenolic units (Barrière and Riboulet 2007) and a higher frequency of branched structures in lignin (Bouvier d'Yvoire et al. 2013). This is of special interest as it makes lignin more reactive to alkaline or oxidative treatments, a property that can be exploited in the production of cellulosic ethanol. Thus, reducing CAD activity can be regarded as a solid strategy for improving saccharification efficiency in maize. However, the impact on lignin quantity, lignin structure, saccharification efficiency and even plant fitness depends on the residual CAD activity and on the genetic background (Halpin et al. 1998b; Pedersen et al. 2005; Vermerris et al. 2010; Fornalé et al. 2012).

According to available literature reports, reduced CAD activity does have a large effect on plant growth in species investigated up to now. Most experiments were conducted in growth chambers, greenhouses or on individual plants in the field. Field tests with maize *bm1* hybrids showed that this mutation did not affect dry matter yield or have an effect on early season growth. However, Sorghum *bmr-6* mutants displayed reduced dry matter yield (up to 32%), regrowth following mowing, height and tillering (summarized in Pedersen et al. 2005). The results were however dependent on the environment and the genetic background, as field experiments in different locations with the *bmr-6* mutation in different varieties rendered different results (Casler et al. 2003; Pedersen et al. 2005). Therefore, as the cross-talk between lignin biosynthesis and plant growth and development is not completely known, a close monitoring of plant growth and biomass production is advisable in plants in which the lignin production or its composition has been altered.

Maize plants with reduced CAD activity can be obtained by genetic transformation. Downregulation of lignin genes using RNA interference (RNAi) has been successfully applied previously in maize (Park et al. 2012; Fornalé et al. 2012; Li et al. 2013). To obtain transgenic maize plants, Agrobacterium-mediated transformation is the method of choice since it allows to generate a high number of independent events with single or low copy numbers of the inserted construct (Zhao et al. 2000; Dai et al. 2001; Shou et al. 2004; Frame et al. 2006b; Frame et al. 2006a; Ishida et al. 2007). Agrobacterium-mediated transformation is also expected to favor stable transgene expression in progeny generations (Meyer and Saedler 1996). Ideally, the genetic modifications are studied in an inbred line of agricultural value as this allows straightforward subsequent phenotypic analysis and combination of different traits by gene-stacking, facilitating the implementation into breeding programs. Although high-efficiency transformation has been achieved for the inbred line A188 (Negrotto et al. 2000) and related genotypes such as the Hi II hybrid (Huang and Wei 2005; Frame et al. 2006a), the majority of maize genotypes, including so-called elite inbreds, cannot be transformed efficiently (Ishida et al. 2007). Furthermore, the A188 inbred line is of little agronomical value (Huang and Wei 2005). Thus, improving the saccharification efficiency by downregulation of the *ZmCAD2* gene in an elite inbred line with agronomic value such as B104

(Frame et al. 2006a; Coussens et al. 2012), provides clear opportunities for breeding applications (Huang and Wei 2005).

Alternatively, to study gene functions in maize insertional mutagenesis by endogenous transposable elements is a frequently used approach (Yi et al. 2009). Mutant screens are conducted in lines containing active transposable elements (TE) of known sequence, and the TE is then used as a tag to identify the genomic DNA sequences flanking the insertion site, which correspond to the disrupted gene. Several TE systems have been extensively used and each has advantages and disadvantages. One of the two systems most commonly used in maize, is *Mutator* (*Mu*) (Brutnell 2002). *Mu* is a high-copy system with typically 50 to 200 copies per individual genome (Walbot and Warren 1988). Forward mutation rates are high, making mutant generation efficient. As a consequence, *Mu* has been used widely for both forward and reverse genetics (Bensen et al. 1995; May et al. 2003; Fernandes et al. 2004; Settles et al. 2004; McCarty et al. 2005; Settles et al. 2007; Williams-Carrier et al. 2010; McCarty et al. 2013). Although identification of the particular *Mu* element responsible for the mutation of interest can be challenging, high-throughput nested PCR-based approaches followed by sequencing have been developed (Settles et al. 2007; Yi et al. 2009). Recently, next generation sequencing-based screening (Mu-seq) has advanced the applicability of *Mu* even further (McCarty et al. 2013). The UniformMu Transposon Resource database contains information of a maize *Mu* population in the W22 genetic background that is publically available and searchable at the Maize Genomics Database website (MaizeGDB.org). In addition, the *Mu* system has been used by breeding companies in which the mutated genes are introgressed in elite breeding lines (Brutnell 2002; Barrière et al. 2013). Thus, *Mu* could form an efficient system to study genetic perturbations of biosynthetic pathways such as lignin formation.

3. Objectives

CAD activity reduction is a solid strategy for improving saccharification efficiency in maize. Expected effects of an alteration of the lignin quantity and its structure depend on the residual CAD activity and the genetic background. We targeted the *ZmCAD2* gene for downregulation by means of RNAi in the elite maize inbred line B104. The effects of downregulation of *ZmCAD2* on lignin quantity, lignin structure and saccharification efficiency were investigated as well the effect on overall plant growth. Alternatively, we evaluated a transposon insertion mutant in *ZmCAD2* for cell wall characteristics and saccharification efficiency. The phenotypic consequences of the mutation in *ZmCAD2* were evaluated in the background of an elite breeding line, proprietary of Biogemma (France). Additionally, the systems-wide effects of *ZmCAD2* perturbation on the transcriptome and metabolome level was studied to obtain further insight in the metabolic framework of lignin biosynthesis.

4. Results

4.1 CAD activity reduction in inbred line B104 using a transgenic approach

The maize main *CAD* gene that is involved in internode lignification is *ZmCAD2* and was targeted for downregulation using RNAi. Three regions in the *ZmCAD2* cDNA sequence were selected with high specificity for *ZmCAD2* to minimize possible off-targets, and cloned into the pBb7GW-I-WG-UBIL vector as a hairpin construct (Figure 35A and B). The three constructs, each thus targeting a

different region of the *ZmCAD2* gene, were named RNAi1, RNAi2 and RNAi3. By means of Agrobacterium-mediated transformation of immature embryos, multiple independent transgenic lines were obtained for the three RNAi constructs; 8, 17 and 10 for RNAi1, RNAi2 and RNAi3 respectively (Supplementary Table 6). A fifty-fifty segregation of hemizygous transgenic and azygous is expected for the progeny of a backcross between a primary transformant and the wildtype B104 inbred line, in the case of a single locus T-DNA insertion. For 7 out of the 35 transgenic lines a segregation pattern was observed that deviated significantly from this expected values ($p>0.05$, Supplementary Table 6). All transgenic lines for which at least four transgenic or control pants were present (Supplementary Table 6), regardless of the segregation pattern, were screened for reduced CAD activity, as described by Fornalé et al. (2012), six weeks after sowing. CAD activity varied from not significantly different to a maximal reduction of 75% in line 107-18 which was transformed with RNAi3 (Figure 49).

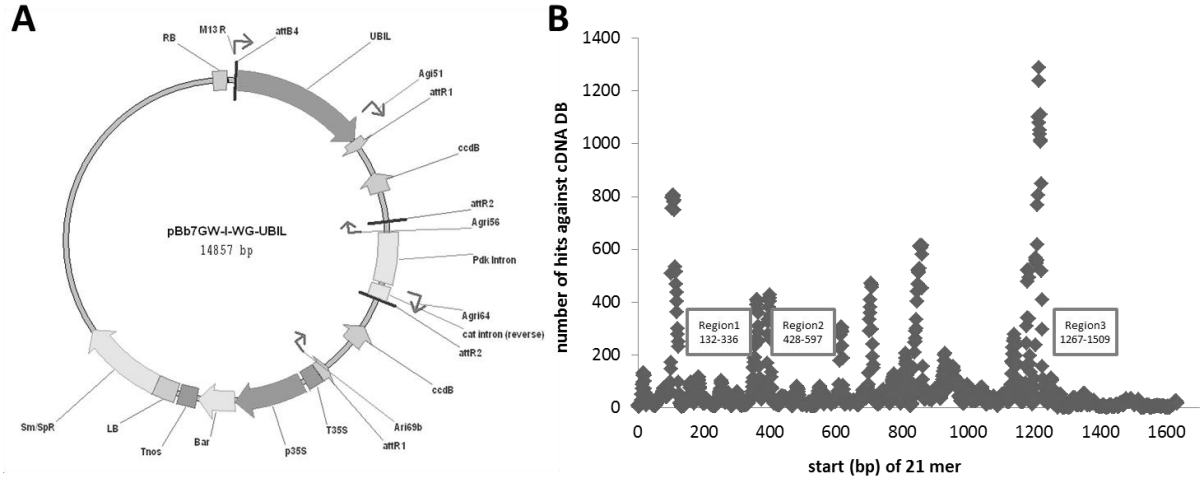


Figure 48. Monocot-specific expression vector (A) and the selection of three regions in the *ZmCAD2* cDNA sequence (B) for RNAi-mediated downregulation of the ZmCAD2 gene in maize. The hairpin vector is available at http://gateway.psb.ugent.be/; RB: right border, UBIL: maize ubiquitin promoter, attR1 and attR2: recognition sites for site-specific recombination with the target region of the gene of interest, T35S: CaMV 35S terminator, p35S: CaMV 35S promoter, Bar: basta resistance gene, Tnos: nopaline synthesis terminator, LB: left border, Sm/SpR: spectinomycin resistance gene. The graph in B shows the blast output of all possible 21-mers against the maize cDNA database (ZmAGPv1). The number of hits from the blast output is shown and does not indicate necessarily perfect matches. Three regions were selected based upon lowest number of hits indicating specificity for the gene of interest thereby reducing the number of possible off-targets.

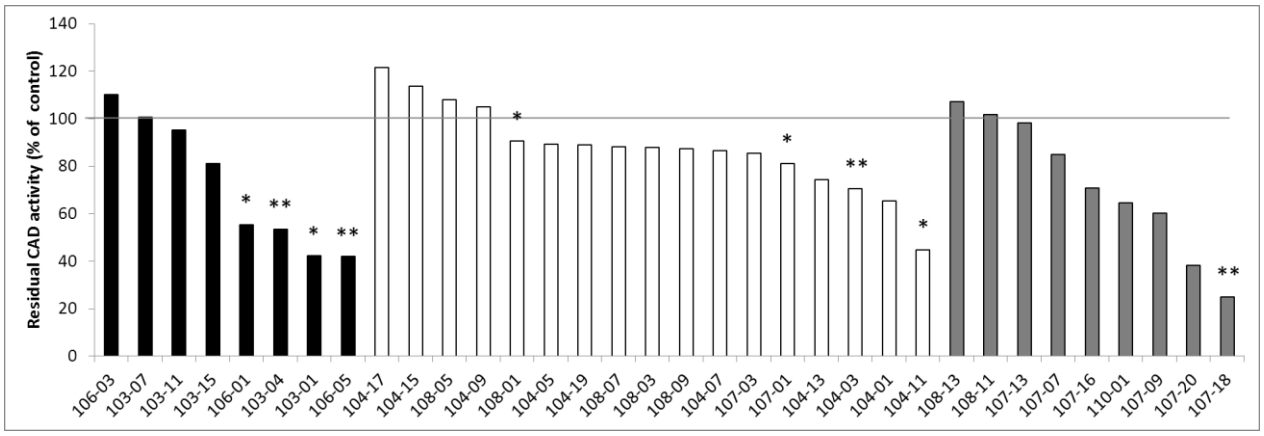


Figure 49. Coniferyl alcohol conversion in transgenic plants expressed as percentage of control plants of the corresponding segregating population. Black, white and grey bars represent plants transformed with RNAi1, RNAi2 and RNAi3 respectively. *: p<0.05; **: p<0.01.

For each RNAi construct, the line with the lowest residual activity was chosen, irrespective of a correct segregation pattern and provided that there were enough seeds left from the same seed stock for the remaining experiments. Ultimately, 103-01, 104-11 and 107-20 were chosen for RNAi1, RNAi2 and RNAi3 respectively. The mean CAD activity reduction over two independent experiments for these lines was 52%, 59% and 65% for 103-01, 104-11 and 107-20 respectively (Figure 49 and Supplementary figure 8). These three lines were further studied for expression of the hairpin construct, expression of *ZmCAD2*, cell wall composition, biomass yield and leaf growth.

4.1.1 Expression levels of ZmCAD2 and relation to reduced CAD enzyme activity

Expression levels of the target gene *ZmCAD2* as well as the hairpin were analyzed by qRT-PCR at two time points in development: two weeks and eight weeks after sowing (see Supplementary figure 11 for representative plants). Transcripts of the hairpin construct were detected in samples of transgenic plants and were absent or negligible in the controls (Figure 50A and Figure 51A). At two weeks after sowing, no downregulation of *ZmCAD2* could be observed (Figure 50B). At eight weeks after sowing, downregulation was observed in internodes of 103-01, leaves and internode of 104-11 and leaves of 107-20 (Figure 51B). At eight weeks after sowing, significant CAD activity reduction was found in leaves of 104-11 and internodes of 107-20 (Figure 52).

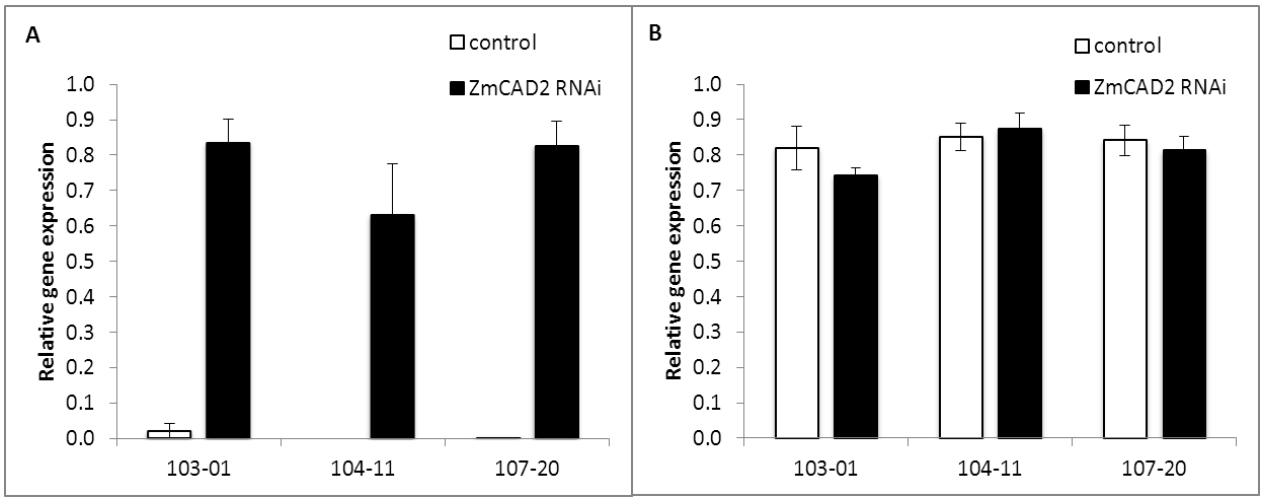


Figure 50. Expression of the hairpin (A) and *ZmCAD2* (B) in leaves two weeks after sowing in control and *ZmCAD2* RNAi plants in the segregating population of three independent lines (103-01, 104-11 and 107-20). Expression values were rescaled on the maximal expression level. Error bars represent standard errors over four biological replicates.

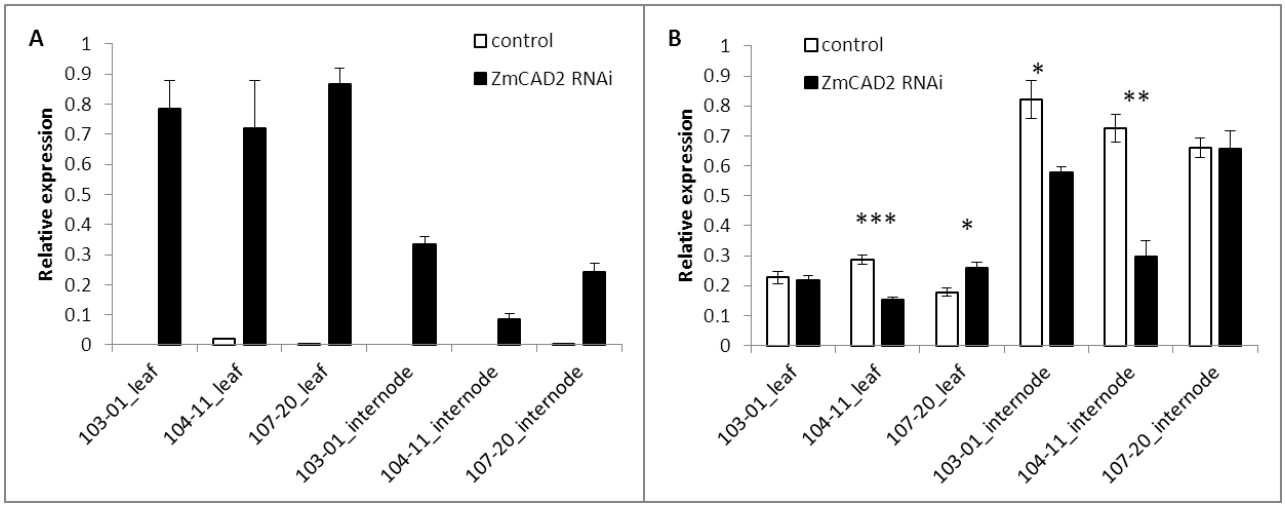


Figure 51. Expression of the hairpin (A) and *ZmCAD2* (B) in mature leaves and internodes, eight weeks after sowing in control and *ZmCAD2* RNAi plants in the segregating population of three independent lines (103-01, 104-11 and 107-20). Expression values were rescaled on the maximal expression level. Error bars represent standard errors over four biological replicates.

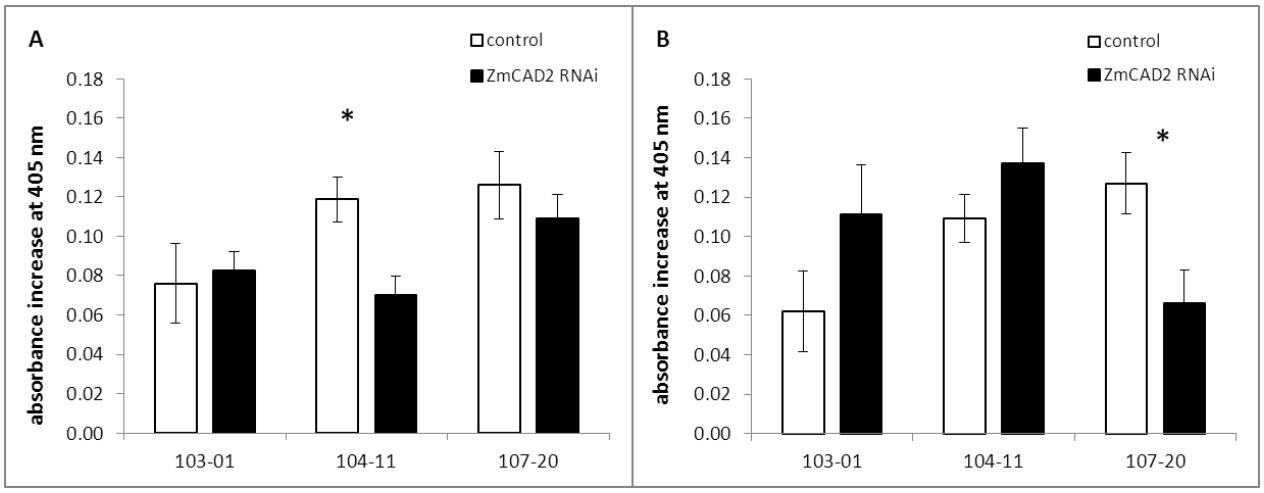


Figure 52. Coniferyl alcohol conversion in leaves (A) and internodes (B) of control and ZmCAD2 RNAi plants 8 weeks after sowing in the segregating population of three independent transgenic lines. *: p<0.05.

4.1.2 Growth and biomass accumulation, lignin content and saccharification efficiency in zmcad2 and control plants

In Sorghum, the perturbation of *ZmCAD2* resulted in delayed development and biomass reduction in previous studies (summarized in Pedersen et al. 2005). To investigate this in the maize *ZmCAD2* RNAi lines, growth rates of the fourth leaf and biomass yield at maturity were determined. The growth kinetics of the fourth leaf were determined using LEAF-E (see Chapter 3). None of the three investigated lines displayed abnormal leaf growth rates nor had altered total aboveground biomass and organ-specific yields (Table 11).

Table 11. Biomass yield and leaf growth analysis in control and ZmCAD2 RNAi plants in the segregating population of three independent transgenic lines. Values are displayed as mean ± SE over biological replicates. Number of biological replicates were shown in Supplementary Table 7. DW: dry weight. Leaf growth parameters were determined using LEAF-E (see Chapter 3).

	103-01			104-11			107-20		
	control	ZmCAD2 RNAi	p-value	control	ZmCAD2 RNAi	p-value	control	ZmCAD2 RNAi	p-value
leaf DW (g)	31 ± 3	33 ± 1	0.396	34 ± 2	36 ± 2	0.505	31 ± 2	33 ± 3	0.568
stem DW (g)	57 ± 6	62 ± 3	0.350	63 ± 3	65 ± 4	0.683	63 ± 4	63 ± 6	0.891
cob DW (g)	18 ± 8	23 ± 4	0.577	37 ± 7	58 ± 11	0.148	47 ± 10	32 ± 6	0.263
total aboveground plant DW (g)	107 ± 10	119 ± 3	0.146	133 ± 8	159 ± 13	0.150	142 ± 12	128 ± 12	0.453
days to germination (d)	6.00 ± 0.08	6.1 ± 0.1	0.429	5.7 ± 0.1	5.6 ± 0.1	0.668	6.1 ± 0.1	6.07 ± 0.07	0.652
leaf #4 length (mm)	668 ± 9	670 ± 10	0.880	667 ± 14	685.7 ± 10.5	0.289	673 ± 9	685 ± 7	0.319
Leaf #4 elongation rate [mm(°Cd^{-1})]	4.28 ± 0.08	4.4 ± 0.1	0.237	4.3 ± 0.1	4.3 ± 0.1	0.856	4.50 ± 0.09	4.48 ± 0.04	0.822

The analysis of lignin content, lignin composition, cellulose content and saccharification efficiency was performed on stem material from mature plants and results were summarized in Table 12. Lignin levels, determined by the acetyl bromide method, were similar in all three lines and did not differ significantly from control plants. In the transgenic line 107-20, cellulose quantity was significantly reduced whereas the amount of cell wall residue per dry weight was increased. Saccharification assays were performed with two different chemical pretreatments: an acid pretreatment with 1M HCl and an alkaline pretreatment with 1M NaOH. Both treatments make the cellulose more accessible to enzymatic hydrolysis but with a different mode of action: HCl removes

the hemicellulose fraction while NaOH removes the lignin fraction. First, the recalcitrance to chemical pretreatment was determined. In 104-11, the residual biomass after pretreatment with 1M HCl was significantly higher than in the control. Next, stem biomass with and without the above stated pretreatments were subjected to enzymatic hydrolysis to determine the efficiency of cellulose degradation. The glucose release was significantly lower for stem biomass of 103-01 without pretreatment and of 104-11 with 1M NaOH pretreatment.

Table 12. Cell wall characteristics in control and ZmCAD2 RNAi plants in the segregating population of three independent transgenic lines. Values are displayed as mean ± SE over four biological replicates. CWR: cell wall residue, DW; dry weight.

	103-01			104-11			107-20		
	control	ZmCAD2 RNAi	p-value	control	ZmCAD2 RNAi	p-value	control	ZmCAD2 RNAi	p-value
acetyl bromide lignin (% CWR)	10.20 ± 0.08	10.1 ± 0.2	0.712	9.68 ± 0.07	9.8 ± 0.2	0.752	9.4 ± 0.2	9.31 ± 0.10	0.871
cellulose (% CWR)	41 ± 4	38 ± 2	0.603	51.7 ± 2	50.0 ± 1.3	0.504	49 ± 2	40 ± 2	**0.030**
CWR (% of DW)	58 ± 2	62 ± 3	0.309	54.1 ± 0.9	58 ± 2	0.077	52.6 ± 0.5	54.1 ± 0.3	**0.033**
CWR + 1M HCl (% of DW)	39 ± 2	44 ± 2	0.124	35.8 ± 0.6	38.6 ± 1.0	**0.048**	36 ± 1	37 ± 1	0.472
CWR + 1M NaOH (% of DW)	36 ± 2	40 ± 2	0.215	33.2 ± 0.5	35.6 ± 0.9	0.059	35 ± 1	34 ± 1	0.640
glucose release per DW after 48h on CWR (%)	4.4 ± 0.2	3.2 ± 0.2	**0.002**	4.1 ± 0.1	4.0 ± 0.2	0.603	4.1 ± 0.2	4.4 ± 0.2	0.376
glucose release per residu after 48h on CWR (%)	7.6 ± 0.4	5.1 ± 0.1	0.002	7.6 ± 0.3	6.9 ± 0.6	0.280	7.7 ± 0.3	8.0 ± 0.4	0.614
glucose release per DW after 48h CWR + 1M HCl (%)	4.7 ± 0.3	4.6 ± 0.3	0.853	5.4 ± 0.3	5.0 ± 0.2	0.294	4.3 ± 0.2	4.5 ± 0.3	0.623
glucose release per residu after 48h CWR + 1M HCl (%)	12.1 ± 0.6	11 ± 1	0.310	15 ± 1	12.9 ± 0.6	0.135	11.9 ± 0.4	11.9 ± 0.6	0.949
glucose release per DW after 48h CWR + 1M NaOH (%)	14.7 ± 0.8	16 ± 1	0.380	15.7 ± 0.4	15.3 ± 0.5	0.503	13.6 ± 0.7	13.3 ± 0.3	0.699
glucose release per residu after 48h CWR + 1M NaOH (%)	40.6 ± 0.7	41 ± 5	0.972	47.3 ± 0.8	42.8 ± 0.6	**0.004**	39 ± 1.0	38.7 ± 0.3	0.986

A lignin composition analysis was performed using thioacidolysis (Table 13). The analysis of lignin's main components *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) showed no significant changes for lines 103-01 and 107-20. In 104-11 the amount of H, S as well as H+G+S was significantly reduced, while the amount of G was not significantly altered. Nevertheless, the relative abundances of these three components remained the same as in the control plants indicating that no specific lignin unit was affected in the 104-11 transgenics. This analysis might indicate an altered extractability of the lignin by reduction of the amount of β-o-4 bonds, which are specifically cleaved by thioacidolysis (Ralph et al. 2008).

Taken together, despite the reduction in *ZmCAD2* expression and CAD activity, the cell wall analysis does not show major alterations in transgenics compared to the controls. No reduction in lignin content nor consistent higher saccharification efficiency using the different pretreatments was observed. In addition, no aldehyde forms of G and S could be detected in *ZmCAD2* RNAi lines or control samples (data not shown). The presence of these compounds is typical for *CAD* deficiency (Kim et al. 2002b).

Table 13. Lignin composition analysis using thioacidolysis in control and ZmCAD2 RNAi plants in the segregating population of three independent transgenic lines. Values are displayed as mean ± SE over four biological replicates. H: *p*-hydroxyphenyl, G: guaiacyl, S: syringyl.

	103-01			104-11			107-20		
	control	ZmCAD2 RNAi	p-value	control	ZmCAD2 RNAi	p-value	control	ZmCAD2 RNAi	p-value
H units per DW (µmol/g)	0.2 ± 0.1	0.35 ± 0.09	0.308	0.36 ± 0.01	0.16 ± 0.07	**0.033**	0.11 ± 0.02	0.09 ± 0.05	0.837
G units per DW (µmol/g)	6 ± 2	13 ± 3	0.071	10.0 ± 0.5	8 ± 1	0.107	8 ± 2	6.7 ± 0.8	0.479
S units per DW (µmol/g)	12 ± 2	24 ± 5	0.056	22.6 ± 0.9	15 ± 2	**0.012**	16 ± 4	15 ± 3	0.784
H+G+S units per DW (µmol/g)	18 ± 4	37 ± 7	0.061	33.0 ± 0.6	23 ± 3	**0.018**	24 ± 6	22 ± 3	0.679
S/G	1.9 ± 0.2	1.84 ± 0.04	0.788	2.3 ± 0.2	1.93 ± 0.06	0.136	1.96 ± 0.06	2.2 ± 0.1	0.157
%H (in H+G+S)	0.9 ± 0.3	1.0 ± 0.2	0.923	1.10 ± 0.04	0.7 ± 0.3	0.136	0.16 ± 0.09	0.3 ± 0.1	0.501
%G (in H+G+S)	35 ± 2	34.9 ± 0.4	0.888	30 ± 2	33.9 ± 0.7	0.116	33.8 ± 0.7	31 ± 1	0.149
%S (in H+G+S)	65 ± 2	64.2 ± 0.6	0.886	69 ± 2	65.4 ± 0.9	0.182	66.1 ± 0.7	68 ± 1	0.172

4.2 A transposon insertion mutant for *ZmCAD2*

In parallel to the transgenic approach, preliminary phenotypic near infrared reflectance spectroscopy (NIRS) quality estimations and transcriptome data was available through collaboration for a transposon insertion mutant in the *ZmCAD2* gene. This *zmcad2* mutant potentially had improved characteristics for conversion to cellulosic ethanol. Therefore, this promising *zmcad2* mutant was investigated more closely for cell wall properties and saccharification efficiency. The *zmcad2* mutant was isolated by transposon tagging with the mutator element by the Biogemma company. The transposon is situated in the last exon of the *ZmCAD2* gene. The *ZmCAD2* gene, corresponding to [redacted] in the first gene annotation model (refgenv1), was however reannotated and named [redacted] in the second model (refgenv2). Sequence analysis performed by Biogemma showed that the transposon in *zmcad2* plants occurred between positions 3499 and 3500 based on the refgenv1 gene model and between positions 3497 and 3498 based on the refgenv2 gene model (see addendum to chapter 5). For each analysis, *zmcad2* and control plants, were compared. These plants were the result of three generations of selfing after five generations of backcrossing (BC5S3) of the original mutant plant with an elite line of the Limagrain company. To clarify the origin of the mutants selected by transposon tagging, a scheme is provided in addendum to chapter 5.

4.2.1 Estimation of cell wall characteristics and (bio)chemical validation

Ground stem material from field grown maize plants was subjected to near infrared reflectance spectroscopy (NIRS). The NIRS data were used for general phenotype discovery and the statistically processed results are summarized here. The estimated parameters were related to forage quality and cover a wide range of maize cell wall characteristics. Forage quality parameters have been shown to be predictive for cellulosic ethanol yield in maize (Andersen et al. 2008; Lorenz et al. 2009), Sorghum (Han et al. 2013) and switchgrass (Sarath et al. 2011; Vogel et al. 2013). These NIRS estimations might thus be informative in describing possible phenotypes caused by the mutation in *ZmCAD2* with regard to cellulosic ethanol production. A selection of NIRS parameters describing cell wall and forage quality parameters relevant for saccharification efficiency is listed in Table 14. The full list of determined NIRS parameters can be found in Supplementary Table 8.

Estimations of the total cell wall content and the cell wall constituents hemicellulose and lignin were based on the analysis system of van Van Soest et al. (1991). This method determines the amount of neutral detergent fiber (NDF), acid detergent fiber (ADF) and acid detergent lignin (ADL). According to NIRS, the *zmcad2* mutant displays a reduction in total cell wall content, and lignin, determined as ADL (Table 14). Furthermore, cellulose content, calculated as ADF-ADL, was also reduced (-11%). The reduction in ADL was consistent with a reduction in the amount of extractable lignin units *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) based upon the nitrobenzene method (Roadhouse and MacDougall 1956). In contrast, the hemicellulose content, calculated as NDF-ADF, was not significantly altered.

The ester ferulic acid (FA) levels was not significantly altered in the mutant whereas the ether ferulic acid rate was significantly decreased. FA is known to be ester-linked to the hemicellulose but can also be incorporated into the lignin by ether bonds (Molinari et al. 2013). In contrast, both 5-5 and 8-O-4 linked diferulic acid (diFA), which are involved in crosslinking different hemicellulose chains but also hemicellulose to lignin (Ishii 1997; Hatfield et al. 1999; MacAdam and Grabber 2002; Jung 2003; Grabber 2005; Harris and Trethewey 2010), were significantly increased in *zmcad2* plants. Furthermore, the maize *zmcad2* mutant shows significantly improved *in vitro* digestibility values. These include dry matter digestibility (IVDMD) and cell wall digestibility (IVNDFD).

Table 14. NIRS estimation of biomass quality parameters in *zmcad2* mutant and control plants. P-values were calculated based upon student t-test in five biological replicates for each group. NDF: neutral detergent fiber, ADF: acid detergent fiber, ADL: acid detergent lignin, IVDMD: *in vitro* dry matter digestibility, IVNDFD: *in vitro* NDF digestibility.

NIRS Parameter	units	control	*zmcad2*	fold change (%)	p-value
NDF	% of Dry matter	55.04	50.73	-7.8	**0.001**
ADF	% of Dry matter	32.44	28.05	-13.6	**0.0004**
ADL	% of Dry matter	3.66	2.55	-30.4	**0.0002**
ester ferulic acid	% of Dry matter	7.11	6.69	-5.9	0.135
ether ferulic acid	% of Dry matter	8.99	8.53	-5.1	**0.015**
5-5 diFerulic acid	% of Dry matter	0.17	0.21	24.4	**0.003**
8-O-4 diFerulic acid	% of Dry matter	0.31	0.36	16.5	**0.001**
ester *p*-coumaric acid	% of Dry matter	15.29	11.43	-25.3	**0.00001**
p-Hydroxybenzaldehyde	% of Dry matter	1.80	0.91	-49.3	**0.00001**
Vanillin	% of Dry matter	8.81	6.92	-21.4	**0.0002**
Syringaldehyde	% of Dry matter	8.68	5.63	-35.2	**0.000002**
IVDMD	% of Dry matter	50.05	59.04	18.0	**0.00002**
IVNDFD	% of Dry matter	29.91	40.97	37.0	**0.00001**

In contrast to the 30% decrease in ADL lignin content predicted by NIRS, no significant changes were detected when lignin content was determined using the acetyl bromide method (Table 15), which was unexpected. On the other hand, analysis by thioacidolysis confirmed a reduction in abundance of lignin units (Table 15). The most abundant subunit was S, comprising 56 % of the total of H, G and S in the stem of control plants. The amount of S was reduced by 50% in the *zmcad2* mutant. G units made up 44% of the total of lignin units and was reduced by 47% in the *zmcad2* mutant. The minor component, H, could not be detected in *zmcad2* mutant plants suggesting that abundances were below the detection limit. As the relative abundance of G and S was shifted in

favor of G units in *zmcad2* mutant lignin, the S/G ratio decreased from 1.27 in control to 1.19 in the *zmcad2* mutant. With thioacidolysis, we also determined the amount of FA. The FA detected here is ether-linked to lignin, but only as end-group. The FA amount was lower in *zmcad2* but the reduction was only significant (-27%) for the *β*-O-4 FA-I form (referred to as A1G by Ralph et al., 2008) expressed per dry weight (Table 15). Furthermore, G and S aldehydes were detected in *zmcad2* mutant samples which are normally not detectable or very low in abundance in wildtype plants (Supplementary figure 10). The presence of these compounds is typical for *CAD* deficiency (Kim et al. 2002b). Consistent with the NIRS prediction of improved digestibility, the saccharification efficiency of stem biomass was improved significantly in *zmcad2* mutants compared to control plants (Figure 53). Using 1M HCl as pretreatment, 17% more glucose was released after 24h hydrolysis. Without acid pretreatment, 20% more glucose was released after 24h hydrolysis (Figure 53).

Table 15. Lignin amount and lignin composition in *cad* mutant and control plants. DW: dry weight, H: *p*-hydroxyphenyl, G: guaiacyl, S: syringyl, FA: ferulic acid, n.d.: not detected. Values are shown as mean ± SE over five biological repeats.

	control			*zmcad2* mutant			fold change (%)	p-value
acetyl bromide lignin (% DW)	4.9	±	0.3	4.7	±	0.1	-4	0.608
H units per DW (µmol/g)	0.10	±	0.01	n.d.	±	-	-100	-
G units per DW (µmol/g)	9.2	±	0.1	4.4	±	0.2	-52	**0.0000004**
S units per DW (µmol/g)	11.7	±	0.2	5.3	±	0.3	-55	**0.0000005**
β-O-4 FA I per DW (µmol/g)	0.43	±	0.03	0.31	±	0.03	-27	**0.028**
β-O-4 FA II per DW (µmol/g)	0.39	±	0.04	0.36	±	0.04	-15	0.565
H units per lignin (µmol/g)	2.1	±	0.3	n.d.	±	-	-100	-
G units per lignin (µmol/g)	179	±	8	95	±	6	-47	**0.00004**
S units per lignin (µmol/g)	227	±	8	114	±	8	-50	**0.00002**
β-O-4 FA I per lignin (µmol/g)	8.3	±	0.5	6.7	±	0.7	-19	0.136
β-O-4 FA IIper lignin (µmol/g)	8	±	1	7.6	±	0.7	-28	0.967
H in H+G+S (%)	0.50	±	0.06	n.d.	±	-	-100	-
G in H+G+S (%)	43.9	±	0.4	45.6	±	0.3	4	**0.011**
S in H+G+S (%)	55.6	±	0.4	54.4	±	0.3	-2	0.051
S/G	1.27	±	0.02	1.19	±	0.02	-6	**0.021**

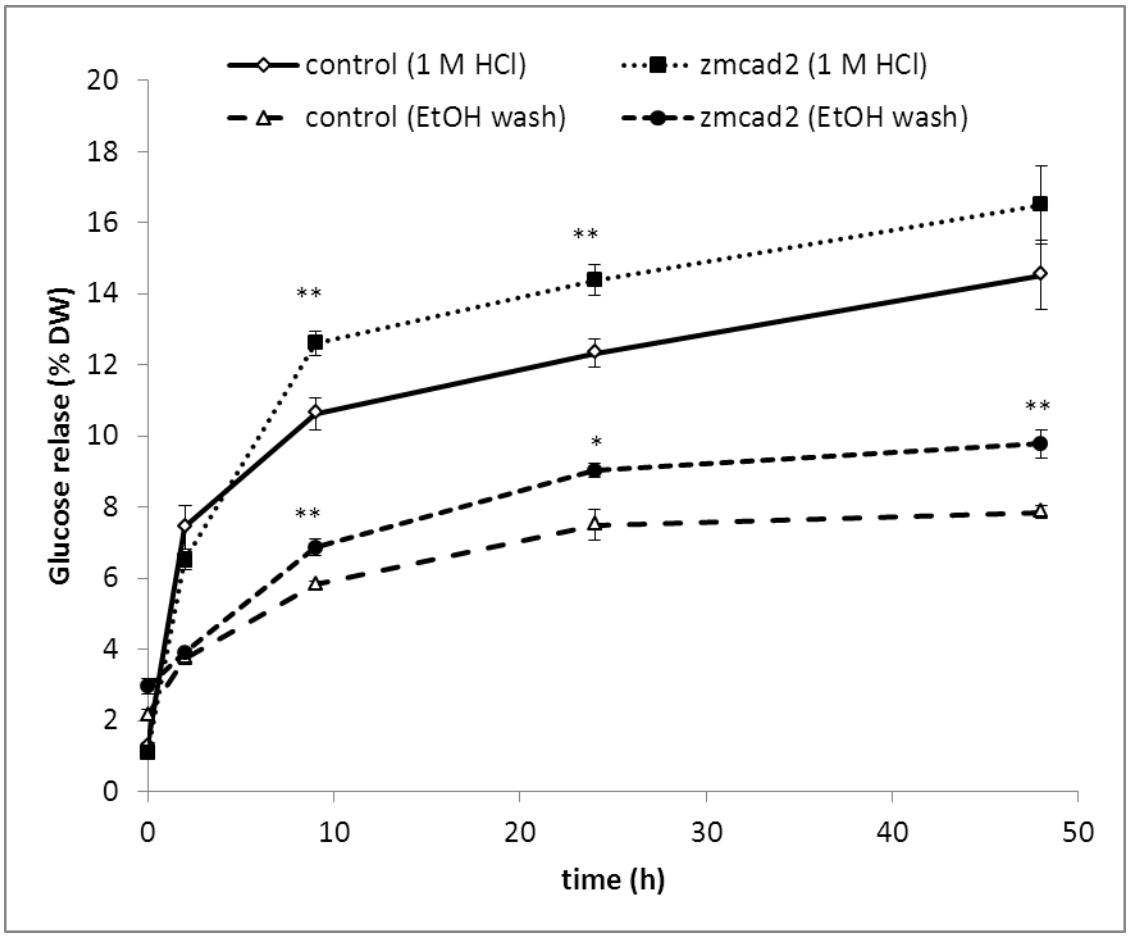


Figure 53. Saccharification efficiency in *cad* mutant and control stems pretreated with ethanol only (EtOH wash) and ethanol and acid (1 M HCl) expressed as glucose release per dry weight (DW). Error bars represent standard errors over five biological replicates. *: p<0.05, **: p<0.01.

