29

High-biomass sorghums as a feedstock for renewable fuels and chemicals

Tallyta N. Silva[1,2] and Wilfred Vermerris[2,3,4]

[1]Graduate program in Genetics and Genomics, [2]UF Genetics Institute [3]Department of Microbiology & Cell Science, [4]Florida Center for Renewable Chemicals and Fuels, University of Florida, Gainesville, FL, USA

Present address for Tallyta N. Silva: Joint BioEnergy Institute, Emeryville, CA, USA; Environmental Genomics and Systems Biology Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA

Corresponding author: Wilfred Vermerris, wev@ufl.edu

Abstract

High-biomass sorghums are intended for use in biorefineries that convert vegetative biomass into renewable fuels and chemicals. The majority of plant biomass consists of cell walls, a complex matrix of cellulose, hemicellulosic polysaccharides and lignin. In the biorefinery, the biomass is ground and then subjected to a thermo-chemical pretreatment (elevated temperature, high pressure, chemical catalysts) that disrupts the cell wall matrix and solubilizes some of the cell wall polymers, followed by enzymatic saccharification of the cellulose, which produces D-glucose. The glucose and sometimes also the xylose derived from the hydrolysis of the hemicellulosic polysaccharides, are subsequently converted to fuels or other useful chemicals by microbial biocatalysts. The genetic improvement of high-biomass sorghums has as its ultimate goals to maximize the yield of fermentable sugars on a per-hectare basis, to minimize the inputs of fertilizer, irrigation, fungicides and pesticides and to reduce the environmental footprint. Breeding strategies thus need to target biomass yield, biomass composition, maturity, pest and disease resistance and nutrient use efficiency. This chapter reviews the genetic basis of these traits and their potential application in breeding programs.

Keywords: biofuel – *brown midrib* – cellulose – dwarf – hybrid – maturity

Contents

29.1. Introduction

29.1.1. Definition and use of high-biomass sorghum

As the name implies, high-biomass sorghums (*Sorghum bicolor* (L.) Moench) are cultivated for the purpose of generating biomass. In the case of sorghum, biomass refers to the vegetative parts of the plant: stems, leaves, and tillers. These vegetative parts are also referred to as lignocellulosic biomass, because the bulk of the biomass consists of cellulose, hemicellulosic polysaccharides and lignin.

High-biomass sorghums are intended for industrial use. In principle, the biomass can be used for combustion to generate heat and/or electricity, but it is difficult for sorghum biomass to compete with woody biomass for calorific value due to both the lower bulk density and the lower concentration of lignin, and because of the relatively high concentration of silica, which generates undesirable slag in the boilers. On the other hand, sorghum biomass is an excellent feedstock for the production of renewable fuels and chemicals. In that case, the most common use is to generate fermentable sugars from the cell wall polysaccharides that make up approximately 75% of the dry biomass. This use sets high-biomass sorghum apart from two other types of sorghum that are cultivated to generate substantial amounts of biomass: forage sorghums, used as fodder, which need to be palatable and digestible, and sweet sorghums, which have juicy stems rich in soluble sugars (glucose, fructose, sucrose) that were historically used for the production of syrup.

The interest in high-biomass sorghums as a separate type of sorghum is recent compared to the other types of sorghum (grain, forage, sweet), and results from the desire for dedicated bioenergy crops that cannot be used as a source of food, unlike sorghum grain and the juice of sweet sorghums. The desire to limit the use of food crops for the production of biofuels is driven

by concerns both ethical (Chakravorty et al., 2009; Rosegrant & Msangi, 2014) and environmental (Searchinger et al., 2008) in nature, especially in light of estimates that by 2050, food production will need to increase by 60% to meet the demand of an increasing world population (FAO, 2009). The use of lignocellulosic biomass can be a sustainable alternative to fossil fuels as long as the cultivation of dedicated bioenergy crops does not compete for prime agricultural land used for food production and does not lead to conversion of 'natural' areas with important ecological functions (e.g. tropical rainforest) (Lambin & Meyfroidt, 2011). In the United States, the 2007 Energy Independence and Security Act (EISA, 2007) mandates that the volume of renewable fuels from lignocellulosic biomass and agricultural waste grow from around 379 million liters (100 million U.S. gallons) in 2010 to 61 billion liters (16 billion U.S. gallons;) in 2022 (Schnepf & Yacobucci, 2013). Current production (2020) is well below this target due to limited commercial production, the low price of petroleum and consumer interest in electrical vehicles. OECD/FAO (2015) forecasted global expansion in fuel ethanol production to 134 billion liters (billion = 10^9), of which 1.7 billion liters are anticipated to be cellulosic ethanol, which is more realistic.

Since dedicated biomass production represent a relatively new use of sorghum, that is currently only occurring on a small scale relative to the production of grain, there is only limited information available on biomass production and yield. Yields vary substantially as a function of genotype, environment and management. Dry matter yields as high as 60 dry ton/ha have been reported under optimal conditions, with availability of water representing an important factor (Olson et al., 2012; Snider et al., 2012). For cultivation under suboptimal conditions (limited irrigation or rain-fed; limited fertilizer inputs), yields in the range of 15-20 dry ton/ha appear to be realistic (Hao et al., 2014; Snider et al., 2012).

29.1.2. Biomass processing at the biorefinery

The processing of biomass sorghums for the production of renewable fuels and chemicals occurs at a biorefinery (Ragauskas et al., 2006). This is a specialized facility that processes biomass feedstocks typically from a range of less than 80 km to minimize the cost of transportation. The biomass is first ground and then subjected to a thermo-chemical pretreatment that makes the cellulose in the plant cell walls accessible. Several types of pretreatment exist (reviewed by Constant et al. 2016; Hu and Ragauskas 2012; Pu et al. 2015), with dilute acid pretreatment and alkaline pretreatment representing the two most common procedures for biomass from grasses. During dilute acid pretreatment the biomass is exposed to high temperature (160-200ºC) and pressure with sulfuric or phosphoric acid as catalyst (Selig et al., 2007; Van Rijn et al., 2018). The hemicellulosic polysaccharides are hydrolyzed and the lignin is displaced so that it no longer occludes the cellulose. After adjusting the pH, cellullolytic enzymes are added to the pretreated biomass to convert the cellulose to D-glucose, a step referred to as enzymatic saccharification. The glucose is subsequently fermented to fuels (e.g. ethanol, butanol) or chemicals (e.g. lactic acid, butyric acid) depending on the microbial biocatalyst selected. Commonly used microbes are baker's yeast (*Saccharomyces cerevisiae*), *Pichia stipitis*, *Escherichia coli*, and *Clostridium* spp. (Huang et al., 2009; Karimi et al., 2006; Lan & Liao, 2013; Yu et al., 2007). Some of these microbes are able to co-ferment D-glucose and D-xylose generated from the hydrolysis of hemicellulosic polysaccharides. The solid residues remaining after fermentation are rich in lignin.

Alkaline pretreatment, performed in sodium hydroxide at temperatures between 60-120ºC, or with the use of ammonia under pressure, dissolves the lignin and hydrolyzes some of the

hemicellulosic polysaccharides (Mcintosh & Vancov, 2010). After solid-liquid separation (which removes much of the lignin) and pH adjustment of the polysaccharide-rich solid fraction, the enzymatic saccharification and fermentation steps are similar as described for the dilute acid pretreatment.

Techno-economic analyses of converting biomass to renewable fuels and chemicals have indicated that several factors contribute to the relatively high cost of production. In addition to the cost of the thermo-chemical pretreatment and the cellulolytic enzymes, the feedstock itself represents a major cost (Aden et al., 2002; Van Rijn et al., 2018; Valdivia et al., 2016). This means that ways to produce the crop more efficiently will have a direct impact on the competitiveness of renewable fuels and chemicals. Sections 2 and 3 of this chapter will review the different approaches that can be pursued to accomplish this.

29.1.3. Sorghum versus other bioenergy crops

High-biomass sorghums are part of a portfolio of crops that can be used as dedicated biomass crops, and that include several other grasses, including switchgrass (*Panicum virgatum* L.), miscanthus (*Miscanthus* species and interspecific hybrids), elephantgrass or napier grass (*Pennisetum purpureum* Schumach.), energy cane (*Saccharum* spp.), giant reed (*Arundo donax* L.), as well as the woody species poplar (*Populus* spp.), pine (*Pinus* spp.), willow (*Salix* spp.) and eucalypt (*Eucalyptus* spp.).

In this list, sorghum is the only annual crop among perennials. Obvious benefits of using perennial crops compared to annual crops are that, on average, they require less fertilizer due to their ability to relocate minerals from the aboveground parts of the plants to the roots at the end of the growing season, which also limits nutrient losses from the soil. Furthermore, they need to be established just once before offering several harvests over a period of multiple years, and they provide a means of controlling erosion and soil health (microbiomes), because the root system survives during the winter. On the other hand, drawbacks associated with the use of these crops are the producer's need to commit to a particular crop and specific genotype for multiple years, even when more productive genotypes may be released during that period, and the need to wait for the crop to establish itself, during which time no harvest occurs.

In addition to the flexibility associated with sorghum being an annual crop, there are several other advantages that make sorghum of particular interest as a high-biomass grass compared to the other species listed above. One of those benefits is that, unlike the other bioenergy grasses with the exception of switchgrass, sorghum is a seed-propagated crop, which makes the establishment of the crop easier and less labor-intensive than the use of vegetative cuttings to establish the crop. As a result of sorghum's long history as a cereal crop, there is an established supply chain consisting of breeding companies, seed producers, and distributors, that ensures pure, high-quality seed. The large genetic diversity within the species (Motlhaodi et al., 2017; Wang et al., 2009), can be accessed via several large germplasm collections at the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT, Hyderabad, India), at the Institute of Crop Science operated by the Chinese Academy of Agricultural Sciences, and at the U.S. National Plant Germplasm System managed by the U.S. Department of Agriculture (USA). Sorghum breeders can request seed from accessions of interest and screen parents for new breeding populations that are adapted to specific environmental conditions.