4.2.2 Analysis of differential gene expression shows a stress response in the zmcad2 mutant

The perturbation of one gene in the lignin pathway can trigger altered expression levels of many more genes within that pathway, related pathways as well as regulatory genes (Bennetzen and Hake 2009). The analysis of these systems-wide consequences of *ZmCAD2* perturbation might lead towards a better understanding of the genetic basis of lignin biosynthesis. To investigate these systems-wide responses, gene expression was quantified using a custom Nimblegen microarray in the ear internode from field grown *zmcad2* and control plants in four developmental stages: V10 (ten visible leaf collars), S (silking), S+7d (seven days after silking) and S+14d (fourteen days after silking). The microarray contained 37,670 probes, including probes for mitochondrial and chloroplast DNA and miRNA encoding genes.

A very high number of genes showed significant differential expression (Table 16). This included both the developmental as the mutational effect. Almost all probes (19,570 out of 23,912 genes) were differentially expressed over development. As the internode was sampled both in vegetative

(V10) as in reproductive stages (S, S+7d and S+14d), this large number of differentially expressed genes was to be expected. However, 53% of those genes that were differentially expressed over development showed a significantly different developmental expression pattern in the *zmcad2* mutant. As the perturbation is situated in lignin formation which is a process of tissue differentiation and thus developmentally regulated, this interaction of mutational and developmental effect was expected. Only a small amount of genes (110 or 0.6% of differentially expressed genes) were significantly altered in expression in the *zmcad2* mutant without interaction with development (Table 16).

Table 16. Differential genes and probes in microarray expression analysis in internodes of *zmcad2* mutant and control plants over development.

ANOVA (p_FDR<0.05)	# diff. probes	# diff genes	# diff. probes per total # diff probes (%)	# diff genes per total # diff genes (%)	# diff probes per total # probes in array (%)	# diff genes per total # genes in array (%)
***zmcad2* mutation effect**	113	110	0.5	0.6	0.3	0.4
***zmcad2* mutation and development effect**	12132	10544	50.7	53.6	32.2	34.0
development effect	11666	9026	48.8	45.9	31.0	29.1
non-significant	13758	11309	-	-	36.5	36.5

To investigate which biological processes could be affected in the zmcad2 mutant internodes, an overrepresentation analysis was performed using the differentially expressed genes between *zmcad2* and control plants in each stage. Significant over or under representation in zmcad2 internodes could be detected in upregulated gene expression in stages S and S+7d and in downregulated genes in stages S and S+14d (Figure 54). No significant overrepresentation or underrepresentation of GO categories could be detected in the V10 stage with the applied statistical methods. According to the Pageman analysis, the highest overrepresentation in differentially expressed genes was situated in the categories “stress response”, “cytochrome P450”, and “RNA”. A milder overrepresentation was found in the categories “photosynthesis”, “amino acid metabolism” and “hormone metabolism”. Furthermore, a mild underrepresentation was found for genes encoding “transporters”. For the “stress response”, the highest enrichment was observed for “response to abiotic stress” and “response to heat”. Surprisingly, no category related to secondary cell wall formation or phenylpropanoid biosynthesis was overrepresented in this analysis. Highly differentially expressed genes from the three main overrepresented categories, namely “stress response”, “cytochrome P450” and “RNA; regulation of transcription” were further investigated (Supplementary Table 9).

For “response to stress”, differentially expressed genes were related to disease resistance including genes encoding pathogenesis related proteins and genes related to abiotic stress such as genes encoding heat shock proteins. Cytochrome P450 proteins are known to be involved in various stress responses in plants (Morant et al. 2003; Lapierre et al. 2004; Narusaka et al. 2004; Shinozaki and Yamaguchi-Shinozaki 2007; Ehlting et al. 2008; Cheng et al. 2010; Mao et al. 2013) and members of

this group act in the phenylpropanoid pathway (Werck-Reichhart 1995; Pan et al. 2009). The highly overrepresented category of "RNA ; regulation of transcription" contains several transcription factors and ethylene, cytokinin, auxin and heat responsive genes. The various affected biological categories might thus have 'stress response' as the common factor.

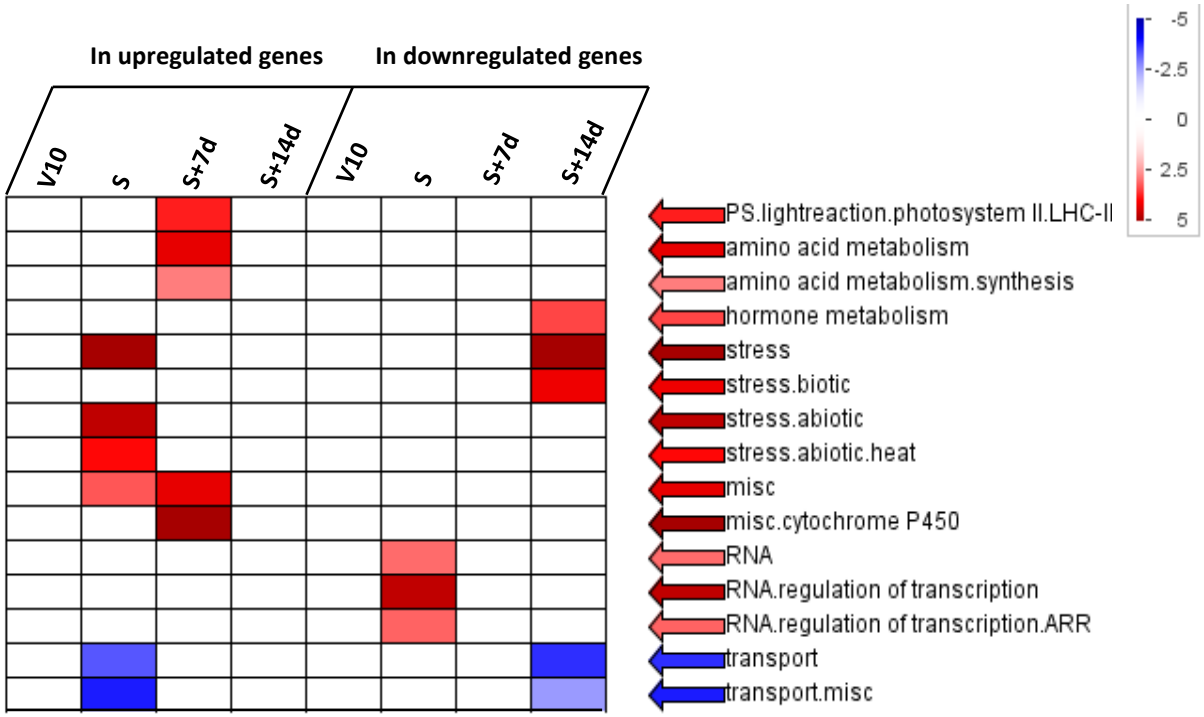


Figure 54. Significant overrepresentation of GO categories in *zmcad2* mutant stems in four developmental stages V10 (10 visible leaf collars), S (silking), S+7d (seven days after silking) and S+14d (fourteen days after silking) using Pageman analysis in MapMan. The differentially expressed transcripts in each bin exceeding the value of 0.6 (log2 scale) were tested for overrepresentation and underrepresentation with two hypergeometric tests (one for upregulated and one for downregulated transcripts) and resulting p-values were corrected for multiple testing using the false discovery rate (FDR) (Benjamini and Yekutieli 2001)**. For both upregulated and downregulated genes, the overrepresentation of genes is colored in red and the underrepresentation is colored blue. The scale indicates the z scores (transformed p-values), thus, a p-value of 0.05 is assigned to a value of 1.96. PS: photosynthesis.**

4.2.3 Genes involved in internode lignification are upregulated in zmcad2 plants

Highly similar to what was reported in Chapter 5, expression of genes putatively involved in lignification follow the developmental profile of increase in expression from V10 to S stage (Supplementary Table 10). To investigate whether *ZmCAD2* perturbation affected the expression of genes within the same pathway, fold changes in gene expression in *zmcad2* mutant compared to control samples were determined (Table 17). The list of all investigated phenylpropanoid genes can be found in Supplementary Table 11.

In contrast to the V10 stage of the *zmc4h1* mutant and control plants (see Chapter 5), no major changes in lignin gene expression levels were detected in the *zmcad2* mutant, except *ZmCAD2* itself ([redacted], Table 17), which was highly downregulated, as expected. As the transposon is situated in an exon, proper transcription of the *ZmCAD2* gene was thus abolished. In the reproductive stages S, S+7d and S+14d, several genes were upregulated in the *zmcad2* mutant. The

highest upregulated genes were *ZmC4H1* and *ZmF5H1* in the S+7d stage. The perturbation of *ZmCAD2* thus appears to cause upregulation of lignin genes. In addition, two other CAD genes ([redacted] and [redacted]), only moderately and lowly expressed in control internodes, show strong upregulation in the *zmcad2* mutant (Supplementary Table 11).

Table 17. Log2 fold changes in expression levels of phenylpropanoid genes with high and increasing expression over development in *zmcad2* mutant as compared to control. Fold changes are color coded according to higher (red) and lower (blue) in *zmcad2* mutant samples in four stages of development: V10 (ten visible leaf collars), S (silking), S+7d (seven days after silking) and S+14d (fourteen days after silking). The expression level was labeled according to maximal expression (log2): "very high" >=15, 15>"high">=13, 13>"moderate">=11, 11>"low". Changes in expression levels of all phenylpropanoid genes can be found in Supplementary Table 11.

				log2 fold change (mutant vs. control)			
protein function	gene name	probe name	expression level	V10	S	S+7d	S+14d
PAL	Hidden		very high	-0.24	0.14	0.18	0.07
PAL			high	0.06	0.34	0.22	0.31
PAL			high	-0.13	0.25	0.16	-0.03
PAL			high	-0.14	0.15	0.34	0.15
PAL			high	-0.08	0.03	0.15	0.22
PAL			high	-0.06	0.39	0.04	-0.30
C4H			high	-0.04	0.55	0.97	0.23
4CL			moderate	-0.26	0.48	-0.75	0.12
HCT			moderate	-0.13	0.32	-0.17	0.09
HCT			moderate	-0.07	-0.09	-0.08	0.15
C3H			high	-0.14	0.39	0.60	-0.03
CCoAOMT			high	0.02	-0.11	0.20	0.67
CCoAOMT			high	-0.27	0.24	-0.08	0.36
CCoAOMT			high	-0.01	0.08	-0.11	0.19
CCR			high	0.20	0.00	0.10	0.44
F5H			moderate	-0.03	-0.12	0.98	0.26
COMT			high	-0.18	0.41	0.28	0.28
CAD			high	-2.28	-1.43	-1.04	-0.82

4.2.4 The metabolic profile of zmcad2 is highly affected

In contrast to the modest fold changes on the expression levels, principal component analysis (PCA) on all peaks quantified by metabolic profiling revealed major changes in *zmcad2* mutant samples compared to its control. Similar to what was observed in the transcriptome data, the difference between *zmcad2* and control samples was most obvious for the plants in S, S+7d and S+14d but was also present, yet less clearly, at the V10 stage (Figure 55).

To visualize the metabolic changes as a result from *ZmCAD2* perturbation, the relative abundances of 53 tentatively identified compounds and transcripts of lignin and flavonoid biosynthetic genes in the internodes of control and *zmcad2* plants were mapped on the metabolic pathway (as given in Figure 41 of Chapter 5, resulting in the metabolic map of Figure 56). The specific changes in phenolic compound abundance of these compounds is discussed below.

The two amino acids tryptophan, synthesized by a series of enzymatic steps from chorismate, and phenylalanine, the substrate of the general phenylpropanoid biosynthesis pathway, were slightly less abundant at the V10 stage but were more abundant in the reproductive S+7d stage. Like tryptophan, the benzoxazinoids are synthetized from chorismate, an intermediate of the shikimate pathway. The glycosylated forms of the toxic biochemicals HMBOA, DIBOA and DIMBOA (Jonczyk et al. 2008; Frey et al. 2009; Dick et al. 2012) were highly abundant in internodes at all investigated developmental stages and were even increased in abundance in *zmcad2* plants. In contrast, the intermediate benzoxazinoid HBOA and its glycosylated form were less abundant.

The hexosylated phenylpropanoic acids, such as *p*-coumaroyl hexose, *p*-hydroxybenzoyl hexose, feruloyl hexose, vanilloyl hexose and syringoyl hexose, were more abundant in *zmcad2* plants as compared to control plants. In contrast, the phenylpropanoic acid 4-O-hexosides, namely *p*-hydroxybenzoic acid, vanillic acid and syringic acid 4-O-hexosides were less abundant. Furthermore, caffeoyl quinate was more abundant in *zmcad2* plants as compared to control plants but *p*-coumaroyl caffeoyl quinate was less abundant. As was observed in the *zmc4h1* mutants in Chapter 5, the flavonoids tricin and tricin derivatives were increased in *zmcad2* plants as compared to control. Yet, tricin coupled to a pentose, hexuronic acid and syringic acid was highly decreased in abundance. A 6-C-hexose-8-C-pentose-apigenin was slightly increased in abundance in *zmcad2* plants. Several flavonolignans, such as G(8-O-4')tricin hexose, G(8-O-4)G(8-O-4')tricin, S(8-O-4)tricin and G(8-O-4)S(8-O-4')tricin were decreased in *zmcad2* plants. Interestingly, a yet not fully identified compound, S(8-O-4')demethoxytricin + 234 + 76, is highly upregulated upon *ZmCAD2* perturbation. Demethoxytricin is in fact chrysoeriol, a compound closely related to tricin. The identified oligolignols comprised one dilignol, two trilignols and one tetralignol. The abundance of these oligolignols, intermediate coupling products of the lignin polymer, increased with the age of the plants, consistent with the lignification pattern in the maize internode over development (Morrison and Kessler 1994; Morrison and Jung 1998). These compounds were less abundant in *zmcad2* plants or even undetectable. In accordance with the decreased abundance of oligolignols, all identified neolignans were decreased in abundance in the mutant.

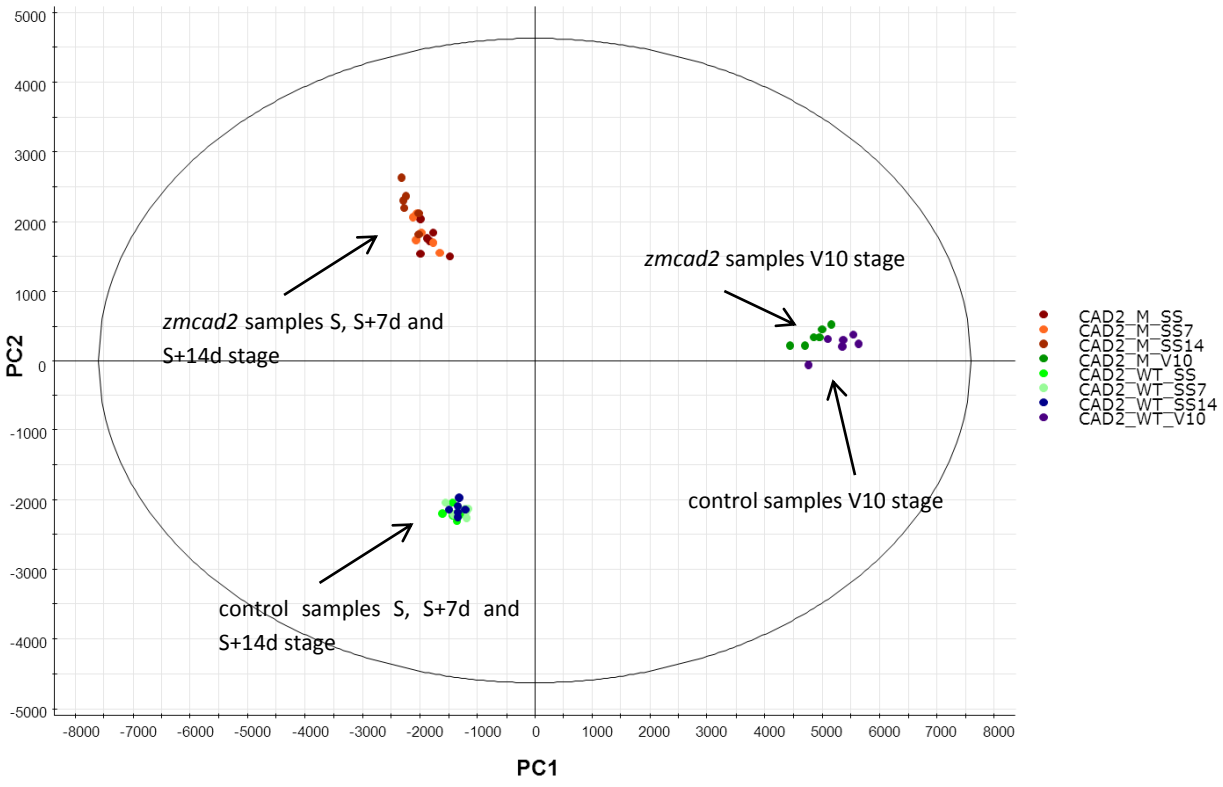


Figure 55. PCA based on all detected peaks by metabolic profiling in *zmc4h1* mutant and control samples in four developmental stages. Each dot represents six biological replicates. V10 (ten visible leaf collars), SS (silking), SF+7 (seven days after silking) and SS+14 (fourteen days after silking). Principal component (PC) 1 and 2 account for 61% and 23% of the variation, respectively.

Table 18. Abundances of tentatively identified phenolic compounds in internodes of control and *zmcad2* internodes in four stages of development: V10, S, S+7d and S+14d (ten visible leaf collars, silking, seven days after silking and fourteen days after silking respectively). Compounds were determined using by UPLC-MS^2 in negative mode. Values for peak area are mean values of six biological replicates. Significantly altered (student t-test) fold changes were depicted in bold. rt: retention time, *p*-CA: *p*-coumaric acid, H: *p*-hydroxyphenyl, G: guaiacyl, S: syringyl, G': guaiacyl aldehyde, S': syringyl aldehyde.

			Mean peak area control				Mean peak area *zmcad2*				Fold change			
Accepted description	**m/z**	**rt**	**V10**	**S**	**S+7d**	**S+14d**	**V10**	**S**	**S+7d**	**S+14d**	**V10**	**S**	**S+7d**	**S+14d**
aminoacids														
Hidden			3449	1939	1588	2719	2610	2205	3537	2228	**-1.32**	1.14	**2.23**	-1.22
			7448	22365	21710	39639	6020	26702	56376	69075	**-1.24**	1.19	**2.60**	**1.74**
			5875	7807	7686	8594	2089	2807	2897	3484	**-2.81**	**-2.78**	**-2.65**	**-2.47**
			38202	113890	85620	78505	29445	85879	65681	59888	-1.30	-1.33	**-1.30**	**-1.31**
			862	164081	156732	160792	1262	77225	84990	81573	1.46	**-2.12**	**-1.84**	**-1.97**
			16793	16	24	73	33348	798	568	968	**1.99**	**50.34**	**23.26**	**13.18**
			4302	0	1	2	6102	21	18	13	1.42	**inf up**	26.40	6.16
			32171	1884	1808	2088	36607	2181	1289	1274	1.14	1.16	**-1.40**	-1.64
			13	784	475	187	11	1051	80	312	-1.22	1.34	**-5.92**	1.67
			0	4458	3723	3167	2	3027	2563	2378	**inf up**	**-1.47**	**-1.45**	-1.33
			5	1792	1180	2192	10	3813	3616	4637	2.25	**2.13**	**3.07**	**2.12**
			127	1045	860	786	183	423	267	530	1.44	**-2.47**	**-3.22**	**-1.48**
			75634	1348	1214	721	80005	42677	41701	47534	1.06	**31.66**	**34.34**	**65.91**
			39780	3324	2382	1175	43850	2483	963	198	1.10	-1.34	**-2.47**	**-5.94**
			584168	3320	2977	2774	576189	9770	7103	3391	-1.01	**2.94**	**2.39**	1.22
			995508	59671	51220	52336	918115	73089	77736	54138	-1.08	1.22	**1.52**	1.03
			399928	2923	2502	2595	276527	2527	3198	1694	**-1.45**	-1.16	1.28	**-1.53**
			150404	0	8	8	123473	129	84	56	-1.22	**inf up**	10.99	**6.63**
			20015	5	2	5	10361	0	0	0	**-1.93**	**inf down**	**inf down**	**inf down**
			9218	0	0	1	7019	0	0	0	-1.31	n.d.	n.d.	**inf down**
			22280	1227	1206	842	24251	0	0	0	1.09	**inf down**	**inf down**	**inf down**
			483	0	0	0	100	0	0	0	**-4.84**	n.d.	n.d.	n.d.
			20186	1602	1496	1929	30408	6289	4405	5024	1.51	**3.93**	**2.95**	**2.60**
			911045	483	769	2677	1213417	159366	123106	89653	**1.33**	**330.29**	**160.01**	**33.49**
			1586	0	2	2	3316	71	62	38	**2.09**	**inf up**	40.44	**19.92**
			31203	10305	9383	7039	27052	8421	3098	2598	-1.15	-1.22	**-3.03**	**-2.71**
			1280167	86277	76505	41197	978477	242922	170788	129955	**-1.31**	**2.82**	**2.23**	**3.15**
			2146	23542	24393	26187	2033	19786	14046	14352	-1.06	**-1.19**	**-1.74**	**-1.82**
			57268	67207	65005	59864	32736	92033	97489	97094	**-1.75**	**1.37**	**1.50**	**1.62**
			44237	2958	2482	1211	26281	3039	560	767	-1.68	1.03	**-4.43**	**-1.58**
			5888	456	469	78	7035	14	11	3	1.19	**-32.31**	**-43.85**	**-25.38**
			349263	188560	174996	173244	348558	277273	292601	352227	-1.00	**1.47**	**1.67**	**2.03**
			62492	108779	113338	130662	59386	84442	88513	124330	-1.05	**-1.29**	**-1.28**	-1.05
			123551	8987	10697	11544	109450	14920	16876	16742	-1.13	**1.66**	**1.58**	**1.45**
			9	513	916	1599	4	713	1981	4112	-2.07	1.39	**2.16**	**2.57**
			3461	492	924	1453	2047	676	1555	2094	-1.69	1.37	1.68	1.44
			8407	24717	22969	21782	6160	22588	22957	23603	**-1.36**	**-1.09**	-1.00	1.08
			26412	1375	1254	1058	26872	2412	2418	2909	1.02	**1.75**	**1.93**	**2.75**
			10246	3069	3117	3387	8553	3465	5699	9630	-1.20	1.13	**1.83**	**2.84**
			380050	101664	92694	95220	348772	185304	163201	178985	-1.09	**1.82**	**1.76**	**1.88**
			25470	3827	3689	5082	18043	13202	8589	11464	**-1.41**	**3.45**	**2.33**	**2.26**

Table 8. continued

			Mean peak area control				Mean peak area *zmcad2*				Fold change			
Accepted description	**m/z**	**rt**	**V10**	**S**	**S+7d**	**S+14d**	**V10**	**S**	**S+7d**	**S+14d**	**V10**	**S**	**S+7d**	**S+14d**
oligolignols														
Hidden			0	2088	2747	4305	0	0	7	0	n.d.	**inf down**	**-395.18**	**inf down**
			0	690	1068	1706	0	0	0	0	n.d.	**inf down**	**inf down**	**inf down**
			1263	38704	42917	47195	0	4256	9636	13639	**inf down**	**-9.09**	**-4.45**	**-3.46**
			97	27904	31563	36276	0	1722	4360	6607	**inf down**	**-16.20**	**-7.24**	**-5.49**
			0	1463	1400	2277	0	51	136	264	n.d.	**-28.93**	**-10.27**	**-8.64**
			3	43036	39704	36666	16	22407	31000	21749	4.94	**-1.92**	**-1.28**	**-1.69**
			50	1024	1053	814	44	282	670	345	-1.14	**-3.64**	**-1.57**	**-2.36**
			0	6116	7002	9710	0	1395	3504	4808	n.d.	**-4.38**	**-2.00**	**-2.02**
			0	14430	16716	23465	0	4352	8487	10873	n.d.	**-3.32**	**-1.97**	**-2.16**
			0	22294	25366	34074	0	7274	14337	19034	n.d.	**-3.06**	**-1.77**	**-1.79**
			0	4968	4300	4323	0	3834	3530	4137	n.d.	**-1.30**	**-1.22**	-1.04
			0	11431	10598	11060	3	9201	9242	9831	**inf up**	**-1.24**	-1.15	**-1.13**
			15	12717	11587	13495	5	4078	4486	7133	-3.35	**-3.12**	**-2.58**	**-1.89**
			62	30782	26820	30338	0	7831	7880	13846	**inf down**	**-3.93**	**-3.40**	**-2.19**
			98	212020	184340	131301	70	260290	293908	252013	-1.40	**1.23**	**1.59**	**1.92**
			1873	41685	35967	30053	3412	57542	49730	37606	**1.82**	**1.38**	**1.38**	**1.25**
			0	7578	7777	9699	0	508	971	973	n.d.	**-14.93**	**-8.01**	**-9.96**
			120	65574	56315	50843	149	76217	72012	70561	1.24	1.16	**1.28**	**1.39**
			156	11638	14169	21590	209	129903	158248	156793	1.34	**11.16**	**11.17**	**7.26**
			0	1242	1111	1080	0	1902	1931	1833	n.d.	**1.53**	**1.74**	**1.70**
			141	15589	14111	12041	0	14918	15147	12545	**inf down**	-1.04	1.07	1.04
			15	21886	19293	15980	0	17301	19514	17228	**inf down**	**-1.27**	1.01	1.08
			0	18045	14078	11647	0	21677	17005	16183	n.d.	1.20	**1.21**	**1.39**
			0	14157	11084	8483	0	14344	12212	12480	n.d.	1.01	1.10	**1.47**
			0	6370	5360	6071	0	4153	5645	3088	n.d.	**-1.53**	1.05	**-1.97**
			0	1194	1052	1263	0	835	1605	918	n.d.	-1.43	**1.53**	-1.38
			94	140173	119195	104054	189	122957	128077	126308	2.00	-1.14	1.07	**1.21**
			0	3004	2702	1871	0	2241	2981	2458	n.d.	-1.34	1.10	1.31
			0	8717	9042	9062	0	6499	4606	6314	n.d.	**-1.34**	**-1.96**	**-1.44**
			0	18052	16833	13210	0	7119	3334	2609	n.d.	**-2.54**	**-5.05**	**-5.06**
			89	12015	10799	8155	9	4931	2579	1905	-9.86	**-2.44**	**-4.19**	**-4.28**
			0	13026	12984	13313	0	4940	6355	5513	n.d.	**-2.64**	**-2.04**	**-2.41**
			0	9762	9366	9870	0	3078	3810	4169	n.d.	**-3.17**	**-2.46**	**-2.37**
			0	28371	28119	28184	0	9883	13136	14065	n.d.	**-2.87**	**-2.14**	**-2.00**
			0	22914	22930	22638	0	7455	9803	11198	n.d.	**-3.07**	**-2.34**	**-2.02**
			0	265	292	678	0	23	113	141	n.d.	**-11.71**	**-2.58**	**-4.81**
			0	542	724	948	0	71	89	152	n.d.	**-7.64**	**-8.10**	**-6.24**
			0	2423	2407	2838	0	438	938	886	n.d.	**-5.53**	**-2.57**	**-3.20**
			0	828	730	1107	0	118	291	252	n.d.	**-7.05**	**-2.51**	**-4.39**
			8	321	540	808	275	8652	10767	11650	**32.82**	**26.97**	**19.95**	**14.42**

5. Discussion

A total of 35 independent transgenic lines in the B104 inbred background were obtained of which 28 lines had a single locus insertion. This is consistent with the observations that Agrobacterium-mediated transformation results in low-copy integration of the T-DNA into the genome (Zhao et al. 2000; Dai et al. 2001; Shou et al. 2004; Frame et al. 2006b; Frame et al. 2006a; Ishida et al. 2007). Reduction in CAD activity ranged from not significantly different to a reduction of 75% in the transgenic plants compared to their respective controls. In three selected transgenic lines with reduced CAD activity, downregulation of the target gene was confirmed in internodes or leaves of plants eight weeks after sowing (V9 stage, approximately one meter high), but no downregulation could be detected in the fourth leaf when it just appeared from the pseudostem. These results suggest that either dsRNA hairpin or target gene expression levels have to exceed a certain threshold before target mRNA levels are efficiently degraded, as suggested before (Lindbo et al. 1993; Eamens et al. 2008).

After analysis of lignin quantity, lignin composition, leaf growth rates and biomass accumulation, we could conclude that the three investigated lines did not display a phenotype consistent with previous reports of CAD perturbation (Halpin et al. 1998b; Fornalé et al. 2012; Chen et al. 2012; Barrière et al. 2013). This is surprising, since the level of CAD activity reduction in the three fully characterized lines (46%, 63% and 67% for 103-01, 104-11 and 107-20 respectively) was comparable to what was reported for RNAi-mediated downregulation of *ZmCAD2* in maize hybrids by Fornalé et al. (2012). *ZmCAD2* RNAi plants of Fornalé et al. (2012) had only a mild phenotype, i.e. no reduction in stem lignin content, no accumulation of cinnamaldehyde and only 8% increase in cellulosic ethanol production from stem biomass. With this data and previous reports, it was suggested that the CAD activity reduction needs to drop below a minimum threshold before becoming rate-limiting (Anterola and Lewis 2002; Saathoff et al. 2011; Fornalé et al. 2012).

It thus seems that the downregulation of *ZmCAD2* in maize by means of RNAi is only effective to a limited extent. As *ZmCAD2* is highly expressed in the stem (Guillaumie et al. 2007; Riboulet et al. 2009), perhaps the level of hairpin transcript is insufficient for effective downregulation. In our study, the Ubiquitin (UBI) promoter (Christensen and Quail 1996) from maize was used to drive overexpression of the hairpin in the pBb7GW-I-WG-UBIL vector (Figure 35). Also in the study of Fornalé et al. (2012), the UBI promoter was used in the vector pAHC25 (Christensen and Quail 1996). Unfortunately, Coussens et al. (2012) showed that the UBI promoter is only of limited strength even compared to CaMVp35S, which was demonstrated previously to be of poor strength in monocots (Schledzewski and Mendel 1994). Moreover, promoters isolated from Brachypodium genes (*BdUBI10* and *BdEF1a*) outperformed the maize UBI and p35S promoters and are thus expected to be more efficient in overexpression or silencing of gene expression in maize and will be used in future experiments (Coussens et al. 2012). As alternative for using the classical RNAi approach, artificial miRNA (amiRNA) holds great promise for highly specific or simultaneous downregulation of lignin genes (Warthmann et al. 2008; Ossowski et al. 2008). In Brachypodium, the *BdCAD1* gene was downregulated using amiRNA and resulted in the brown-midrib phenotype, demonstrating its effectiveness (Trabucco et al. 2013). Previously, no brown-reddish coloration could be obtained for *CAD* perturbation using a transgenic approach (Ralph et al. 2001; Anterola and Lewis 2002; Chen et al. 2003; Saathoff et al. 2011; Fu et al. 2011; Fornalé et al. 2012), as the level of

downregulation was probably not comparable to that of mutants (Halpin et al. 1998b; Zhang et al. 2006; Tsuruta et al. 2010; Bouvier d'Yvoire et al. 2013). One of the benefits of transforming an inbred line is that introduced traits can be studied in a stable genetic background. Moreover, combinations of different traits (trait stacking) can be readily made by simple crosses and investigated with straightforward analysis of the progeny. Therefore, optimization of transformation vectors and downregulation efficiency are desirable. Taken together, recommendations for future experiments for downregulation of lignin genes are the use of alternative promoters for maize UBI and p35S, such as the Brachypodium promoters of *BdUBI10* and *BdEF1a*, and the use of amiRNA to specifically target multiple family members at once.

In contrast to the RNAi lines, the transposon insertion mutant for *ZmCAD2* has a phenotype with reduced lignin quantity, reduced S/G ratio, accumulation of cinnamyl aldehydes and enhanced saccharification efficiency. The accumulation of cinnamyl aldehydes in the stem is consistent with previous reports of reduced CAD activity (Halpin and Knight 1994; Baucher et al. 1996; Vailhé and Andrée 1998; Halpin et al. 1998b; Ralph et al. 2001; Palmer et al. 2008), and has been regarded as a marker compound for CAD deficiency (Kim et al. 2000; Ralph et al. 2001). The reduction in S/G ratio has also been observed in CAD downregulated tobacco, alfalfa and maize plants (Vailhé and Andrée 1998; Baucher et al. 1999; Fornalé et al. 2012) and Brachypodium and Sorghum CAD mutants (Sattler et al. 2009; Bouvier d'Yvoire et al. 2013) but no changes in S/G ratio were observed in the maize bm1 mutant (Halpin et al. 1998b) and CAD downregulated poplar (Baucher et al. 1996). We observed a reduction in ADL lignin of 30% but no significant changes were detected using the acetyl bromide method. ADL typically underestimates actual lignin levels in grasses, yielding lignin concentrations that are only one half to one quarter of acetyl bromide and Klason methods (Hatfield and Fukushima 2005), whereas the acetyl bromide method may overestimate lignin (Voelker et al. 2010). The acetyl bromide method might therefore not the best choice for detecting only modest changes in lignin quantity. We therefore plan to perform a Klason lignin measurement. In CAD downregulated poplar, tobacco and maize plants, the lignin content was reported to be unaltered (Baucher et al. 1996; Vailhé and Andrée 1998; Fornalé et al. 2012) or reduced up to 30% in CAD mutants in Sorghum, maize and Brachypodium (Pillonel et al. 1991; Halpin et al. 1998b; Bouvier d'Yvoire et al. 2013). The reduction in CAD activity thus appears to be dependent on the residual CAD activity as the mutants have the lowest residual CAD activity. Mutations in CAD were shown previously to be associated with a brown coloration of the stem (Halpin et al. 1998b; Sattler et al. 2009). However, no brown-red coloration of stems or leaves was observed in our transposon insertion mutant (personal communication with Biogemma). Until present, the origin of this coloration remains elusive.

In addition to the biochemical data, micro-array and metabolic profiling was performed on mutant and control samples in four developmental stages. Gene expression has been studied previously by suppression subtractive hybridization (SSH) and micro-array in the *bm1* mutant (Shi et al. 2006). This study showed differential gene expression in several metabolic and signaling pathways with most differentially ESTs in carbohydrate, energy and amino acid metabolism. Moreover, they found that all genes related to flavonoid, stilbene and lignin biosynthesis, thus the phenylpropanoid pathway, were downregulated except for one P450 protein encoding gene. After blasting the sequence we found out that this was in fact a *p-coumarate 3-hydroxylase* (*C3H*) gene. In our dataset,

we see exact the opposite: genes in the phenylpropanoid pathway were modestly but significantly upregulated upon *ZmCAD2* perturbation. This difference might originate from different sampling stages and methods. Shi et al. (2006) harvested stem (internode, node and leaf sheath) material from five and seven week old plants whereas we harvested the ear internode in V10, S, S+7d and S+14d stages. Although it is difficult to compare the different conditions, V10 is probably a more advanced developmental stage than the five or seven week old greenhouse-grown plants in Shi et al. (2006). According to our observations changes in transcript levels of phenylpropanoid genes became only apparent after the switch from vegetative to reproductive phase. The highest upregulation was detected for *ZmC4H1* and *ZmF5H1*. As the mutated *ZmCAD2* gene was highly downregulated, two other *CAD* genes showed upregulation in *zmcad2* samples. The highest upregulated *CAD* gene ([redacted]) is phylogenetically related to the rice *OsCAD1* and *OsCAD4* and Arabidopsis *AtCAD1* genes. These genes are part of a subgroup in the CAD family that arose from an angiosperm-specific duplication event and are likely to have gained new functions (Zhang et al. 2006). Besides the differential expression of phenylpropanoid genes, the overrepresentation analysis indicated stress responsive genes as significantly overrepresented, as well as cytochrome P450 genes and genes involved in RNA, regulation of transcription. Nevertheless, when examining the annotation of the individual genes, these seemingly different biological processes might all have the stress response in common. Cytochrome P450 proteins have various roles in plant defense and response to stress situations (Morant et al. 2003; Lapierre et al. 2004; Narusaka et al. 2004; Shinozaki and Yamaguchi-Shinozaki 2007; Ehlting et al. 2008; Cheng et al. 2010; Mao et al. 2013). Also, the genes that were highly differentially expressed in "RNA, regulation of transcription" correspond to hormone and stress responsive transcription factors. Taken together, the *zmcad2* plants might express a constant state of increased stress and defense response, as was reported for many other lignin mutants or transgenic lines (Rohde et al. 2004; Sibout et al. 2005; Dauwe et al. 2007; Leplé et al. 2007). A similar response was also detected in the maize *zmc4h1* mutant, described in chapter5.

In order to get a better insight into the alterations in metabolite abundances as a result of lignin perturbation, we conducted a phenolic profiling of control and *zmcad2* internodes. Downregulation of *CAD* transcript levels results in the incorporation of the hydroxycinnamyl aldehyde monolignol precursors into the lignin polymer, similarly to what was observed in *CAD*-deficient tobacco and poplar plants by NMR (Ralph et al. 2001). The fact that these lignin intermediates are transported to the cell wall and can be incorporated into the lignin polymer along with the traditional H, G and S units, led to the exploration of novel opportunities to design the lignin polymer according to agricultural needs. This concept of lignin engineering provides interesting possibilities for the improvement of the processing efficiency of plant biomass for pulping, forage digestibility and biofuels (Vanholme et al. 2008). In addition to the accumulation of hydroxycinnamyl aldehydes in the lignin polymer, by means of PCA we have shown that the whole metabolome of *zmcad2* plants is highly affected. This was already present from the vegetative stage V10 but became highly apparent in the reproductive stages S, S+7d and S+14d, correlating with the lignification process (Morrison and Kessler 1994; Morrison and Jung 1998). Although only a limited number of compounds could be tentatively identified to date, we were able to show that the phenolic metabolism is altered upon *ZmCAD2* perturbation. There is clearly a reduction in oligolignols, neolignans and flavonolignans that are derived from the monolignols *p*-coumaryl, coniferyl and syringyl alcohol. Incorporation of

the aldehyde precursor forms into the lignin polymer was shown by thioacidolysis. However, we also detected additional changes in phenolic metabolism that indicate possible detoxification routes for accumulating metabolites. To visualize these metabolic changes as a result of *ZmCAD2* perturbation, the relative abundances of 53 tentatively identified compounds and transcripts of lignin and flavonoid biosynthetic genes in S+7d internodes of control and *zmccad2* plants were mapped on the metabolic pathways (as given in Figure 41 in Chapter 5), resulting in the metabolic map as in Figure 56. Phenolic profiling by UPLC-MS revealed an increased abundance of DIBOA, DIMBOA and derivates, tryptophan, phenylalanine, phenylpropanoic acid hexosides and the flavonoids apigenin and tricin. In contrast, *p*-CA derived benzenoids and phenylpropanoic acids, and phenylpropanoic acid 4-O-hexosides were decreased in abundance. The observed differential responses in the different metabolic classes might be related to the presence of the defense response in the case of the flavonoids, the possibility of incorporation into the lignin polymer as compensation for the lack of the conventional monolignols or different detoxification routes at different locations in the plant cell. In order to understand better these metabolic shifts in *zmcad2* plants and lignification in general, the construction of a library of phenolic compounds in maize is of great importance.

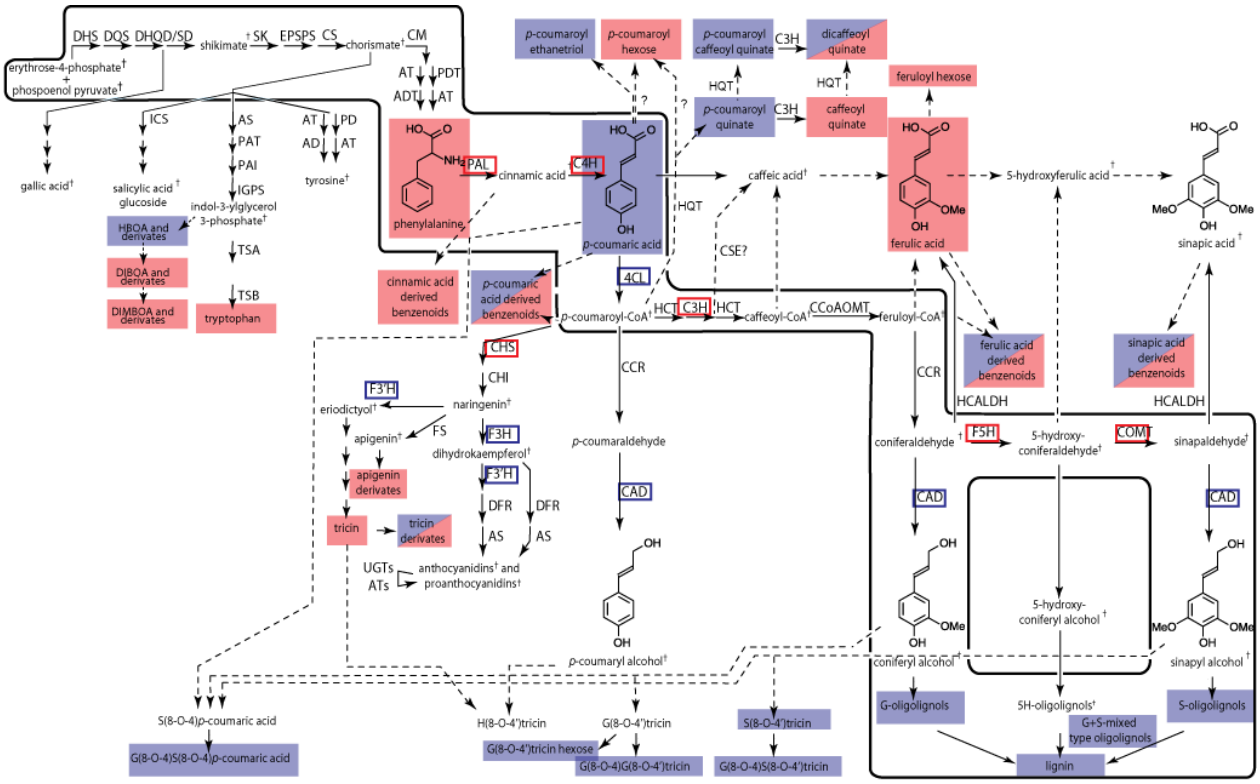


Figure 56. Metabolic shifts in phenolic metabolism at the S+7d stage, as a result of a mutation of *ZmCAD2*. The relative increase and decrease in transcript and metabolite abundances in the mutants, as compared to control, was mapped manually on the pathway as discrete features. Differences in metabolites are indicated via boxes, where red represents a significant increase and blue a significant decrease in abundance. Differences in transcript abundance are indicated with right-angled, framed boxes: significant increases and decreases are visualized via red and blue boxes with solid borders. †: the compound was not detected in the chromatograms of maize control and *zmcad2* internodes. The metabolic map of phenolic metabolism in the internode of maize was

based on the metabolic map of the Arabidopsis stem (Vanholme et al. 2012b) **and adapted for maize using maize literature on flavonoid biosynthesis** (Sharma et al. 2012)**, benzoxazinoids** (Jonczyk et al. 2008) **and maize phenolic profiling (this study).**

6. Conclusions

Despite the reduction in CAD activity and *ZmCAD2* downregulation in the transgenic maize lines, no obvious lignin-related phenotype could be detected. Most likely, a higher reduction was needed which could be attained using stronger promoters to drive the hairpin, such as p*BdUBI10* and p*BdEF1a*. The CAD activity needs to drop below a certain threshold before changes in lignin quantity and composition become apparent. A clear phenotype was observed for a transposon insertion mutant for *ZmCAD2*. The analysis of differential gene expression in four developmental stages provided insight into the feedback mechanisms that are at play in *zmcad2* plants. This feedback mechanism involved a general stress response. It is believed that either the altered metabolic flux or the integrity of the cell wall triggers this stress response in lignin mutants. This should be investigated further in order to prevent yield losses which are sometimes associated with lignin mutants. Understanding the feedback mechanisms on both the transcript and the metabolome level might provide insights into this matter. Currently, the amount of known compounds in the maize metabolome is very limited. The construction of a mass spec database for maize is thus of great importance. Additional to understanding feedback mechanisms, elucidating the structure of hundreds of compounds that are accumulating in the *zmcad2* mutant metabolome could be useful for lignin engineering. Taken together, the micro-array data in combination with metabolic profiling in four developmental stages could form sufficient novelty for publication. The micro-array dataset used in this study could also be valuable for identifying maize genes involved in the lignification process by means of co-expression analysis. The powerful correlation network analysis tool CORNET (De Bodt et al. 2010) could be used for that purpose, provided that its functionalities become fully extended to maize

7. Material and Methods

7.1 Transgenic maize plants

7.1.1 Maize transformation

Using a pearl script, the cDNA sequence of *ZmCAD2* was cut up *in silico* into all possible 21-mers. These oligo sequences were blasted against the full maize cDNA database (AGP_v1) using the CLC bio Genomics workbench (CLC bio, Aarhus, Denmark) and three regions of 150-250 bp (for RNAi1, RNAi2 and RNAi3) were identified with few off-target hits and thus considered as suitable for target-specific downregulation of the gene of interest (Figure 35B). These three regions were cloned into a monocot-specific destination vector pBb7GW-I-WG-UBIL (Figure 35A) using primer sequences in Supplementary Table 12. The vector was transformed into the Agrobacterium strain EHA101. Embryogenic callus of inbred line B104 was used for transformation according to Coussens et al. (2012). An schematic representation of the transformation protocol is shown in Supplementary figure 9. In total, 35 independent transgenic lines were generated for all three RNAi constructs; 8, 17 and 10 for RNAi1, RNAi2 and RNAi3 respectively (Supplementary Table 6).

7.1.2 Plant material

After four months, basta resistant shoots were transferred to soil and grown until maturity. The plants were maintained in a greenhouse at minimum 25°C during day and 23°C during night in a 16h/8h rhythm. Supplementary light was added when natural light intensity was below 200 W/m^2 using high-pressure sodium vapour lamps. Fertilizer was added with the irrigation water; conductivity Ec = 1mS/cm; water soluble fertilizer Poly-feed (Haifa, Belgium) (N, P_2O_5, K_20; 20:5:20 + 3 MgO). Flowering transgenic plants were used as male and female in a backcross to the wildtype B104 line to produce a 50% hemizygous offspring. After seven months, transgenic seed stocks were obtained. For screening of the primary transformants, 30 BC1 seeds per line were sown in 2l pots and grown for six weeks. Transgenic and wildtype progeny plants were selected based on basta resistance using basta leaf painting (Yao et al. 2006). The two youngest, fully expanded leaves were harvested from each plant and the midrib was removed. Leaves of four plants per genotype were pooled and ground together for protein extraction and CAD activity assay.

Three lines were selected based on CAD activity reduction and the number of seeds available for further experiments. For further experiments, the remainder of BC1 seed was used.

For each of the three selected lines, 60 seeds were planted in 2l pots and used as follows:

1. Two weeks after sowing, 30 plants per line were genotyped based on basta resistance with the ammonium multiwell assay, modified from De Block et al. (1995) and Rasco-Gaunt et al. (1999) to use in maize. All glassware was rinsed with MilliQ water (Millipore, Billerica, MA, USA). Fresh leaf discs (approximate diameter 4 mm) were incubated in 200 µl of fresh incubation medium (50 mM potassium phosphate buffer, pH 5.8, 2% sucrose, 25 mg/l glufosinate, 0.1 mg/l 2,4-dichlorophenoxyacetic acid, and 0.1% Tween 20) in a 96-well plate for 24–48 h at 21 C under continuous light (100 µE/m2 s1). Then, 50 µl of the incubation medium was added to 100 µl of reagent A (25 g/l sodium tartrate, 25 g/l trisodium citrate, 34 g sodium salicylate, and 0.0012 g/l sodium nitroprusside) and, subsequently, to 100 µl reagent B (5 g/l NaOH and 3.20 ml sodium hypochlorite solution containing 12% active chlorine) and poured into the plate. The plate was shaken for 5 min at room temperature in the dark and then incubated for 15 min at 37 °C. Blue and white wells corresponded to basta-sensitive control and basta-resistant transgenic plants, respectively. From these 30 genotyped plants, 12 basta resistant and 12 basta sensitive plants were selected for expression analysis. The bottom 1 cm section of the fourth leaf was sampled in three plants and pooled as one biological replicate. This way, four biological replicates were used per genotype for RNA extraction.
2. The other 30 plants were used for measuring leaf growth rates using LEAF-E as described in Chapter 3. Leaf length measurements of the fourth leaf were taken as soon as it appeared from the pseudostem, until full length. After the measurements, also these 30 plants were genotyped using basta leaf painting as described above. For expression analysis and CAD activity assay the two youngest fully expanded leaves were harvested eight weeks after sowing and pooled as one biological replicate. Four biological replicates were used for transgenic and control plants. The eighth and ninth internode were sampled and pooled from the same plants and used for expression and CAD activity assays. The rest of the plants

were repotted in 10L pots and grown until maturity. Dry weight of stems, cobs and leaves was determined. Whole stems were ground using a cutting mill (Fristch, Lelystad, Netherlands) sieve of 0.5 mm. This ground stem material was used for lignin analysis and saccharification efficiency.

7.1.3 CAD activity assay

Enzyme activity assays were performed in different plant tissues (leaf and internode, as detailed above) of transgenic and control plants according to Fornalé et al., (2012) with minor modifications. Frozen tissues were homogenized at 4C in five volumes of extraction buffer (100 mM Tris-HCl pH 7.5, 2% (w/v) PEG 6000, 5 mM DTT, 2% (w/v) PVPP) and centrifuged three times at 10 000 g for 10 min at 4°C to remove cell debris. Protein content was determined by the method of (Bradford 1976) and 300 µg of protein extract were used for the enzymatic assay. Triplicates of each sample were incubated at 30°C in the presence of 100 mM Tris-HCl pH 8.8, 100 µM coniferyl alcohol and 200 µM NADP. Increase in absorbance at 400 nm (indicative of the conversion of coniferyl alcohol to coniferyl aldehyde) was recorded.

7.1.4 Expression analysis using qRT-PCR

qRT-PCR was conducted to determine the expression levels of the transgene and of the target gene *ZmCAD2*. For transgene expression primer sequences were designed, one forward primer in the cloned *ZmCAD2* fragment (one for each construct) and one common reverse primer in the intron sequence of the pBb7GW-I-WG-UBIL vector (Figure 35A) with Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/primer3/) using the standard settings (Supplementary Table 13). The qRT-PCR primers for *ZmCAD2* endogene expression levels were chosen in a region of the cDNA sequence that was not cloned in the RNAi constructs (Supplementary Table 13).

For expression analysis, leaf and internode samples were ground with a Mixer Mill MM 400 and Tungsten carbide 25 ml grinding jars (Retsch, Haan, germany). RNA was extracted using RNeasy kit (Qiagen, Valencia, CA) and a DNase treatment was performed using DNA-free™ (Ambion, Life technologies, Carlsbad, California, U.S.). Extracted RNA was quantified using a nanodrop® ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and diluted so that a total of 400 ng RNA was used for cDNA synthesis using the First strand cDNA synthesis kit (Thermo Scientific, Thermo Fisher Scientific, Waltham, MA, USA). A ten times diluted cDNA sample was used for RT-qPCR using the SensiFAST™ SYBR No-ROX Kit Cat. No. BIO-98020 (BIOLINE) on a Lightcycler 480 (Roche, Basel, Switzerland). Samples were run in technical quadruplicates on the LC480 with following protocol: 1 activation cycle of 10 min at 95°C; 45 amplification cycles of 10 s at 95°C, 10 s at 60 °C and 10 s at 72 °C; 1 melting curve cycle measuring from 65 to 95 °C. Fluorescence values were exported from the lightcycler program whereupon Ct values, normalization factors and primer efficiencies were calculated based upon Ramakers et al. (2003) using *ZmEF1a* and *Zm18S* as reference genes.

7.1.5 Lignin analysis

Lignin analysis was performed as described in addendum to chapter 5.

7.1.6 Cellulose analysis

Aliquots of 5 mg ground stem material were subjected to a sequential extraction to obtain a purified CWR, as described in addendum to chapter 5, lignin analysis. To estimate the amount of cellulose, we used a colorimetric method (DuBois et al. 1956; Leplé et al. 2007). The CWR was incubated with 2 M TFA (trifluoroacetic acid) and 20 µl inositol (5 mg ml-1) for 2 h at 99°C while shaking (750 rpm). After incubation, the remaining pellet was washed three times with water and twice with acetone and dried under vacuum. Concentrated sulfuric acid (150 µl) and 30 µl 5% (w/v) phenol (freshly made in water) were added to the dried pellet and incubated for 1 h at 90°C with gentle shaking (500 rpm). After centrifugation for 3 min at 23477 g, a 50 µl aliquot of the supernatant was diluted 20 times with MilliQ water (Millipore, Billerica, MA, USA) to measure the absorbance at 493 nm. The amount of cellulose was calculated back from a standard curve of Avicel®PH-101 (FMC BioPolymer, Philadelphia, PA, USA).

7.1.7 Saccharification assay

Aliquots of 10 mg of dry stem material were used. The biomass was either pretreated with 1 ml 1M HCl at 80°C for 2h, while shaking (850 rpm) or nor pretreated with acid. The extract was removed and the pretreated material was washed three times with water to obtain a neutral pH. Subsequently, the material was incubated with 1 ml 70% (v/v) ethanol overnight at 55°C. The remaining biomass was washed three times with 1 ml 70% (v/v) ethanol, once with 1 ml acetone, dried under vacuum for 45 min and weighed. The pretreated ethanol-extracted residue was dissolved in 1 ml acetic acid buffer solution (pH 4.8) and incubated at 50°C. The enzyme mix added to the dissolved material contained the Accelerase 1500® mix (Genencor, Danisco). The enzyme mix was first desalted over an Econo-Pac 10DG column (Bio-Rad, Hercules, CA, USA), stacked with Bio-gel® P-6DG gel (Bio-Rad) according to the manufacturer's guidelines. The enzyme mix was further diluted 10-fold and the activity of the diluted enzyme mix was measured with a filter paper assay (Xiao et al. 2004). To each biological sample, dissolved in acetic acid buffer (pH 4.8), the enzyme mix with an activity of 0.002 filter paper units was added. After a short spinning to remove droplets from the lid of the reaction tubes, 20 µl aliquots of the supernatant were taken after 0h, 4h, 7h, 24h and 48h incubation at 50°C and 10 fold diluted with acetic acid buffer (pH4.8). The concentration of glucose in these diluted samples was measured indirectly with a spectrophotometric color reaction (glucose oxidase-peroxidase; GOD-POD) A 100 ml aliquot of the reaction mix from this color reaction contained 50 mg ABTS, 44.83 mg GOD (Sigma-Aldrich, St. Louis, MO, USA) and 173 µl of 4% (w/v) POD (Roche Diagnostics, Brussels, Belgium) in acetic acid buffer (pH 4.5). To measure the concentration of glucose, 50 µl of the diluted samples was added to 150 µl GOD_POD solution and incubated for 30 min at 37 °C. The absorbance was measured spectrophotometrically at a wavelength of 405 nm. The concentration in the original sample was calculated with a standard curve based on known D-glucose (Sigma-Aldrich) concentrations.

7.2 Maize *ZmCAD2* mutant plants

The material and methods section of the transposon insertion mutant was as described in addendum to chapter 5 as material and methods were identical.

8. References

Andersen J, Zein I, Wenzel G (2008) Characterization of phenylpropanoid pathway genes within European maize (Zea mays L.) inbreds. BMC plant … 14:1–14. doi: 10.1186/1471-2229-8-2

Anterola AM, Lewis NG (2002) Trends in lignin modification: a comprehensive analysis of the effects of genetic manipulations/mutations on lignification and vascular integrity. Phytochemistry 61:221–94.

Barrière Y, Chavigneau H, Delaunay S (2013) gene underlie the maize brown-midrib1 (bm1) phenotype with similar effects on lignin characteristics and have potential interest for bioenergy production. Maydica vol58:6–21.

Barrière Y, Riboulet C (2007) Genetics and genomics of lignification in grass cell walls based on maize as model species. … Genomes Genomics

Baucher M, Bernard-Vailhé M a, Chabbert B, et al. (1999) Down-regulation of cinnamyl alcohol dehydrogenase in transgenic alfalfa (Medicago sativa L.) and the effect on lignin composition and digestibility. Plant Mol Biol 39:437–47.

Baucher M, Chabbert B, Pilate G, et al. (1996) Red Xylem and Higher Lignin Extractability by Down-Regulating a Cinnamyl Alcohol Dehydrogenase in Poplar. Plant Physiol 112:1479–1490.

Benjamini Y, Yekutieli D (2001) The control of the false discovery rate in multiple testing under dependency. Ann Stat 29:1165–1188. doi: 10.2307/2674075

Bennetzen J, Hake S (2009) Handbook of Maize: Genetics and Genomics.

Bensen RJ, Johal GS, Crane VC, et al. (1995) Cloning and characterization of the maize An1 gene. Plant Cell 7:75–84.

Block M, Sonville A, Debrouwer D (1995) The selection mechanism of phosphinothricin is influenced by the metabolic status of the tissue. Planta. doi: 10.1007/BF00191569

Bouvier d'Yvoire M, Bouchabke-Coussa O, Voorend W, et al. (2013) Disrupting the cinnamyl alcohol dehydrogenase 1 gene (BdCAD1) leads to altered lignification and improved saccharification in Brachypodium distachyon. Plant J 73:496–508. doi: 10.1111/tpj.12053

Bradford MM (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal Biochem 72:248–54.

Brutnell TP (2002) Transposon tagging in maize. Funct Integr Genomics 2:4–12. doi: 10.1007/s10142-001-0044-0

Casler MD, Pedersen JF, Undersander DJ (2003) Forage Yield and Economic Losses Associated with the Brown-Midrib Trait in Sudangrass. Crop Sci 43:782. doi: 10.2135/cropsci2003.0782

Chabannes M, Barakate A (2001) Strong decrease in lignin content without significant alteration of plant development is induced by simultaneous down-regulation of cinnamoyl CoA reductase. Plant … 28:257–270.

Chen L, Auh C-K, Dowling P, et al. (2003) Improved forage digestibility of tall fescue (Festuca arundinacea) by transgenic down-regulation of cinnamyl alcohol dehydrogenase. Plant Biotechnol J 1:437–49. doi: 10.1046/j.1467-7652.2003.00040.x

Chen W, VanOpdorp N, Fitzl D, et al. (2012) Transposon insertion in a cinnamyl alcohol dehydrogenase gene is responsible for a brown midrib1 mutation in maize. Plant Mol Biol 80:289–97. doi: 10.1007/s11103-012-9948-4

Cheng DW, Lin H, Takahashi Y, et al. (2010) Transcriptional regulation of the grape cytochrome P450 monooxygenase gene CYP736B expression in response to Xylella fastidiosa infection. BMC Plant Biol 10:135. doi: 10.1186/1471-2229-10-135

Cheong YH, Chang HS, Gupta R, et al. (2002) Transcriptional profiling reveals novel interactions between wounding, pathogen, abiotic stress, and hormonal responses in Arabidopsis. Plant Physiol 129:661–677. doi: 10.1104/pp.002857

Christensen a H, Quail PH (1996) Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. Transgenic Res 5:213–8.

Coussens G, Aesaert S, Verelst W, et al. (2012) Brachypodium distachyon promoters as efficient building blocks for transgenic research in maize. J Exp Bot 63:4263–73. doi: 10.1093/jxb/ers113

Dai S, Zheng P, Marmey P, et al. (2001) Comparative analysis of transgenic rice plants obtained by Agrobacterium-mediated transformation and particle bombardment. Mol Breed 7:25–33. doi: 10.1023/A:1009687511633

Dauwe R, Morreel K, Goeminne G, et al. (2007) Molecular phenotyping of lignin-modified tobacco reveals associated changes in cell-wall metabolism, primary metabolism, stress metabolism and photorespiration. Plant J 52:263–85. doi: 10.1111/j.1365-313X.2007.03233.x

De Bodt S, Carvajal D, Hollunder J, et al. (2010) CORNET: a user-friendly tool for data mining and integration. Plant Physiol 152:1167–79. doi: 10.1104/pp.109.147215

Dick R, Rattei T, Haslbeck M, et al. (2012) Comparative analysis of benzoxazinoid biosynthesis in monocots and dicots: independent recruitment of stabilization and activation functions. Plant Cell 24:915–28. doi: 10.1105/tpc.112.096461

DuBois M, Gilles KA, Hamilton JK, et al. (1956) Colorimetric Method for Determination of Sugars and Related Substances. Anal Chem 28:350–356. doi: 10.1021/ac60111a017

Eamens A, Wang M-B, Smith NA, Waterhouse PM (2008) RNA silencing in plants: yesterday, today, and tomorrow. Plant Physiol 147:456–68. doi: 10.1104/pp.108.117275

Ehlting J, Sauveplane V, Olry A, et al. (2008) An extensive (co-)expression analysis tool for the cytochrome P450 superfamily in Arabidopsis thaliana.

BMC Plant Biol 8:47. doi: 10.1186/1471-2229-8-47

Eudes A, Pollet B, Sibout R, et al. (2006) Evidence for a role of AtCAD 1 in lignification of elongating stems of Arabidopsis thaliana. Planta 225:23–39. doi: 10.1007/s00425-006-0326-9

Fernandes J, Dong Q, Schneider B, et al. (2004) Genome-wide mutagenesis of Zea mays L. using RescueMu transposons. Genome Biol 5:R82.

Fornalé S, Capellades M, Encina A, et al. (2012) Altered lignin biosynthesis improves cellulosic bioethanol production in transgenic maize plants down-regulated for cinnamyl alcohol dehydrogenase. Mol Plant 5:817–30. doi: 10.1093/mp/ssr097

Frame BR, McMurray JM, Fonger TM, et al. (2006a) Improved Agrobacterium-mediated transformation of three maize inbred lines using MS salts. Plant Cell Rep 25:1024–34. doi: 10.1007/s00299-006-0145-2

Frame BR, Paque T, Wang K (2006b) Maize (Zea mays L.). Methods Mol Biol 343:185–199.

Frey M, Schullehner K, Dick R, et al. (2009) Benzoxazinoid biosynthesis, a model for evolution of secondary metabolic pathways in plants. Phytochemistry 70:1645–1651. doi: 10.1016/j.phytochem.2009.05.012

Fu C, Xiao X, Xi Y, et al. (2011) Downregulation of Cinnamyl Alcohol Dehydrogenase (CAD) Leads to Improved Saccharification Efficiency in Switchgrass. BioEnergy Res 4:153–164. doi: 10.1007/s12155-010-9109-z

Grabber JH (2005) How Do Lignin Composition, Structure, and Cross-Linking Affect Degradability? A Review of Cell Wall Model Studies. Crop Sci 45:820. doi: 10.2135/cropsci2004.0191

Guillaumie S, San-Clemente H, Deswarte C, et al. (2007) MAIZEWALL. Database and developmental gene expression profiling of cell wall biosynthesis and assembly in maize. Plant Physiol 143:339–363.

Halpin C, Holt K, Chojecki J, et al. (1998) Brown-midrib maize (bm1)–a mutation affecting the cinnamyl alcohol dehydrogenase gene. Plant … 14:545–53. doi: 10.1046/j.1365-313X.1998.00153.x

Halpin C, Knight M (1994) Manipulation of lignin quality by downregulation of cinnamyl alcohol dehydrogenase. Plant …

Han K-J, Pitman WD, Kim M, et al. (2013) Ethanol production potential of sweet sorghum assessed using forage fiber analysis procedures. GCB Bioenergy 5:358–366. doi: 10.1111/j.1757-1707.2012.01203.x

Harris PJ, Trethewey J a. K (2010) The distribution of ester-linked ferulic acid in the cell walls of angiosperms. Phytochem Rev 9:19–33. doi: 10.1007/s11101-009-9146-4

Hatfield R, Fukushima RS (2005) Can Lignin Be Accurately Measured? Crop Sci 45:832. doi: 10.2135/cropsci2004.0238

Hatfield RD, Ralph J, Grabber JH (1999) Cell wall cross-linking by ferulates and diferulates in grasses. J Sci Food Agric 79:403–407. doi: 10.1002/(SICI)1097-0010(19990301)79:3<403::AID-JSFA263>3.0.CO;2-0

Huang X, Wei Z (2005) Successful Agrobacterium-mediated genetic transformation of maize elite inbred lines. Plant Cell Tissue Organ Cult 187–200. doi: 10.1007/s11240-005-5772-8

Ishida Y, Hiei Y, Komari T (2007) Agrobacterium-mediated transformation of maize. Nat Protoc 2:1614–21. doi: 10.1038/nprot.2007.241

Ishii T (1997) Structure and functions of feruloylated polysaccharides. Plant Sci 127:111–127. doi: 10.1016/S0168-9452(97)00130-1

Jonczyk R, Schmidt H, Osterrieder A, et al. (2008) Elucidation of the final reactions of DIMBOA-glucoside biosynthesis in maize: characterization of Bx6 and Bx7. Plant Physiol 146:1053–63. doi: 10.1104/pp.107.111237

Jung H-JG (2003) Maize stem tissues: ferulate deposition in developing internode cell walls. Phytochemistry 63:543–549. doi: 10.1016/S0031-9422(03)00221-8

Kiedrowski S, Kawalleck P, Hahlbrock K, et al. (1992) Rapid activation of a novel plant defense gene is strictly dependent on the Arabidopsis RPM1 disease resistance locus. EMBO J 11:4677–4684.

Kim H, Ralph J, Lu F, et al. (2002) Identification of the structure and origin of thioacidolysis marker compounds for cinnamyl alcohol dehydrogenase deficiency in angiosperms. J Biol Chem 277:47412–9. doi: 10.1074/jbc.M208860200

Kim H, Ralph J, Yahiaoui N, et al. (2000) Cross-coupling of hydroxycinnamyl aldehydes into lignins. Org Lett 2:2197–2200. doi: 10.1021/ol005906o

Kim S-J, Kim M-R, Bedgar DL, et al. (2004) Functional reclassification of the putative cinnamyl alcohol dehydrogenase multigene family in Arabidopsis. Proc Natl Acad Sci U S A 101:1455–1460.

Lapierre C, Pilate G, Pollet B, et al. (2004) Signatures of cinnamyl alcohol dehydrogenase deficiency in poplar lignins. Phytochemistry 65:313–321. doi: 10.1016/j.phytochem.2003.11.007

Lauter FR (1996) Root-specific expression of the LeRse-1 gene in tomato is induced by exposure of the shoot to light. Mol Gen Genet 252:751–754.

Leplé J-C, Dauwe R, Morreel K, et al. (2007) Downregulation of cinnamoyl-coenzyme A reductase in poplar: multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. Plant Cell 19:3669–91. doi: 10.1105/tpc.107.054148

Li X, Chen W, Zhao Y, et al. (2013) Downregulation of caffeoyl-CoA O-methyltransferase (CCoAOMT) by RNA interference leads to reduced lignin production in maize straw. Genet Mol Biol 36:540–6. doi: 10.1590/S1415-47572013005000039

Lindbo JA, Silva-Rosales L, Proebsting WM, Dougherty WG (1993) Induction of a Highly Specific Antiviral State in Transgenic Plants: Implications for Regulation of Gene Expression and Virus Resistance. Plant Cell 5:1749–1759. doi: 10.1105/tpc.5.12.1749

Lorenz AJ, Anex RP, Isci A, et al. (2009) Forage quality and composition measurements as predictors of ethanol yield from maize

(Zea mays L.) stover. Biotechnol Biofuels 2:5. doi: 10.1186/1754-6834-2-5

MacAdam JW, Grabber JH (2002) Relationship of growth cessation with the formation of diferulate cross-links and p-coumaroylated lignins in tall fescue leaf blades. Planta 215:785–93. doi: 10.1007/s00425-002-0812-7

Mansell RL, Gross GG, Stöckigt J, et al. (1974) Purification and properties of cinnamyl alcohol dehydrogenase from higher plants involved in lignin biosynthesis. Phytochemistry 13:2427–2435. doi: 10.1016/S0031-9422(00)86917-4

Mao G, Seebeck T, Schrenker D, Yu O (2013) CYP709B3, a cytochrome P450 monooxygenase gene involved in salt tolerance in Arabidopsis thaliana. BMC Plant Biol 13:169. doi: 10.1186/1471-2229-13-169

May BP, Liu H, Vollbrecht E, et al. (2003) Maize-targeted mutagenesis: A knockout resource for maize. Proc Natl Acad Sci U S A 100:11541–11546.

McCarty DR, Latshaw S, Wu S, et al. (2013) Mu-seq: sequence-based mapping and identification of transposon induced mutations. PLoS One 8:e77172. doi: 10.1371/journal.pone.0077172

McCarty DR, Settles AM, Suzuki M, et al. (2005) Steady-state transposon mutagenesis in inbred maize. Plant J 44:52–61. doi: 10.1111/j.1365-313X.2005.02509.x

Meyer P, Saedler H (1996) HOMOLOGY-DEPENDENT GENE SILENCING IN PLANTS. Annu Rev Plant Physiol Plant Mol Biol 47:23–48. doi: 10.1146/annurev.arplant.47.1.23

Molinari HBC, Pellny TK, Freeman J, et al. (2013) Grass cell wall feruloylation: distribution of bound ferulate and candidate gene expression in Brachypodium distachyon. Front Plant Sci 4:50. doi: 10.3389/fpls.2013.00050

Morant M, Bak S, Møller BL, Werck-Reichhart D (2003) Plant cytochromes P450: tools for pharmacology, plant protection and phytoremediation. Curr Opin Biotechnol 14:151–62.

Morrison T, Jung H (1998) Cell-wall composition of maize internodes of varying maturity. Crop Sci.

Morrison T, Kessler J (1994) Activity of two lignin biosynthesis enzymes during development of a maize internode. J … 133–139.

Narusaka Y, Narusaka M, Seki M, et al. (2004) Crosstalk in the responses to abiotic and biotic stresses in Arabidopsis: analysis of gene expression in cytochrome P450 gene superfamily by cDNA microarray. Plant Mol Biol 55:327–42. doi: 10.1007/s11103-004-0685-1

Negrotto D, Jolley M, Beer S, et al. (2000) The use of phosphomannose-isomerase as a selectable marker to recover transgenic maize plants (Zea mays L.) via Agrobacterium transformation. Plant Cell Rep 19:798–803. doi: 10.1007/s002999900187

Ossowski S, Schwab R, Weigel D (2008) Gene silencing in plants using artificial microRNAs and other small RNAs. Plant J 53:674–690. doi: 10.1111/j.1365-313X.2007.03328.x

Palmer N a, Sattler SE, Saathoff AJ, et al. (2008) Genetic background impacts

soluble and cell wall-bound aromatics in brown midrib mutants of sorghum. Planta 229:115–27. doi: 10.1007/s00425-008-0814-1

Pan Y, Michael TP, Hudson ME, et al. (2009) Cytochrome P450 monooxygenases as reporters for circadian-regulated pathways. Plant Physiol 150:858–78. doi: 10.1104/pp.108.130757

Park S-H, Mei C, Pauly M, et al. (2012) Downregulation of Maize Cinnamoyl-Coenzyme A Reductase via RNA Interference Technology Causes Brown Midrib and Improves Ammonia Fiber Expansion-Pretreated Conversion into Fermentable Sugars for Biofuels. Crop Sci 52:2687. doi: 10.2135/cropsci2012.04.0253

Pedersen JF, Vogel KP, Funnell DL (2005) Impact of Reduced Lignin on Plant Fitness. Crop Sci 45:812. doi: 10.2135/cropsci2004.0155

Pillonel C, Mulder MM, Boon JJ, et al. (1991) Involvement of cinnamyl-alcohol dehydrogenase in the control of lignin formation in Sorghum bicolor L. Moench. Planta 185:538–44. doi: 10.1007/BF00202964

Quirino BF, Normanly J, Amasino RM (1999) Diverse range of gene activity during Arabidopsis thaliana leaf senescence includes pathogen-independent induction of defense-related genes. Plant Mol Biol 40:267–278. doi: 10.1023/A:1006199932265

Raes J, Rohde A, Christensen JH, et al. (2003) Genome-wide characterization of the lignification toolbox in Arabidopsis. Plant … 133:1051–1071. doi: 10.1104/pp.103.026484.role

Ralph J, Kim H, Lu F, et al. (2008) Identification of the structure and origin of a thioacidolysis marker compound for ferulic acid incorporation into angiosperm lignins (and an indicator for cinnamoyl CoA reductase deficiency). Plant J 53:368–79. doi: 10.1111/j.1365-313X.2007.03345.x

Ralph J, Lapierre C, Marita JM, et al. (2001) Elucidation of new structures in lignins of CAD- and COMT-deficient plants by NMR. Phytochemistry 57:993–1003.

Ramakers C, Ruijter JM, Deprez RHL, Moorman AF. (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data. Neurosci Lett 339:62–66. doi: 10.1016/S0304-3940(02)01423-4

Rasco-Gaunt S, Riley A, Lazzeri P, Barcelo P (1999) A facile method for screening for phosphinothricin (PPT)-resistant transgenic wheats. Mol Breed 5:255–262. doi: 10.1023/A:1009689906936

Riboulet C, Guillaumie S, Méchin V (2009) Kinetics of phenylpropanoid gene expression in maize growing internodes: relationships with cell wall deposition. Crop … 211–223. doi: 10.2135/cropsci2008.03.0130

Roadhouse EF, MacDougall D (1956) A study of the nature of plant lignin by means of alkaline nitrobenzene oxidation.

Rohde A, Morreel K, Ralph J (2004) phenotyping of the pal1 and pal2 mutants of Arabidopsis thaliana reveals far-reaching consequences on phenylpropanoid, amino acid, and

carbohydrate metabolism. Plant Cell … 16:2749–2771. doi: 10.1105/tpc.104.023705.phospho-enol-pyruvate

Saathoff A, Sarath G, Chow E, et al. (2011) Downregulation of cinnamyl-alcohol dehydrogenase in switchgrass by RNA silencing results in enhanced glucose release after cellulase treatment. PLoS One 6:e16416. doi: 10.1371/journal.pone.0016416

Sarath G, Dien B, Saathoff AJ, et al. (2011) Ethanol yields and cell wall properties in divergently bred switchgrass genotypes. Bioresour Technol 102:9579–85. doi: 10.1016/j.biortech.2011.07.086

Sattler SE, Saathoff AJ, Haas EJ, et al. (2009) A nonsense mutation in a cinnamyl alcohol dehydrogenase gene is responsible for the Sorghum brown midrib6 phenotype. Plant Physiol 150:584–95. doi: 10.1104/pp.109.136408

Schledzewski K, Mendel RR (1994) Quantitative transient gene expression: Comparison of the promoters for maize polyubiquitin1, rice actin1, maize-derivedEmu andCaMV 35S in cells of barley, maize and tobacco. Transgenic Res 3:249–255. doi: 10.1007/BF02336778

Settles AM, Holding DR, Tan BC, et al. (2007) Sequence-indexed mutations in maize using the UniformMu transposon-tagging population. BMC Genomics 8:116. doi: 10.1186/1471-2164-8-116

Settles AM, Latshaw S, McCarty DR (2004) Molecular analysis of high-copy insertion sites in maize. Nucleic Acids Res 32:e54.