The result of sorghum's comparatively long history as a crop combined with the genetic diversity within the species is that sorghum has been adapted to a wide range of environments, including a range in latitudes that is hard to match by the other bioenergy species mentioned earlier. Even though energy cane, which are *Saccharum* genotypes selected for biomass production rather than sugar yield (Matsuoka et al., 2014), are considered to have the most efficient photosynthesis and annual biomass accumulation potential among the cultivated grasses (Slewinski, 2012), currently available *Saccharum* genotypes are only productive in tropical and subtropical environments. The same is true for available germplasm of *Pennisetum purpureum*, which is even more sensitive to low temperatures than *Saccharum* (Burner et al., 2017). Miscanthus and switchgrass, on the other hand, are cold-tolerant perennials. Switchgrass biomass yields (5-11 ton dry matter/ha; (Schmer et al., 2008), tend to be lower than those of sorghum. While miscanthus yields of up to 40 ton dry matter/ha have been reported in warm climates, yields are subject to large genotype-by-environment effects (Clifton-Brown et al., 2001). Giant reed is a riparian species that propagates vegetatively and is considered invasive in many environments. Aside from the need to have adequate water supply, the genetic diversity with this species is minimal (Ahmad et al., 2008; Saltonstall et al., 2010), turning any commercial cultivation essentially into a monoculture and its associated risks of susceptibility to sudden outbreaks of pests or diseases.

In addition to being adapted (or adaptable) to tropical, subtropical and temperate climates, sorghum is also adapted to a wide variety of environmental conditions, which includes tolerance to high temperatures, periods of drought, and periods of water-logging (e.g. low-lying areas near rivers). Sorghum can also grow on a variety of soils, including soils rich in minerals and with different pH values. These features make sorghum an excellent candidate for cultivation on low-productivity lands, which tend to be avoided for the cultivation of most food crops due to a combination of low yield and sub-optimal quality, and therefore reduced market value. The cultivation of sorghum on low-productivity land also addresses a major concern raised over the cultivation of bioenergy crops in general, namely the competition for land with food crops (the food versus fuel debate; (Chakravorty et al., 2009; Rosegrant & Msangi, 2014), or the land use change that occurs when natural areas are converted to farm land to accommodate bioenergy crops (Searchinger et al., 2008). Furthermore, it also increases the chances that sorghum will retain its relevance as greater variation in precipitation and temperature at a given location are anticipated as a result of climate change.

Therefore, improvement of high-biomass sorghum has the potential to enhance the efficiency of producing biofuels and renewable chemicals and contribute to a more sustainable production of these commodities.

29.2. High-biomass sorghum ideotype

This section describes the ideotype of high-biomass sorghums, which is a compilation of the phenotypic traits that would make high-biomass sorghums maximally compatible with a biorefinery operation. From the perspective of maximizing the efficiency of the supply chain leading to renewable fuels and products, the feedstock production needs to be geared towards maximizing the yield of fermentable sugars derived from a hectare of high-biomass sorghum at the lowest possible cost. This can be accomplished by developing genotypes that are efficient with

inputs (water, fertilizer), which are functions of both canopy and root system architectures, and that have resistance against the major pests and diseases in the region of production, so that the use of pesticides and fungicides can be minimized. Furthermore, the ideal genotypes yield large amounts of biomass with a composition that requires relatively mild thermo-chemical pretreatment conditions, and whereby low enzyme loadings are sufficient for enzymatic saccharification of cell wall polysaccharides. Each of these traits is discussed in further detail below.

29.2.1 Biomass yield

Plant biomass yield is largely determined by stem yield, which in turn is determined by stem volume. Stem volume can be maximized by increasing plant height and stem diameter. Plant height is controlled by maturity (photoperiod sensing) and the activity of plant growth regulators. With the availability of the sorghum genome sequence (Paterson et al., 2009) and the tools that ensued, much progress has been made in understanding the genetic basis of maturity and height control. This knowledge has direct benefits for the breeding of high-biomass sorghums. The target ideotype for high-biomass sorghum is a plant taller than 3.5 m (Braconnier et al., 2011). Under optimal conditions (genotype, environment, management), dry biomass yields as high as 60 Mg/ha have been reported (Olson et al., 2012; Snider et al., 2012).

Table 29.1 Comparison of maximum sorghum biomass yields reported in in the United States

Publication	Water source	Plot details	Location	Dry biomass yield (Mg/ha)
McCollum et al. (2005)	Irrigated	296,400 seeds/ha	Texas	27.4
Venuto & Kindiger, (2008)	Rainfed	1.5 × 7.5 m^2, rows 20 cm apart; 22.5 kg seed/ha$^+$	Oklahoma	27 (average); 40.3 (maximum)
Olson et al. (2012)	(Limited) irrigation	1.5 × 50 m^2; rows 76 cm apart; 132,000 plants/ha	Texas	49.5 (limited irrigation)* 59 (irrigation)*
Snider et al. (2012)	Rainfed	rows 19 cm apart; 1.5 m of two center rows harvested; 116,000 seeds/ha	Alabama	61.1
Packer & Rooney (2014)	Rainfed	6.7 m long rows, 75 cm apart. The center 1.5 m of the row was harvested; 150,000 plants/ha	Texas	32.4 (average) 41.3 (maximum)
Meki et al. (2017)	Limited irrigation	15 × 15 m^2, rows 23 cm apart; 180,000 seeds/ha	Texas	37.9

$^+$At an average 1000-seed weight of 25g, this represents 900,000 seeds/ha; *extrapolated from 9 plants

Due to the limited commercial production of high-biomass sorghums, biomass yields reported in the literature are mostly based on small plots, or from relatively small samples (e.g. sections of 1.5 m) from medium-sized plots at well-maintained research sites, which may overestimate the biomass yield feasible on a commercial scale. In addition, there is no standard

procedure to determine biomass yields. As can be gleaned from Table 29.1, plot sizes, seeding densities, and row spacing vary considerably.

Snider et al. (2012) examined the effect of a number of these parameters on biomass yield. Venuto and Kindiger (2008) determined that a single late harvest generally resulted in a greater biomass yield than an early harvest followed by a second harvest of the ratoon crop.

29.2.1.1 Plant height as a function of maturity

Having originated in equatorial Africa, sorghum is originally a short-day plant that naturally exhibits considerable photoperiod sensitivity. Naturally photoperiod-sensitive sorghum genotypes need short days in order to make the transition from the vegetative to the reproductive phase. Under long days, these genotypes remain vegetative and continue to elongate. The date of planting is thus an important factor to consider when producing photoperiod-sensitive biomass sorghum. In an evaluation of two biomass sorghum genotypes planted on different dates, Meki et al. (2017) observed that these sorghum genotypes behaved like photoperiod-insensitive short-day grain sorghums when planting occurred during a time of the year with short days (less than 12 hours of daylight). The plants flowered after approximately 90 days and reached heights less than 2 m. In contrast, if the planting was carried out under long days, these genotypes flowered later in the season and reached heights greater than 3 m. According to Mullet (2017), a large portion of the estimated 40,000 accessions of the sorghum world germplasm exhibit delayed flowering under long-day conditions.

29.2.1.2 Genetic control of maturity: *Ma* genes

Increased plant height is often correlated with late flowering because flowering terminates apical growth in most Poaceae. The genes known for controlling floral initiation in sorghum are called Maturity (*Ma*) genes. Plants with recessive alleles in the maturity genes are photoperiod insensitive and behave like a long-day plant. Studies to manipulate these alleles allowed the domestication of sorghum and its spread to temperate regions, such as the U.S. The first maturity loci identified were *Ma1*, *Ma2*, *Ma3* and *Ma4* (Quinby, 1967), followed by the identification of *Ma*5 and *Ma6* (Rooney & Aydin, 1999) in forage and biomass sorghum.

Ma1 is the main photoperiod sensitivity locus and is located at position 40.3 Mb on chromosome 6 (Thurber et al., 2013). The incorporation of the recessive *ma1* allele was instrumental for the success of grain sorghum production in the U.S., since the resulting earlier flowering allows more time for grain filling, and reduces the risk of frost damage in temperate regions (Klein et al., 2008). For high-biomass sorghum production, however, the dominant *Ma1* allele is preferred, because the delayed flowering it causes results in taller plants. As a result of an interspecific cross of *Sorghum bicolor* with *Sorghum propinquum*, the *Ma1* locus was shown to be genetically linked to the *Dwarf2* locus (see 2.1.4) on chromosome 6 and explains around 55% of variation in plant height and around 86% of flowering time variation (Lin et al., 1995).

Quinby (1974) observed that *Ma1/ma1* heterozygotes flower later than either homozygotes, but only when *ma2* is recessive. Based on positional cloning, Murphy et al. (2011) suggested that *Ma1* encodes Pseudoresponse Regulator Protein 37 (SbPRR37; Sb06g012260), a repressor of flowering in long days (Fig. 29.1). During short days, *SbPRR37* has its expression peak in the morning, while in long days the peaks are in the morning and evening (Murphy et al.,

2011). However, according to Cuevas et al. (2016), the association of *SbPRR37* to *Ma1* was confounded with the presence of *SbFT12*, also a floral suppressor, and many other annotated genes in that chromosomal region that were previously unknown. These authors associated *Ma1* to *SbFT12* based on fine mapping, association genetics, mutant complementation and evolutionary analysis.