Sharma M, Chai C, Morohashi K, et al. (2012) Expression of flavonoid 3'-hydroxylase is controlled by P1, the regulator of 3-deoxyflavonoid biosynthesis in maize. BMC Plant Biol 12:196. doi: 10.1186/1471-2229-12-196

Shi C, Koch G, Ouzunova M, et al. (2006) Comparison of maize brown-midrib isogenic lines by cellular UV-microspectrophotometry and comparative transcript profiling. Plant Mol Biol 62:697–714. doi: 10.1007/s11103-006-9049-3

Shinozaki K, Yamaguchi-Shinozaki K (2007) Gene networks involved in drought stress response and tolerance. J Exp Bot 58:221–7. doi: 10.1093/jxb/erl164

Shou H, Frame BR, Whitham SA, Wang K (2004) Assessment of transgenic maize events produced by particle bombardment or Agrobacterium-mediated transformation. Mol Breed 201–208. doi: 10.1023/B:MOLB.0000018767.64586.53

Sibout R, Eudes A, Mouille G, et al. (2005) CINNAMYL ALCOHOL DEHYDROGENASE-C and -D are the primary genes involved in lignin biosynthesis in the floral stem of Arabidopsis. Plant Cell 17:2059–76. doi: 10.1105/tpc.105.030767

Trabucco GM, Matos D a, Lee SJ, et al. (2013) Functional characterization of cinnamyl alcohol dehydrogenase and caffeic acid O-methyltransferase in Brachypodium distachyon. BMC Biotechnol 13:61. doi: 10.1186/1472-6750-13-61

Tsuruta S, Ebina M, Kobayashi M, et al. (2010) Structure and expression profile of the

cinnamyl alcohol dehydrogenase gene and its association with lignification in the sorghum (Sorghum bicolor (L.) Moench) bmr-6 mutant. Breed Sci 60:314–323. doi: 10.1270/jsbbs.60.314

Vailhé B, Andrée M (1998) Effect of downregulation of cinnamyl alcohol dehydrogenase on cell wall composition and on degradability of tobacco stems. J. … 505:

Valério L, Carter D, Rodrigues JCJ, et al. (2003) Down regulation of cinnamyl alcohol dehydrogenase, a lignification enzyme, in Eucalyptus camaldulensis. Mol Breed 12:157–167. doi: 10.1023/A:1026070725107

Van Soest PJ, Robertson JB, Lewis B a (1991) Methods for dietary fiber, neutral detergent fiber, and nonstarch polysaccharides in relation to animal nutrition. J Dairy Sci 74:3583–97. doi: 10.3168/jds.S0022-0302(91)78551-2

Vanholme R, Morreel K, Ralph J, Boerjan W (2008) Lignin engineering. Curr Opin Plant Biol 11:278–85. doi: 10.1016/j.pbi.2008.03.005

Vanholme R, Storme V, Vanholme B, et al. (2012) A systems biology view of responses to lignin biosynthesis perturbations in Arabidopsis. Plant Cell 24:3506–29. doi: 10.1105/tpc.112.102574

Vermerris W, Sherman DM, McIntyre LM (2010) Phenotypic plasticity in cell walls of maize brown midrib mutants is limited by lignin composition. J Exp Bot 61:2479–90. doi: 10.1093/jxb/erq093

Voelker SL, Lachenbruch B, Meinzer FC, et al. (2010) Antisense down-regulation of 4CL expression alters lignification, tree growth, and saccharification potential of field-grown poplar. Plant Physiol 154:874–86. doi: 10.1104/pp.110.159269

Vogel KP, Mitchell RB, Sarath G, et al. (2013) Switchgrass Biomass Composition Altered by Six Generations of Divergent Breeding for Digestibility. Crop Sci 53:853. doi: 10.2135/cropsci2012.09.0542

Walbot V, Warren C (1988) Regulation of Mu element copy number in maize lines with an active or inactive Mutator transposable element system. Mol Gen Genet 211:27–34.

Warthmann N, Chen H, Ossowski S, et al. (2008) Highly specific gene silencing by artificial miRNAs in rice. PLoS One 3:13–15. doi: 10.1371/journal.pone.0001829

Werck-Reichhart D (1995) Cytochromes P450 in phenylpropanoid metabolism. Drug Metabol Drug Interact 12:221–43.

Williams-Carrier R, Stiffler N, Belcher S, et al. (2010) Use of Illumina sequencing to identify transposon insertions underlying mutant phenotypes in high-copy Mutator lines of maize. Plant J 63:167–177. doi: 10.1111/j.1365-313X.2010.04231.x

Xiao Z, Storms R, Tsang A (2004) Microplate-based filter paper assay to measure total cellulase activity. Biotechnol Bioeng 88:832–7. doi: 10.1002/bit.20286

Yao Q, Cong L, Chang JL, et al. (2006) Low copy number gene transfer and stable expression in a commercial wheat cultivar via particle bombardment. J Exp

Bot 57:3737–46. doi: 10.1093/jxb/erl145

Yi G, Luth D, Goodman TD, et al. (2009) High-throughput linkage analysis of Mutator insertion sites in maize. Plant J 58:883–92. doi: 10.1111/j.1365-313X.2009.03821.x

Zhang K, Qian Q, Huang Z, et al. (2006) GOLD HULL AND INTERNODE2 encodes a primarily multifunctional cinnamyl-alcohol dehydrogenase in rice. Plant … 140:972–983. doi: 10.1104/pp.105.073007.cies

Zhao ZY, Cai T, Tagliani L, et al. (2000) Agrobacterium-mediated sorghum transformation. Plant Mol Biol 44:789–798.

9. Supplemental figures and tables

Supplementary Table 6. Segregation analysis on independent transgenic lines based on the resistance and sensitivity to the herbicide basta. Resistant means transgenic and sensitive means segregating control plants. A p-value of less than 0.05 in the chi-square test was used for selecting lines deviating from the 0.5 frequency if segregating transgenic and control plants.

transgenic line	RNAi construct	basta sensitive	basta resistent	p value of chi square test for 0.5 frequency
103-01	RNAi_1	4	26	0.00006
103-04	RNAi_1	8	22	0.011
103-07	RNAi_1	4	25	0.0001
103-11	RNAi_1	16	14	0.715
103-15	RNAi_1	16	14	0.715
106-01	RNAi_1	17	12	0.353
106-03	RNAi_1	23	7	0.003
106-05	RNAi_1	13	16	0.577
104-01	RNAi_2	9	21	0.028
104-03	RNAi_2	14	15	0.853
104-05	RNAi_2	11	19	0.144
104-07	RNAi_2	7	21	0.008
104-09	RNAi_2	18	12	0.273
104-11	RNAi_2	15	15	1.000
104-13	RNAi_2	12	15	0.564
104-15	RNAi_2	20	10	0.068
104-17	RNAi_2	13	16	0.577
104-19	RNAi_2	12	18	0.273
108-01	RNAi_2	17	11	0.257
108-03	RNAi_2	10	16	0.239
108-05	RNAi_2	15	13	0.705
108-07	RNAi_2	14	12	0.695
108-09	RNAi_2	15	14	0.853
107-01	RNAi_2	15	12	0.564
107-03	RNAi_2	7	8	0.796
108-11	RNAi_3	15	14	0.853
108-13	RNAi_3	17	13	0.465
107-07	RNAi_3	17	13	0.465
107-09	RNAi_3	10	19	0.095
107-11	RNAi_3	2	24	0.00002
107-13	RNAi_3	12	18	0.273
107-18	RNAi_3	15	14	0.853
107-20	RNAi_3	17	13	0.465
110-01	RNAi_3	15	15	1.000
107-16	RNAi_3	8	17	0.072

Supplementary Table 7. Number of biological replicates for biomass and growth measurements in control and ZmCAD2 RNAi plants in the segregating population of three independent transgenic lines

	103-01		104-11		107-20	
	control	ZmCAD2 RNAi	control	ZmCAD2 RNAi	control	ZmCAD2 RNAi
leaf DW (g)	10	7	7	9	5	10
stem DW (g)	10	7	7	9	5	10
cob DW (g)	10	7	7	9	5	10
total aboveground plant DW (g)	10	7	7	9	5	10
days to germination (d)	30	28	28	28	30	27
leaf #4 length (mm)	19	15	10	13	12	13
Leaf #4 elongation rate [mm(°Cd^{-1})]	19	15	10	13	12	13

Supplementary Table 8. Full list of estimated NIRS parameters for the *zmcad2* transposon mutant and corresponding control plants. IVDMD and IVNDFD from Table 14 correspond to CASEAUF and FDP respectively in this table.

NIRS Parameter	units	control	*zmcad2*	fold change (%)	p-value
Acid detergent Fiber (ADF)	% of Dry matter	32.44	28.05	-13.6	**0.0004**
Neutral Detergent Fiber (NDF)	% of Dry matter	55.04	50.73	-7.8	**0.001**
acid detergent lignin (ADL)	% of Dry matter	3.66	2.55	-30.4	**0.0002**
Raw cellulose rate (CB)	% of Dry matter	27.94	24.98	-10.6	**0.0004**
Organic matter content (MO)	% of Dry matter	0.94	0.93	-1.1	**0.013**
Analytic dry matter (MSA)	gram	94.45	94.31	-0.1	0.617
% Dry matter digestibility Aufrère enzymatic method (CASEAUF)	% of Dry matter	50.05	59.04	18.0	**0.00002**
% Organic matter digestibility Aufrère enzymatic method (CASEMOAUF)	% of Dry matter	48.14	57.46	19.4	**0.00001**
% Dry matter digestibility (DMSAUF)	% of Dry matter	58.31	63.09	8.2	**0.0005**
Organic matter digestibility (DMO)	% of Dry matter	62.17	68.10	9.5	**0.0001**
Digestible part of the cell wal (FDP)	% of Dry matter	29.91	40.97	37.0	**0.00001**
Soluble sugars rate (SSR)	% of Dry matter	14.08	16.59	17.8	**0.001**
In Vitro Digestibility of the "Non-Starch, soluble Carbohydrates part" (DINAG)	% of Dry matter	41.88	50.89	21.5	**0.00001**
Dry matter protein rate (MPT)	% of Dry matter	7.87	9.01	14.5	0.078
energy value for milk production /kg of dry matter (UFLMS)	French Milk Feed Unit for Energy	0.71	0.81	13.4	**0.0004**
energy value for meat production /kg of dry matter (UFVMS)	French Meat Feed Unit for Energy	0.59	0.70	18.9	**0.0003**
Organic matter protein rate (MATO)	% of organic matter	83.91	97.26	15.9	0.061
ester ferulic acid rate (FE25)	% of Dry matter	7.11	6.69	-5.9	0.135
ether ferulic acid rate (FE170)	% of Dry matter	8.99	8.53	-5.1	**0.015**
5-5 diFerulic acid rate (DIF55)	% of Dry matter	0.17	0.21	24.4	**0.003**
8-O-4 diFerulic acid rate (DIF8O4)	% of Dry matter	0.31	0.36	16.5	**0.001**
ester *p*-coumaric acid rate (CO25)	% of Dry matter	15.29	11.43	-25.3	**0.00001**
p-Hydroxybenzaldehyde. H sub-units rate (H4N)	% of Dry matter	1.80	0.91	-49.3	**0.00001**
Syringaldehyde. S sub-units rate (SAN)	% of Dry matter	8.68	5.63	-35.2	**0.000002**
Vanillin. G sub-units rate (VAN)	% of Dry matter	8.81	6.92	-21.4	**0.0002**
Quality parameter 1		1.24	1.76	41.9	**0.001**
Quality parameter 2		2.04	2.98	46.0	**0.0007**

Supplementary Table 9. Biological function and fold changes in the *zmcad2* of differentially expressed genes in the functional categories 'stress response', 'cytochrome P450' and 'RNA, regulation of transcription'. Log2 fold changes were shown higher than |1| for stages S (silking), S+7d (seven days after silking) and S+14d (fourteen days after silking).

		log2 fold change		
MapMan Bin Name	Transcript ID	S	S+7d	S+14d
stress.biotic	Hidden	-1.3422	-2.2105	-2.7577
stress.biotic		1.0271	1.4644	0.8704
stress.biotic		-1.2378	-2.8854	-3.2999
stress.biotic		0.9794	-0.9172	1.3784
stress.biotic		0.835	-0.6769	-1.2063
stress.biotic		-0.6447	-0.7368	-1.2885
stress.biotic		-0.8973	-1.8938	-1.9141
stress.biotic		1.3579	1.9175	1.2438
stress.biotic.respiratory burst		0.277	1.0129	0.4821
stress.biotic.PR-proteins		-1.3112	-1.1283	-0.703
stress.biotic.PR-proteins		0.9542	1.2315	0.7286
stress.biotic.PR-proteins		-0.4836	-1.08	-0.8236
stress.abiotic.heat		1.9311	1.4935	1.4134
stress.abiotic.heat		1.0116	-1.7814	0.3802
stress.abiotic.heat		0.5799	-1.09	0.3458
stress.abiotic.heat		0.6884	-1.0702	0.6018
stress.abiotic.heat		0.9348	-1.1604	0.4435
stress.abiotic.heat		0.4173	-0.7994	1.0212
stress.abiotic.heat		1.0466	0.9853	0.9468
stress.abiotic.heat		1.6689	-1.1792	0.7731
stress.abiotic.drought/salt		0.3917	1.2398	0.6607
stress.abiotic.drought/salt		0.7219	1.0433	0.5616
misc.cytochrome P450		-0.8021	-1.4329	-1.4752
misc.cytochrome P450		1.5855	1.1824	1.5805
misc.cytochrome P450		-1.0703	-0.6749	-1.1306
misc.cytochrome P450		1.045	0.5265	0.7172
misc.cytochrome P450		-1.1414	-0.652	-0.6304
misc.cytochrome P450		-1.0081	-2.2765	-1.2771
misc.cytochrome P450		-1.172	-2.0837	-1.9395
RNA.regulation of transcription.AP2/EREBP, APETALA2/Ethylene-responsive element binding protein family		-1.2281	0.7897	1.3949
RNA.regulation of transcription.AP2/EREBP, APETALA2/Ethylene-responsive element binding protein family		0.9627	0.4628	-1.2546
RNA.regulation of transcription.ARR		-1.4117	0.9215	1.0548
RNA.regulation of transcription.bHLH,Basic Helix-Loop-Helix family		-0.9623	-2.3266	-1.2458
RNA.regulation of transcription.C2H2 zinc finger family		-0.3974	0.4758	1.008
RNA.regulation of transcription.HB,Homeobox transcription factor family		1.4152	1.8729	1.6368
RNA.regulation of transcription.HB,Homeobox transcription factor family		-1.0032	-0.665	-0.485
RNA.regulation of transcription.HB,Homeobox transcription factor family		0.3999	1.0684	1.0046
RNA.regulation of transcription.HSF,Heat-shock transcription factor family		1.1136	-0.8284	1.4273
RNA.regulation of transcription.MADS box transcription factor family		0.4005	1.0792	0.4684
RNA.regulation of transcription.MYB domain transcription factor family		-1.0237	-0.6689	-0.3867
RNA.regulation of transcription.MYB-related transcription factor family		-1.4155	-1.955	-1.2976
RNA.regulation of transcription.SBP,Squamosa promoter binding protein family		0.9244	1.0782	0.6018
RNA.regulation of transcription.Aux/IAA family		-0.3194	1.5413	1.5804
RNA.regulation of transcription.B3 transcription factor family		-0.7496	-1.5745	-1.2735
RNA.regulation of transcription.FHA transcription factor		2.0779	2.4108	1.8377
RNA.regulation of transcription.putative transcription regulator		1.2697	1.1932	1.4601
RNA.regulation of transcription.unclassified		-0.8118	-1.6175	-0.829
RNA.regulation of transcription.unclassified		-1.7983	-1.8541	-1.3796

Supplementary Table 10. Expression levels and representation of change in expression over development of putative phenylpropanoid genes involved in lignification of maize wildtype internodes. Increasing and decreasing expression was illustrated by subtracting the mean expression per gene over development from the mean normalized expression value per gene in each stage and color code all subsequent values for higher (red) and lower (blue) expression than average. The expression level was labeled according to maximal expression: "very high" >=15, 15>"high">=13, 13>"moderate">=11, 11>"low". V10 (ten visible leaf collars), S (silking), S+7d (seven days after silking) and S+14d (fourteen days after silking).

			mean normalized expression (log2)					[(normalized expression)-(mean per gene)] (log2)			
protein name	gene name	probe name	V10	S	S+7d	S+14d	expression level	V10	S	S+7d	S+14d
PAL	Hidden		13.97	15.08	14.61	14.64	very high	-0.61	0.50	0.04	0.07
PAL			13.14	14.29	13.87	14.00	high	-0.69	0.46	0.05	0.18
PAL			14.46	14.53	14.54	14.51	high	-0.05	0.02	0.03	0.00
PAL			13.84	14.98	14.65	14.77	high	-0.72	0.42	0.09	0.21
PAL			13.46	14.72	14.04	14.12	high	-0.62	0.63	-0.05	0.04
PAL			12.42	13.12	12.95	13.29	high	-0.52	0.18	0.00	0.35
PAL			14.37	14.74	14.08	14.04	high	0.06	0.44	-0.22	-0.27
PAL			10.70	10.86	10.80	11.00	moderate	-0.14	0.02	-0.04	0.16
PAL			10.13	9.84	9.81	9.86	low	0.22	-0.07	-0.09	-0.05
PAL			9.37	9.43	9.29	9.46	low	-0.02	0.05	-0.10	0.07
PAL			9.68	9.68	9.66	9.94	low	-0.06	-0.06	-0.08	0.20
C4H			12.91	12.98	12.66	13.41	high	-0.08	-0.01	-0.33	0.42
C4H			11.50	11.49	11.41	11.93	moderate	-0.08	-0.10	-0.17	0.35
C4H			9.67	9.69	9.80	9.96	low	-0.11	-0.09	0.02	0.18
4CL			11.32	11.18	11.10	11.39	moderate	0.07	-0.07	-0.15	0.14
4CL			12.34	12.65	12.14	11.55	moderate	0.17	0.48	-0.03	-0.62
4CL			12.59	12.93	12.42	11.58	moderate	0.21	0.55	0.04	-0.80
4CL			11.12	10.02	9.86	9.92	moderate	0.89	-0.21	-0.37	-0.31
4CL			12.13	10.69	10.53	10.64	moderate	1.14	-0.31	-0.47	-0.36
4CL			11.50	11.57	11.70	11.71	moderate	-0.12	-0.05	0.08	0.09
4CL			9.24	9.37	9.18	9.30	low	-0.03	0.10	-0.09	0.03
4CL			9.35	9.43	9.44	9.50	low	-0.08	0.00	0.01	0.07
C3H			13.94	13.72	13.26	14.00	high	0.21	-0.01	-0.47	0.27
C3H			10.67	10.03	9.85	9.75	low	0.60	-0.04	-0.23	-0.33
HCT			11.03	12.37	11.78	11.74	moderate	-0.70	0.64	0.05	0.01
HCT			11.45	12.42	12.32	12.22	moderate	-0.65	0.32	0.21	0.11
HCT			11.75	12.38	12.61	12.35	moderate	-0.53	0.11	0.34	0.08
CCoAOMT			13.05	14.48	14.01	13.44	high	-0.69	0.73	0.26	-0.31
CCoAOMT			12.76	14.25	13.89	13.26	high	-0.78	0.71	0.35	-0.28
CCoAOMT			12.55	14.27	14.07	13.33	high	-1.00	0.72	0.51	-0.23
CCoAOMT			12.49	13.28	12.99	12.88	high	-0.42	0.37	0.08	-0.03
CCoAOMT			11.99	12.97	13.18	13.29	high	-0.87	0.11	0.32	0.43
CCoAOMT			12.12	12.91	13.02	13.16	high	-0.68	0.11	0.22	0.36
CCoAOMT			12.52	13.32	13.46	13.44	high	-0.66	0.13	0.27	0.26
CCoAOMT			10.60	10.92	11.00	11.38	moderate	-0.38	-0.05	0.02	0.41
CCoAOMT			11.82	12.77	12.54	12.77	moderate	-0.66	0.29	0.07	0.30
CCoAOMT			10.12	10.92	10.73	11.06	moderate	-0.59	0.21	0.03	0.35
CCoAOMT			9.37	9.32	9.46	9.50	low	-0.04	-0.09	0.05	0.09
CCoAOMT			10.11	10.17	10.01	10.07	low	0.02	0.08	-0.08	-0.03
CCR			13.47	12.03	11.85	12.33	high	1.05	-0.39	-0.57	-0.09
CCR			14.52	13.04	13.02	13.45	high	1.02	-0.47	-0.49	-0.06
CCR			12.07	13.35	13.37	13.08	high	-0.90	0.38	0.40	0.12
CCR			11.29	10.08	10.17	10.32	moderate	0.83	-0.39	-0.30	-0.14
CCR			12.03	10.47	10.16	10.57	moderate	1.22	-0.34	-0.64	-0.24
CCR			11.40	12.01	11.99	12.06	moderate	-0.47	0.14	0.12	0.20
CCR			12.69	12.54	12.59	12.46	moderate	0.12	-0.03	0.02	-0.11
CCR			10.24	9.39	9.32	9.37	low	0.66	-0.19	-0.26	-0.21
CCR			9.17	10.55	9.86	9.90	low	-0.70	0.68	-0.01	0.03
CCR			9.81	10.03	10.05	10.09	low	-0.19	0.03	0.05	0.10
CCR			9.24	9.31	9.30	9.15	low	-0.01	0.06	0.05	-0.10
CCR			9.68	9.86	9.59	9.77	low	-0.05	0.14	-0.13	0.04
CCR			10.10	9.73	9.51	9.45	low	0.40	0.04	-0.19	-0.25
F5H			9.74	11.02	10.76	11.28	moderate	-0.96	0.32	0.06	0.58
F5H			9.65	10.51	10.49	10.80	low	-0.72	0.15	0.13	0.44
COMT			13.18	14.10	13.95	13.96	high	-0.62	0.31	0.15	0.16
COMT			12.31	13.33	13.14	13.24	high	-0.70	0.32	0.14	0.24
CAD			13.63	12.64	12.90	12.44	high	0.73	-0.26	0.00	-0.46
CAD			13.67	12.46	12.13	11.97	high	1.11	-0.10	-0.43	-0.59
CAD			12.76	13.86	13.39	13.22	high	-0.55	0.55	0.08	-0.09
CAD			12.62	12.14	12.15	12.62	moderate	0.24	-0.24	-0.23	0.24
CAD			9.49	11.20	10.30	10.33	moderate	-0.84	0.87	-0.03	0.00
CAD			10.35	10.38	10.42	10.47	low	-0.06	-0.02	0.02	0.06
CAD			10.48	9.77	9.73	9.74	low	0.55	-0.16	-0.20	-0.19
CAD			9.68	9.59	9.72	9.68	low	0.01	-0.08	0.06	0.01
CAD			9.52	9.65	9.84	9.59	low	-0.13	0.00	0.19	-0.06
CAD			10.05	10.10	9.88	10.47	low	-0.08	-0.03	-0.24	0.35

Supplementary Table 11. Log2 fold changes in expression levels of putative phenylpropanoid genes involved in lignification. Fold changes are color coded according to overexpression (red) and downregulation (blue) in *cad* mutant samples in four stages of development: V10 (ten visible leaf collars), S (silking), S+7d (seven days after silking) and S+14d (14d after silking).

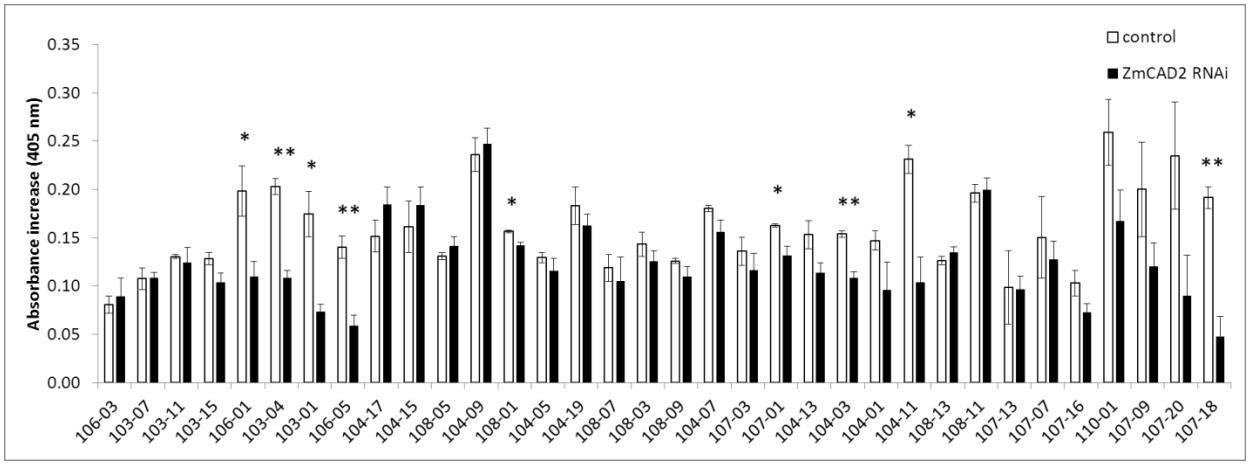
				log2 fold change (mutant vs. control)			
protein name	gene name	probe name	expression level	V10	S	S+7d	S+14d
PAL	Hidden		very high	-0.24	0.14	0.18	0.07
PAL			high	-0.08	0.03	0.15	0.22
PAL			high	0.08	0.09	0.05	0.08
PAL			high	-0.13	0.25	0.16	-0.03
PAL			high	-0.14	0.15	0.34	0.15
PAL			high	-0.06	0.39	0.04	-0.30
PAL			high	0.06	0.34	0.22	0.31
PAL			moderate	-0.18	0.34	0.05	0.00
PAL			low	0.09	0.49	0.01	0.20
PAL			low	-0.03	0.15	0.09	-0.02
PAL			low	0.04	0.79	0.08	-0.05
C4H			high	-0.04	0.55	0.97	0.23
C4H			moderate	0.02	0.38	0.57	0.31
C4H			low	-0.09	0.54	0.07	-0.06
4CL			moderate	0.01	-0.14	0.15	-0.12
4CL			moderate	-0.31	0.36	-0.68	0.07
4CL			moderate	-0.26	0.48	-0.75	0.12
4CL			moderate	0.02	-0.06	0.31	0.05
4CL			moderate	-0.09	-0.16	0.20	-0.04
4CL			moderate	-0.12	-0.02	0.28	0.15
4CL			low	-0.09	-0.11	0.18	-0.01
4CL			low	0.00	0.03	0.06	0.03
C3H			high	-0.14	0.39	0.60	-0.03
C3H			low	0.07	0.03	0.21	0.19
HCT			moderate	-0.27	-0.03	0.02	0.04
HCT			moderate	-0.13	0.32	-0.17	0.09
HCT			moderate	-0.07	-0.09	-0.08	0.15
CCoAOMT			high	0.02	-0.11	0.20	0.67
CCoAOMT			high	-0.09	-0.08	-0.13	0.48
CCoAOMT			high	-0.03	-0.26	-0.39	0.62
CCoAOMT			high	-0.01	0.08	-0.11	0.19
CCoAOMT			high	0.07	0.44	0.02	0.32
CCoAOMT			high	-0.26	0.23	0.11	0.15
CCoAOMT			high	-0.27	0.24	-0.08	0.36
CCoAOMT			moderate	0.05	0.29	0.65	-0.04
CCoAOMT			moderate	-0.19	0.52	0.66	0.66
CCoAOMT			moderate	-0.19	0.13	0.23	0.18
CCoAOMT			low	-0.08	0.19	0.03	-0.01
CCoAOMT			low	0.01	-0.31	0.19	-0.03
CCR			high	-0.12	0.12	0.50	-0.14
CCR			high	0.00	0.29	0.23	-0.20
CCR			high	0.20	0.00	0.10	0.44
CCR			moderate	0.04	0.18	0.06	-0.06
CCR			moderate	0.21	0.07	0.50	-0.18
CCR			moderate	0.08	-0.13	0.14	0.10
CCR			moderate	-0.08	-0.07	0.09	0.14
CCR			low	0.17	-0.14	0.05	-0.07
CCR			low	-0.03	-0.86	0.46	0.26
CCR			low	-0.10	-0.14	0.04	0.05
CCR			low	-0.01	-0.12	-0.02	0.08
CCR			low	-0.10	-0.45	0.20	-0.17
CCR			low	0.00	-0.04	-0.01	0.05
F5H			moderate	-0.03	-0.12	0.98	0.26
F5H			low	-0.07	0.04	0.11	0.17
COMT			high	-0.18	0.41	0.28	0.28
COMT			high	-0.10	0.30	0.19	0.04
CAD			high	0.05	0.02	-0.07	0.02
CAD			high	-0.09	-0.03	0.43	0.20
CAD			high	-2.28	-1.43	-1.04	-0.82
CAD			moderate	-0.02	0.20	0.62	0.16
CAD			moderate	-0.20	-0.98	-0.14	-0.29
CAD			low	0.02	0.18	0.22	-0.07
CAD			low	0.09	0.15	0.03	-0.07
CAD			low	-0.05	0.16	-0.10	0.17
CAD			low	-0.03	0.16	-0.18	0.18
CAD			low	0.02	0.37	1.03	0.84

Supplementary Table 12. Primer sequences used for cloning of three regions in the *ZmCAD2* cDNA sequence

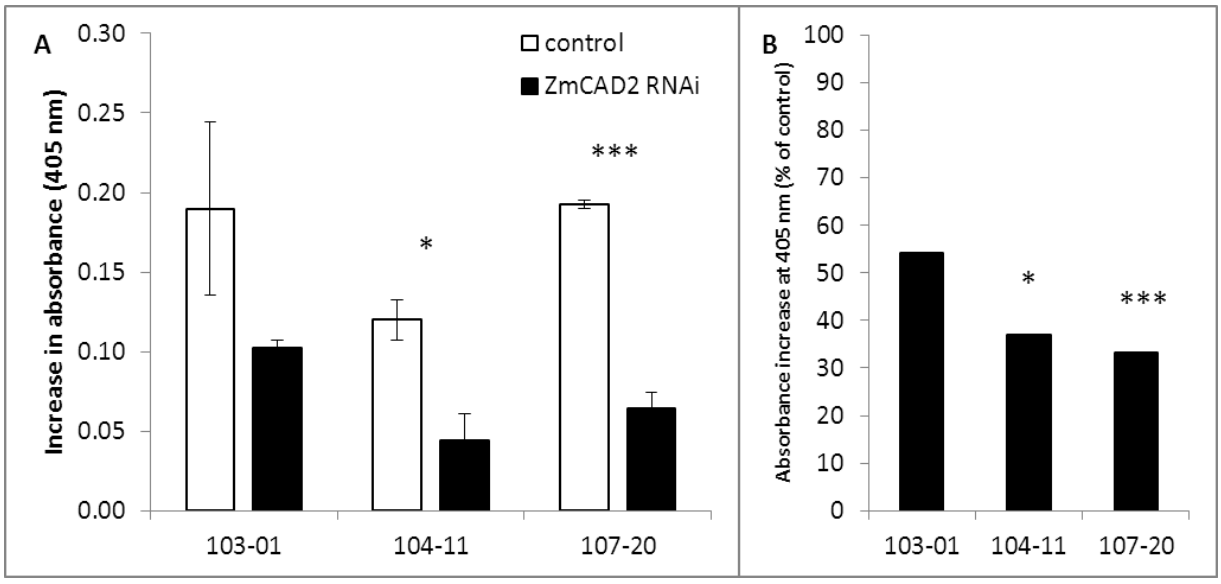
Construct	Primer FW	Primer RV	Product size (bp)
RNAi1	CCGATCCCGAATCGAATG	ATACTTTGAAGCCCCGAGGT	204
RNAi2	GGTGATCGTTGGGTGCTG	CGGGATCTTCACCACAAACT	169
RNAi3	AGTCGAACCACGTCTGTGC	CAAATCCAGATTGCTGGTGA	242

Supplementary Table 13. Primer sequences used for qRT-PCR analysis of transgene and *ZmCAD2* endogene expression levels.

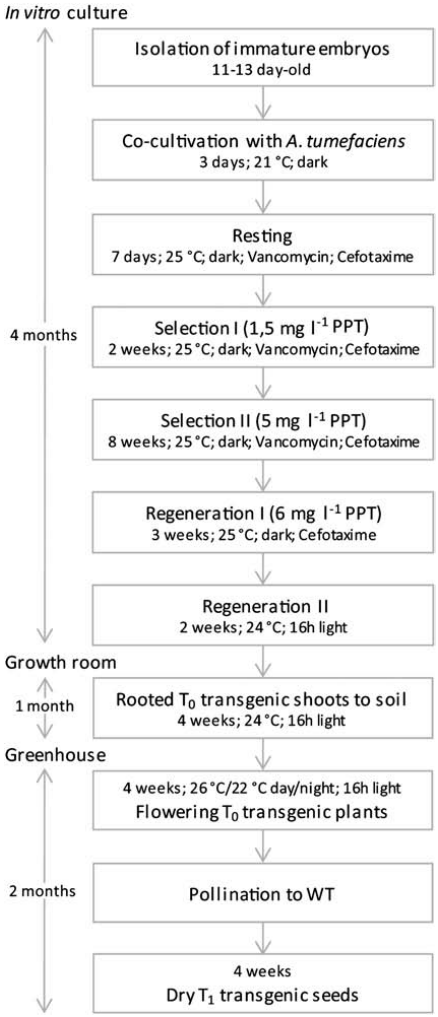
Target	Primer FW	Primer RV	Product size (bp)
RNAi1	ACATCCACCAGGCCAAGA	CTTCGTCTTACACATCACTTGTCA	
RNAi2	GTCGTCGACCAGAAGTTTGT	CTTCGTCTTACACATCACTTGTCA	
RNAi3	GGTGTCAGTCTCACCAGCAA	CTTCGTCTTACACATCACTTGTCA	
ZmCAD2	CGACTCGCTGGACTACATCA	TTCAGTTCTGCGTCGACAAG	223
ZmEF1a	AGTCCGTTGAGATGCACCATG	CACATACCCACGCTTCAGATCC	107
Zm18S	ACCTTACCAGCCCTTGACATATG	GACTTGACCAAACATCTCACGAC	118



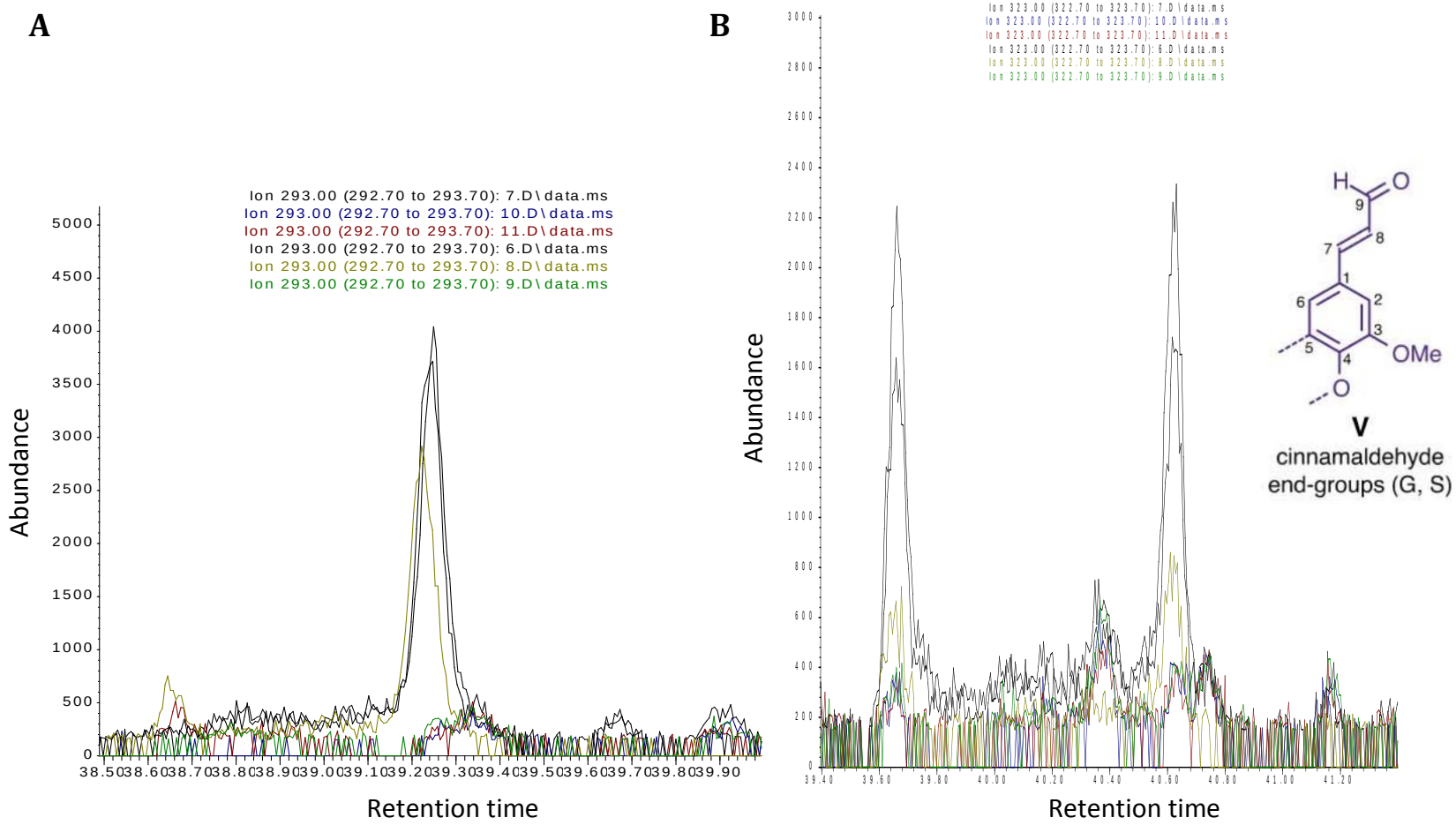
Supplementary figure 7. Coniferyl alcohol conversion measured as the increase in absorbance at 405 nm in control and *ZmCAD2* RNAi plants in the segregating population for independent transgenic each line. *: p<0.05; **: p<0.01.



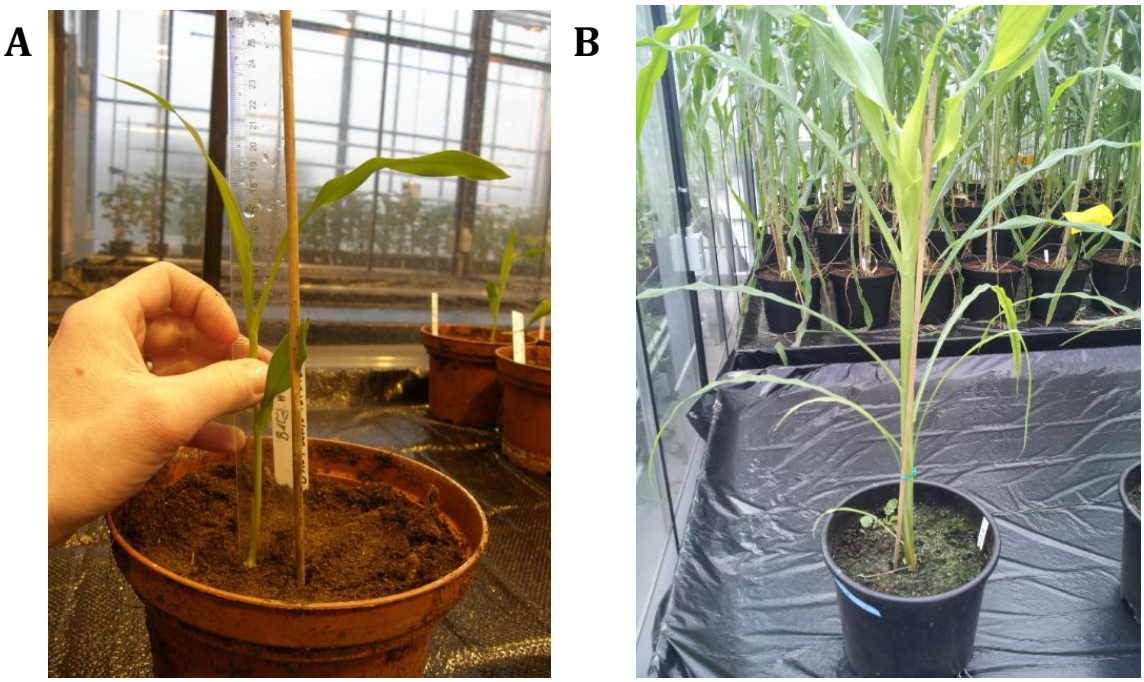
Supplementary figure 8. Coniferyl alcohol conversion measured as the increase in absorbance at 405 nm (A) and expressed as percentage of control (B) in control and *ZmCAD2* RNAi plants in the segregating population for three independent transgenic lines (103-01, 104-11 and 107-20). *: p<0.05; ***: p<0.001.



Supplementary figure 9. Scheme for Agrobacterium-mediated transformation of the maize B104 inbred line at VIB (from Coussens et al., 2012)



Supplementary figure 10. Detection of guaiacyl with m/z 293 (A) and syringyl with m/z 323 (B) aldehydes in GCMA detection of the thioacidolysis analysis of ground stem material of *cad* mutant (samples 6, 7 and 8) and control (samples 9,10 and 11). The aldehydes are below the detection limit (A) or very low in abundance (B) in control samples but accumulate in the *cad* mutant samples. Proposed structure of detected compounds is modified from Kim et al., 2012.



Supplementary figure 11. Representative stages of maize pants used for expression analysis in the fourth leaf when it appears from the pseudostem (A) and at V9, nine visible leaf collars for internodes and mature leaves (B).

Chapter 6: Targeting the *Cinnamyl Alcohol Dehydrogenase 2* gene in maize for improved saccharification efficiency

Chapter 7: Gibberellic acid overproduction as strategy for improved maize bioenergy feedstock

Chapter 7: Gibberellic acid overproduction as strategy for improved maize bioenergy feedstock

Chapter 7: Gibberellic acid overproduction as strategy for improved maize bioenergy feedstock[2]

Seed stocks of maize plants overexpressing *AtGA20ox1* were obtained from Systems Biology of Yield group at VIB-PSB. My contribution to this work involved the experimental planning as well as harvest of the plant material, cell wall analysis, microscopy on internode sections, gene expression analysis and plant height, internode width and organ-specific weight measurements. Both for harvesting the plant material and lab analysis I received great assistance from technicians at ILVO, Growth and Development and the Bioenergy and Systems Biology of Yield groups at VIB-PSB. Furthermore, I wrote the outcome of this research topic in the form of paper in preparation for submission to the Plant Biotechnology Journal.

1. Summary

The overproduction of gibberellic acid (GA) results into bigger plants and organs in several plant species, a trait that could be used for crop improvement. However, when using this trait for the improvement of feedstock for bioethanol production, the conversion efficiency of the cellulose in the cell wall to fermentable sugars is of great importance. In that respect, biomass accumulation as well as cell wall characteristics such as cellulose and lignin content and saccharification efficiency in GA overproducing maize plants were investigated.

GA overproducing maize plants accumulated more stem biomass but less seed yield than control plants. The stems are longer but also more slender. Biochemical analyses demonstrated the accumulation of more cell wall residue but also more lignin and cellulose. As a result, stems of GA overproducing plants contain more lignin and cellulose per dry weight. The altered cell wall composition results in a reduced saccharification efficiency of dry whole-stem material when pretreated with acid, but not with alkali. Cell wall analysis as well as expression analysis of lignin biosynthetic genes in developing stems revealed that cellulose and lignin are deposited earlier in development in the GA overproducing plants as compared to control plants. In addition, this early lignification resulted in altered monomeric composition with lower syringyl/guaiacyl ratio in developing stems of GA overproducing plants.

This study shows that strategies to improve stem biomass quantity can influence cell wall properties. The fact that biomass quantity and quality can be interconnected is important for developing strategies to improve lignocellulosic feedstock for bioethanol production.

[2] Manuscript written according guidelines of Plant Biotechnology Journal

2. Introduction

2.1 Maize as bioenergy feedstock

Bioenergy grasses are defined as members of the grass family (*Poaceae*) that employ C4 metabolism, capable of producing high yield in the form of lignocellulose, fermentable juice, or fermentable grain (Vermerris et al. 2007; Feltus and Vandenbrink 2012). Preferably, bioenergy grasses are perennial and require no or only low amounts of fertilizer. In hot arid conditions C4 plants convert energy more efficiently into biomass than C3 plants and have up to 60% higher water and nutrient use efficiency (Heaton et al. 2008a). As a consequence, C4 plants have the potential to produce exceptionally high grain yield as well as stem and leaf biomass yield and are promising as biomass-producing plants for cellulosic biofuels (Jakob et al. 2009). However, in more moderate climates, also C3 forage grasses such as perennial ryegrass, can form a potential feedstock for the production of bioenergy (Wang and Brummer 2012; Tonini et al. 2012). Major bioenergy grasses include *Saccharum* spp. (sugarcane), *Sorghum bicolor* (sorghum), *Miscanthus* spp. (Miscanthus), and *Panicum virgatum* (switchgrass) (Feltus and Vandenbrink 2012).

Also *Zea mays* (maize), is often regarded as a bioenergy grass, despite its annual growth behavior and the need for significant water and fertilizer input. Maize can be regarded as an excellent model for bioenergy research for several reasons. First, maize is the second most cultivated crop globally, and has the biggest production worldwide (FAO Statistics Division 2013a) and will likely remain a dominant crop in the foreseeable future. Additionally, maize is grown as an important forage crop, especially in Europe (Argillier et al. 1998; Bosch et al. 1999; Boon et al. 2012). When maize is grown as a dual-purpose crop, the corn stover, which are the dry residues remaining after the grain harvest, can provide an abundant source of biomass for cellulosic ethanol production (Vermerris et al. 2007), even when sufficient amount of biomass is left on the field to prevent a reduction in soil fertility (van der Weijde et al. 2013). Second, maize is, together with sugarcane, the predominant bioenergy crop, based on the volume of biofuels produced (Vermerris 2011). Yet, at present these so-called first generation biofuels are produced from the easily accessible sugars in the starch and fermentable juice, and are considered unsustainable since they compete with the food chain and provide little or no greenhouse gas reduction once all impacts of cultivation and processing are taken into account (European Academies Science Advisary Council 2012). Nevertheless, their well-established production chains can supply the next generation of biorefineries with large amounts of agricultural residues (Waclawovsky et al. 2010; van der Weijde et al. 2013). Third, maize has a long history as a genetic model, with a remarkable genetic diversity and many useful genetic and genomic resources (Carpita and McCann 2008; Barrière et al. 2009; Schnable et al. 2009b; Vermerris 2011). Forth, despite the historic focus on grain yield, there exists an extensive knowledge base of the cell wall composition and architecture and the underlying genetic mechanisms in maize. This knowledge can be used to advance breeding for (vegetative) biomass yield and quality for the purpose of bioenergy feedstock as well as forage (Carpita and McCann 2008). Moreover, fundamental genetic knowledge obtained from maize can be applied to related grass species, such as Miscanthus, Sorghum and switchgrass (Carpita and McCann 2008; Vermerris 2011).

2.2 Quantity versus quality

The improvement of lignocellulosic crops as feedstock for the production of bioethanol currently focuses either on enhanced biomass yield or on improved biomass quality. However, focusing on one of these traits holds the risk of improving one aspect and unwillingly affecting negatively the other, with no net benefit as result. For example, in the past decades, breeding efforts in Europe achieved a great improvement in maize whole plant yield with a ~4.4 t/ha increase overall. However, the introduction of hybrids with favorable alleles for grain production, pest resistance and stalk breakage resistance has led to unfavorable effects on plant digestibility. Since the 1950s, cell wall digestibility has declined steadily, leading to a substantially reduced feeding value of elite maize hybrids (Barrière et al. 2006). This is until cell wall digestibility was included as a quality parameter for new maize breeding varieties (1998 in France) after which efforts were undertaken for the breeding of more digestible forage maize (Pichon et al. 2006; Riboulet et al. 2008; Brenner et al. 2010; Jung 2011; Boon et al. 2012). Another example illustrating the need for adapted breeding strategies is the proposed use of *Miscanthus* x *giganteus* for the production of cellulosic ethanol. *M.* x *giganteus*, a sterile triploid hybrid of a cross between *M. sinensis* and *M. sacchariflorus* (Hodkinson et al. 2002) is, in contrast to maize and switchgrass, relatively well adapted to more temperate climates (Farage et al. 2006; Wang et al. 2008) and has been grown successfully in Europe (Clifton-Brown et al. 2001). Due to its impressive biomass yield (15-20 t/ha; Christian et al. 2008; Van Hulle et al. 2010), and low fertilizer requirements *M.* x *giganteus* holds great promise for the production as bioenergy crop (Heaton et al. 2008b; Heaton 2010; Slavov et al. 2013). However, its mineral content and cell wall composition are not optimal for conversion to ethanol (Jakob et al. 2009; Heaton 2010). The relatively high lignin content in *M.* x *giganteus* hinders the enzymatic conversion of the biomass into fermentable sugars (de Vrije 2002; Allison et al. 2009; Van Hulle et al. 2010; Brosse et al. 2012). The high concentrations of alkali minerals results in the formation of ashes with low-melting points and can lead to blockage and slagging of combustion equipment that is used for thermal conversion of the residual biomass after fermentation (Jenkins et al. 1998; Allison et al. 2009; Sannigrahi and Ragauskas 2011). Breeding Miscanthus genotypes with lower lignin and ash content, resulting in more efficient biomass conversion, will further improve its suitability as a biofuel crop (Jakob et al. 2009). Lignin content is one of the major determinants of saccharification efficiency, but besides the production of lignin monomers, enzymes in the lignin biosynthetic pathway are involved in the biosynthesis of salicylates, coumarins, hydroxycinnamic amides, pigments, UV light protectants, antioxidants, flavonoids, isoflavonoids, anthocyannins and tannins, among others (Zabala et al. 2006; Vogt 2010; Vanholme et al. 2012a). Therefore, the genetic perturbation of these genes will affect cell wall lignification but most likely also other biochemical routes (Chen 2011). Some adverse effects of lignin perturbation on plant fitness were summarized in Pedersen et al. (2005). Depending on the genotype and the environmental conditions, potential effects on agricultural fitness include reduced grain yield, reduced dry matter yield, reduced early season vigor and increased lodging in maize and Sorghum brown midrib mutants.

Considering the possible pleiotropic effects of high biomass yield on saccharification efficiency and vice versa, breeding strategies for the improvement of crops for the production of cellulosic ethanol should integrate both traits for future breeding programs.

2.3 GA levels affect organ size

Plant biomass yield is determined by a number of factors, such as the efficiencies of the capture of solar energy and the conversion of the captured solar energy into vegetative tissues that constitute the bulk of plant biomass (Zhu et al. 2008). Biomass used for cellulosic biofuel production is primarily from vegetative tissues, such as stems and leaves. In principle, a prolonged or increased vegetative meristem activity could lead to the production of more vegetative tissues and thus more biomass (Demura and Ye 2010). Vegetative meristem activity, and thus the size of organs such as the maize leaf or the Arabidopsis root, is under hormonal control (Wolters and Jürgens 2009; Ubeda-Tomas and Bennett 2010; Dudits et al. 2011; Nelissen et al. 2012). In the Arabidopsis root, meristem activity is controlled by the antagonistic action of auxin and cytokinin, where auxin promotes cell division while cytokinin promotes cell differentiation (Dello Ioio et al. 2008). In the maize leaf, the accumulation of auxin and cytokinin in the division zone suggest a similar mechanism. In addition, a narrow peak of gibberellic acid (GA) was detected at the transition zone between the division and the elongating zone, showing that GA drives cell division at the distal end of the division zone in the maize leaf (Nelissen et al. 2012). In tobacco stems, internode elongation seems to be dependent on bioactive GA's that are transported from the leaves by the phloem (Dayan et al. 2012). The mode of action of GA in stems is more related to cell expansion than cell division, suggesting that GA can play a different role in stems as compared to leaves (Nelissen et al. 2012).

The effect of GA on stem height and whole plant growth rates under different environmental conditions has been extensively studied in several plant species (Juska 1958; Norcia et al. 1964; Paleg et al. 1965; Cleland and Briggs 1969; Baker 1987; Lambers et al. 1995; Lester et al. 1997; Peng et al. 1999; Bultynck 2002; Srivastava 2002; Sasaki et al. 2002; Zawaski and Busov 2014) and external application of GA is being applied in agriculture (Sponsel 2010). It has even been suggested that variation in endogenous GA levels is responsible for the variation in whole plant growth rates in hybrid maize (Rood et al. 1990). Several studies have demonstrated that enhanced endogenous GA levels can result into taller plants and larger organs in Arabidopsis (Huang et al. 1998; Coles et al. 1999), tobacco (Biemelt et al. 2004), rice (Oikawa et al. 2004), potato (Carrera et al. 2000), tomato (García-Hurtado et al. 2012), citrus (Fagoaga et al. 2007) and poplar (review by Dubouzet et al. 2013). In maize, the overexpression of *AtGA20ox* under the control of a constitutive promoter resulted in plants with longer leaves and stems (Nelissen et al., 2012). By overproducing GA, the division zone of the leaf becomes larger and thus produces more cells making the leaves longer.

On the other hand, the reduction in endogenous GA levels has formed the basis for the green revolution, when mutants in GA biosynthesis were used for breeding wheat and rice varieties with shorter stems (Peng et al. 1999; Sasaki et al. 2002; Evenson and Gollin 2003). The shorter and stronger stem conferred a higher resistance to lodging, allowing the plants to bear the weight of improved grain yields that were accomplished by breeding efforts and the increased use of fertilizers and pesticides (Peng et al. 1999; Hedden 2010). These GA biosynthesis mutants had shorter stems with improved harvest index (dry weight of the seed divided by the total dry weight). This was explained by a shift in the way grasses partition structural carbohydrates in the plant, since more carbohydrates were allocated to the seeds (Slewinski 2012).

2.4 Enhanced GA levels can result into higher lignin accumulation

Since long, the involvement of GA in fiber differentiation and lignin formation in phloem fibers and xylem has been observed (Aloni 1979; Roberts et al. 1988; Aloni et al. 1990). Increased levels of GA can induce increases of xylem lignification in tobacco, and altered S/G ratios in the lignin of poplar (Israelsson et al. 2003; Biemelt et al. 2004; Mauriat and Moritz 2009). Furthermore, GA can affect the length of xylem fibers, which are highly lignified (Ridoutt et al. 1996; Israelsson et al. 2005; Mauriat and Moritz 2009). If this is also the case in grasses, alterations of GA levels can have consequences for the saccharification efficiency of the biomass.

A link between lignin biosynthesis and GA biosynthesis exists through KNOX transcription factors (Mele et al. 2003; Hay and Tsiantis 2010). Genetic analyses in Arabidopsis showed that KNOX expression maintains high cytokinin (CK) and low GA levels in the shoot apical meristem to promote cell division over cell differentiation by activating CK biosynthesis, activating GA catabolism, repressing GA biosynthesis and repressing lignin accumulation. KNOX transcription factors directly regulate the gibberellin catabolism gene *ga2ox1* (Bolduc and Hake 2009) and can directly bind promoter regions of the lignin biosynthetic genes *COMT1*, *CCoAOMT1* and the peroxidase encoding *AtPrx2* (Mele et al. 2003; Hay and Tsiantis 2010). GA overproduction could disturb the CK/GA balance, thereby initiating a feedback loop that alters lignin biosynthetic gene expression. However, a GA spraying experiment on wildtype tobacco hypocotyls challenges this hypothesis of a direct effect of GA levels on lignin biosynthetic gene expression, as lignin formation was induced without detectable increase in lignin gene expression (Biemelt et al. 2004). This suggests that upregulation of lignin biosynthetic genes occurs only on a long-term basis and has thus to be considered as an indirect effect of enhancing endogenous GA levels (Biemelt et al. 2004).

3. Objectives

Increasing the amount of bioactive GA has been proposed as a good strategy for biomass enhancement. On the other hand, this strategy of yield improvement has potentially pleiotropic effects regarding saccharification efficiency caused by increased lignification. This study aims to evaluate the potential use of *AtGA20ox* overexpression in maize as strategy for the improvement as feedstock for bioethanol production. Previously, in maize the effect of GA overproduction was studied on leaf growth only. Here, plant morphology, organ-specific yield contribution and whole-plant biomass yield of GA overexpressing maize plants were determined. Given the potential impact of elevated GA levels on lignin biosynthesis, biomass quality traits relevant to saccharification efficiency were examined. Furthermore, lignin and cellulose deposition as well as lignin composition and lignin biosynthetic gene expression were monitored during stem development to get insights into the mechanisms that can mediate the interaction between GA overexpression and cell wall properties.

4. Results

4.1 GA overproduction leads to altered plant morphology and biomass accumulation.

Maize plants that overproduce GA were obtained by overexpression of the GA biosynthesis gene *GA20ox1* (Nelissen et al. 2012). These plants grow on average 37% or 1 m taller than control plants at maturity (silking stage) (Figure 57A). Already in the V10 stage, when 10 leaf collars are visible, the transgenic plants are 45% taller than the non-transgenic control plants. This effect of elongated stems in *GA20ox* overexpressing plants has also been observed in Arabidopsis (Huang et al. 1998; Coles et al. 1999), tobacco (Biemelt et al. 2004), rice (Oikawa et al. 2004), potato (Carrera et al. 2000), tomato (García-Hurtado et al. 2012), citrus (Fagoaga et al. 2007) and poplar (review by Dubouzet et al. 2013). In contrast, the transgenic maize plants have a reduced stem diameter. The diameter of the ninth internode is smaller in the transgenics throughout development, with a reduction of 21% or 4 mm at silking stage (Figure 57). Like the increase in stem length, this reduction in stem diameter was observed already from the V10 stage on (-20% or -3.3 mm).

The fact that the increased plant height is compensated for by a reduction in stem diameter immediately raises the question whether biomass yield is altered in the transgenic plants. The total and organ-specific biomass yield of the transgenic and non-transgenic control plants at fully mature dry stage is shown in Figure 58. The total aboveground plant biomass, expressed as dry weight, was reduced by 12% in the GA overproducing plants (Figure 58). The leaf as well as grain fractions were reduced by 17% and 55% respectively. In contrast, the weight of stems and cobs without seeds was increased by 32% and 37% respectively. Thus, whereas the grain contributes most to total aboveground plant biomass in control plants (37% of total biomass, cf. 19% in GA overproducing plants), stem weight is the major contributor to total biomass in GA overproducing plants (42% cf. 28% in control plants). The decrease in grain yield is caused by a lower number of seeds as the weight per 100 seeds was not significantly altered (Table 19 and Supplementary figure 17). The decrease in total aboveground biomass and grain yield has a negative influence on the agricultural use of these plants. Nevertheless, in terms of second generation bioethanol production, the increased stem biomass, the major source of cellulose, is a desirable feature.

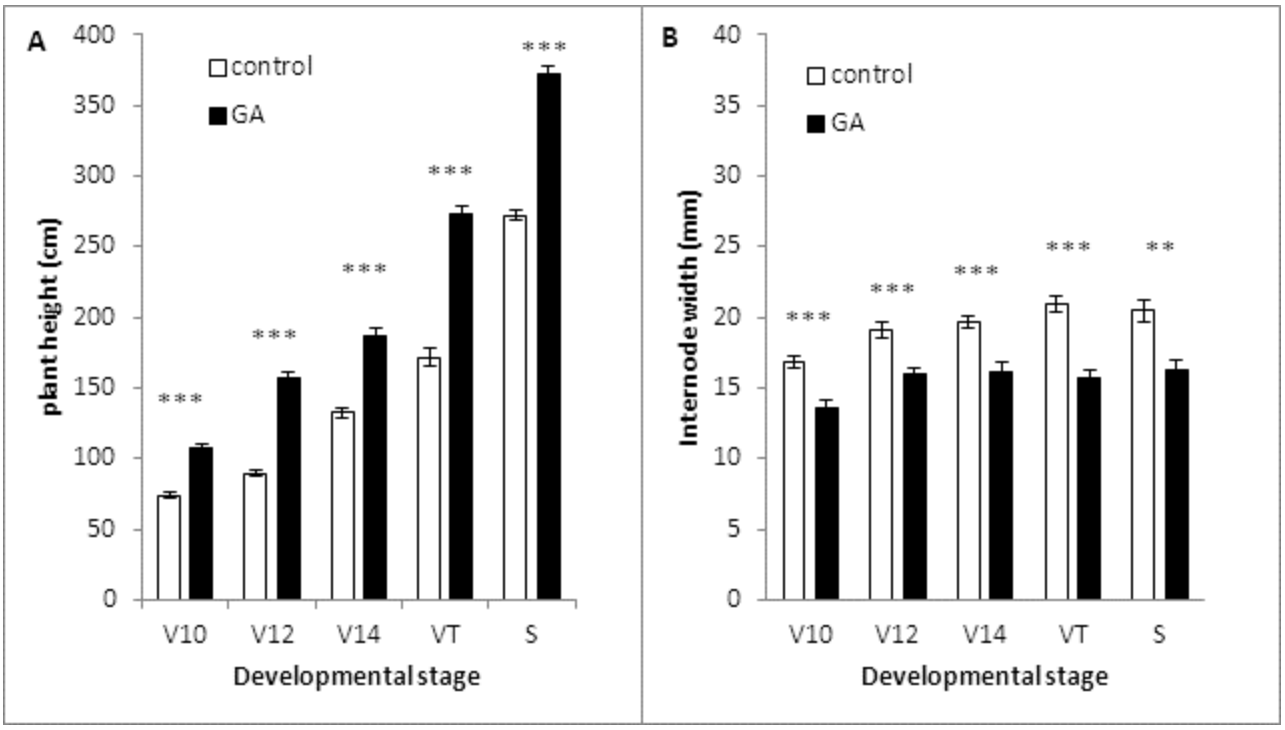


Figure 57. Plant height (A) and internode width (B) of GA overproducing (GA) and control plants (control) over development. V10, V12, V14: 10, 12 and 14 leaf collars visible respectively; VT: male flowering (tasseling); S: female flowering (silking); Plant height is measured as height of the youngest leaf in V10, V12 and V14 stages and as height of the tassel in VT and S stages. Internode width is the width of the ninth internode, counted from the bottom; **: p<0.01, ***: p<0.001. Error bars represent standard errors for 9 biological repeats in V10, V12, V14 and VT stages and 7 biological repeats for S stage.

Table 19. Seed weight and count in control and GA overproducing maize plants. Values for weight per 100 seeds and number of seeds per cob are means over nine biological replicates and p-values result from a student t-test.

	weight/100 seeds (g)			# seeds per cob		
	mean	stderr	p	mean	stderr	p
control	26.2	0.8	0.209	355	20	0.00009
GA	24.8	0.8		200	22	

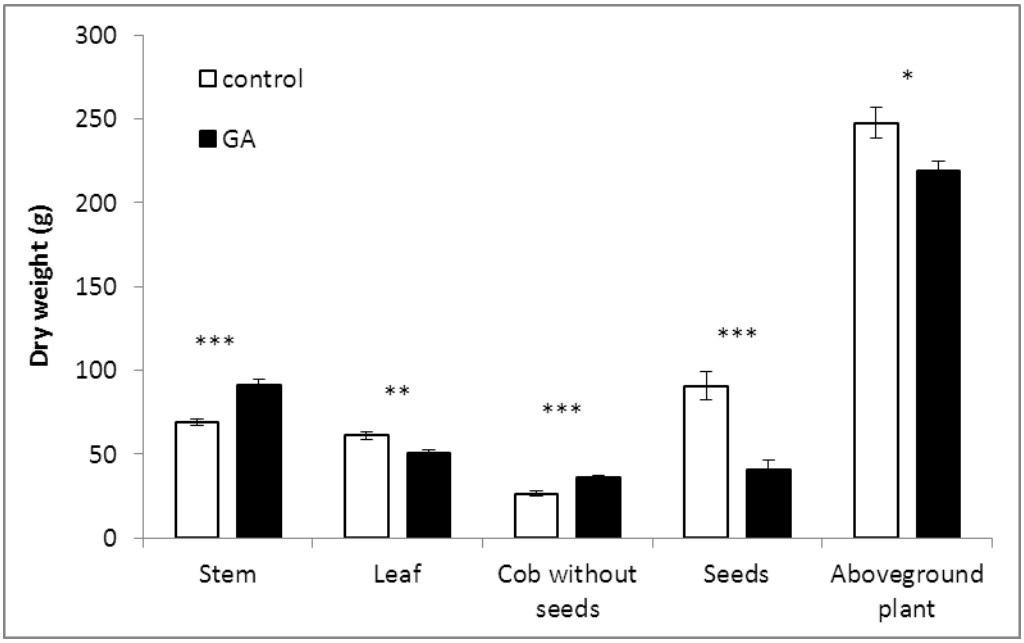


Figure 58. Dry weight measurements of mature plant parts of GA overproducing (GA) and control plants. *: p<0.05, **: p<0.01, ***: p<0.001. Error bars represent standard errors for 13 biological repeats for control plants and 15 biological repeats for GA overproducing plants.

4.2 GA overproduction affects saccharification efficiency

In a next step we determined whether the combined increase of stem height and reduced stem diameter observed in GA overexpressing plants was associated with altered cell wall characteristics. To this end, the saccharification efficiency and cell wall components cellulose and lignin were determined in fully mature, dry stems of GA overproducing and control plants.

The saccharification efficiency of stems from GA overproducing and control plants was tested using two different pretreatments: acid (1M HCl) and alkali (1M NaOH). A first observation is that biomass of GA overproducing stems is more recalcitrant to the tested pretreatments than control stem biomass (Figure 59). Using an acid or alkali pretreatment step, an 8% and 6% increase in residual biomass respectively was observed for the stems of GA overproducing plants. Since alkali pretreatment removes the lignin in the cell wall, the cellulose is highly exposed to enzymatic hydrolysis. The acid pretreatment removes part of the hemicellulose but leaves the lignin in the cell wall intact. The highest enzymatic glucose release is generally obtained after alkali pretreatment (Pedersen et al. 2011). As expected, also here the alkali pretreatment induced the highest glucose release upon enzymatic hydrolysis (Figure 60). After 48 hours of hydrolysis, on average 84% of the alkali pretreated biomass was converted to glucose compared to 23% after acid pretreatment. Using the alkali pretreatment, no significant differences in saccharification efficiency were detected between the GA overproducing and control stem biomass. In contrast, a significant decrease (-23%) in glucose yield after 48h of incubation was observed using the acid pretreatment (Figure 60), suggesting altered cell wall composition such as increased lignification in stems of GA overexpressing plants.

The observed decrease in glucose yield after acid pretreatment might be just caused by the higher amount of residual biomass after pretreatment (Figure 59) in GA overproducing compared to control plants. To eliminate this possibility, the results for glucose yield were also expressed per DW, the initial stem biomass before pretreatment. The results were however similar: a significant decrease (-17%) in glucose yield after 48h of incubation using the acid pretreatment and no differences using the alkali pretreatment (Supplementary figure 12). This shows that the decrease in glucose yield is not due to the higher amount of residual biomass after the acid pretreatment but likely reflects compositional changes in the cell wall in GA overproducing plants.

The effect of the pretreatment and the saccharification efficiency is very dependent on the composition of the biomass, with a major negative impact of the presence of lignin (Vogel and Jung 2001; Boudet et al. 2003; Hendriks and Zeeman 2009; Van Acker et al. 2013; Zeng et al. 2014). Since here, differences in pretreatment and in saccharification efficiency are observed, it is reasonable to assume that compositional changes such as in the lignin content in the stem of GA overproducing maize plants are, at least in part, responsible. In that respect, the cell wall fraction, cellulose content, lignin content, lignin composition and saccharification efficiency were determined on ground stem material of transgenic and control plants. The cell wall fraction or cell wall residue (CWR) per unit dry weight (DW) was significantly higher in GA overproducing plants (Figure 61). Also the cellulose content and lignin content were significantly higher in GA overproducing plants (Figure 61A). Lignin content per CWR was 6% higher and lignin per DW 19% higher in transgenic plants. The fact that the increase in lignin content is more pronounced when expressed per DW is due to the higher CWR/DW ratio in the GA overproducing plants (Figure 61). Stem sections of the ear internode of GA overproducing plants at S+14d stage show a more intense staining with phloroglucinol (Figure 62), indicative for the presence of more lignin. The more intense staining with phloroglucinol seems to be located in the sclerenchyma surrounding xylem and phloem in the vascular bundle and in the collenchyma (Figure 62). Biemelt et al. (2004) showed that GA20ox1 overexpressing tobacco plants apparently formed more xylem cells. However, stem cross sections of maize plants overexpressing GA20ox1, do not indicate the presence of more xylem cells. Consistent with literature (Grabber et al. 2004b), the relative abundance of the three general lignin subunits, *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S), showed that mainly G (~35%) and S (~60%) units were present, with only low levels of H (<5%), in the maize internode. The monomeric composition of lignin in GA overproducing plants showed a significant increase of the minor lignin component H, but not of G or S units (Figure 61B).

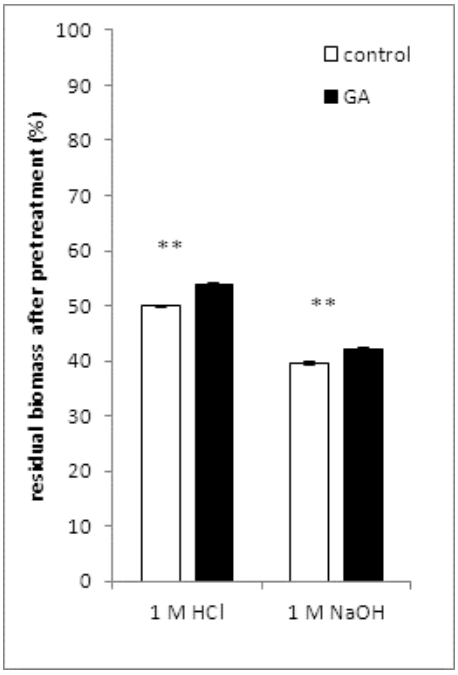


Figure 59. Residual biomass after pretreatment with either acid or alkali of GA overproducing (GA) and control plants; Error bars represent standard errors of 3 biological repeats. **: p<0.01

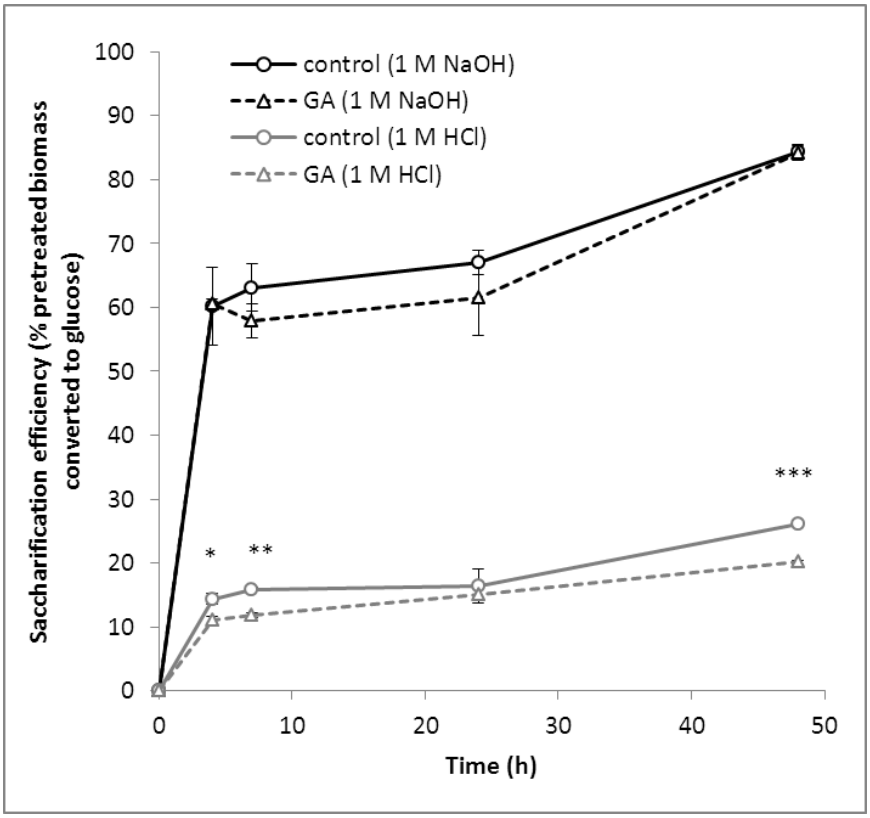


Figure 60. Saccharification efficiency of ground stem material of non-transgenic and transgenic plants with 1 M HCl and 1 M NaOH pretreatment, expressed as glucose yield per pretreated biomass weight. Error bars represent standard errors of three biological replicates. *: p<0.05, **: p<0.01, ***:p<0.001.

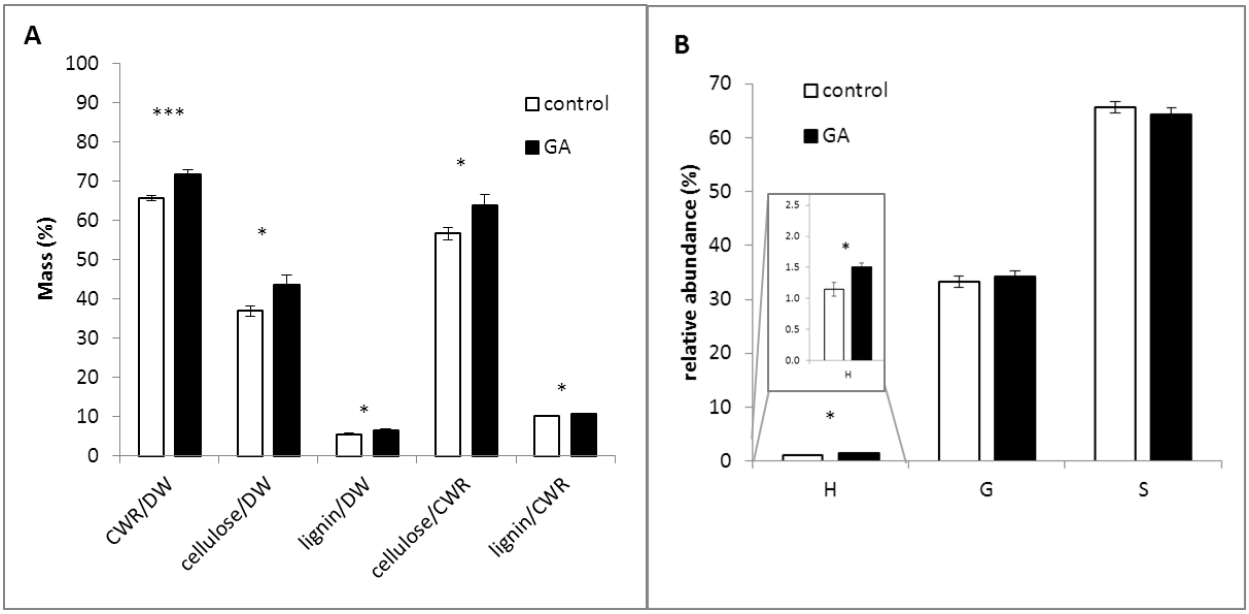


Figure 61. Cellulose levels, lignin levels, cell wall fraction (A) and lignin composition (B) in mature stems of GA overproducing (GA) and control plants. Error bars represent standard errors of 10 biological repeats. *: p<0.05, ***: p<0.001

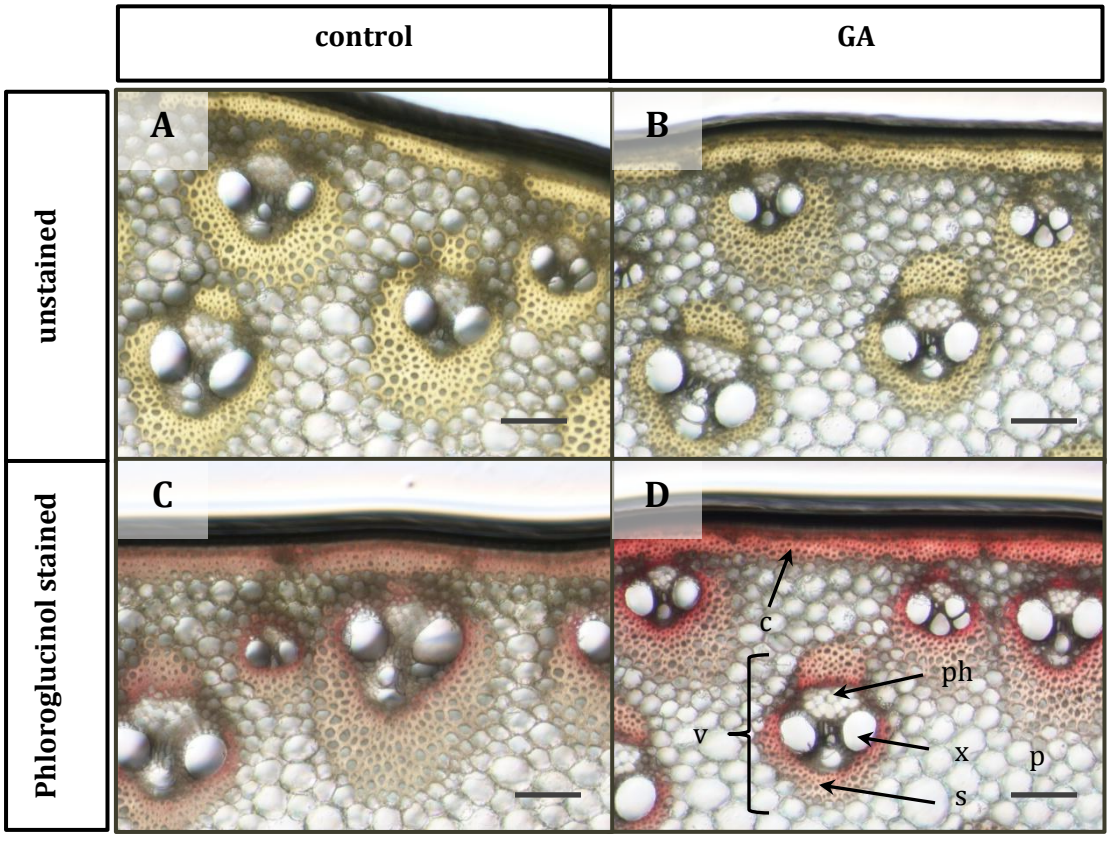


Figure 62. Stem sections of control and GA overproducing plants (GA) at S+14d (14 days after silking) stage, unstained (A and B respectively) and stained with phloroglucinol (C and D respectively). v = vascular bundle, x = xylem vessel, ph = phloem, s = sclerenchyma fibres, p = parenchyma, c = collenchyma. Scale bar is 100 µm.

4.3 Cell wall biosynthesis is altered in stems of GA overproducing plants

Increased lignification in stems of poplar and tobacco as a result of GA spraying or enhanced GA biosynthesis have been described before (Israelsson et al. 2003; Biemelt et al. 2004; Mauriat and Moritz 2009). Nevertheless, the characterization of cell wall composition and lignin biosynthesis of GA overproducing plants in different stages of development were not performed to date. Here, the effect of enhanced GA levels on the cell wall composition was studied in a developmental series of the maize stem from early vegetative until fourteen days after silking.

4.3.1 The growing internode of GA enhanced plants contains a higher cell wall fraction

In the developing internode, the onset of secondary cell wall formation is marked by the increase of hemicellulose and cellulose deposition and is followed by the formation of lignin (Morrison and Kessler 1994; Matos et al. 2013). Cellulose, hemicellulose and lignin are deposited inside the primary cell wall of cells that have ceased to elongate (Alberts et al. 2002). In this way, cell walls in a developing internode gradually grow thicker. This trend of increasing CWR per unit DW (CWR/DW) is clearly visible in the ninth internode of maize control plants in V10, V12, V14 (10, 12 and 14 visible leaf collars respectively) and S (silking) stage (Figure 63). The CWR/DW ratio drops between S and S+14d (14d after silking) stage. This accumulation of dry matter that is not part of CWR can be related to sucrose accumulation in stems and relocation of carbon from leaves to the cob situated above the investigated internode (Setter and Meller 1984; Dwyer et al. 1995; Jung and Casler 2006; de Souza et al. 2013).