The influence of *Ma2*, *Ma3* and *Ma4* in photoperiod sensitivity is much smaller than *Ma1*. However, *Ma2* is in complex interaction with *Ma1*. For instance, *Ma1/ma1* shows overdominant late flowering compared to *Ma1/Ma1* in the presence of *ma2/ma2*, but *Ma1/ma1* and *Ma1/Ma1* cannot be distinguished from each other in the presence of a dominant *Ma2* allele (Quinby, 1974). There are sorghum varieties that flower late under long days with the recessive alleles *ma1*, *ma2* or *ma3* (Pao & Morgan, 1986).

Childs et al. (1997) showed that *Ma3* encodes the apoprotein of phytochrome B, a photoreceptor involved in photoperiod sensing and repression of flowering. It represses the expression of the sorghum ortholog of the maize gene *Teosinte branched1* and responds to light signals inducing growth of axillary buds (Kebrom et al., 2006). These authors demonstrated that the mutant allele $ma3^{R}$ harbored a frameshift mutation in the *phytochrome B* gene.

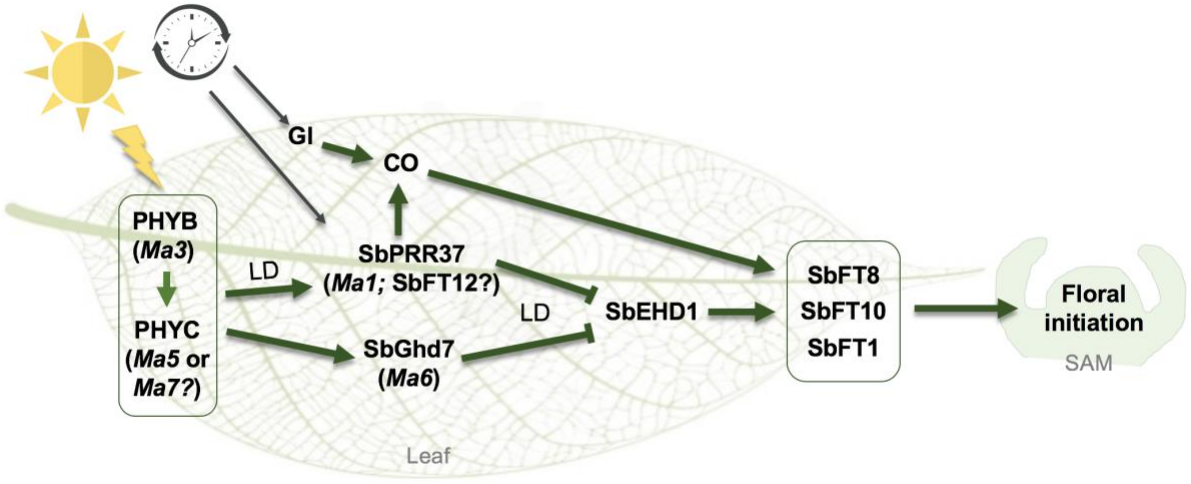

Fig. 29.1 Simplified scheme of flowering regulation in sorghum, based on information from Yang et al. (2014), Cuevas et al. (2016), Wolabu et al. (2016) and Mullet (2017). Light signaling on phytochrome B (PHYB) and phytochrome C (PHYC) induces the expression of *SbPRR37* and *SbGhd7*, which are floral repressors. PHYB stabilizes and interacts with PHYC. Expression of *SbPRR37* and *SbGhd7* leads to repression of *SbEHD1*, and subsequently, of the genes encoding florigens (*SbFT1*, *SbFT8* and *SbFT10*). *SbFT1*, *SbFT8* and *SbFT10* correspond to *SbCN15*, *SbCN12* and *SbCN8*, respectively (Yang et al. (2014). The result will be a delay in floral initiation, which occurs in the shoot apical meristem (SAM). During each day-night cycle, expression of *GIGANTEA* (*GI*) is regulated, which in turn, regulates the activity of the floral-inducing gene *CONSTANS (CO)*. Next, the expression of *SbFT8* and *SbFT10* increases, resulting in floral initiation. According to Wolabu et al. (2016), *SbFT1* shows the same expression pattern of *SbFT8* and *SbFT10*, but at lower levels, indicating regulation through the same pathway.

As mentioned before, the selection of recessive mutant alleles of *Ma* genes has been performed to adapt sorghum to temperate climates with long days during the growing season. However, the dominant allele of *Ma4* has been associated to early flowering. *Ma4* exists as a dominant allele in grain sorghum. Interestingly, at high temperatures *ma4* behaves as *Ma4* (Quinby, 1966).

When a sorghum plant contains dominant *Ma5* and *Ma6* alleles, the floral initiation is inhibited, independently of day length (Childs et al., 1997). According to Yang et al. (2014), *Ma5* may encode phytochrome C, which has been shown to influence flowering time in rice in long days. *Ma6* encodes the sorghum ortholog of the rice Grain Number, Plant Height and Heading Date7 (*SbGhd7*; Sb06g000570), a repressor that down-regulates *Early Heading Date1* (*SbEHD1*), *Centroradialis12* (*SbCN12)* and *Centroradialis18 (SbCN18)*, delaying flowering in long days. The manipulation of these alleles by breeding has great potential for development of improved high-biomass sorghum (Murphy et al., 2014; Yang et al., 2014). According to Murphy et al. (2014), dominant alleles of *SbGhd7* and *SbPRR37* have additive effects in biomass sorghum, delaying flowering for approximately 175 days until daylight length is less than 12.3 h.

Mullet et al. (2010) mentioned maturity gene *Ma7*, which interacts with *Ma5* and *Ma6*. According to these authors, candidate genes in the *Ma7* interval include two MADS-box genes and a gene encoding phytochrome C. However, further studies are necessary to elucidate the potential of *Ma7* for breeding high-biomass sorghum.

29.2.1.3 Plant height as a function of phytohormones and growth regulators

The plant hormones auxin, gibberellins and brassinosteroids are known to control plant height, and mutations that affect the biosynthesis of these hormones, their transport or their sensing tend to result in shorter plants that may also exhibit altered morphologies. Auxin is the plant hormone associated with apical dominance and is transported from the shoot apical meristem to lower parts of the plant. The maize (*Zea mays* L.) *brachytic2* mutant and the sorghum *dwarf3* mutant are, as their names imply, short plants resulting from a defective auxin transporter (Multani et al., 2003) (see also section 2.1.4).The gibberellins are cyclic diterpenoids that regulate many biological processes including stem elongation. In maize, the discovery that five *dwarf* loci encode enzymes involved in the biosynthesis of gibberellin (Bensen et al., 1995; Coe et al., 1988; Fujioka et al., 1988; Phinney and Spray, 1982) provided early evidence for the importance of this hormone in plant growth. Ordonio et al. (2015) demonstrated that loss-of-function mutations in four sorghum genes involved in the early steps of gibberellin synthesis, resulted in dwarf plants with bent culms (stems). These findings suggest that plants that produce more gibberellin may make more biomass. Indeed, a study by Okuno et al. (2014) on rice suggests this may be true. These authors evaluated rice mutants that produce higher levels of gibberellin and reported greater lodging resistance due to larger culm diameters and/or increased lignin concentration, and increased biomass yield. Therefore, if similar mutants or genetic variants were available in sorghum, they may have potential to enhance biomass production.

Brassinosteroids are polyhydroxylated steroidal plant hormones involved in stem elongation (Taiz and Zeiger, 2010). Mutants in Arabidopsis, pea, tomato and rice (Bishop et al., 1999; Li et al., 1996; Tanabe et al., 2005) in which brassinosteroid biosynthesis is compromised show dwarfism. There is, however, only a limited understanding of the role of brassinosteroids in

sorghum. Mantilla Perez et al. (2014) identified 26 sorghum candidate genes related to brassinosteroid biosynthesis and signaling and performed association mapping with plant architecture traits. The authors concluded that the overall phenotypic variation in plant height explained by markers/genes associated with brassinosteroid synthesis and signaling pathways was only 6%. Additional studies are necessary to validate the functionality of sorghum genes predicted to be involved in brassinosteroid biosynthesis and signaling. Nonetheless, the involvement of the *Dwarf1* gene (Section 2.1.4) in brassinosteroid signaling implies this class of hormones plays an important role in controlling plant height in sorghum.

29.2.1.4. Dwarfing genes

The best-known sorghum genes determining plant height are known as *Dwarfing* (*Dw*) genes, which influence internode and apex elongation. Four unlinked *Dw* genes have been identified in sorghum, *Dw1* through *Dw4* (Quinby, 1974; 1975). Dominant alleles at all four loci result in tall plants. During the domestication of sorghum to produce grain in temperate regions, the recessive alleles, mainly *dw1*, *dw2* and *dw3*, were selected to obtain shorter plants that are compatible with mechanical harvesting (Klein et al., 2008). The four *Dw* loci act in an additive fashion to control height, so that height can be reduced from over 3 m for a plant harboring one or two dominant *Dw* alleles at each of the four *Dw* loci to just 60 cm when a plant harbors homozygous recessive *dw* alleles at all four dwarfing loci (Quinby, 1967).

The *Dw1* locus is located at position ~57 Mb on chromosome 9 and is now known to act as a positive modulator of brassinosteroids signaling by inhibiting *Brassinosteroid Insensitive2* (*BIN2*), a negative regulator (Morris et al., 2012; Hirano et al., 2017; Thurber et al., 2013). Association mapping in sorghum conversion lines detected significant association between plant height and flowering time in *Dw1* (Thurber et al., 2013). *Dw2* has been mapped to chromosome 6 and is linked to the maturity gene *Ma1* (Quinby, 1974, 1975; Lin et al., 1995). It encodes a protein kinase homologous to KIPK, a member of the protein kinase family in Arabidopsis (Hilley et al., 2017). Besides plant height, *Dw2* also influences panicle length, seed weight and leaf area (Graham & Lessman, 1966; Pereira & Lee, 1995).