In GA overproducing plants, the CWR/DW ratio also increases with stem maturation and is in all investigated stages, except V14, significantly higher than in stems of control plants (Figure 63). Perhaps in this way, the possible loss in strength of the more slender stem is compensated by the presence of more and/or thicker cell walls.

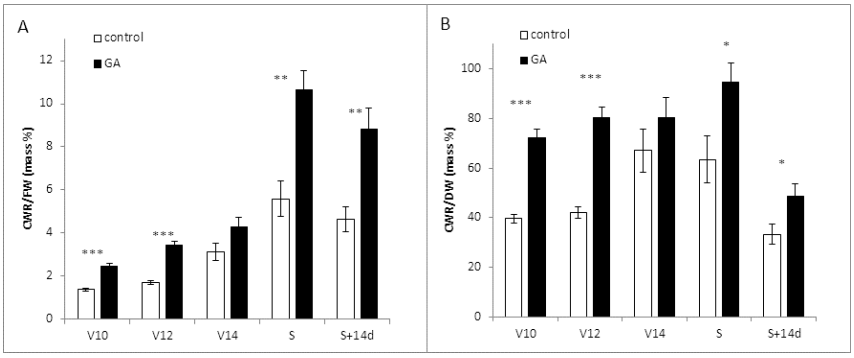


Figure 63. Cell wall accumulation in the developing ninth internode of control and GA overproducing plants (GA); cell wall residue (CWR) per fresh weight (FW) (A) and per dry weight (DW) (B); V10, V12, V14 (10, 12 and 14 leaf collars visible respectively), S (silking) and S+14d (14d after silking) indicate the different developmental stages; Error bars represent standard errors of 9, 9, 8, 7 and 6 biological replicates in stages V10, V12, V14, S and S+14d respectively; *: p<0.05, **: p<0.01, ***:p<0.001.

4.3.2 Cellulose and lignin accumulate earlier in development in internodes of GA overproducing plants

The amount of cellulose per fresh and dry weight in internodes of both control and GA overproducing plants increase as the internode matures, but drops again after silking (Figure 64A and B), consistent with the cell wall measurements (Figure 63). Cellulose deposition, expressed as percentage of the cell wall, increases substantially between V10 and V12 stage (10 and 12 visible leaf collars, respectively) in control plants (Figure 64C). Thereafter, the cellulose content decreases slightly and remains constant at around 55% of CWR from the V14 stage (14 visible leaf collars) on. Thus, the major accumulation of cellulose in the cell wall occurs between V10 and V12 stage. The cellulose content, expressed as a fraction of the FW or the DW biomass is significantly higher in GA overproducing plants for all developmental stages except S+14d (Figure 64A and B). At V10, the cellulose content is even five times higher than in control plants. Strikingly, the cellulose fraction constitutes already 52% of the cell wall at V10 in transgenic plants, close to the average cellulose percentage of more mature internodes (Figure 8C).

Cellulose deposition precedes the initiation of lignin formation (Morrison and Kessler 1994). This is clearly visible in the developmental series of the growing internodes of maize (Figure 65). In control plants, the lignin content as percentage of the cell wall increases rapidly between V12 and V14 and further increases until the transition to reproductive phase (Figure 65A). At S stage, around 10% of the cell wall consists of lignin. Lignification of the cell wall occurs earlier in development in the GA enhanced plants (Figure 65A). In V10 and V12 stage, respectively 50% and 63% more lignin is present in internode cell walls of transgenic plants. At V14 stage, the cell walls contain 9.4 % lignin already, compared to 7.5% in control plants. Since the cell wall fraction is significantly higher in GA overproducing plants, the overall lignin content is also significantly increased in V10, V12, V14 and S stage (Figure 65B).

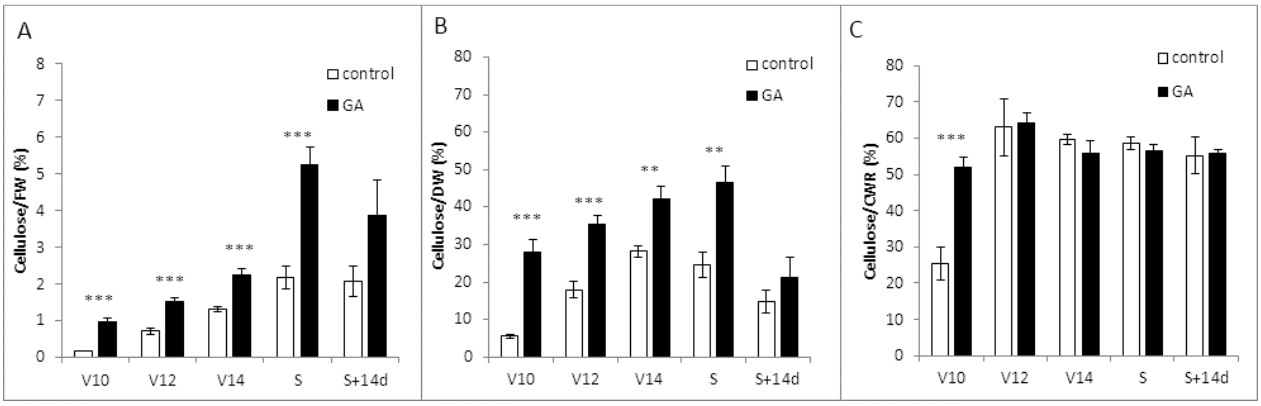


Figure 64. Cellulose deposition in the developing ninth internode of control and GA overproducing plants (GA); cellulose per fresh weight (FW) (A), per dry weight (DW) (B) and per cell wall residue (CWR) (C); V10, V12, V14 (10, 12 and 14 leaf collars visible respectively), S (silking) and S+14d (14d after silking) indicate the different developmental stages; Error bars represent standard errors of 9, 9, 8, 7 and 6 biological replicates in stages V10, V12, V14, S and S+14d respectively; *: p<0.05, **: p<0.01, ***:p<0.001.

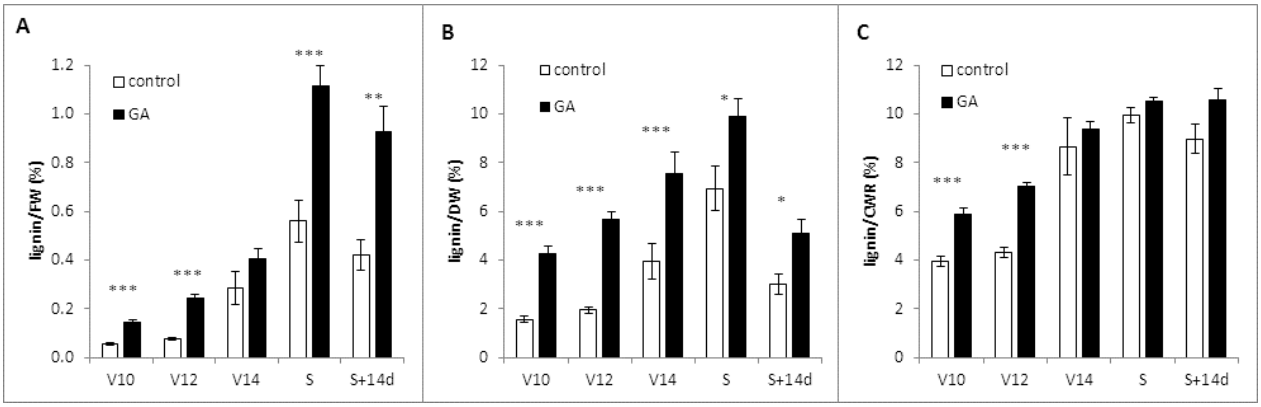


Figure 65. Lignin deposition in the developing ninth internode of control and GA overproducing plants (GA); lignin per fresh weight (FW) (A), per dry weight (DW) (B) and per cell wall residue (CWR) (C); V10, V12, V14 (10, 12 and 14 leaf collars visible respectively), S (silking) and S+14d (14d after silking) indicate the different developmental stages; Error bars represent standard errors of 9, 9, 8, 7 and 6 biological replicates in stages V10, V12, V14, S and S+14d respectively; *: p<0.05, **: p<0.01, ***:p<0.001.

4.3.3 Shift in expression of lignin biosynthetic genes supports early accumulation of lignin

To further support the higher lignin accumulation in internodes of GA enhanced plants, the expression pattern of genes involved in monolignol biosynthesis (*PAL*, *C4H*, *4CL*, *HCT*, *C3H*, *CCoAOMT*, *CCR*, *COMT*, *F5H* and *CAD*) and subsequent polymerization by peroxidases (*POX*) and laccases (*LAC*) was investigated. For the above listed gene families, those family members were investigated with a putative role in lignification, based on expression analysis studies that used the MAIZEWALL database (Guillaumie et al. 2007; Riboulet et al. 2009). For the gene families of which multiple members were analyzed, the expression data was provided in Supplementary figure 13 and 3. The expression kinetics of the most representative member of each gene family (Figure 10) followed one of the following profiles; 1: steadily increasing expression towards the most mature stage (*ZmPAL1*, *ZmC4H1*, *ZmC3H1*, *ZmCCR1* and *ZmCAD2*); 2: Increasing expression, coming to a maximum value around transition to reproductive stage and thereafter decreasing again (*ZmHCT2*, *ZmCCoAOMT3*, and *ZmCOMT*); 3: steadily decreasing expression towards the most mature stage (*POX39*); 4: steady expression with peak expression at one of the mature stages (*ZmF5H1* and *ZmLAC4*).

The expression pattern of most genes thus followed the profile of internode lignification: increasing lignin deposition in the cell wall in the vegetative stages and reaching a maximum around the transition to the reproductive phase. The *POX39* gene (profile 3), was highest expressed in the vegetative stage and expression decreased towards the mature stage, suggesting perhaps a different role than acting in lignification. The expression profiles were preserved in GA overproducing plants but most genes were upregulated in the early vegetative stages after which expression levels were similar to or lower than for control plants (Figure 67). This suggests that the maximal expression of lignin biosynthesis genes was shifted to an earlier point in development in GA overproducing plants compared to control plants. Interestingly, *ZmF5H1* which is involved specifically in the formation of syringyl (S) units, was upregulated throughout development in the transgenics (Figure 67).

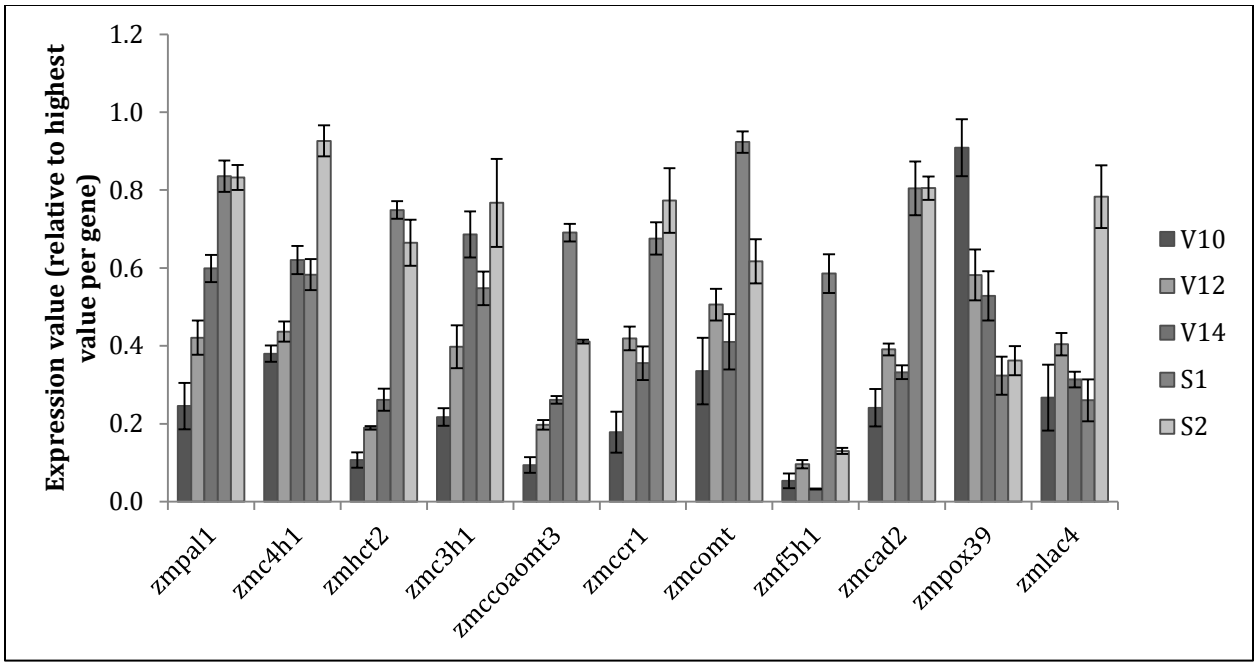


Figure 66. Expression profile of lignin biosynthetic genes in the developing maize ninth internode of control plants. V10, V12, V14 (10, 12 and 14 leaf collars visible respectively), S (silking) and S+14d (14d after silking) indicate the different developmental stages. Per gene family the most representative member is shown. Error bars represent standard errors of four biological replicates.

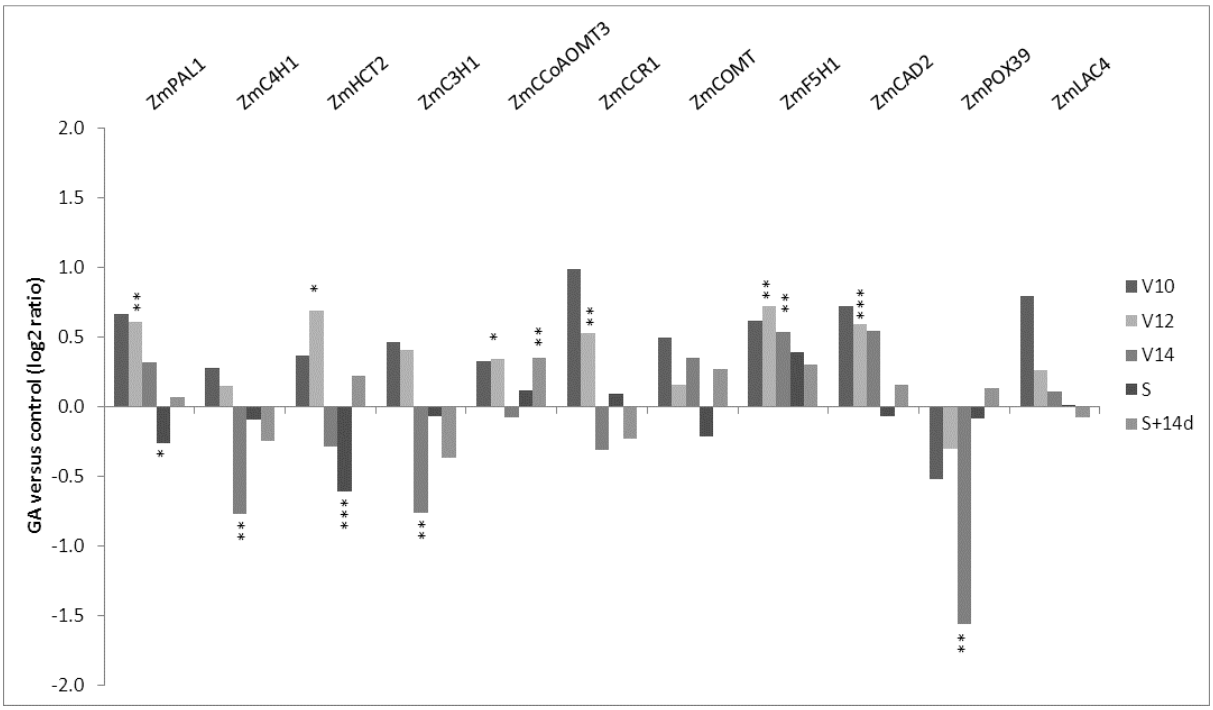


Figure 67. Ratio of lignin biosynthetic gene expression in GA enhanced plants versus control plants in the developing maize ninth internode. V10, V12, V14 (10, 12 and 14 leaf collars visible respectively), S (silking) and S+14d (14d after silking) indicate the different developmental stages. Per gene family the most representative member is shown. *: p<0.05, **: p<0.01, ***:p<0.001.

4.3.4 Lower S/G ratio in the internodes of GA overproducing plants

Besides the lignin amount, the lignin composition is of great importance for the degradability of the cell wall (Li et al., 2010; Van Acker et al., 2013, Papa et al., 2013; Guo et al., 2001). The relative abundance of the extractable lignin units, *p*-hydroxyphenyl (H), guaiacyl (G) and syringyl (S) in GA overproducing and control plants (Figure 68) is consistent with the compositional analysis of dry stems at fully mature stage (Figure 61B). The H unit comprised only a minor component in the lignin polymer (<5%) and was mainly present in young internodes (Figure 68A). The amount of H units was not altered in the GA enhanced plants. Around 30% of the lignin consists of G units and there was a clear increase in abundance in the GA enhanced plants in V12 and V14 stage (Figure 68B). The most abundant units in the lignin of the developing ninth internode of maize plants were S units and counted up to 60-70% of total monomers present. The fraction of S units was reduced in the transgenics in V12 and V14 stage (Figure 68C). Accordingly, S/G, which is often related to cell wall digestibility, was reduced significantly in the transgenics in V12 and V14 stage (Figure 68D).

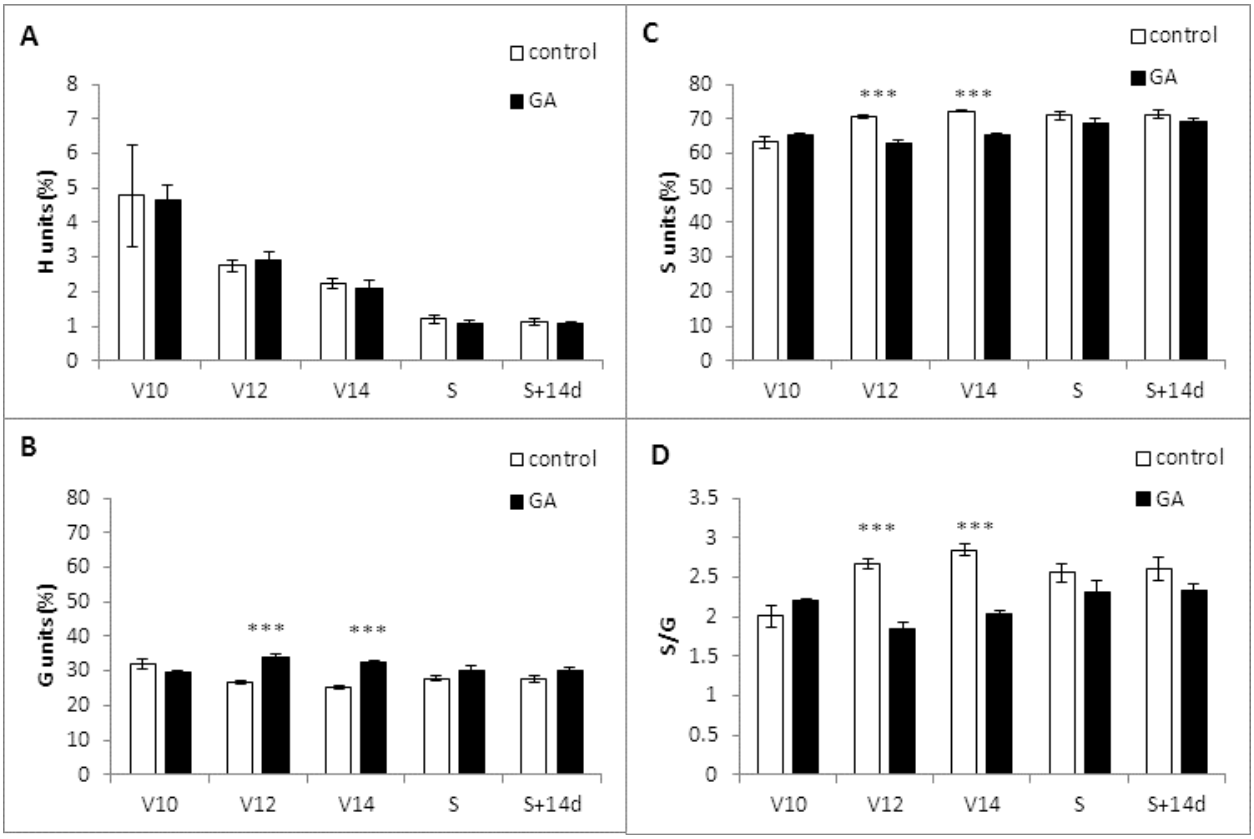


Figure 68. Lignin composition in the ninth internode of control and GA overproducing (GA) plants over development. The relative abundance of H, G and S units (A,B and C respectively) as well as the S/G ratio (D) was shown. V10, V12, V14 (10, 12 and 14 leaf collars visible respectively), S (silking) and S+14d (14d after silking) indicate the different developmental stages; Error bars represent standard errors of 5 biological replicates; ***:p<0.001.

5. Discussion

In this study, we investigated the effect of enhanced GA levels in maize plants, for the purpose of engineering varieties with enhanced growth and biomass quantity and quality traits for the production of second generation biofuels.

The total biomass of GA overproducing plants was lower than that of control plants and this could be attributed to a reduction in grain yield mainly. The seeds are not used for second generation biofuels, but seed setting could be important for the production of seeds for cultivation or maize used for dual-purpose. The decrease in grain weight arose from a lower seed number since the average seed weight was not significantly altered. The possibility of a reduced fertility of the transgenics cannot be excluded here, but this lower seed setting could also be due to the particular set-up of the experiment. A dedicated experiment under controlled conditions to estimate fertility parameters, in combination with a field trial conducted on agronomical scale would shed light on this aspect. The leaf dry weight of GA overproducing plants at maturity was also reduced, despite their increased length (Nelissen et al. 2012). The decrease in grain and leaf weight is partially compensated by an increase in stem weight in GA overproducing plants. The dry stems thus become the major contributors to aboveground plant biomass in GA overproducing plants, a highly desirable trait for lignocellulosic crops (Feltus and Vandenbrink 2012). The distribution of biomass in the different organs of the plant is referred to as biomass allocation. Plants are capable of altering biomass allocation upon changing environmental conditions or show a different allocation pattern in different planting densities (Poorter et al. 2012). In this study we thus observed an effect of GA overproduction on biomass allocation in the aerial parts of the maize plant. Interestingly, the allocation pattern in these maize plants overproducing GA show the opposite effects as the wheat and rice GA mutants that formed the basis of the green revolution. Whereas GA mutants had an improved harvest index, these maize GA overproducing plants have a lower harvest index. Thus, the lower seed yield might arise from a shift in carbon allocation, less to the seeds and more to the stem. The stem of these maize GA overproducing plants might form a strong competing sink with the ear, resulting in a reduced flux of carbon to the ear and less developing seeds on the cob. In *GA20ox1* overexpressing trees (*Populus* spp.) (Eriksson et al. 2000), a similar effect of increased fraction of stem biomass was observed but the slight increase in the fraction of leaf biomass observed in poplar contrasts to the reduction in leaf biomass in GA overproducing maize plants in this study.

In addition, the stems of GA overproducing plants are much taller but are also more slender. This shows that besides the altered biomass allocation pattern, GA overproducing plants have altered stem morphology. Although the characteristics of the stem were not investigated at the cellular level, cells in the internode of transgenic plants might be longer but thinner, explaining the observations of longer but more slender internodes. This phenomenon is generally addressed as a "compensation mechanism", and has been observed in many cases when plant or organ size was altered (Tsuge et al. 1996; De Veylder et al. 2001; Tsukaya 2002a; Kim et al. 2002a; Tsukaya 2002b; Beemster et al. 2003; Narita et al. 2004; Boudolf et al. 2004; Tsukaya 2005; Horiguchi et al. 2005; Tsukaya 2006; Barrôco et al. 2006; Horiguchi et al. 2006; Ferjani et al. 2007). The mechanisms underlying this compensation effect are largely unknown but currently of great interest to researchers (Horiguchi et al. 2011).

The developing stem of GA overproducing plants accumulates more DW than control plants, a phenomenon that was also previously described in poplar and tobacco overexpressing *GA20ox* (Eriksson et al. 2000; Biemelt et al. 2004). The increase in DW in tobacco plants was attributed to a higher degree of lignification (Biemelt et al. 2004). However, in our analysis, whole-stems of GA overproducing mature dry plants contain more lignin as well as cellulose and have a higher CWR. This indicates the presence of more and/or thicker secondary cell wall, perhaps to compensate for the loss in strength of the more slender stem. The cell walls contain high amounts of cellulose and the presence of more cell walls is thus a desirable trait for bioenergy crops (Demura and Ye 2010; Yang et al. 2013). On the other side, the altered cell wall composition resulted in a higher recalcitrance to acid and alkali pretreatment, and a reduced saccharification efficiency of dry whole-stem material when pretreated with acid, but not with alkali. Recalcitrance to both acid pretreatment and enzymatic digestion is directly proportional to lignin content (Chen and Dixon 2007). This indicates that the higher lignin content influences the saccharification efficiency of stem biomass of GA overproducing plants negatively. On the other hand, when saccharification efficiency and biomass accumulation are combined, as was done by Chen et al. (2007), the decrease in saccharification efficiency using acid pretreatment by 23% would be compensated by the increase in stem biomass of 32%. Using an alkali pretreatment, no decrease in saccharification efficiency was observed, resulting in a major increase in glucose yield when combined with the increase in stem biomass yield of GA overproducing plants.

In the developing internode, the onset of secondary cell wall formation involves cellulose deposition first, followed by lignin biosynthesis. This is in accordance with what was described by Morrison et al. (1994). Later on, both cellulose and lignin levels continue to rise with the accumulation of more and/or thicker secondary cell walls. Remarkably, the cell wall fraction (CWR/DW) drops in the S+14d stage, perhaps as a consequence of sucrose accumulation (Setter and Meller 1984; Dwyer et al. 1995; Jung and Casler 2006; de Souza et al. 2013). In the GA overproducing plants, CWR, lignin and cellulose levels are higher over development suggesting a higher and faster formation of secondary cell walls in the developing internode. These findings are in accordance with the measurements on fully mature dry stems. The earlier onset of lignification is supported by lignin biosynthetic gene expression measurements indicating that peak expression occurs earlier in GA overproducing plants.

The majority of lignin biosynthetic genes for the expression analysis using qRT-PCR were chosen based on expression data from previous studies that used the MAIZEWALL array for expression analysis (Guillaumie et al. 2007; Riboulet et al. 2009). This macroarray contains gene-specific tags for 735 cell wall-related genes that were studied in different organs and during internode development in maize. We chose those members in the lignin biosynthetic multigene families that had a high expression in the maize stem (Guillaumie et al. 2007) and more specifically the maize internode over development (Riboulet et al. 2009). Compared with the expression study of lignin genes in the maize internode over development by Riboulet et al. (2009), our analysis considered gene expression over a much longer period in development. Riboulet et al. (2009) started sampling internodes at VT stage, followed by S, S+8d and S+15d. In the present study, lignin biosynthetic gene expression was investigated in internodes at V10, V12, V14, S and S+14 stage, thus covering also the vegetative phase of plant development. The expression profiles of the investigated lignin

biosynthetic genes correlated well with the process of lignification in our experiment, showing that starting internode sampling early in development is relevant for studying expression patterns of lignin biosynthetic gene expression.

The lignin in mature maize stems is reported to be composed of approximately 4% of H units, 35% of G units and 61% of S units (Halpin et al. 1998b). We found a very similar composition in our analysis of stems of control plants (1% of H units, 33% of G units and 66% of S units). When lignin is deposited in the cell wall of a young, growing internode, it is enriched in G lignin as compared to the fully mature internode (Vallet and Chabbert 1996; Morrison and Jung 1998; Chen et al. 2002; Grabber et al. 2004a). In our analysis, the S/G ratio increased from V10 to V14 stage as the internode matures, in accordance with literature. As compared to control plants, the S/G ratio is significantly lower in GA overproducing plants at V12 and V14 stage as a result of a decrease in the relative abundance of S and an increase of G units. In S-lignin biosynthesis, ferulate-5-hydroxylase (F5H) is the rate-limiting enzyme (Ruegger et al. 1999). The expression analysis in the developing ninth internode showed an upregulation of the *ZmF5H1* gene expression in the GA overproducing plants (Figure 67). However in the lignin monomeric composition, a lower S/G ratio was found, which is intuitively contradictory as increased *F5H* expression would imply more S units. We demonstrated that lignin deposition occurs earlier in development in GA overproducing plants. Since the internodes of GA overproducing plants elongate so fast and to such height, we expect that a high demand for lignin monomers is present. Because of the high demand, possibly the rate-limiting enzyme *F5H* is not able to convert enough G-lignin into S-lignin in such a short time interval, leading to an overrepresentation of G-lignin. In contrast, the slower elongation of the non-transgenic control internode allows for availability of relatively more S to form the lignin polymer. Although this is only a hypothesis, the higher expression of the *ZmF5H1* gene in the GA overproducing plants (Figure 67) fits this scenario since the high demand for lignin monomers would put a stress on this enzymatic step and drive *ZmF5H1* expression higher. According to literature, changes in S/G ratio can reflect an altered extractability or degradability of the lignin polymer (Jung and Casler 1991; Jung and Buxtono 1994; Baucher et al. 1999; Jung et al. 1999). However, debate remains involving the role of the different lignin subunits in saccharification efficiency, since an increased as well as a decreased S/G ratio have been proposed to be beneficial (Grabber and Ralph 1997; Feltus and Vandenbrink 2012). Most likely the altered S/G ratios reflect other compositional changes in the lignin that influence saccharification efficiency (Grabber and Ralph 1997), being the effect of an altered S/G ratio indirect. The altered S/G ratio, albeit only significant in two developmental stages, could thus explain the higher recalcitrance of stem cell walls from the transgenic plants to enzymatic degradability after acid pretreatment.

The question has been raised whether the altered cell wall properties are a direct effect of enhanced GA levels (Biemelt et al. 2004). Perhaps the alterations are caused by 'sensing' a need for compensation for the more slender stem, but this cannot be answered with the current experimental setup. However, our results demonstrate that enhancing GA levels in maize to improve stem biomass quantity can influence cell wall properties. With this case study we thus demonstrated that biomass quantity and quality are interconnected. This is important for the development of future breeding strategies to improve lignocellulosic feedstock for bioethanol production.

6. Experimental procedures

6.1 Plant material

As described by Nelissen et al. (2012), constitutive overexpression of *GA20-oxidase* was attained by cloning *AtGA20ox1* gene (At4g25420) behind the maize UBI1 promoter (Christensen and Quail 1996) in the vector pMBb7Fm21GW-UBIL (http://gateway.psb.ugent.be/) and introduced it into maize B104 inbred line by *Agrobacterium tumefaciens* transformation of immature embryos as described (Coussens et al. 2012). A segregating population of the backcross of a transgenic plant heterozygous for the *AtGA20ox1* overexpression construct to the wildtype B104 plant was used for the experiments described here. All experiments were executed in a greenhouse at minimum 25°C during day and 23°C during night in a 16h/8h rhythm. Supplementary light was added when natural light intensity was below 200 W/m^2 using high-pressure sodium vapor lamps. Fertilizer was added with the water supply; conductivity Ec = 1mS/cm; water soluble fertilizer Poly-feed (Haifa, Belgium) (N, P_2O_5, K_20; 20:5:20 + 3 MgO). Transgenic and control plants in the segregating population were identified based on basta resistance using basta leaf painting (Yao et al. 2006).

A total of 160 progeny plants were used for the analysis of plant morphology, cell wall characterization and expression analysis over development. Nine transgenic and nine control plants were selected randomly in V10, V12, V14 and VT (10, 12 and 14 visible leaf collars and tasseling respectively) stage and seven in S and S+14d (silking and 14 days after silking respectively) stage. The height of the youngest visible leaf in V10, V12 and V14 stage and height of the tassel in VT, S and S+14d was measured using a foldable meter. The width of the ninth internode, counted from the base, was measured along its broadest diameter using a vernier caliper (0.01 mm accuracy). The ninth internode was dissected from the stem in V10, V12, V14, S and S+14d stages and cut transversally in three parts of similar lengths, immediately frozen in liquid nitrogen and stored at -80°C. The top part of the internode (developmentally older than the bottom part; Scobbie et al. 1993) was used for biochemical cell wall characterization and the bottom part for expression analysis. This was considered convenient for our purpose since in this way, expression analysis of lignin biosynthetic genes can be evaluated earlier on in development than cell wall characteristics. Thirty progeny plants were used for organ-specific biomass, cell wall and saccharification analysis at fully mature, completely dry stage.

An independent set of 30 progeny plants were grown until maturity. Fertilization was achieved by shaking the stems for pollen shed every morning until no more pollen was produced. The plants were allowed to dry off completely whereupon leaves, stems, cobs and seeds were separated weighted. Whole stems were ground using a cutting mill (Fristch, Lelystad, Netherlands) with sieve of 0.5 mm and used for cell wall characterization and saccharification analysis.

6.2 Expression analysis using qRT-PCR

Genes involved in lignin biosynthesis and secondary cell wall formation were selected from the maizewall database (Guillaumie et al. 2007; http://www.polebio.lrsv.ups-tlse.fr/MAIZEWALL/).

These include 16 phenylpropanoid genes, three peroxidases, one laccase and four transcription factors (Supplementary Table 14). Gene-specific primers from the maizewall database were tested for specificity and primer efficiency >1.7 in qRT-PCR (described below). If not satisfactory, new primers were designed using Primer3 (http://bioinfo.ut.ee/primer3-0.4.0/primer3/) with standard settings (Supplementary Table 14). Internode samples were ground with a Mixer Mill MM 400 and Tungsten carbide 25 ml grinding jars (Retsch, Haan, germany). RNA was extracted using RNeasy kit (Qiagen, Valencia, CA) and a DNase treatment was performed using DNA-free™ (Ambion, Life technologies, Carlsbad, California, U.S.). Extracted RNA was quantified using the nanodrop® ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA) and diluted so that a total of 400 ng RNA was used for cDNA synthesis using the First strand cDNA synthesis kit (Thermo Scientific, Thermo Fisher Scientific, Waltham, MA, USA). A ten times diluted cDNA sample was used for RT-qPCR using KIT on a Lightcycler 480 (Roche, Basel, Switzerland). Samples were run in technical triplicates on the LC480 with following protocol: 1 activation cycle of 10 min at 95°C; 45 amplification cycles of 10 s at 95°C, 10 s at 60 °C and 10 s at 72 °C; 1 melting curve cycle measuring from 65 to 95 °C. Fluorescence values were exported from the lightcycler program whereupon Ct values, normalization factors and primer efficiencies were calculated according to Ramakers et al. (2003) using *ZmEF1a* and *Zm18S* as reference genes.

6.3 Cell wall fraction in ninth internode over development

For the biochemical analysis of cell wall components such as cellulose, lignin and cell wall residue (CWR), fresh internode samples were frozen in liquid nitrogen for use. To get a correction factor that allowed to estimate the CWR per unit dry matter, the dry to fresh weight ratio was determined for three stages over development (V10, V14 and S+14d) using the samples that were reserved for expression analysis (bottom part of the same 9th internode). The dry weight accumulation in the transgenics and non-transgenics was modeled using a polynomial and the respective equations were used to convert the cell wall residue values per fresh weight into cell wall residue per dry weight values.

6.4 Lignin analyses

Aliquots of 5 mg ground stem material were subjected to a sequential extraction to obtain a purified CWR. The extractions were done in 2-ml vials, each time for 30 min, at near boiling temperatures for water (98°C), ethanol (76°C), chloroform (59°C) and acetone (54°C). The remaining CWR was dried under vacuum. The lignin was quantified according to a modified version of the acetyl bromide method (Dence 1992), optimized for small amounts of plant tissue. The dried CWR was dissolved in 0.1 freshly made 25% acetyl bromide in glacial acetic acid and 4 µl 60% perchloric acid. The solution was incubated for 30 min at 70°C while shaking (850 rpm). After incubation, the slurry was centrifuged at 14000 rpm for 15 min, To the supernatant, 0.2 ml of 2M sodium hydroxide and 0.5 ml glacial acetic acid was added. The pellet was washed with 0.5 ml glacial acetic acid. The supernatant and the washing phase were combined and the final volume was adjusted to 2 ml with glacial acetic acid. After 20 min at room temperature, the absorbance at 280 nm was measured with a nanodrop® ND-1000 spectrophotometer (Thermo Scientific, Wilmington, DE, USA). The lignin concentrations were calculated by means of the Bouguer-Lambert-Beer law: $A = \varepsilon \times l \times c$ (A = absorbance, ε = extinction coefficient, l = path length, c = concentration), with $\varepsilon = 20.48\ L\ g^{-1}\ cm^{-1}$ (Fukushima and Hatfield 2004) and l = 0.1 cm.

The lignin composition was investigated with thioacidolysis as previously described (Robinson and Mansfield 2009). The monomers involved in β-O-4 ether bonds, released upon thioacidolysis, were detected with gas chromatography (GC) as their trimethylsilyl (TMS) ether derivates on a Hewlett-Packard HP 6890 Series system (Agilent, Santa Clara, SA, USA) coupled with a HP-5973 mass-selective detector. The GC conditions were as described (Robinson et al., 2009). The quantitative evaluation was based on the specific prominent ions for each compound. A summary of the specific ions for each specific compound can be found in Supplementary Table 15. Response factors for H, G and S units were taken from (Yue et al. 2012).

6.5 Cellulose analysis

Aliquots of 5 mg ground stem material were subjected to a sequential extraction to obtain a purified CWR, as described above. To estimate the amount of cellulose, we used a colorimetric method (based on DuBois et al. 1956; Leplé et al. 2007). The CWR was incubated with 2 M TFA (trifluoroacetic acid) and 20 µl inositol (5 mg ml-1) for 2 h at 99°C while shaking (750 rpm). After incubation, the remaining pellet was washed three times with water and twice with acetone and dried under vacuum. Concentrated sulfuric acid (150 µl) and 30 µl 5% (w/v) phenol (freshly made in water) were added to the dried pellet and incubated for 1 h at 90°C with gentle shaking (500 rpm). After centrifugation for 3 min at 23,477 g, a 50 µl aliquot of the supernatant was diluted 20 times with MilliQ water (Millipore, Billerica, MA, USA) to measure the absorbance at 493 nm. The amount of cellulose was calculated back from a standard curve of Avicel®PH-101 (FMC BioPolymer, Philadelphia, PA, USA).