The first cloned dwarfing gene in sorghum was *Dw3*. It was identified as the gene encoding PGP1/PGP19, an auxin transporter, and is the ortholog of maize *Brachytic2* and Arabidopsis *PGP1* Multani et al., 2003). Mutations in *dw3* are caused by an unstable insertion of a retrotransposon in the gene, which reverts to its wild-type allele at a frequency of 0.5-1 % (Multani et al., 2003). This is the reason why large fields of sorghum that contain the *dw3* mutation will contain a small but noticeable number of taller plants that are otherwise phenotypically identical. The *Dw3* gene is located on chromosome 7 (Brown et al., 2008).*Dw4* has not been cloned yet, but it is known to be unlinked to the other three dwarfing loci (Quinby & Karper, 1954). Morris et al. (2013) identified a potential location of the *dw4* locus at ~6.6 Mbp on chromosome 6 based on the location of the next most significant peak in genome-wide association study on height and heterozygosity scan.

29.2.2 Biomass composition

Sorghum biomass consists predominantly of cell walls. Cell walls are a complex matrix in which cellulose microfibrils form the main structural component, held in place by a network of hemicellulosic polysaccharides, in grasses predominantly glucuronoarabinoxylans (GAX), and a small amount of pectin (Carpita & Gibeaut, 1993). Secondary cell walls, present in the xylem and

sclerenchyma fibers, also contain lignin, an aromatic polymer that provides mechanical strength and the hydrophobic coating needed to facilitate the transport of water. In addition to these structural components, sorghum biomass also contains some residual proteins from when the plants were metabolically active as well as some starch, synthesized from excess D-glucose, and minerals, notably silica. Although the exact biomass composition varies according to the genotype, developmental stage at harvesting time and environmental conditions, biomass is generally composed of approximately 45% cellulose, 20-25% hemicellulosic polysaccharides, 18-22% lignin, 5% starch, 5-8% minerals, 3-5% pectin, 3-5% protein (Castro et al., 2017; Rooney et al., 2007).

29.2.2.1 Cellulose

Cellulose is the main structural component of the plant cell wall and the primary source of the D-glucose that fermentative microorganisms can convert to biofuels or other chemicals in the biorefinery. Cellulose is produced by cellulose synthases (CesA) that are associated with the plasma membrane and that use UDP-D-glucose as substrate for the synthesis of glucan chains (Saxena & Brown., 2005; Somerville, 2006). Addition of a new D-glucose residue is accompanied by the release of a water molecule. Subsequent D-glucose residues are rotated 180º relative to each other. As a consequence, the repeat unit of cellulose is anhydro-cellobiose (Fig. 28.2).

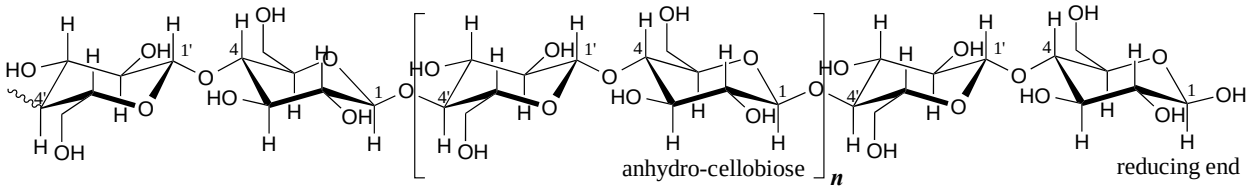

Fig. 29.2 A representation of cellulose, a β-1,4-linked polymer of D-glucose, with anhydro-cellobiose, inside the square brackets, as the repeat unit. A cellulose microfibril consists of 36 glucan strands that are held together by hydrogen bonds.

The catalytic mechanism of cellulose synthase in plants is still being refined (Morgan et al., 2013; Olek et al., 2014; Sethaphong et al., 2013). CesA's are organized in groups of six, with each group consisting of three different subunits. The identity of these subunits differs in the primary versus secondary cell wall. Six clusters of six CesA units form a so-called terminal complex (Mueller & Brown, 1980) that produces 36 glucan strands that together form a cellulose microfibril. Due to the regular structure of cellulose it can be present in crystalline form. The sorghum genome (Paterson et al., 2009) contains 10 *CesA* genes (Vermerris & Saballos, 2013). Even with one of these genes not being expressed (under conditions tested), this large number of *CesA* genes implies some redundancy, and the underlying reason is not yet clear. Between the important structural role of cellulose in the plant cell wall, and the redundancy in *CesA* genes, there is an inherent risk associated with the modification of the expression of *CesA* genes.

Murray et al. (2008) used a biparental mapping population derived from the grain sorghum BTx623 and the sweet sorghum 'Rio' to map a number of traits related to bioenergy production, including cellulose content. Based on data from multiple locations and years, they identified a QTL on chromosome 3 associated with cellulose content, but the underlying gene(s) were not identified,

as this study predated the release of the sorghum genome sequence and high-density molecular markers.

The role of genetic variation on cellulose crystallinity was investigated by Vandenbrink et al. (2012). These authors used a set of 20 genotypes identified in a population of 386 diverse sorghum genotypes that had been shown earlier to vary for the yield of fermentable sugars obtained after enzymatic saccharification (Vandenbrink et al., 2010). The selected genotypes were grown in two locations in two different years. The crystallinity index (CI), measured using X-ray diffraction, varied among the genotypes and was negatively correlated ($r^2 = 0.25$) with the yield of fermentable sugars after 24 h of enzymatic saccharification. The correlation between the CI values from plants harvested at the two locations was, however, weak ($r^2 = 0.07$), and lower than the correlation in the yield of fermentable sugars ($r^2 = 0.31$). These combined observations suggest significant environmental and/or genotype × environment effects on CI and saccharification yields.

29.2.2.2 Hemicellulosic polysaccharides

Hemicellulosic polysaccharides comprise a set of hexose- and pentose-based polymers that are distinct from cellulose in that they are sensitive to degradation in low concentrations of acids, and that vary substantially in structure and composition among plant species (Carpita & Gibeaut, 1993). The cell walls of grasses contain as their main hemicellulosic polysaccharide glucuronoarabinoxylan (GAX; Fig. 29.3), a polymer with a backbone consisting of D-xylose residues and substituted with L-arabinose and D-glucuronic acid residues. L-Arabinose residues can be substituted with ferulate, a hydroxycinnamic acid that enables crosslinking of neighboring GAX molecules, as well as GAX and lignin. In contrast, the cell walls of most angiosperm dicots contain xyloglucan as the main hemicellulosic polysaccharide (Carpita & Gibeaut, 1993).

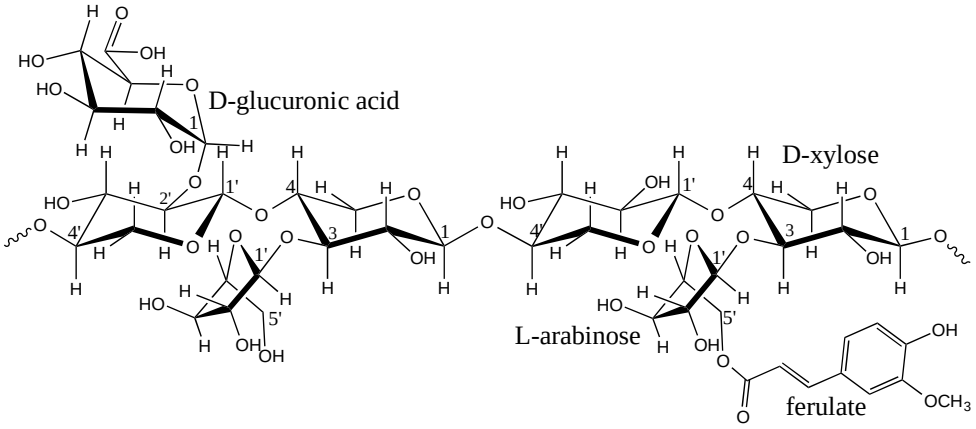

Fig. 29.3 A representation of glucuronoarabinoxylan, the predominant hemicellulosic polysaccharide in sorghum and other grasses. The backbone of D-xylose residues is substituted with L-arabinose and D-glucuronic acid. In this depiction, a ferulate molecule is esterified to the L-arabinose residue.

Grasses are unique in that they also produce mixed-linkage β-glucans, but this is a transiently produced polymer during the development of the primary cell wall, and is no longer present when the plants are mature. The backbone of several hemicellulosic polysaccharides is synthesized by cellulose synthase like (CSL) enzymes, encoded by *Csl* genes. For example, the mixed-linkage β-glucans are synthesized by CSLF (Burton et al., 2006). As the name implies,

these enzymes share structural similarity with CesA's. A main difference is that CSLs are not associated with the plasma membrane. Instead, hemicellulosic polysaccharides are synthesized in the Golgi complex, rather than in the cell wall. Even though CSLs were initially hypothesized to be responsible for the synthesis of the xylan backbone of GAX, recent evidence points to a member of the glycosyl transferase (GT) family, GT47, as the enzyme responsible for this role (Zhang et al., 2014).

29.2.2.3 Lignin

Even though lignin is an important component for the functioning of the secondary plant cell wall, at the biorefinery it is primarily perceived to be an undesirable component that needs to be removed. In addition to forming a physical barrier that occludes cellulose, the cellulolytic enzymes used to generate fermentable monosaccharides from cellulose adhere to lignin irreversibly (Zeng et al., 2014). As a consequence, higher enzyme loadings are needed for the enzymatic saccharification of plant tissues rich in lignin than based strictly on the amount of cellulose present in those tissues (Zeng et al., 2012).