6.6 Saccharification assay

Aliquots of 20 mg of dry stem material were used. The biomass was pretreated with 1 ml of either 1M HCl or 1M of NaOH at 80°C for 2h, while shaking (850 rpm). The supernatant was removed and the pellet containing pretreated material was washed three times with water to obtain a neutral pH. Subsequently, the material was incubated in 1 ml 70% (v/v) ethanol overnight at 55°C. The remaining biomass was washed three times with 1 ml 70% (v/v) ethanol, once with 1 ml acetone, dried under vacuum for 45 min and weighed. The pretreated ethanol-extracted residue was dissolved in 1 ml acetic acid buffer solution (pH 4.8) and incubated at 50°C. Accelerase® 1500 (Genencor, Denmark) enzyme mix was first desalted over an Econo-Pac 10DG-column (Bio-Rad, Hercules, CA, USA), stacked with Bio-gel® P-6DG gel (Bio-rad) according to the manufacturer's guidelines. The activity of the enzyme mix was measured with a filter paper assay (Xiao et al., 2004). To each sample, dissolved in acetic acid buffer (pH 4.8), the enzyme mix with an activity of 0.04 filter paper units was added. After a short spinning to remove droplets from the lid of the reaction tubes, 20 µl aliquots of the supernatant were taken after 0h, 4h, 7h, 24h and 48h incubation at 50°C and 10 fold diluted with acetic acid buffer (pH4.8). The concentration of glucose in these diluted samples was measured indirectly with a spectrophotometric color reaction (glucose oxidase-peroxidase; GOD-POD) A 100 ml aliquot of the reaction mix from this color reaction contained 50 mg ABTS, 44.83 mg GOD (Sigma-Aldrich, St. Louis, MO, USA) and 173 µl of 4% (w/v) POD (Roche Diagnostics, Brussels, Belgium) in acetic acid buffer (pH 4.5). To measure the concentration of glucose, 50 µl of the diluted samples was added to 150 µl GOD_POD solution and incubated for 30 min at 37 °C. The absorbance was measured spectrophotometrically at a wavelength of 405 nm. The concentration in the original sample was calculated with a standard

curve based on known D-glucose (Sigma-Aldrich) concentrations. Glucose release was then expressed per unit dry weight or unit biomass left after pretreatment.

6.7 Microscopy on stem sections

Stem sections of the ear internode were used for microscopic analysis. Stem pieces of 5 cm were fixated using freshly made fixating agent (2.5% formaldehyde in 0.05 M acetic acid buffer) for two days at 4°C while shaking. The fixated stems were dehydrated by removing the fixating agent and replacing with increasing concentrations of ethanol (10%, 30%, 50% and 70%), each time incubated for 2h at 4°C while shaking. Sections of 200 µm were made using a vibroslicer. Stem sections were stained using Wiesner staining reagent (1% phloroglucinol w/v in 100 ml 95% ethanol (v/v) and 16 ml 37% (v/v) HCl). Stained and unstained sections were visualized using Olympus BX51 microscope (Shinjuku, Tokyo, Japan).

6.8 Statistical analysis

For analyses in which multiple technical replicates were involved (three technical replicates for lignin, cellulose and qRT-PCR analysis and four technical replicates for saccharification analysis), average values were used for statistical analysis. Statistical analyses consisted of Student t-test comparisons of transgenic / control samples. All analyses were carried out in the software package STATISTICA version 11 (Statsoft Inc., USA).

7. References

Alberts B, Johnson A, Lewis J, et al. (2002) Molecular biology of the cell, 4th editio. Ann Bot 91:1616.

Allison GG, Morris C, Hodgson E, et al. (2009) Measurement of key compositional parameters in two species of energy grass by Fourier transform infrared spectroscopy. Bioresour Technol 100:6428–33. doi: 10.1016/j.biortech.2009.07.015

Aloni R (1979) Role of Auxin and Gibberellin in Differentiation of Primary Phloem Fibers. PLANT Physiol 63:609–614. doi: 10.1104/pp.63.4.609

Aloni R, Tollier MT, Monties B (1990) The Role of Auxin and Gibberellin in Controlling Lignin Formation in Primary Phloem Fibers and in Xylem of Coleus blumei Stems. Plant Physiol 94:1743–7.

Argillier O, Barrière Y, Dardenne P (1998) Genotypic variation for in vitro criteria and relationships with in vivo digestibility in forage maize hybrids. Plant … 441:

Baker J (1987) Experiment With the Plant Growth Hormone Gibberellin. Honor. Theses

Barrière Y, Alber D, Dolstra O, et al. (2006) Past and prospects of forage maize breeding in Europe. II. History, germplasm evolution and correlative agronomic changes. Maydica 51:435–449.

Barrière Y, Méchin V, Riboulet C, et al. (2009) Genetic and genomic approaches for improving biofuel production from maize. Euphytica 170:183–202. doi: 10.1007/s10681-009-9923-6

Barrôco RM, Peres A, Droual A-M, et al. (2006) The cyclin-dependent kinase inhibitor Orysa;KRP1 plays an important role in seed development of rice. Plant Physiol 142:1053–64. doi: 10.1104/pp.106.087056

Baucher M, Bernard-Vailhé M a, Chabbert B, et al. (1999) Down-regulation of cinnamyl alcohol dehydrogenase in transgenic alfalfa (Medicago sativa L.) and the effect on lignin composition and digestibility. Plant Mol Biol 39:437–47.

Beemster GTS, Fiorani F, Inzé D (2003) Cell cycle: the key to plant growth control? Trends Plant Sci 8:154–8. doi: 10.1016/S1360-1385(03)00046-3

Biemelt S, Tschiersch H, Sonnewald U (2004) Impact of altered gibberellin metabolism on biomass accumulation, lignin biosynthesis, and photosynthesis in transgenic tobacco plants. Plant Physiol 135:254–265. doi: 10.1104/pp.103.036988.254

Bolduc N, Hake S (2009) The maize transcription factor KNOTTED1 directly regulates the gibberellin catabolism gene ga2ox1. Plant Cell 21:1647–58. doi: 10.1105/tpc.109.068221

Boon EJMC, Struik PC, Engels FM, Cone JW (2012) Stem characteristics of two forage maize (Zea mays L.) cultivars varying in whole plant digestibility. IV. Changes during the growing season in anatomy and chemical composition in relation to fermentation characteristics of a lower internode. NJAS - Wageningen

J Life Sci 59:13–23. doi: 10.1016/j.njas.2011.05.001

Bosch L, Casañas F, Sánchez Bell E, et al. (1999) Forage maize in mild-temperate zones: breeding strategies for the near future. trends Agron 2:1–12.

Boudet AM, Kajita S, Grima-Pettenati J, Goffner D (2003) Lignins and lignocellulosics: a better control of synthesis for new and improved uses. Trends Plant Sci 8:576–81. doi: 10.1016/j.tplants.2003.10.001

Boudolf V, Vlieghe K, Beemster GTS, et al. (2004) The plant-specific cyclin-dependent kinase CDKB1;1 and transcription factor E2Fa-DPa control the balance of mitotically dividing and endoreduplicating cells in Arabidopsis. Plant Cell 16:2683–92. doi: 10.1105/tpc.104.024398

Brenner EA, Zein I, Chen Y, et al. (2010) Polymorphisms in O-methyltransferase genes are associated with stover cell wall digestibility in European maize (Zea mays L.). BMC Plant Biol 10:27. doi: 10.1186/1471-2229-10-27

Brosse N, Dufour A, Meng X, et al. (2012) Miscanthus: a fast-growing crop for biofuels and chemicals production. Biofuels, Bioprod Biorefining 6:580–598. doi: 10.1002/bbb.1353

Bultynck L (2002) Leaf expansion and biomass allocation in wild wheat (Aegilops) species.

Carpita NC, McCann MC (2008) Maize and sorghum: genetic resources for bioenergy grasses. Trends Plant Sci 13:415–20. doi: 10.1016/j.tplants.2008.06.002

Carrera E, Bou J, García-Martínez JL, Prat S (2000) Changes in GA 20-oxidase gene expression strongly affect stem length, tuber induction and tuber yield of potato plants. Plant J 22:247–56.

Chen F, Dixon R a (2007) Lignin modification improves fermentable sugar yields for biofuel production. Nat Biotechnol 25:759–61. doi: 10.1038/nbt1316

Chen L, Auh C, Chen F, et al. (2002) Lignin deposition and associated changes in anatomy, enzyme activity, gene expression, and ruminal degradability in stems of tall fescue at different developmental stages. J Agric Food Chem 50:5558–65.

Chen Y (2011) Pleiotropic effects of genes involved in cell wall lignification on agronomic characters.

Christensen a H, Quail PH (1996) Ubiquitin promoter-based vectors for high-level expression of selectable and/or screenable marker genes in monocotyledonous plants. Transgenic Res 5:213–8.

Christian DG, Riche a. B, Yates NE (2008) Growth, yield and mineral content of Miscanthus×giganteus grown as a biofuel for 14 successive harvests. Ind Crops Prod 28:320–327. doi: 10.1016/j.indcrop.2008.02.009

Cleland C, Briggs W (1969) Gibberellin and CCC effects on flowering and growth in the long-day plant Lemna gibba G3. Plant Physiol 503–507.

Clifton-Brown JC, Lewandowski I, Andersson B, et al. (2001) Performance of 15 Genotypes at Five Sites in Europe. Agron

J 93:1013. doi: 10.2134/agronj2001.9351013x

Coles JP, Phillips a L, Croker SJ, et al. (1999) Modification of gibberellin production and plant development in Arabidopsis by sense and antisense expression of gibberellin 20-oxidase genes. Plant J 17:547–56.

Coussens G, Aesaert S, Verelst W, et al. (2012) Brachypodium distachyon promoters as efficient building blocks for transgenic research in maize. J Exp Bot 63:4263–73. doi: 10.1093/jxb/ers113

Dayan J, Voronin N, Gong F, et al. (2012) Leaf-induced gibberellin signaling is essential for internode elongation, cambial activity, and fiber differentiation in tobacco stems. Plant Cell 24:66–79. doi: 10.1105/tpc.111.093096

De Souza AP, Arundale R a., Dohleman FG, et al. (2013) Will the exceptional productivity of Miscanthus x giganteus increase further under rising atmospheric CO 2 ? Agric For Meteorol 171-172:82–92. doi: 10.1016/j.agrformet.2012.11.006

De Veylder L, Beemster GT, Beeckman T, Inzé D (2001) CKS1At overexpression in Arabidopsis thaliana inhibits growth by reducing meristem size and inhibiting cell-cycle progression. Plant J 25:617–26.

De Vrije T (2002) Pretreatment of Miscanthus for hydrogen production by Thermotoga elfii. Int J Hydrogen Energy 27:1381–1390. doi: 10.1016/S0360-3199(02)00124-6

Dello Ioio R, Nakamura K, Moubayidin L, et al. (2008) A genetic framework for the control of cell division and differentiation in the root meristem. Science 322:1380–4. doi: 10.1126/science.1164147

Demura T, Ye Z-H (2010) Regulation of plant biomass production. Curr Opin Plant Biol 13:299–304. doi: 10.1016/j.pbi.2010.03.002

Dence CW (1992) The Determination of Lignin. In: Stephen YL, Dence CW (eds) Methods Lignin Chem. Springer Berlin Heidelberg, pp 33–61

DuBois M, Gilles KA, Hamilton JK, et al. (1956) Colorimetric Method for Determination of Sugars and Related Substances. Anal Chem 28:350–356. doi: 10.1021/ac60111a017

Dubouzet JG, Strabala TJ, Wagner A (2013) Potential transgenic routes to increase tree biomass. Plant Sci 212:72–101. doi: 10.1016/j.plantsci.2013.08.006

Dudits D, Abrahám E, Miskolczi P, et al. (2011) Cell-cycle control as a target for calcium, hormonal and developmental signals: the role of phosphorylation in the retinoblastoma-centred pathway. Ann Bot 107:1193–202. doi: 10.1093/aob/mcr038

Dwyer LM, Andrews CJ, Stewart DW, et al. (1995) Carbohydrate Levels in Field-Grown Leafy and Normal Maize Genotypes. Crop Sci 35:1020. doi: 10.2135/cropsci1995.0011183X003500040016x

Eriksson ME, Israelsson M, Olsson O, Moritz T (2000) Increased gibberellin biosynthesis in transgenic trees promotes growth, biomass production

and xylem fiber length. Nat Biotechnol 18:784–8. doi: 10.1038/77355

European Academies Science Advisary Council (2012) The current status of biofuels in the European Union, their environmental impacts and future prospects; EASAC policy report 19. 47. doi: ISBN: 978-3-8047-3118-9

Evenson RE, Gollin D (2003) Assessing the impact of the green revolution, 1960 to 2000. Science 300:758–62. doi: 10.1126/science.1078710

Fagoaga C, Tadeo FR, Iglesias DJ, et al. (2007) Engineering of gibberellin levels in citrus by sense and antisense overexpression of a GA 20-oxidase gene modifies plant architecture. J Exp Bot 58:1407–20. doi: 10.1093/jxb/erm004

FAO Statistics Division (2013) FAOSTAT 2011. http://faostat3.fao.org/faostat-gateway/go/to/download/Q/QC/E. Accessed 28 Dec 2013

Farage PK, Blowers D, Long SP, Baker NR (2006) Low growth temperatures modify the efficiency of light use by photosystem II for CO2 assimilation in leaves of two chilling-tolerant C4 species, Cyperus longus L. and Miscanthus x giganteus. Plant Cell Environ 29:720–8.

Feltus FA, Vandenbrink JP (2012) Bioenergy grass feedstock: current options and prospects for trait improvement using emerging genetic, genomic, and systems biology toolkits. Biotechnol Biofuels 5:80. doi: 10.1186/1754-6834-5-80

Ferjani A, Horiguchi G, Yano S, Tsukaya H (2007) Analysis of leaf development in fugu mutants of Arabidopsis reveals three compensation modes that modulate cell expansion in determinate organs. Plant Physiol 144:988–99. doi: 10.1104/pp.107.099325

Fukushima RS, Hatfield RD (2004) Comparison of the acetyl bromide spectrophotometric method with other analytical lignin methods for determining lignin concentration in forage samples. J Agric Food Chem 52:3713–20. doi: 10.1021/jf035497l

García-Hurtado N, Carrera E, Ruiz-Rivero O, et al. (2012) The characterization of transgenic tomato overexpressing gibberellin 20-oxidase reveals induction of parthenocarpic fruit growth, higher yield, and alteration of the gibberellin biosynthetic pathway. J Exp Bot 63:5803–13. doi: 10.1093/jxb/ers229

Grabber J, Ralph J (1997) p-Hydroxyphenyl, guaiacyl, and syringyl lignins have similar inhibitory effects on wall degradability. J Agric … 8561:2530–2532.

Grabber JH, Ralph J, Lapierre C, et al. (2004a) Genetic and molecular basis of grass cell-wall biosynthesis and degradability. III. Towards a forage grass ideotype. C R Biol 327:467–479. doi: 10.1016/j.crvi.2004.03.004

Grabber JH, Ralph J, Lapierre C, et al. (2004b) Genetic and molecular basis of grass cell-wall degradability. I. Lignin–cell wall matrix interactions. C R Biol 327:455–465. doi: 10.1016/j.crvi.2004.03.004

Guillaumie S, San-Clemente H, Deswarte C, et al. (2007) MAIZEWALL. Database and developmental gene expression profiling

of cell wall biosynthesis and assembly in maize. Plant Physiol 143:339–363.

Halpin C, Holt K, Chojecki J, et al. (1998) Brown-midrib maize (bm1)–a mutation affecting the cinnamyl alcohol dehydrogenase gene. Plant … 14:545–53. doi: 10.1046/j.1365-313X.1998.00153.x

Hay A, Tsiantis M (2010) KNOX genes: versatile regulators of plant development and diversity. Development 137:3153–65. doi: 10.1242/dev.030049

Heaton E (2010) Miscanthus: a promising biomass crop. Adv …. doi: 10.1016/B978-0-12-381518-7.00003-0

Heaton E a, Flavell RB, Mascia PN, et al. (2008a) Herbaceous energy crop development: recent progress and future prospects. Curr Opin Biotechnol 19:202–9. doi: 10.1016/j.copbio.2008.05.001

Heaton E a., Dohleman FG, Long SP (2008b) Meeting US biofuel goals with less land: the potential of Miscanthus. Glob Chang Biol 14:2000–2014. doi: 10.1111/j.1365-2486.2008.01662.x

Hedden P (2010) Green Revolution Genes. In: Taiz L, Zeiger E (eds) Plant Physiol., Fifth. Sinauer associates, Los Angeles, p 782

Hendriks a TWM, Zeeman G (2009) Pretreatments to enhance the digestibility of lignocellulosic biomass. Bioresour Technol 100:10–8. doi: 10.1016/j.biortech.2008.05.027

Hodkinson TR, Chase MW, Lledó MD, et al. (2002) Phylogenetics of Miscanthus, Saccharum and related genera (Saccharinae, Andropogoneae, Poaceae) based on DNA sequences from ITS nuclear ribosomal DNA and plastid trnLintron and trnL-F intergenic spacers. J Plant Res 115:381–92. doi: 10.1007/s10265-002-0049-3

Horiguchi G, Fujikura U, Ferjani A, et al. (2006) Large-scale histological analysis of leaf mutants using two simple leaf observation methods: identification of novel genetic pathways governing the size and shape of leaves. Plant J 48:638–44. doi: 10.1111/j.1365-313X.2006.02896.x

Horiguchi G, Kim G-T, Tsukaya H (2005) The transcription factor AtGRF5 and the transcription coactivator AN3 regulate cell proliferation in leaf primordia of Arabidopsis thaliana. Plant J 43:68–78. doi: 10.1111/j.1365-313X.2005.02429.x

Horiguchi G, Nakayama H, Ishikawa N, et al. (2011) ANGUSTIFOLIA3 plays roles in adaxial/abaxial patterning and growth in leaf morphogenesis. Plant Cell Physiol 52:112–24. doi: 10.1093/pcp/pcq178

Huang S, Raman AS, Ream JE, et al. (1998) Overexpression of 20-oxidase confers a gibberellin-overproduction phenotype in Arabidopsis. Plant Physiol 118:773–81.

Israelsson M, Eriksson ME, Hertzberg M, et al. (2003) Changes in gene expression in the wood-forming tissue of transgenic hybrid aspen with increased secondary growth. Plant Mol Biol 52:893–903. doi: 10.1023/A:1025097410445

Israelsson M, Sundberg B, Moritz T (2005) Tissue-specific localization of

gibberellins and expression of gibberellin-biosynthetic and signaling genes in wood-forming tissues in aspen. Plant J 44:494–504. doi: 10.1111/j.1365-313X.2005.02547.x

Jakob K, Zhou F, Paterson AH (2009) Genetic improvement of C4 grasses as cellulosic biofuel feedstocks. Vitr Cell Dev Biol - Plant 45:291–305. doi: 10.1007/s11627-009-9214-x

Jenkins B., Baxter L., Miles T. (1998) Combustion properties of biomass. Fuel Process Technol 54:17–46. doi: 10.1016/S0378-3820(97)00059-3

Jung H (2011) Forage digestibility: the intersection of cell wall lignification and plant tissue anatomy. Int Adv Rumin Nutr Res … 162–174.

Jung HG, Casler MD (1991) Relationship of lignin and esterified phenolics to fermentation of smooth bromegrass fibre. Anim Feed Sci Technol 32:63–68. doi: 10.1016/0377-8401(91)90010-P

Jung HG, Casler MD (2006) Maize Stem Tissues. Crop Sci 46:1793. doi: 10.2135/cropsci2005.02-0085

Jung H-JG, Buxtono DR (1994) Forage quality variation among maize inbreds: Relationships of cell-wall composition and in-vitro degradability for stem internodes. J Sci Food Agric 66:313–322. doi: 10.1002/jsfa.2740660308

Jung H-JG, Ni W, Chapple CCS, Meyer K (1999) Impact of lignin composition on cell-wall degradability in anArabidopsis mutant. J Sci Food Agric 79:922–928. doi: 10.1002/(SICI)1097-0010(19990501)79:6<922::AID-JSFA307>3.0.CO;2-9

Juska F (1958) Some effects of gibberellic acid on turf grasses. Golf Course Report 25–28.

Kim G-T, Shoda K, Tsuge T, et al. (2002) The ANGUSTIFOLIA gene of Arabidopsis, a plant CtBP gene, regulates leaf-cell expansion, the arrangement of cortical microtubules in leaf cells and expression of a gene involved in cell-wall formation. EMBO J 21:1267–79. doi: 10.1093/emboj/21.6.1267

Lambers H, Nagel O, Arendonk J Van (1995) The control of biomass partitioning in plants from favourable''and stressful''environments: a role for gibberellins and cytokinins. Bulg J Plant Physiol 21:24–32.

Leplé J-C, Dauwe R, Morreel K, et al. (2007) Downregulation of cinnamoyl-coenzyme A reductase in poplar: multiple-level phenotyping reveals effects on cell wall polymer metabolism and structure. Plant Cell 19:3669–91. doi: 10.1105/tpc.107.054148

Lester DR, Ross JJ, Davies PJ, Reid JB (1997) Mendel's stem length gene (Le) encodes a gibberellin 3 beta-hydroxylase. Plant Cell 9:1435–43. doi: 10.1105/tpc.9.8.1435

Matos D a, Whitney IP, Harrington MJ, Hazen SP (2013) Cell Walls and the Developmental Anatomy of the Brachypodium distachyon Stem Internode. PLoS One 8:e80640. doi: 10.1371/journal.pone.0080640

Mauriat M, Moritz T (2009) Analyses of GA20ox- and GID1-over-expressing aspen suggest that gibberellins play two distinct roles in wood formation. Plant J

58:989–1003. doi: 10.1111/j.1365-313X.2009.03836.x

Mele G, Ori N, Sato Y, Hake S (2003) The knotted1-like homeobox gene BREVIPEDICELLUS regulates cell differentiation by modulating metabolic pathways. Genes Dev 2088–2093. doi: 10.1101/gad.1120003.2088

Morrison T, Jung H (1998) Cell-wall composition of maize internodes of varying maturity. Crop Sci.

Morrison T, Kessler J (1994) Activity of two lignin biosynthesis enzymes during development of a maize internode. J … 133–139.

Narita NN, Moore S, Horiguchi G, et al. (2004) Overexpression of a novel small peptide ROTUNDIFOLIA4 decreases cell proliferation and alters leaf shape in Arabidopsis thaliana. Plant J 38:699–713. doi: 10.1111/j.1365-313X.2004.02078.x

Nelissen H, Rymen B, Jikumaru Y (2012) A local maximum in gibberellin levels regulates maize leaf growth by spatial control of cell division. Curr Biol 1–5. doi: 10.1016/j.cub.2012.04.065

Norcia LN, Evans JD, Faulkner DN, Rohrbaugh LM (1964) Growth changes in pea plants induced by gibberellic acid and their effects on the concentrations of tissue lipids. Biochem J 90:633–7.

Oikawa T, Koshioka M, Kojima K, et al. (2004) A role of OsGA20ox1 , encoding an isoform of gibberellin 20-oxidase, for regulation of plant stature in rice. Plant Mol Biol 55:687–700. doi: 10.1007/s11103-004-1692-y

Paleg L, Kende H, Ninnemann H, Lang A (1965) Physiological effects of gibberellic acid. 8. Growth retardants on barley endosperm. Plant Physiol 40:165–9.

Pedersen JF, Vogel KP, Funnell DL (2005) Impact of Reduced Lignin on Plant Fitness. Crop Sci 45:812. doi: 10.2135/cropsci2004.0155

Pedersen M, Johansen KS, Meyer AS (2011) Low temperature lignocellulose pretreatment: effects and interactions of pretreatment pH are critical for maximizing enzymatic monosaccharide yields from wheat straw. Biotechnol Biofuels 4:11. doi: 10.1186/1754-6834-4-11

Peng J, Richards D, Hartley N (1999) “Green revolution”genes encode mutant gibberellin response modulators. Nature 400:256–61. doi: 10.1038/22307

Pichon M, Deswartes C, Gerentes D, et al. (2006) Variation in lignin and cell wall digestibility in caffeic acid O-methyltransferase down-regulated maize half-sib progenies in field experiments. Mol Breed 18:253–261. doi: 10.1007/s11032-006-9033-2

Poorter H, Niklas K, Reich P (2012) Biomass allocation to leaves, stems and roots: meta-analyses of interspecific variation and environmental control. New … 193:30–50. doi: 10.1111/j.1469-8137.2011.03952.x

Ramakers C, Ruijter JM, Deprez RHL, Moorman AF. (2003) Assumption-free analysis of quantitative real-time polymerase chain reaction (PCR) data.

Neurosci Lett 339:62–66. doi: 10.1016/S0304-3940(02)01423-4

Riboulet C, Guillaumie S, Méchin V (2009) Kinetics of phenylpropanoid gene expression in maize growing internodes: relationships with cell wall deposition. Crop … 211–223. doi: 10.2135/cropsci2008.03.0130

Riboulet C, Lefevre B, Denoue D, Barriere Y (2008) Genetic variation in maize cell wall for lignin content, lignin structure, p-hydroxycinnamic acid content, and digestibility in set of 19 lines at silage harvest maturity. Maydica 53:11–19.

Ridoutt BG, Pharis RP, Sands R (1996) Fibre length and gibberellins A1 and A20 are decreased in Eucalyptus globules by acylcyclohexanedione injected into the stem. Physiol Plant 96:559–566. doi: 10.1111/j.1399-3054.1996.tb00227.x

Roberts LW, Gahan PB, Aloni R (1988) Vascular Differentiation and Plant Growth Regulators. doi: 10.1007/978-3-642-73446-5

Robinson AR, Mansfield SD (2009) Rapid analysis of poplar lignin monomer composition by a streamlined thioacidolysis procedure and near-infrared reflectance-based prediction modeling. Plant J 58:706–14. doi: 10.1111/j.1365-313X.2009.03808.x

Rood SB, Buzzell RI, Major DJ, Pharis RP (1990) Gibberellins and Heterosis in Maize: Quantitative Relationships. Crop Sci 30:281. doi: 10.2135/cropsci1990.0011183X003000020008x

Ruegger M, Meyer K, Cusumano JC, Chapple C (1999) Regulation of ferulate-5-hydroxylase expression in Arabidopsis in the context of sinapate ester biosynthesis. Plant Physiol 119:101–10.

Sannigrahi P, Ragauskas AJ (2011) Characterization of Fermentation Residues from the Production of Bio-Ethanol from Lignocellulosic Feedstocks. J Biobased Mater Bioenergy 5:514–519. doi: 10.1166/jbmb.2011.1170

Sasaki A, Ashikari M, Ueguchi-Tanaka M, et al. (2002) Green revolution: a mutant gibberellin-synthesis gene in rice. Nature 416:701–2. doi: 10.1038/416701a

Schnable PS, Ware D, Fulton RS, et al. (2009) The B73 maize genome: complexity, diversity, and dynamics. Science 326:1112–5. doi: 10.1126/science.1178534

Setter TIML, Meller VH (1984) Reserve Carbohydrate in Maize Stem. Plant Physiol 75:617–622.

Slavov G, Allison G, Bosch M (2013) Advances in the genetic dissection of plant cell walls: tools and resources available in Miscanthus. Front Plant Sci 4:217. doi: 10.3389/fpls.2013.00217

Slewinski TL (2012) Non-structural carbohydrate partitioning in grass stems: a target to increase yield stability, stress tolerance, and biofuel production. J Exp Bot 63:4647–70. doi: 10.1093/jxb/ers124

Sponsel V (2010) Commercial uses of gibberellins. In: Taiz L, Zeiger E (eds) Plant Physiol., Fifth edit. Sinauer associates, Los Angeles, p 782

Srivastava LM (2002) Plant Growth and Development: Hormones and Environment, reprint. 772.

Tonini D, Hamelin L, Wenzel H, Astrup T (2012) Bioenergy production from perennial energy crops: a consequential LCA of 12 bioenergy scenarios including land use changes. Environ Sci Technol 46:13521–30. doi: 10.1021/es3024435

Tsuge T, Tsukaya H, Uchimiya H (1996) Two independent and polarized processes of cell elongation regulate leaf blade expansion in Arabidopsis thaliana (L.) Heynh. Development 122:1589–600.

Tsukaya H (2002a) Interpretation of mutants in leaf morphology: genetic evidence for a compensatory system in leaf morphogenesis that provides a new link between cell and organismal theories. Int Rev Cytol 217:1–39.

Tsukaya H (2002b) The leaf index: heteroblasty, natural variation, and the genetic control of polar processes of leaf expansion. Plant Cell Physiol 43:372–8.

Tsukaya H (2005) Leaf shape: genetic controls and environmental factors. Int J Dev Biol 49:547–55. doi: 10.1387/ijdb.041921ht

Tsukaya H (2006) Mechanism of leaf-shape determination. Annu Rev Plant Biol 57:477–96. doi: 10.1146/annurev.arplant.57.032905.105320

Ubeda-Tomas S, Bennett MJ (2010) Plant development: size matters, and it's all down to hormones. Curr Biol 20:R511–3. doi: 10.1016/j.cub.2010.05.013

Vallet C, Chabbert B (1996) Histochemistry of lignin deposition during sclerenchyma differentiation in alfalfa stems. Ann … 625–632.

Van Acker R, Vanholme R, Storme V, et al. (2013) Lignin biosynthesis perturbations affect secondary cell wall composition and saccharification yield in Arabidopsis thaliana. Biotechnol Biofuels 6:46. doi: 10.1186/1754-6834-6-46

Van der Weijde T, Alvim Kamei CL, Torres AF, et al. (2013) The potential of C4 grasses for cellulosic biofuel production. Front Plant Sci 4:107. doi: 10.3389/fpls.2013.00107

Van Hulle S, Roldán-ruiz I, Bockstaele E Van, Muylle H (2010) Comparison of Different Low-Input Lignocellulosic Crops as Feedstock for Bio-ethanol Production. In: Huyghe C (ed) Sustain. use Genet. Divers. Forage Turf Breed. Springer Netherlands, Dordrecht, pp 365–368

Vanholme R, Morreel K, Darrah C, et al. (2012) Metabolic engineering of novel lignin in biomass crops. New Phytol 196:978–1000. doi: 10.1111/j.1469-8137.2012.04337.x

Vermerris W (2011) Survey of genomics approaches to improve bioenergy traits in maize, sorghum and sugarcane. J Integr Plant Biol 53:105–19. doi: 10.1111/j.1744-7909.2010.01020.x

Vermerris W, Saballos A, Ejeta G, et al. (2007) Molecular breeding to enhance ethanol production from corn and sorghum stover. Crop … 47:S–142. doi: 10.2135/cropsci2007.04.0013IPBS

Vogel K, Jung H (2001) Genetic modification of herbaceous plants for feed and fuel. CRC Crit Rev Plant Sci 37–41.

Vogt T (2010) Phenylpropanoid biosynthesis. Mol Plant 3:2–20. doi: 10.1093/mp/ssp106

Waclawovsky AJ, Sato PM, Lembke CG, et al. (2010) Sugarcane for bioenergy production: an assessment of yield and regulation of sucrose content. Plant Biotechnol J 8:263–76. doi: 10.1111/j.1467-7652.2009.00491.x

Wang D, Naidu SL, Portis AR, et al. (2008) Can the cold tolerance of C4 photosynthesis in Miscanthus x giganteus relative to Zea mays be explained by differences in activities and thermal properties of Rubisco? J Exp Bot 59:1779–87. doi: 10.1093/jxb/ern074

Wang Z-Y, Brummer EC (2012) Is genetic engineering ever going to take off in forage, turf and bioenergy crop breeding? Ann Bot 110:1317–25. doi: 10.1093/aob/mcs027

Wolters H, Jürgens G (2009) Survival of the flexible: hormonal growth control and adaptation in plant development. Nat Rev Genet 10:305–17. doi: 10.1038/nrg2558

Yang F, Mitra P, Zhang L, et al. (2013) Engineering secondary cell wall deposition in plants. Plant Biotechnol J 11:325–35. doi: 10.1111/pbi.12016

Yao Q, Cong L, Chang JL, et al. (2006) Low copy number gene transfer and stable expression in a commercial wheat cultivar via particle bombardment. J Exp Bot 57:3737–46. doi: 10.1093/jxb/erl145

Yue F, Lu F, Sun R-C, Ralph J (2012) Syntheses of lignin-derived thioacidolysis monomers and their uses as quantitation standards. J Agric Food Chem 60:922–8. doi: 10.1021/jf204481x

Zabala G, Zou J, Tuteja J, et al. (2006) Transcriptome changes in the phenylpropanoid pathway of Glycine max in response to Pseudomonas syringae infection. BMC Plant Biol 6:26. doi: 10.1186/1471-2229-6-26

Zawaski C, Busov V (2014) Roles of Gibberellin Catabolism and Signaling in Growth and Physiological Response to Drought and Short-Day Photoperiods in Populus Trees. PLoS One 9:e86217. doi: 10.1371/journal.pone.0086217

Zeng Y, Zhao S, Yang S, Ding S-Y (2014) Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. Curr Opin Biotechnol 27:38–45. doi: 10.1016/j.copbio.2013.09.008

Zhu X-G, Long SP, Ort DR (2008) What is the maximum efficiency with which photosynthesis can convert solar energy into biomass? Curr Opin Biotechnol 19:153–9. doi: 10.1016/j.copbio.2008.02.004

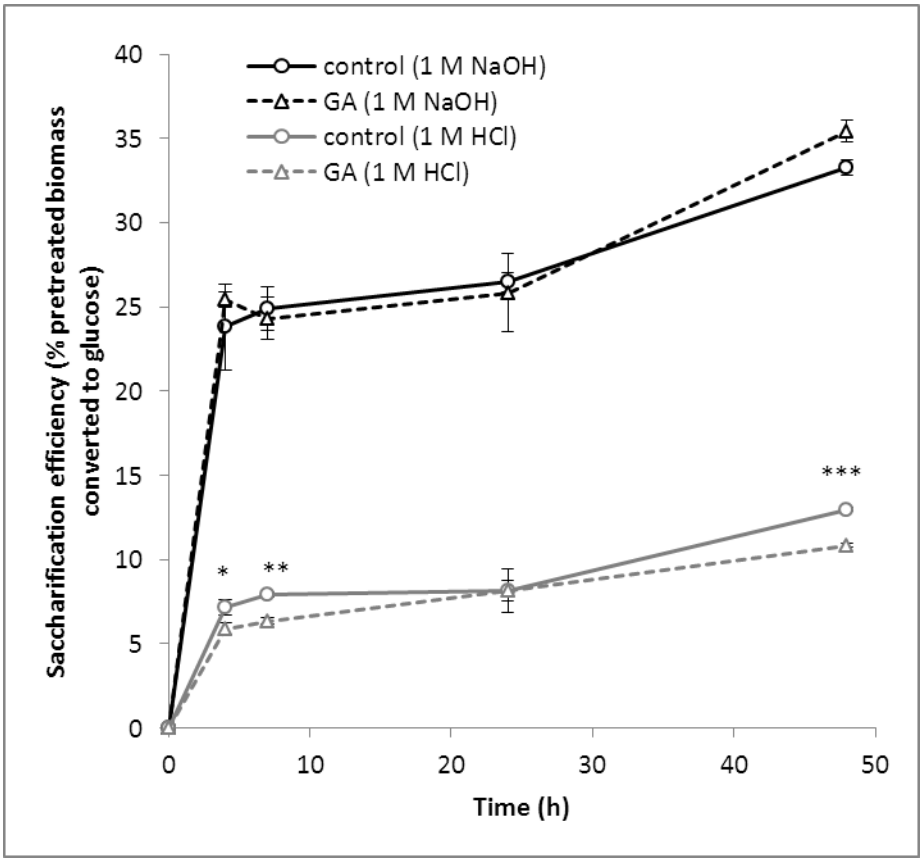
8. Supplemental figures and tables

Supplementary Table 14. Selection of genes for qPCR based upon expression profile described in Riboulet et al., 2009

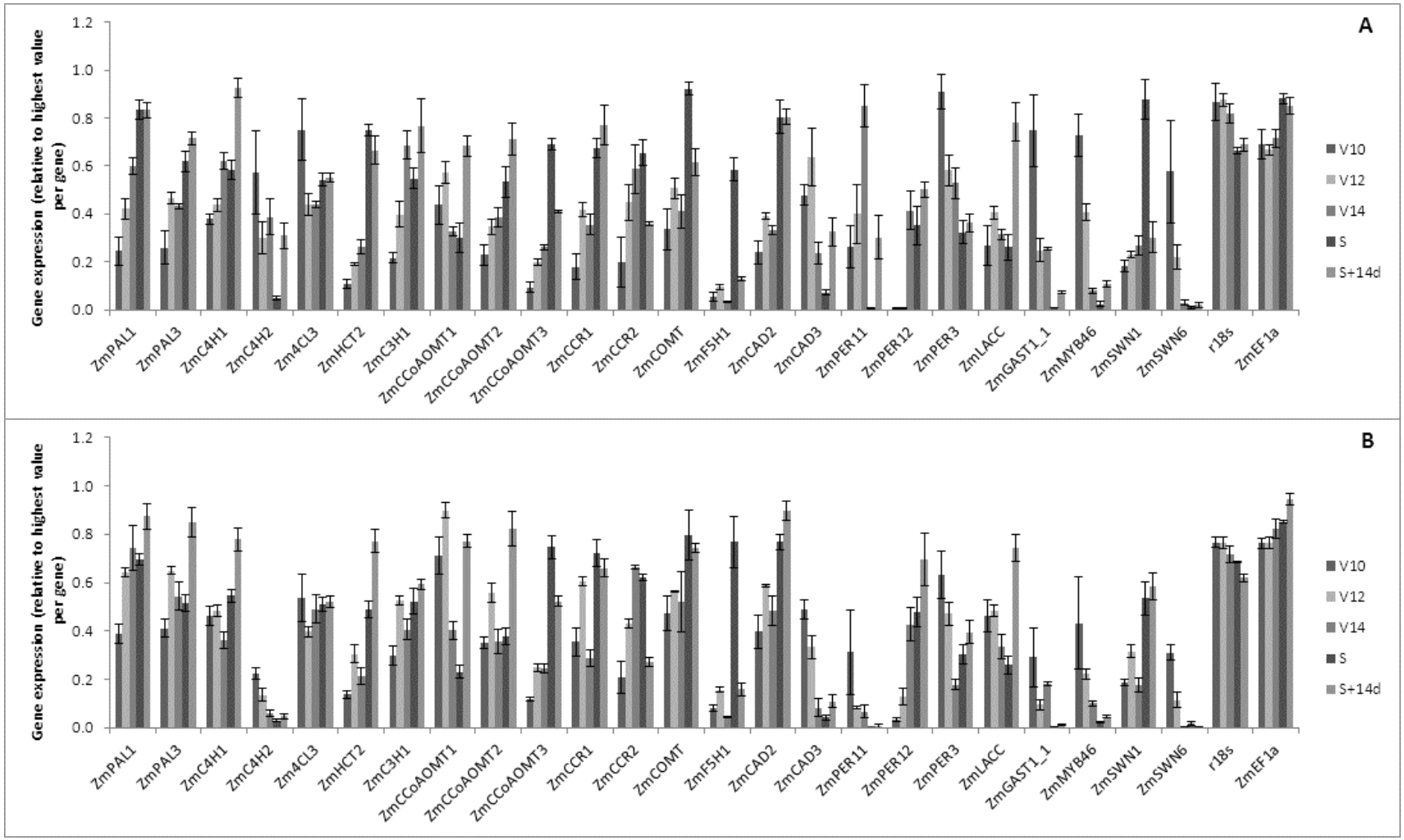
gene	Plaza 2.5 code / contig nr / GRMZM code	Expression pattern (Riboulet et al., 2009)	primer FW	primer RV	amplicon size
ZmPAL1	Hidden	Mature internode	Hidden	Hidden	171
ZmPAL3		Young stems			192
ZmC4H1		Mature internode			140
ZmC4H2		Mature internode			153
Zm4CL3		Mature internode			229
ZmHCT2		Mature internode			247
ZmC3H1		Mature internode			242
ZmCCoAOMT1		Mature internode			185
ZmCCoAOMT2		Mature internode			250
ZmCCoAOMT3		Mature internode/young stems			178
ZmCCR1		Mature internode/young stems			167
ZmCCR2		Young stems			140
ZmCOMT		Mature internode			250
ZmF5H1_A		Mature internode/young stems/leaves			160
ZmCAD2		Mature internode/young stems			223
ZmCAD3		leaves			193
ZmPER1		Mature internode			238
ZmPER2		Mature internode			223
ZmPER3		Young stems			209
ZmLAC		Mature internode			224
ZmGAST1_1		-			206
ZmMYB46		-			190
ZmSWN1		-			157
ZmSWN6		-			242
ZmEF1a		-			107
Zm18S		-			118

Supplementary Table 15. List of specific prominent ions used to extract the ion-specific chromatograms and quantify the different lignin units, released during thioacidolysis.