The severity of the thermo-chemical pretreatments discussed earlier can be reduced by the reduction in lignin concentration *in planta* and/or by altering the subunit composition of the lignin via genetic means. The most direct way to accomplish this is by modulating the flux through the metabolic pathways leading to lignin (Vermerris & Abril, 2015). Lignin is synthesized in the cell wall from the reaction of monolignol radicals formed by peroxidases in the presence of hydrogen peroxide (H_2O_2) or laccases in the presence of oxygen (O_2). The main monolignols in grasses are coniferyl alcohol and sinapyl alcohol, which give rise to guaiacyl (G) and syringyl (S) residues in the lignin polymer, respectively, typically in a ratio of approximately 1.5: 1(Fig. 29.4). A small amount (3-5%) of *p*-hydroxyphenyl residues is formed from the incorporation of *p*-coumaryl alcohol. The incorporation of sinapyl alcohol in the growing lignin polymer is enabled through the esterification of *p*-coumaric acid to the hydroxyl moiety on C_9 of sinapyl alcohol (Hatfield et al., 2008; Petrik et al., 2014), which explains the abundance of *p*-coumaroyl esters associated with the lignin of grasses. The lignin in grasses, including sorghum, also contains tricin (Fig. 29.4), a flavone that can act as a nucleation site for lignin polymerization (Lan et al., 2016).

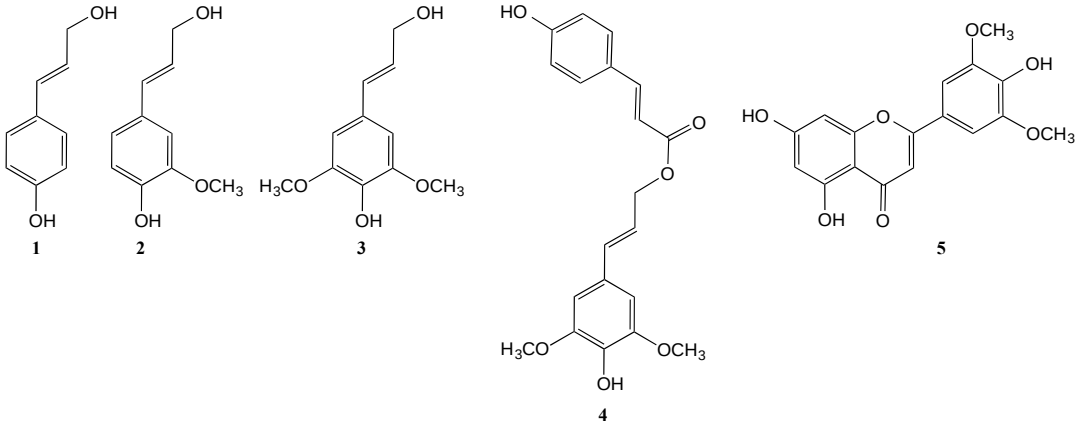

Fig. 29.4 The structure of *p*-coumaryl alcohol (1), coniferyl alcohol (2) and sinapyl alcohol (3), the *p*-coumaroyl ester of sinapyl alcohol (4), and tricin (5).

The *brown midrib* (*bmr*) mutants are the best-known cell wall mutants of sorghum. They were first reported by (Porter et al. (1978), who generated a population of chemically induced mutants. The name of these mutants refers to the characteristic reddish-brown coloration of the the central vein of the leaf. Porter et al. (1978) also determined that some of these mutants were more digestible when used as fodder. The individual mutants were given consecutive numbers regardless of allelic relationships, which complicated genetic studies. Saballos et al. (2008) determined that this collection contained four independent loci, represented by the mutations *bmr2*, *bmr6*, *bmr12* and *bmr19*. In order to reflect the allelic relationships among the collection of *bmr* mutants, these authors proposed that *bmr2*, *bmr6*, *bmr12* and *bmr19* be referred to as reference (*ref*) alleles. For example, by renaming the allelic mutants *bmr2* and *bmr14* as *bmr2-ref* and *bmr2-14*, it would be clear that they represented two different alleles of the same gene, while reflecting the original designation from Porter et al. (1978). An additional set of *bmr* mutants was identified in a mutagenized population developed by Xin et al. (2008). Allelism tests with the *bmr* mutants in this population identified an additional four *bmr* loci (Sattler et al., 2014).

The first sorghum *Brown midrib* gene to be cloned was *Bmr12*, which was shown to encode the enzyme caffeic acid *O*-methyltransferase (COMT), the enzyme, despite its name, responsible for methylating the 5-hydroxyl moiety of 5-hydroxyconiferyl alcohol and 5-hydroxyconiferyl aldehyde, the precursors of sinapyl alcohol (Bout & Vermerris, 2003). The *bmr12* mutation is a nonsense mutation, causing a premature stop codon in the mRNA transcript that would result in a truncated, inactive enzyme. This, combined with the fact that there is only a single *COMT* gene in the sorghum genome (Sb07g003860), explains the strong reduction in syringyl residues in this mutant. The first evidence that sorghum biomass from a *bmr* mutant was more amenable to enzymatic saccharification was provided by Vermerris et al. (2007). Ground stover from the *bmr12* and *bmr6* mutants subjected to enzymatic saccharification for 72 hours resulted in a 25% increase in the amount of glucose and a 75% increase in the amount of xylose relative to wild-type stover. Dien et al. (2009) subsequently showed that combining the *bmr6* and *bmr12* mutations in a double mutant had an additive effect on the efficiency of enzymatic saccharification and ethanol production. The *bmr6* mutation is a null mutation in the *SbCAD2* gene (Saballos et al., 2009; Sattler et al., 2009), which encodes the main cinnamyl alcohol dehydrogenase (CAD) involved in the reduction of hydroxycinnamaldehydes to their corresponding hydroxycinnamyl alcohols. The lignin in the *bmr6* mutant also contains fewer syringyl residues, because CAD has a greater substrate affinity for sinapaldehyde than for coniferaldehyde (Sattler et al., 2009; Jun et al., 2017).

The importance of the S/G ratio on enzymatic saccharification was further demonstrated by Sattler et al. (2012), who examined several additional *bmr12* mutants identified in the EMS-mutagenized population developed by Xin et al. (2008) that were shown to harbor missense mutations. Similar reductions in lignin concentrations relative to the wild-type control were reported for the mutants *bmr12-34* and *bmr12-35*, yet only biomass from the *bmr12-34* resulted in a greater yield of fermentable sugars following enzymatic saccharification. This difference was attributed to the low S/G ratio of 0.08 in the *bmr12-34* mutant, versus 0.42 and 0.63 for the *bmr12-35* and wild-type, respectively.

The impact of other alleles of *bmr6* on the efficiency of enzymatic saccharification was investigated and shown to improve the yield of fermentable sugars relative to the wild-type control (Scully et al., 2016), but without the subtle variation observed with the different *bmr12* alleles.

The *Bmr2* gene encodes the major 4-coumarate CoA ligase involved in lignin biosynthesis (Saballos et al., 2012). This mutation reduces the concentration of lignin in the biomass without a major impact on lignin subunit composition and improves the yield of fermentable sugars following 48 h of enzymatic saccharification by 17%, compared to 25% for *bmr6* and *bmr12* (Saballos et al., 2008).

Biomass from the *bmr* mutants representing the four novel loci reported by Sattler et al. (2014) do not appear to enhance the efficiency of enzymatic saccharification. Combined with the observed reduction in the yield of fermentable sugars for the *bmr19* mutant (relative to the wild-type control) (Saballos et al., 2008), this demonstrates that reduction in lignin concentration is not guaranteed to enhance biomass conversion properties of sorghum.

A recently reported dominant sorghum mutant, *RED for GREEN* (*RG*), has reduced lignin concentrations in the stem and higher lignin concentrations in the leaves (Petti et al., 2013). The name of this mutant refers to the fact that the leaves display a red color resulting from the accumulation of anthocyanins and 3-deoxyanthocyanidins. Enzymatic saccharification for 48 hours of leaf and stem biomass from this mutant showed increased yields of fermentable sugars from stems and decreased yields of sugars from leaves, relative to the wild-type control. An analysis of the lignin subunit composition of the mutant suggested a slight increase in the S/G ratio, but this value could be affected by changes in other cell wall constituents and will need to be experimentally verified with additional analyses.

29.2.3 Traits that enhance sustainable production of biomass

In addition to biomass yield and biomass composition, traits that limit the need for crop inputs (water, fertilizer, fungicides, pesticides) are important to reduce both the environmental footprint and economic cost associated with biomass production, and hence of the renewable fuels and chemicals derived from them. In this respect, relevant traits include: root system architecture, which influences the ability to take up water and nutrients; canopy architecture and stomatal conductance, which influence the photosynthetic and water use efficiency; resistance against microbial pathogens and insect pests, which tend to be more efficient and effective than chemical and biological methods of control. Genetic studies have identified genes and quantitative trait loci (QTL) affecting these various traits, described in other chapters of this book, that can be exploited in biomass sorghum breeding programs. This section provides a brief summary on resistance against anthracnose, a major disease of sorghum in warm and humid areas around the world that affects all parts of the plant, and that substantially reduces biomass yield and biomass quality in susceptible high-biomass sorghum genotypes.

In sorghum, anthracnose is caused by the fungus *Colletotrichum sublineola* Henn. ex Sacc. & Trotter, a hemi-biotrophic fungus whose hyphae initially grow in between cells, but then penetrate cells and kill them (Crouch & Beirn, 2009). The fungus first kills the leaves, then moves into the stem pith, where it reduces stem integrity, and ultimately in the panicle, causing losses in grain yield losses as high as 70% (Cota et al., 2017; Thomas et al., 1995)

Several bi-parental mapping studies and genome-wide association studies have identified a number of QTL and useful molecular markers linked to anthracnose resistance that can be used as sources of resistance (Cruet-Burgos et al., 2020; Cuevas et al., 2014; Felderhoff et al., 2016; Klein et al., 2001; Mohan et al., 2010; Perumal et al., 2009; Singh et al., 2006; Upadhyaya et al.,

2013). As a result of the genetic diversity within and between pathogen populations (Prom et al. 2012), it is important to identify heritable anthracnose resistance targeting the environment in which the sorghum will be cultivated. Furthermore, if germplasm is screened for anthracnose in the greenhouse, the inoculum used during the screening needs to be representative of the *C. sublineola* population structure in the field (Cruet-Burgos et al., 2020). Stacking of multiple resistance loci will increase the likelihood newly developed germplasm will display anthracnose resistance in different environments.

29.3. Breeding strategies to enhance biomass yield

The goal of breeding high-biomass sorghums is to maximize the yield of fermentable sugars per hectare and to minimize the cost of production. This section will review two different strategies to ensure high biomass yield. A distinction needs to be made between breeding cultivars and hybrids. Cultivars are inbred lines that are propagated via self-pollination. The progeny are identical to the parents. Cultivars are relatively easy to breed via a number of methods, with the pedigree method commonly used. Hybrids are the progeny of two inbred parents. The two advantages hybrids offer are hybrid vigor (heterosis), which benefits yield, and protection of the intellectual property of the breeder (seed company), since the (proprietary) inbred parents are needed to generate additional seed. The disadvantage is that the production of hybrids requires more upfront effort. This is because sorghum is principally a self-pollinated species. In order to create hybrids, a male-sterile female line, referred to as A-line, needs to be developed. The use of cytoplasmic male sterility has made commercial hybrid sorghum production feasible since the 1950s (Smith & Frederiksen, 2000). The A-line can only be fertilized by a different plant, which is accomplished in commercial seed production by planting strips of A-lines in between strips of fertile lines. For propagation of A-lines, an isogenic fertile B-line is used. For hybrid seed production (on the A-line) a restorer R-line is used as male parent. This line, ideally genetically distinct to maximize heterosis, restores fertility so that the hybrid offspring are able to produce seed. If seed production is not desirable, which may be the case for high-biomass sorghums, it would also be possible to use as male line an inbred parent that does not have the ability to restore fertility. The hybrid plants will then produce panicles that do not produce pollen.

The breeding of hybrid sorghums relies on evaluating the combining ability of different inbred lines, and the ability of one of the parents to restore fertility. For evaluation purposes, crosses between inbreds can initially be made by removing anthers manually or, in warm climates, by placing a plastic bag over a panicle to induce anther dehiscence. Once an inbred appears promising, it can be converted to a male-sterile A-line via at least five backcrosses to a cytoplasmic male sterile line, ideally of similar pedigree. Generation of new A- and B-lines is a lengthy process, and for that reason, it is common to maintain a collection, while the focus is on developing new R-lines that generate superior hybrids.

29.3.1 Photoperiod-sensitive biomass hybrids

Photoperiod-sensitive sorghums planted under long-day conditions have great potential for high biomass yields due to the continuing vegetative growth until day length becomes short enough to induce the transition to the reproductive phase. In principle, any photoperiod sorghum inbred with a suitable plant architecture (e.g. stem diameter, canopy architecture, root system

architecture) has excellent biomass potential. Heights of 5-6 m and dry biomass yields of up to 60 Mg/ha have been reported (Snider et al., 2012).

Production of photoperiod-sensitive hybrid sorghum seed can be complicated when seed production involves photoperiod-sensitive inbred parents and seed quality may be compromised by early-season frost. A solution is to produce the seed in an off-season nursery relatively close to the equator, when day length is short.

An elegant alternative for the breeding of photoperiod-sensitive hybrids was proposed by Rooney and Aydin (1999), who reported that inbred line EBA-3, a photoperiod-*in*sensitive grain sorghum from Argentina, when crossed with other photoperiod-*in*sensitive inbred lines, generated photoperiod-sensitive hybrids. This occurred because the genotype at the maturity loci *Ma5* and *Ma6* (see section 2.1.2) of EBA-3 was *ma5ma5/Ma6Ma6* and most of the other inbreds were *Ma5Ma5/ma6ma6,* which resulted in *Ma5ma5/Ma6ma6* hybrids. After photoperiod response evaluation of F_1, F_2 and $F_{2:3}$ populations, the authors concluded that the two independent loci *Ma5* and *Ma6* interact in complementary dominant epistasis. This finding meant that there is no need for off-season nurseries for seed production. Instead, new photoperiod-sensitive hybrids can be developed from crossing photoperiod-insensitive lines that are homozygous for contrasting alleles at the *Ma5* and *Ma6* loci (Fig. 29.5, Scheme I). This also means that the inbred parent lines can be short. The male lines Tx2909 and Tx2910 are publicly available lines with *ma5*/*Ma6* derived from EBA-3 (Hawkins, 2013).

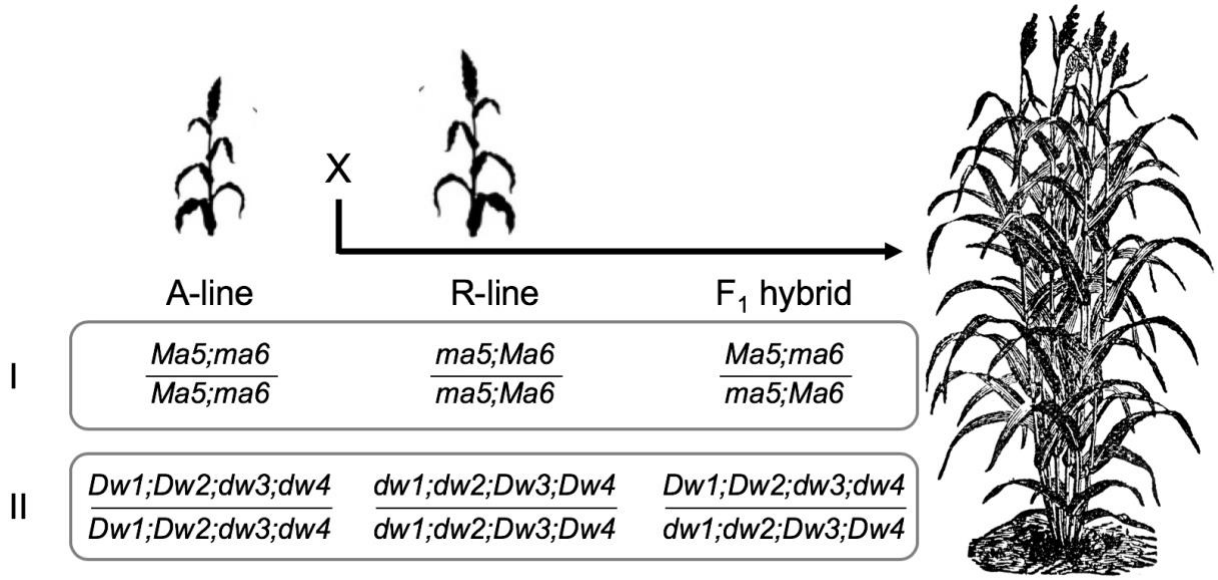

Fig. 29.5 Two ways of using short inbred parent lines that are homozygous recessive at complementary loci to produce tall F_1 biomass hybrids. **Scheme I** depicts the strategy of Rooney and Aydin (1999), with A- and R-lines that are both photoperiod insensitive, so that seed can be produced during the summer season in regions with long days. The F_1 hybrid progeny is photoperiod sensitive, and will continue to grow vegetatively under long days, reaching heights over 4 m. **Scheme II** depicts the strategy by authors Vermerris and Silva, whereby short, photoperiod insensitive two-dwarf inbred lines with contrasting *dwarf* genotypes are crossed to produce an F_1 hybrid heterozygous at all four *dwarf* loci, which is, therefore, tall. Hybrids reaching heights of up to 4 m, that flowered 90-100 days after planting have been generated.

29.3.2 Tall biomass hybrids derived from short inbred parents

Two potential drawbacks associated with the use of photoperiod-sensitive hybrids are their late maturity and extreme height. The consequence of late maturity is that the biomass remains metabolically active for a longer period than photoperiod-insensitive genotypes, which translates in high moisture contents at the time of harvest. This increases the cost of transportation and has the risk of rot during storage of the biomass. The extreme height, a main contributor to biomass yield, can be a disadvantage in areas with a high frequency of strong winds during the growing season, due to the risk of lodging. This risk is elevated in coastal areas in subtropical regions, such as the U.S. Gulf coast, the U.S. Atlantic coast south of North Carolina, Mexico and the islands of the Caribbean, Hawaii, and the Philippines.

An alternative strategy for generating biomass hybrids that addresses these concerns is the use of inbred lines with complementary *dwarf* loci. In this approach, combine-compatible two-dwarf A- and R-lines are crossed to produce zero-dwarf hybrid offspring, because the A- and R-lines are homozygous recessive for contrasting dwarf loci, as illustrated in Fig. 29.5, Scheme II. The authors have used this approach in their high-biomass sorghum breeding program in North Florida, and used an A-line with a height of 1.5 m and an R-line with a height of 1.7 m to generate hybrids that reach heights of up to 4 m, and that flower 90-100 days after planting. Biomass yields appear promising based on initial small-scale evaluations, and multi-location trials are planned.

29.4. Analytical methods to assess biomass quality

The cell wall composition and saccharification efficiency need to be evaluated in order to assess the biomass improvements achieved by breeding. Currently, there are several screening methods, which differ by equipment, throughput, complexity of the data analysis, specificity and sensitivity. The main methods are described in this section.

29.4.1 Near-infrared reflectance (NIR) and Fourier-transform infrared (FTIR) spectroscopy

Near-infrared reflectance (NIR) and Fourier-transform infrared (FTIR) spectroscopy are vibrational spectroscopic techniques that are used to infer the chemical composition of solid and liquid samples based on the amount of light that is absorbed in the near-infrared range of the electromagnetic spectrum (800-2500 nm; NIRS), or the mid-infrared region of the spectrum (2500-4000 nm; FTIR). The light absorbance is associated with specific bend and stretch vibrations of molecular bonds as long as they alter the dipole moment of the molecule (Siesler et al., 2012).