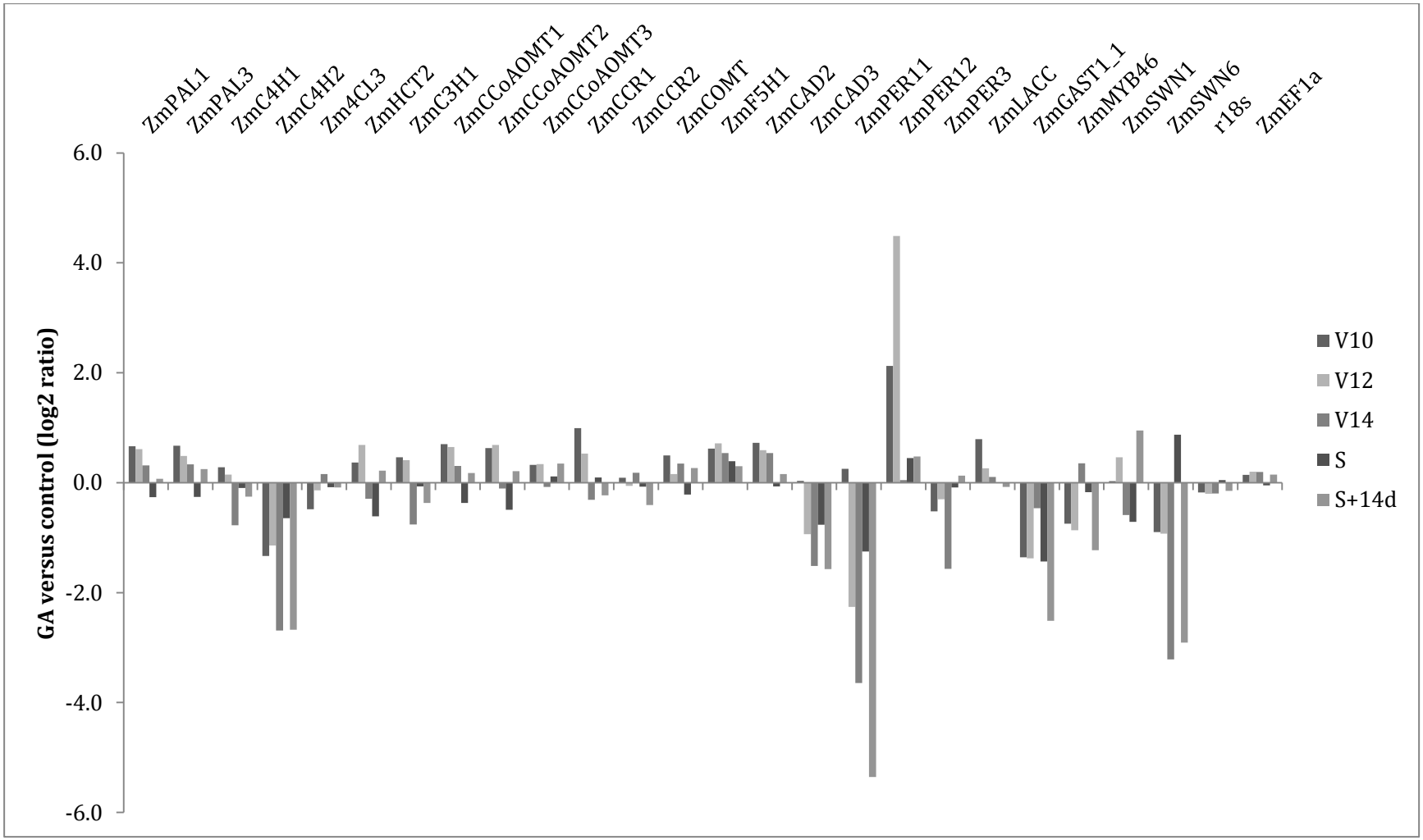
Compound	Target ion	qualifier 1	qualifier 2	qualifier 3
H	239	205	179	-
G	269	235	209	418
S	299	265	239	448



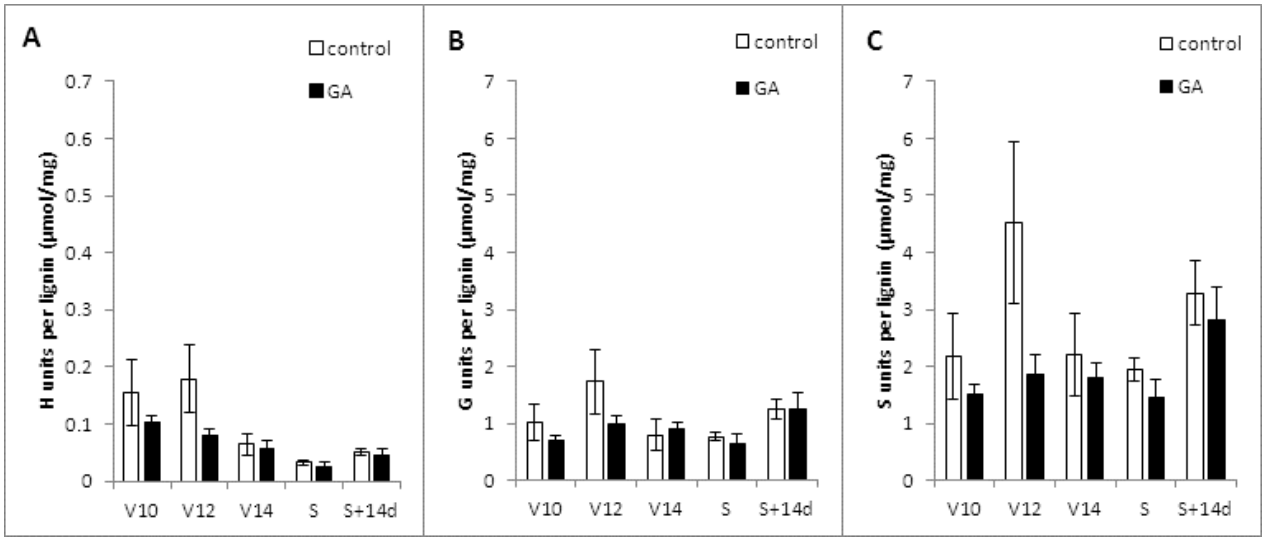
Supplementary figure 12. Saccharification efficiency of ground stem material of non-transgenic and transgenic plants with 1 M HCl and 1 M NaOH pretreatment, expressed as glucose yield per dry weight (DW). Error bars represent standard errors of three biological replicates. *: p<0.05, **: p<0.01, ***:p<0.001.



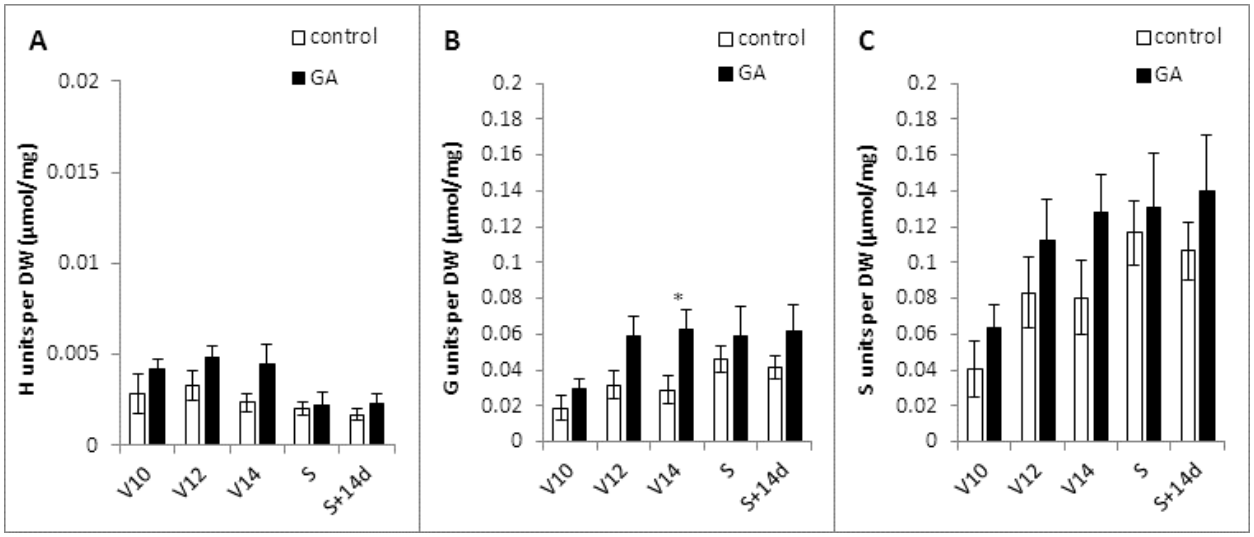
Supplementary figure 13. Expression profile of lignin biosynthetic genes in the developing maize ninth internode of control (A) and GA overproducing (B) plants. V10, V12, V14 (10, 12 and 14 leaf collars visible respectively), S (silking) and S+14d (14d after silking) indicate the different developmental stages; Error bars represent standard errors of four biological replicates.



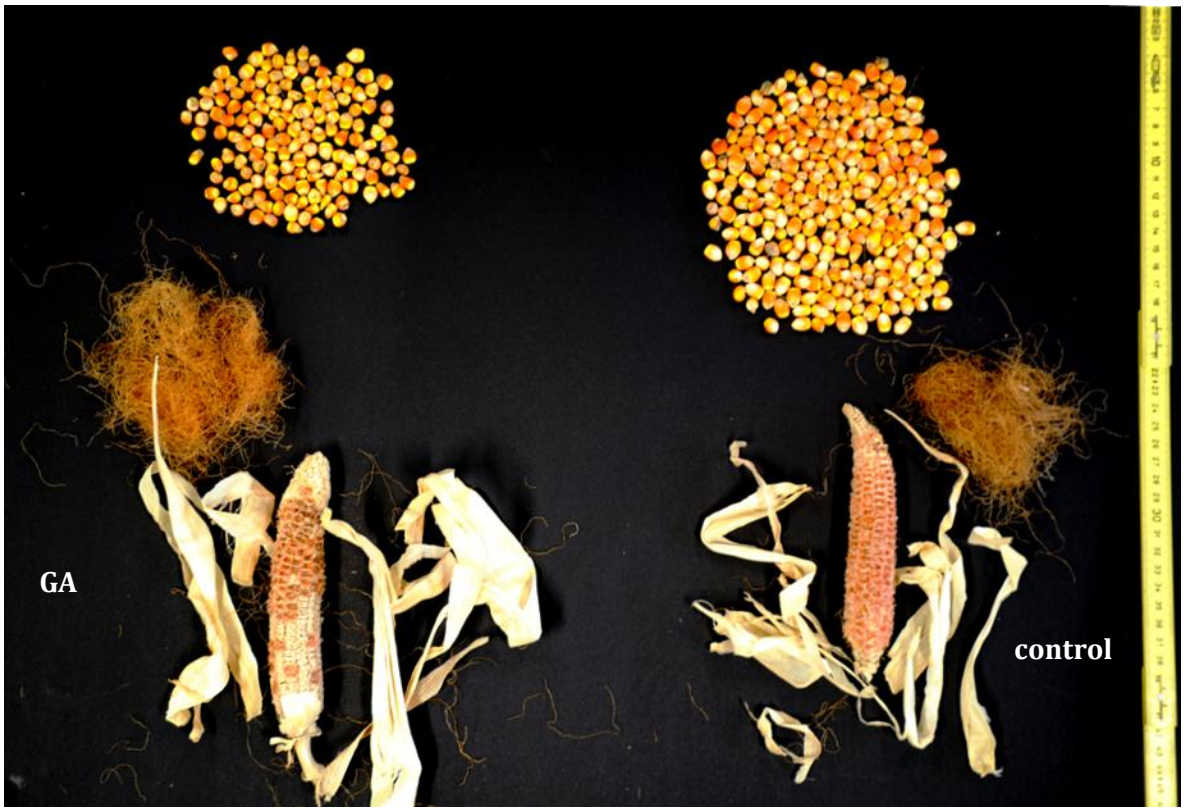
Supplementary figure 14. Lignin biosynthetic gene expression of GA overproducing plants versus control plants in the developing maize ninth internode expressed as log2 ratios. V10, V12, V14 (10, 12 and 14 leaf collars visible respectively), S (silking) and S+14d (14d after silking) indicate the different developmental stages.



Supplementary figure 15. Lignin composition in the ninth internode of control and GA overproducing (GA) plants over development, expressed as abundance of H, G and S units per lignin amount (A,B and C respectively). V10, V12, V14 (10, 12 and 14 leaf collars visible respectively), S (silking) and S+14d (14d after silking) indicate the different developmental stages; Error bars represent standard errors of 5 biological replicates.



Supplementary figure 16. Lignin composition in the ninth internode of control and GA overproducing (GA) plants over development, expressed as abundance of H, G and S units per dry weight (DW)(A,B and C respectively). V10, V12, V14 (10, 12 and 14 leaf collars visible respectively), S (silking) and S+14d (14d after silking) indicate the different developmental stages; Error bars represent standard errors of 5 biological replicates. *:p<0.05.



Supplementary figure 17. Cobs and grain yield of representative GA overproducing (left) and control (right) plants.

Chapter 7: Gibberellic acid overproduction as strategy for improved maize bioenergy feedstock

Chapter 8: General conclusion and perspectives

Chapter8: General conclusions and perspectives

The production of bioethanol from the non-edible part of the plant (second generation bioethanol) is currently economically unattractive due to the high recalcitrance of the lignocellulosic material to enzymatic hydrolysis. Therefore, the biomass production and the degradation efficiency of the biomass into fermentable sugars needs major improvement.

In this PhD research we focused on enhancing biomass production by application of yield enhancing genes and lowering the need for pretreatment by lignin engineering. Fundamental in these fields has been obtained in the dicot model system *Arabidopsis thaliana* (Arabidopsis). Yet, high yielding energy crops such as Miscanthus, switchgrass, Sorghum, sugarcane and maize are all monocot species. We therefore applied a translational approach from Arabidopsis to the monocot crop maize and the small grass model *Brachypodium distachyon* (Brachypodium).

1. Transgenic approach to study the effect of IYGs and lignin perturbation in Brachypodium and maize

Overexpression of the *AtGA20ox1* gene in Brachypodium did not result in increased leaf length. This was however unexpected since overexpression of the same gene in maize did result in increased leaf length and plant height. Although there is no conclusive evidence, the main reason could be that overexpression was too weak to cause a visible phenotype. Another transgenic approach, the downregulation of *BdCAD1* resulted in only moderate reduction of *BdCAD1* transcript abundance. Nevertheless, a correlation was found for hairpin expression, *BdCAD1* endogene downregulation and CAD enzyme activity reduction in one of the seven investigated lines. Although the phenotype of this line was not investigated further, even if present, a strong phenotype was not expected based on the level of downregulation. A similar transgenic approach was undertaken in maize, with the downregulation of the maize *ZmCAD2* gene. Also here, we could confirm the expression of the hairpin construct and the downregulation of *ZmCAD2* in stems of transgenic plants. Three lines with varying levels of mild *ZmCAD2* downregulation were characterized for growth rates, biomass accumulation, lignin quantity and composition and saccharification efficiency. Yet, no phenotype consistent with what was described for *ZmCAD2* downregulation in maize previously was observed. All investigated Brachypodium and maize transgenic lines expressed the transgene but seemingly at levels that were only moderate. At least this is the case for the hairpin constructs, since endogene transcript levels were still largely abundant. In all these constructs, the UBIL promoter was used to drive transgene expression. Meanwhile, it was shown that pUBIL is of only limited strength compared to promoter sequences isolated from the Brachypodium *BdEF1α* and *BdUBI10* genes. The difficulties in detecting phenotypic changes in the transgenic lines might thus be attributed to the limited level of overexpression by the UBIL promoter. The characterization of the two new Brachypodium promoters will be of great advantage for constitutive high overexpression of a transgene. In addition, there is still optimization of the constructs for monocot transformation possible. For example, specific intron sequences could be used that were reported to be effective in

boosting transgene expression levels when inserted at the 5' end of the transgene sequence. A thorough investigation of these effects has, to my knowledge, not been carried out in monocot plants. Furthermore, besides the newly investigated Brachypodium *pBdGLU1* endosperm-specific promoter, several more tissue-specific promoters have been described in literature but their availability for the research community is highly limited. Another limitation of monocot transformation at PSB to date is the choice of selectable markers. For maize, basta selection is successful for the efficient generation of transgenic plants. However, future strategies for combining favorable traits in maize will require gene stacking. The analysis of transgenic lines with multiple transformed constructs would greatly benefit from the use of different selectable markers for the individual transformation vectors. In addition, the use of a selectable marker that does not confer herbicide or antibiotic resistance, such as *phosphomannnose isomerase* (*PMI*) would potentially reduce the recalcitrance towards allowing transgenic plants in field trials or commercialization of a new genetically modified crop. With a plethora of these monocot transformation vectors in house, PSB could explore the use of gene stacking for bioenergy crop improvement more effectively and explore a series of more recent targeted approaches for gene silencing such as artificial miRNA, TALEN, Zinc Finger or CRISPR/Cas9 technologies in maize.

2. Using a TILLING population in Brachypodium to study lignin perturbation

As alternative for gene silencing by genetic transformation, specific mutants for lignin biosynthetic genes were investigated. In Brachypodium, mutants in the *BdCAD1* gene were identified by their *brown stem* phenotype in the BRACHYTILL population. These mutants displayed a 45% increase in saccharification efficiency compared to wildtype plants. Interestingly, the expression of CAD family genes was upregulated in the mutant, pointing towards an auto-regulatory mechanism for CAD gene expression, similar to what was claimed for the 4CL family in Sorghum previously. The same BRACHYTILL collection was screened for the presence of mutations in *BRADI3G05750*, which was assigned the name *Bd4CL1* in this report. Despite the total of 17 mutations, identified in a 735 kb region in the first exon of *Bd4CL1*, we were unable to detect a premature stop codon. Mutations that were predicted to affect protein function were analyzed for improved saccharification efficiency but no significant improvements were detected. 4CL protein activities were not measured so residual 4CL activity might have been present and sufficient for its enzymatic function. Alternatively, there might be functional redundancy playing a role in the 4CL gene family.

3. A systems-wide analysis of lignification and perturbation of lignin biosynthesis in maize

In maize, the effect of two genetic perturbations of genes in the lignin biosynthetic pathway on forage quality parameters, lignin composition and saccharification efficiency was investigated.

The perturbation of *ZmC4H1* in maize resulted in a decrease of the lignin fraction, without altering total amount of cell wall content and the level of cellulose. Interestingly, the reduction in lignin content was compensated for by an increase in hemicellulose content in *zmc4h1* plants. Similarly, perturbation of *ZmCAD2* resulted in decreased lignin content. Discordantly, the amount of cellulose was decreased and the amount of hemicellulose remained unaltered in *zmcad2* plants.

The lignin composition was altered in both *zmc4h1* as *zmcad2* plants. However, as the S/G ratio was increased in *zmc4h1* plants, it was decreased in *zmcad2* plants. The *in vitro* digestibility estimation by NIRS and the saccharification efficiency of stem biomass were improved in *zmc4h1* and *zmcad2* plants compared to their respective controls. As the saccharification assay, performed on stem biomass of *zmcad2* plants was improved, both with and without 1 M HCl pretreatment, *zmc4h1* plants had a higher sugar release only using the 1 M HCl pretreatment. Without the acid pretreatment we could not observe a significant increase in glucose release for *zmc4h1* plants. These two different genetic perturbations in lignin biosynthesis had thus different responses in levels of cellulose, hemicellulose and lignin as well as lignin composition and saccharification efficiency without pretreatment. It is however also interesting to perform additional pretreatments that have a different mode of action. For instance, pretreatment with alkali (1 M NaOH or ammonia) will improve cellulose accessibility by cleaving carbohydrate-lignin linkages and partial removal of lignin whereas acid pretreatment will act by hydrolyzing hemicellulose while leaving the lignin intact. Biomass from *CAD*-deficient plants is known to yield a high saccharification yield using alkaline pretreatment. Besides these chemical methods, also steam explosion, a frequently applied pretreatment using only water and high temperatures, could be tested.

In previous studies, both increased and decreased S/G ratio have been linked to enhanced extractability of cellulose from its complex matrix. However, the altered S/G ratio would reflect other changes in cell wall structure, rather than only altered levels of S and G units. A recent study of cell wall content in a series of Arabidopsis lignin mutants has evidenced that a high S/G ratio has a negative effect on saccharification carried out without pretreatment but had a positive effect on saccharification efficiency with acid pretreatment. This remarkable correlation suggest that cell walls with a high S/G ratio form a matrix in which the matrix polysaccharides (hemicelluloses) render the cellulose less accessible by cellulases. The results that were obtained in this study for the *zmc4h1* maize plants are in line with this hypothesis. This opens perspectives for further investigation of the underlying factors of this altered cell wall matrix in lignin mutants.

It is poorly understood why perturbation of lignin biosynthesis sometimes, but not always, results into a yield penalty. Although plant growth and biomass yield was not monitored in the maize *zmc4h1* and *zmcad2* mutants, *zmc4h1* plants did have a delay in development of one week in field conditions, apparent from the moment of silking. The *zmcad2* plants that were compared with control plants in four developmental stages showed differential expression of genes that are known to be involved in plant defense and response to biotic and abiotic stress. In literature, there are currently two hypotheses for the cause of the stress response in lignin mutants. The first hypothesis is that the accumulation of specific substrates, caused by perturbation of an enzymatic step in lignin biosynthesis and consequent detoxification routes, triggers a signaling cascade that leads to a stress response. The second hypothesis assumes the presence of a sensing mechanism for cell wall integrity. It is suggested that the continuous leaching of carbohydrates from the impaired cell wall as a result of the reduction in the amount of lignin would mimic wounding and lead towards the stress response. Either way, understanding why lignin mutants display this stress response might be key to avoid pleiotropic effects of lignin perturbation in the future.

4. Monitoring grass leaf growth and pleiotropic effects of yield enhancement

From previous studies, we learned that lignin perturbation is likely to cause pleiotropic effects. And vice versa, yield enhancement strategies can have pleiotropic effects on cell wall properties. In order to be able to closely monitor growth potential during the phenotyping of the transgenic plants generated during this PhD, we developed a tool that is readily employed for the analysis of leaf growth measurements. This tool, that we called LEAF-E, is an MS Excel-based function fitting program which is highly user-friendly. This is in contrast to alternative fitting procedures that are embedded in statistical software packages with troublesome extraction of the desired data from the output. This tool has helped in a correct and robust way of interpreting biologically relevant data, such as the maximal leaf elongation rate, from hand-made measurements. Additionally, the procedure was automated in an MS Excel macro making LEAF-E an excellent tool for comparing leaf growth behavior in different genotypes or to analyze the response of specific genotypes to a treatment. Moreover, the method is probably applicable for most C3 and C4 monocot species. We thus anticipate that this method could be used for leaf and plant growth analysis for various experimental setups in this department and beyond.

We investigated the pleiotropic effects of enhanced growth rates on biomass composition and saccharification efficiency using maize plants overexpressing the *AtGA20ox1* gene. These plants were previously characterized with higher growth rates, longer leaves and taller stems. Most interestingly, the biomass yield of the stem fraction was significantly increased in GA overproducing plants. As the stem forms the major source of cellulose for conversion into bioethanol, this is a beneficial characteristic of GA overproduction in maize plants. On the other hand, grain weight was significantly decreased due to a lower number of kernels per cob. A low seed yield is naturally an undesired feature of a crop, especially if these plants would be used for dual-purpose: the grain for food and the non-edible plant parts for bio-ethanol production. These observations were in accordance with altered sink-source relationships in GA overproducing plants as compared to the control plants. It is however recommended to determine the agricultural characteristics in a large scale field trial. When examining the stem biomass for biomass quality traits relevant to saccharification efficiency, we found that stems of GA overproducing pants have more and/or thicker cell walls than wildtype with thus potentially more cellulose to be converted into bioethanol. However, we found that stem material of GA overproducing plants is more recalcitrant to chemical pretreatment and displays a reduced saccharification efficiency when pretreated with acid, but not with alkali. Recalcitrance to both acid pretreatment and enzymatic digestion is directly proportional to lignin content. This characterization of the stem biomass undeniably proved that enhancing GA levels in maize to improve stem biomass quantity can also influence cell wall properties. The fact that biomass quantity and quality are interconnected is important for developing strategies to improve lignocellulosic feedstock for bioethanol production. This also implies that this interconnection can be exploited. For example, crops with delayed or suppressed flowering produce more vegetative biomass and have more easily degradable biomass. Targeting flowering genes in grasses could therefore form a good strategy for bioenergy feedstock improvement.

5. Brachypodium and maize as model for bio-energy research: a comparison

As in this PhD both Brachypodium and maize have been evaluated for their potential function as a model system for bioenergy crop improvement, this work forms an ideal foundation for comparison. For this comparison, I listed some key features of using Brachypodium and maize as model system.

Based on their phylogeny, maize is more closely related to the high yielding grasses *Miscanthus* spp., Sorghum and sugarcane whereas Brachypodium is more closely related to the temperate cereals wheat and barley and forage grasses such as ryegrass. As cellulosic ethanol can be produced from plant biomass from either phylogenetic group (wheat straw, forage grass biomass, sugarcane bagasse, maize stover and dedicated bioenergy grasses such as Miscanthus), both Brachypodium and maize could thus form relevant model systems for bioenergy crop improvement.

There are however major differences in the level of establishment as a model for biological researchers between these two species. Maize, with a huge economic value, has been a genetic model for over a century with a huge research community as a result. Brachypodium on the other hand, was introduced as a completely new model plant in 2001 with a growing, but limited research community. A significant advantage of the wild grass Brachypodium over maize is its size and easy growth requirements, which is why it can be introduced as alternative model in labs working on Arabidopsis without major changes in infrastructure. Another advantage is the high quality of its genome sequence, which is very useful for comparative genomics approaches and natural diversity.

The cultivation of maize, either in the greenhouse or outside on the field, requires skill and significant investment in infrastructure. The maize genome sequence is of poorer quality than Brachypodium, mainly because of its size (2.3Gb versus Brachypodium: 272Mb). It is however the lack of consistency of the gene nomenclature that forms the biggest downside of maize genomics, to my opinion. NCBI EST-sequence names, GRMZM, AC, AY and other codes of two different versions of the genome annotation (Refseq-v1 and v2 or AGP_v1 and v2 from “A Golden Path”) and unrelated contig numbers from a macro-array are all being used in maize literature. Their inter-relation can often only be determined by BLAST searches. And recently, a third version of the maize B73 genome assembly (Refgen_v3 or AGP_v3) has been released. However, this version is not yet integrated in all online tools for BLAST, comparative genomics and mutant accession databases. So, significant improvement can be made in this area.

Taken together, after four years of working with Brachypodium and maize, I was able to evaluate both species as models for bioenergy crop improvement. In my opinion, it is the presence of the transformation facility at PSB, the single stem anatomy and definition of standardized developmental stages, the availability of a fast and easy genotyping method (leaf painting), the value as a crop species and the close phylogenetic relationship with other important C4 bioenergy crops that makes maize the preferred model system for bioenergy crop improvement.

6. Maize as model for systems biology in monocots

A complex biological process such as organ growth regulation or lignification of the secondary cell wall is best studied using a combination of data at the transcript, protein, metabolite and phenotype

level. Approaches that integrate these different types of data and attempt to build a predictive model are termed as 'systems biology' and requires substantial knowledge of the molecular components of the system (e.g. genes, proteins and metabolites) and tools developed to monitor these components. Therefore, systems biology is most suitable for model systems and model organisms. For photosynthesis and spatial regulation of cell division, the developmental gradient in the maize leaf has served as a useful model for a systems biology approach. For lignin biosynthesis, studies using combinations of transcript, metabolome and phenotypic data was, to my current knowledge, limited to the model plant Arabidopsis, which led to insight into the metabolic network of lignification and to the discovery of new genes with a role in lignification, such as *caffeoyl shikimate esterase* (*CSE*).

With the current progress in metabolic profiling of the maize stem, partly discussed in this work, I believe that also maize can form a powerful model for lignification in monocot species. Nevertheless, the current number of available mutants and transgenic lines with perturbations in lignin biosynthesis, optimally in a common genetic background, is insufficient to allow wide scale analysis of the lignification toolbox as was done for Arabidopsis. For this, the generation of maize transgenic lines with lignin perturbations should become a priority.

Finally, single-gene based approaches most likely do not administer the major increase in biomass production and saccharification efficiency as would be necessary to make bioethanol production from lignocellulosic biomass economically profitable. Ultimately, combinations of improved plant growth traits with combinations of improved biomass degradation traits by gene stacking could form the basis of the next generation of bioenergy crops.

Acknowledgements

I dedicate this thesis, the result of more than four years of work as a PhD student, to my wife Liesje and my son Mauro.

Liesje, you are the love of my life. Sometimes I wonder why, but you also love me. I can proudly say that it is thanks to you that I was able to finish this PhD trajectory. You were with me from the start, the IWT defense in November 2009, all the way until now, May 2014, at the public defense of this PhD. You listened patiently to my complaints, you stimulated me to go abroad for meetings and with me, you celebrated the moments of success and you endured the feelings of stress, desperation and frustration. I want you to know that this thesis and the accompanying title of PhD in sciences is as much yours as it is mine.

Mauro, thanks for being born! You are an amazing son. Already at your young age of 1 year old you point at the grass, vegetables, flowers and trees with excitement, an excitement that you share with your father. When I was locked inside the house for a long time while writing this thesis, spending this time with you motivated me to keep writing and correcting until the end.

Furthermore, there is a lot of other people I would like to thank. Forgive me for not all mentioning you by name and instead thanking you as groups but know that my gratitude goes to every one of you. I would like to thank my family: my mother and father, brothers, aunts and uncles, parents in law, brothers and sisters in law and aunts and uncles in law for every time acting interested in my work but most of all for the support and thoughtful motivating words in times with less encouraging results at the working place.

I want to thank my promoter Hilde Muylle for the never ending stream of confidence in my abilities, even when I must have looked pretty clumsy when even simple PCRs started to fail. I remember your excellent interrogation technique that you administered for tackling scientific issues, which mostly leaded to me providing an outcome for my own problem. Your positive attitude, open-mindedness and encouraging mentality to cooperate with external partners are legendary. I cannot think of a better person for guidance of my time as a PhD student.

I want to thank my promoter Prof. Dirk Inzé for providing the opportunity to start and finish this PhD, to be part of your lab with amazingly good scientists and providing the necessary guidance.

I thank Isabel Roldán-Ruiz for teaching me how to pay attention to detail while writing manuscripts and the discussions about the scientific work.

I thank all of my colleagues from ILVO and PSB for the numerous moments of great help, scientific discussions, coffee breaks, celebrations, laughs and joy while working and afterwards and in between as well. You all know that I enjoyed my time with you a lot.

I thank the members of the examination committee, which were in part also members of my PhD guidance committee for reading and correcting the manuscript and providing scientific comments and adjusting the scientific course of the PhD when needed. In particular I would like to thank Wout Boerjan for the confidence that he put in me by making me responsible for several scientific tasks

and communication to the project members in the European Seventh Framework Program named SUNLIBB. I am really grateful for this opportunity to be among the best scientists in our field, contributing to scientific discussions, giving presentations, traveling to exotic places (some more than others) and making new friends. In addition, I am really grateful for the position that I received in your outstanding group.

I also want to thank Richard Sibout from INRA Versailles for the nice discussions and the productive collaboration with your group.

I thank my friends from Volleybal, school and my (now ex-) music band Skatchou Bottos for providing a huge amount of distraction and practice of my social skills.

I have always had a fascination for this wonderful thing called nature and especially in plants and this in combination with an almost compulsive behavior of trying to find out how things work. But if it wasn't for the people around me that shared/liked/encouraged/inspired/instigated this excitement, I wouldn't have been able to follow my passion and become a plant scientist. To end, I would like to thank some of those people for this personally. First of all my father for teaching me plant names from when I was a kid and for sharing the same interests in biology as me, I am sure that there are several genetic factors for this that are worthwhile looking into. Also my excellent teachers of sciences in high school and Hogeschool Gent, especially my teacher biotechnology Katrien Strubbe for her inspiring classes. I really want to thank Liesbeth Vercruyssen for her excellent guidance of my Master thesis at PSB by showing me how a first-rate scientist is expected to work and for the fun times during the numerous and notorious PSB parties.

Curriculum vitae

Wannes Voorend

Address: Beverestraat 24, 9700 Oudenaarde
Tel: +32478576538
E-mail: wannesti@hotmail.com
Date/place of Birth: 12/01/1984; Zottegem

1.1 Education

PhD in Science: Biotechnology, University of Ghent, 2009-current

Title: Enhancing biomass production and accessibility of the cell wall for fermentation in *Brachypodium distachyon* as a model in *Zea mays* as a crop

Funded by IWT
Supervisors: Prof. Dirk Inzé (VIB) and Dr. Hilde Muylle (ILVO)
Cosupervisors: Prof. Wout Boerjan (VIB), Mieke Van Lijsebettens (VIB), Hilde Nelissen (VIB) and Isabel Roldán-ruiz (ILVO)

Master in Applied Sciences: biochemistry, biotechnology (Hogeschool Gent), July 2009

Thesis: Study of transgenic *Arabidopsis thaliana* plants with enhanced root growth and leaf size using phenotypic analysis, qPCR and GFP-fusion proteins (under supervision of Dirk Inzé and Liesbeth Vercruyssen, Systems Biology of Yield lab, VIB Plant systems biology, Technologiepark, Zwijnaarde)

Bridging program from Bachelor to Master in Applied Sciences: biochemistry, biotechnology, July 2008

Bachelor in Chemistry – Biochemistry, biotechnology, July 2007

Thesis: Genetic diversity, mating type, detection and host plant susceptibility of *Phytophthora ramorum* (under supervision of Annelies Vercauteren and Kurt Heungens, Crop protection, Plant unit, ILVO, Burg. van Gansberghelaan, Merelbeke)

Wetenschappen-wiskunde, Atheneum Wispelberg and De Wingerd, Gent, July 2002

1.2 Skills

General skills in research project management and data analysis. Specific expertise and interests in:

Language Skills:

- Dutch: Native
- English: fluent
- French: Good

Teaching Skills:

- 3 Scientists@Work sessions to highschool students

- Supervision of the research projects of 1 master 2 student, 1 master 1 student and 1 professional bachelor student

Experiment design

During my PhD I developed profound experience in setting up experiments for phenotypic and molecular characterization of transgenic/mutated lines. This includes early and late stage genotyping, expression analysis, leaf growth analysis, biomass measurements, protein activity measurements and cell wall characterization (lignin, cellulose, glucose release upon enzymatic hydrolysis, etc.) for greenhouse and growth chamber experiments with up to 200 plants for maize and up to 400 plants for *Brachypodium*.

Data analysis

- Expression data from qPCR and micro-array
- Metabolomics data from GC-MS and LC-MS^n
- Leaf growth data with non-linear regression

1.3 Other skills

- Statistical software: basic knowledge of Statistica and SPSS
- Bioinformatics software: experience with CLC genomics workbench, Ensembl Plants, VectorNTI
- Metabolic analysis software: Mzmine, Masslinx
- Plant transformation: Evaluating and maintaining an *Agrobacterium*-mediated transformation platform for *Brachypodium* starting from callus derived from immature embryos
- Writing and presenting reports
- Former trainer and Volleyball coach and Windsurfing instructor
- Musical skills: Saxophone
- driving license

1.4 Interests

I love playing Volleyball indoor as well as outdoor (Beachvolleyball) and play both sports at a reasonably high level (indoor 1st division, beachvolleybal National circuit). Until 2012 I played the Saxophone in a band named Skatchou Bottos and performed regularly and together produced an EP and a full CD. Until 2010 I did windsurfing as a hobby, was an active member of the club board (Absolut windsurfing, Oudenaarde) and taught windsurfing to kids and adults for 8 summers in a row.

1.5 References

Professor Dirk Inzé (PhD Supervisor)
Department of Plant Systems Biology
Ghent University

Technologiepark 927, 9052 Gent, BELGIUM

Dr. Hilde Muylle (PhD Supervisor)
Plant Science Unit
Institute for Agricultural and Fisheries Research (ILVO)
Caritasstraat 21, 9090 Melle, Belgium

1.6 APPENDIX

Conferences, Presentations and Courses Attended

- "Ensembl Genome Browser" course organized by the VIB Bioinformatics Training & Service (BITS) facility on April, 20th 2010 at the UGent Campus Ardoyen, Gent
- "Introduction to Bioinformatics" course organized by the VIB Bioinformatics Training & Service (BITS) facility on February, 15nd and 16nd 2010 at the UGent Campus De Sterre, Gent.
- "Vector NTI Basics training" organized by the VIB Bioinformatics Training & Service (BITS) facility on October, 1st 2010 at Ghent University, Technology Park Zwijnaarde, Building FVMS, Gent.
- Stay abroad: URGV, Evry and INRA Versailles, France for Screening Brachypodium TILLING population for mutations in *Bd4CL1,* 23-Feb-2011 → 11-Mar-2011
- First European Brachypodium Workshop, 19/10/11 → 21/10/11 - Versailles, France, poster presentation
- Vibes in bioscience, Gent, Belgium, 5-Sep-2012 → 7-Sep-2012 (participant)
- MRP Happy hour by MRP Ghent BioEconomy, Gent, Belgium, 12-Nov-2012 (invited speaker)
- Media training; communicating food science, 20-21 June, 2013, by Maitre, ILVO, Melle, Belgium
- SUNLIBB project progress meetings in Cambridge, UK (September 2011), Versailles, France (September 2012), Maresias, Brazil (March 2012), Wageningen, Netherlands (September 2013) with presentations on project progress
- 9th International Conference on Renewable Resources & Biorefineries, Antwerp, Belgium, 5-Jun-2013 → 7-Jun-2013, Oral presentation

Publications

- Disrupting the cinnamyl alcohol dehydrogenase 1 gene (BdCAD1) leads to altered lignification and improved saccharification in Brachypodium distachyon. Bouvier d'Yvoire, Madeleine; Bouchabke-Coussa, Oumaya; Voorend, Wannes; Antelme, Sébastien; Cézard, Laurent; Legée, Frédéric; Lebris, Philippe; Legay, Sylvain; Whitehead, Caragh; McQueen-Mason, Simon J; Gomez, Leonardo D; Jouanin, Lise; Lapierre, Catherine; Sibout, Richard. The Plant journal : for cell and molecular biology, Vol. 73, No. 3, 2013, p. 496-508.