One of the advantages of NIR spectroscopy is that it can be performed in a non-destructive manner, i.e. on plant tissues collected without killing the plant (typically leaves). Samples used for NIR spectroscopy can be used for subsequent analyses afterwards. Furthermore, acquisition of NIR spectra can be completed within minutes, allowing a high throughput. The method does not use hazardous chemicals and does not require *a priori* knowledge of the nature of the compositional differences among samples.

NIR spectroscopy can penetrate deeper layers of a sample and is sensitive to aromatic compounds such as lignin. For that reason, the technique is normally used for the analysis of adult plants grown for forage or to produce biomass for bioenergy. NIR spectroscopy has been used for

chemical characterization of biomass feedstocks from several species, such as corn, sorghum, rice and Miscanthus (Payne & Wolfrum, 2015; Vermerris et al., 2007; Vermerris & Saballos, 2013).

The use of NIRS for the prediction of biomass composition requires the development of a model in which NIR absorbance values are associated quantitatively with data on the chemical composition obtained with more traditional (wet-chemical) methods. Infrared spectra of complex samples such as sorghum biomass are composed of hundreds of data points. Furthermore, the absorbances from the same functional groups at different wavelengths (so-called overtones) can be strongly correlated, making it difficult to characterize a sample based on a small set of absorbance values at specific wavelengths. Hence, multivariate statistical processes are generally used to develop models that can be used to predict biomass composition based on NIRS. overcome this problem. It is important to build the model with a subset of samples that captures the range in composition within the population of samples and to use highly standardized wet-chemical protocols to determine the biomass compositional.

For cell wall composition analysis of young plants, techniques such as Fourier-transform infrared (FTIR) spectroscopy are preferable. FTIR spectra can be obtained by pressing finely ground samples in a thin potassium bromide disk and measuring the absorbance of infrared light, by drying a cell wall suspension on a BaF_2 microscope slide, or by placing a suspension of cell walls on a gold-plated reflective surface under a microscope with an FTIR spectrometer (Sené et al., 1994; Vermerris et al., 2002; Yong et al., 2005)

The limitations of FTIR are its semiquantitative nature and overlap in absorption and vibrational coupling between chemical bonds corresponding to different cell wall polymers (Alonso-Simón et al., 2011). FTIR spectroscopy has been used to characterize the composition of biomass sorghum (Balogun, et al., 2014) and to identify changes in chemical composition structures after pretreatment and enzymatic hydrolysis (Corredor et al., 2009; Jamaldheen et al., 2018), including in sorghum.

29.4.2 Analytical pyrolysis

Pyrolysis is the thermal degradation of a compound at temperatures above 500°C under anoxic conditions, generating a volatile pyrolysate (Evans & Milne, 1987; Boon, 1989). The pyrolysate generated with the help of a small ceramic oven or heated filament can go directly into a mass spectrometer (Py-MS) or into a gas chromatograph coupled to a mass spectrometer (Py-GC-MS) for identification and quantification. The low-molecular weight compounds of a pyrolysate are break-down products of polysaccharides and lignin (Boon, 1989; Meier & Faix, 1992, Ralph & Hatfield, 1991). Fragments derived from cell wall polysaccharides undergo rearrangements, and although fragments derived from hexoses and pentoses can typically be easily identified, it is generally difficult to determine their exact origin. On the other hand, lignin and hydroxycinnamic acids maintain their substitution pattern on the benzene ring, making it easy to identify them as originating from *p*-hydroxyphenyl, guaiacyl or syringyl residues (Boon, 1989; Ralph & Hatfield, 1991). It is important to be aware that the cell walls from grasses contain substantial amounts of esterified *p*-coumaric acid (Fig. 29.4) and ferulic acid (Fig. 29.3) that result in the pyrolytic formation of 4-vinylphenol and 4-vinylguaiacol, respectively, and that should not be confused with pyrolysis fragments derived from H- and G-residues in the lignin. The application of tetramethylammonium hydroxide (TMAH; 2.5% (v/v) in methanol) to the sample prevents the

decarboxylation reaction leading to the vinyl moieties and enables the distinction in origin (Mulder et al., 1992; Sattler et al., 2014; Vermerris & Boon, 2001).

Due to the differences in the pyrolytic fragments derived from polysaccharides and lignin, it is possible to pyrolyze whole biomass samples. Another advantage is that it requires only small samples (10-1000 µg) (Mulder et al., 1992; Lapierre, 1993). Py-MS is faster than Py-GC-MS, since it requires only a few minutes to analyze one sample compared to 40-60 minutes required by Py-GC-MS, being a better option for high-throughput analyses. However, different fragment ions with the same mass-to-charge (*m/z*) ratio can be resolved only in Py-GC-MS (Vermerris & Saballos, 2013).

29.4.3 High-throughput pretreatment and saccharification assays

The conventional methods for pretreatment and hydrolysis are laborious and time-consuming. Conventionally, raw biomass is subjected to temperatures above 140ºC, then solids and liquids are separated by filtration. After washing the solids, cellulolytic enzymes are added to the mixture and the liquids go through a post-hydrolysis stage. In addition to the time commitment necessary to perform these steps, the methods used to measure the sugar left in the solids are tedious (Studer et al., 2010). Therefore, there has been interest in developing small-scale, high-throughput methods that are fast and automated.

Selig et al. (2010) developed a 96-well multiplate to perform hydro-thermal pretreatment and enzymatic saccharification in a single reactor. The system relies on stackable nickel/gold-plated 96-well aluminum reactor plates and a clamping system that holds up to 20 stacked plates together, allowing for up to 1920 individual sugar analyses per run, all fit to a modified two-gallon Parr reactor. The system has also contiguous steam ports to facilitate steam transport throughout stacks when heated, and water transport for cooling after pretreatment. Each individual reactor plate is sealed with high-temperature aluminum foil tape to prevent evaporative losses, condensation or water incursions. For liquids and solids handling, a Powdernium powder dispensing system (Symyx, Geneva, Switzerland) and a Biomek® FX automated pipetting system are used. The authors could determine amounts of end-products such as glucose and xylose rapidly when the pretreatment and saccharification were performed using this 96-well multiplate system (Selig et al., 2010, 2011).

Studer et al. (2010) reported a similar approach in which a steam heating and water quenching system is applied to a 96-well plate. This method also enabled sequential pretreatment and enzymatic hydrolysis, without the need for solid/liquid separation and solid washing in between. Santoro et al. (2010) developed a custom-designed robot called iWALL that can grind and weigh 1-5 mg of plant tissue samples of more than 243 plants in 16 h. The iWALL has one 96-tube rack of input vials and three 96-tube racks of output vials. An automated workstation is used to perform pretreatment, hydrolysis and sugar analysis, which can be completed in 36 h. The system allows analysis of around 970 biomass samples in a week.In addition to the high throughput and ease of use of the abovementioned systems, they only require small amounts of samples, with Selig et al. (2010) using 5 mg, Santoro et al. (2010) 1-5 mg, and Studer et al. (2010) 2.6 mg of biomass. The sugar yields obtained with these high-throughput systems are similar to conventional methods (Santoro et al., 2010; Studer et al., 2010). These high-throughput screenings have been

used for pretreatment and saccharification of biomass from species such as *Populus*, oilseed rape, maize and wheat straws (Studer et al., 2010; Santoro et al., 2010; Elliston et al., 2015), and are also applicable to the analysis of sorghum biomass.

29.5. Future perspectives

The use of biomass sorghum as a feedstock for the production of renewable fuels and chemicals is an attractive alternative because of the potential to cultivate sorghum on low-productivity land, which minimizes competition with food production. Furthermore, the comparatively low-input requirements, tolerance to biotic and abiotic stresses, and excellent yield potential contribute towards sustainable crop production. In order to achieve the crop's full potential, its genetic diversity needs to be fully exploited and breeding strategies need to be implemented to improve traits such as canopy and root system architectures, cell wall composition and disease and pest resistance. Therefore, the elucidation of metabolic pathways and signaling cascades influencing these traits, as well as identifying the loci that control them, is extremely important. The use of novel techniques, such as genome editing, to discover, study and manipulate these genes can expedite the breeding process. For instance, the use of the CRISPR/Cas9 genome editing system (Jinek et al., 2012; Jiang et al., 2013) has potential to introduce precise changes in genes with the aim of, for example, altering catalytic properties of enzymes. The implementation of this approach, especially when it can be done in a way that avoids the resulting plants being labeled as transgenic, has the potential to lead to novel genetic variation that can complement traditional breeding methods.

Acknowledgements
The authors gratefully acknowledge funding from U.S. Department of Energy's Office of Energy Efficiency and Renewable Energy, Bioenergy Technologies Office and sponsored by the US DOE's International Affairs under award no. DE-PI0000031 for research on high-biomass sorghums that can be produced sustainably, and from the U.S. Department of Energy (BER) grant nos. DE-SC0014439 and DE-SC0019097 for research on anthracnose resistance. Tallyta N. Silva is grateful for financial support from CAPES Foundation (Brazilian Ministry of Education) and the Science without Borders program (BEX-1883-13-8), as well as from the University of Florida Genetics Institute.

29.6. References

Aden A, Ruth M, Ibsen K, Jechura J, Neeves K, Sheehan J, Wallace B, Montague L, Slayton A, Lukas J (2002) Lignocellulosic Biomass to Ethanol Process Design and Economics Utilizing Co-Current Dilute Acid Prehydrolysis and Enzymatic Hydrolysis for Corn Stover. NREL, Golden, Colorado. https://www.nrel.gov/docs/fy02osti/32438.pdf

Ahmad R, Liow PS, Spencer DF, Jasieniuk M (2008) Molecular evidence for a single genetic clone of invasive Arundo donax in the United States. Aquatic Bot 88:113–120 http://doi.org/10.1016/j.aquabot.2007.08.015

Alonso-Simón A, García-Angulo P, Mélida H, Encina A, Álvarez JM, Acebes JL (2011) The use of FTIR spectroscopy to monitor modifications in plant cell wall architecture caused by cellulose biosynthesis inhibitors. Plant Signaling Behavior 6:1104–1110 http://doi.org/10.4161/psb.6.8.15793

Balogun AO, Lasode OA., Li H, McDonald AG (2014) Fourier Transform Infrared (FTIR) study and thermal decomposition kinetics of *Sorghum bicolor* glume and *Albizia pedicellaris* residues. Waste Biomass Valorization 6:109–116 http://doi.org/10.1007/s12649-014-9318-3

Bensen RJ, Johal GS, Crane VC, Tossberg JT, Schnable PS, Meeley RB, Briggs SP (1995) Cloning and characterization of the maize *An1* gene. The Plant Cell 7:75–84 https://doi.org/10.1105/tpc.7.1.75

Bishop GJ, Nomura T, Yokota T, Harrison K, Noguchi T, Fujioka S, Takatsuto S, Jones JDG, Kamiya Y (1999) The tomato DWARF enzyme catalyses C-6 oxidation in brassinosteroid biosynthesis. Proceedings of the National Academy of Sciences USA 96, 1761–1766. https://doi.org/10.1073/pnas.96.4.1761

Bout S, Vermerris W (2003) A candidate-gene approach to clone the sorghum *Brown midrib* gene encoding caffeic acid *O*-methyltransferase. Mol Genetics Genomics 269: 205–214 http://doi.org/10.1007/s00438-003-0824-4

Braconnier S, Gutjard S, Trouche G, Reedy B, Rao S, Schaffert R, Parella R, Zacharias A, Rettenmaier N, Reinhardt G, Monti A, Amaducci S, Marocco A, Snijman W, Terblanche H, Zavala-Garcia F, Janssen R, Rutz D (2011) Definiton of new sorghum ideotypes to meet the increasing demand of biofuels. In: peoceedings of 19th European Biomass Conference and Exhibition, pp. 782-786 http://www.etaflorence.it/proceedings/?detail=7504

Brown PJ, Rooney WL, Franks C, Kresovich S (2008) Efficient mapping of plant height quantitative trait loci in a sorghum association population with introgressed dwarfing genes. Genetics 637:629–637 http://doi.org/10.1534/genetics.108.092239

Burner DM, Hale AL, Viator RP, Belesky DP, Houx JH, Ashworth AJ, Fritschi FB.(2017) Ratoon cold tolerance of Pennisetum, Erianthus, and Saccharum bioenergy feedstocks. Indust Crops Products 109:327–334 http://doi.org/10.1016/j.indcrop.2017.08.020

Burton RA, Wilson SM, Hrmova M, Harvey AJ, Shirley NJ, Medhurst A, Stone BA, Newbigin EJ, Bacic A, Fincher GB (2006) Cellulose synthase–like *CslF* genes mediate the synthesis of cell wall (1,3:1,4)-β-D-glucans. Sci 311:1940–1942. http://doi.org/10.1126/science.1122975

Carpita NC, Gibeaut DM (1993) Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth. The Plant J 3(1):1–30 http://doi.org/10.1111/j.1365-313X.1993.tb00007.x

Castro E, Nieves IU, Rondón V, Sagues WJ, Fernández-Sandoval MT, Yomano LP, York SW, Erickson J, Vermerris W (2017) Potential for ethanol production from different sorghum cultivars. Indust Crops Products 109:367–373 http://doi.org/10.1016/j.indcrop.2017.08.050

Chakravorty U, Hubert MH, Nostbakken L (2009) Fuel Versus Food. Ann Rev Resource Econom 1:645–663 http://doi.org/10.1146/annurev.resource.050708.144200

Childs KL, Miller FR, Cordonnier-Pratt MM, Pratt LH, Morgan PW, Mullet JE (1997) The sorghum photoperiod sensitivity gene, *Ma3*, encodes a phytochrome B. Plant Physiol 113:611–619

Clifton-Brown JC, Lewandowski I, Andersson B, Basch G, Christian DG, Kjeldsen JB, Jorgensen U, Mortensen JV, Riche AB, Schwarz KU, Tayebi K, Teixeira F (2001). Performance of 15 Miscanthus genotypes at five sites in Europe. Agron J 93:1013–1019

Constant S, Wienk HLJ, Frissen AE, Peinder P de, Boelens R, et al. (2016) New insights into the structure and composition of technical lignins: a comparative characterisation study. Green Chemistry 18:2651–2665 http://doi.org/10.1039/c5gc03043

Corredor DY, Salazar JM, Hohn KL, Bean S, Bean B, Wang D (2009) Evaluation and characterization of forage sorghum as feedstock for fermentable sugar production. App Biochem Biotech 158(1):164–179 http://doi.org/10.1007/s12010-008-8340-y

Cota LV, Souza AGC, Costa RV, Silva DD, Lanza FE, Aguiar FM, Figueiredo JEF (2017) Quantification of yield losses caused by leaf anthracnose on sorghum in Brazil. J Phytopath 165:479–485

Crouch JA, Beirn LA (2009) Anthracnose of cereals and grasses. Fungal Diversity 39:19–44

Cruet-Burgos CM, Cuevas HE, Prom LK, Knoll JE, Stutts LR, Vermerris W (2020) Genomic dissection of anthracnose (*Colletotrichum sublineolum*) resistance response in sorghum differential line SC112-14.

G3 [Genes, Genomes, Genetics] 10:1403-1412

Cuevas HE, Prom LK, Erpelding JE (2014) Inheritance and molecular mapping of anthracnose resistance genes present in sorghum line SC112-14. Mol Breed 34:1943–1953 http://doi.org/10.1007/s11032-014-0151-y

Cuevas HE, Zhou C, Tang H, Khadke PP, Das S, Lin YR, Ge Z, Clemente T, Upadhyaya HD, Hash CT, Paterson AH (2016) The evolution of photoperiod-insensitive flowering in sorghum, a genomic model for panicoid grasses. Mol Bio Evolution 33:2417–2428 http://doi.org/10.1093/molbev/msw120

Dien BS, Sarath G, Pedersen JF, Sattler SE, Chen H, Funnell-Harris DL, Nichols NN, Cotta MA (2009) Improved sugar conversion and ethanol yield for forage sorghum (*Sorghum bicolor* (L.) Moench) lines with reduced lignin contents. BioEnergy Res 2:153–164

Elliston A, Wood IP, Soucouri MJ, Tantale RJ, Dicks J, Roberts IN, Waldron KW (2015) Methodology for enabling high-throughput simultaneous saccharification and fermentation screening of yeast using solid biomass as a substrate. Biotech Biofuels 8:1–9 http://doi.org/10.1186/s13068-014-0181-z

Evans RJ, Milne TA (1987) Molecular characterization of the pyrolysis of biomass. 1. Fundamentals Energy Fuels 1:123–137 http://doi.org/10.1021/ef00004a001

FAO (2009) How to feed the world in 2050? Insights from an expert meeting at FAO. http://doi.org/10.1111/j.1728-4457.2009.00312.x

Felderhoff TJ, McIntyre LM, Saballos A, Vermerris W (2016) Using genotyping by sequencing to map two novel anthracnose resistance loci in *Sorghum bicolor*. G3 [Genes Genomes Genetics] 6:1935–1946 http://doi.org/10.1534/g3.116.030510

Fujioka S, Yamane H, Spray CR, Gaskin P, Macmillan J, Phinney B, Takahashi N (1988) Qualitative and quantitative analyses of gibberellins in vegeetatitve shoots of normal, *dwarf-1*, *dwarf-2*, *dwarf-3*, and *dwarf-5* seedlings of *Zea mays* L. Plant Physiol 88:1367–1372

Graham D, Lessman KJ (1966) Effect of height on yield and yield components of two isogenic lines of *Sorghum vulgare*. Crop Sci 6: 372–374 http://doi.org/10.2135/cropsci1966.0011183X000600040024x

Hao B, Xue Q, Bean BW, Rooney WL, Becker JD (2014) Biomass production, water and nitrogen use efficiency in photoperiod-sensitive sorghum in the Texas High Plains. Biomass Bioenergy 62:108–116 http://doi.org/10.1016/j.biombioe.2014.01.008

Hatfield R, Ralph J, Grabber JH (2008) A potential role for sinapoyl *p*-coumarate as a radical transfer mechanism in grass lignin formation. Planta 228:919–928 http://doi.org/10.1007/s00425-008-0791-4

Hawkins EM (2013) Genetic control of flowering time and biomass yield in sorghum. Ph.D. dissertation, University of Illinois. Accessed via: https://core.ac.uk/download/pdf/19529763.pdf

Hilley JL, Weers BD, Truong SK, McCormick RF., Mattison AJ, McKinley B, Morishige DT, Mullet JE (2017) Sorghum Dw2 encodes a protein kinase regulator of stem internode length. Scientific Reports 7: 1–13 http://doi.org/10.1038/s41598-017-04609-5

Hirano K, Kawamura M, Araki-Nakamura S, Fujimoto H, Ohmae-Shinohara K, Yamaguchi M, Fujii A, Sasaki H, Kasuga S, Sazuka T (2017) Sorghum DW1 positively regulates brassinosteroid signaling by inhibiting the nuclear localization of BRASSINOSTEROID INSENSITIVE 2. Scientific Reports 7(126):1–10 http://doi.org/10.1038/s41598-017-00096-w

Hu F, Ragauskas A (2012) Pretreatment and lignocellulosic chemistry. BioEnergy Res 5:1043–1066 http://doi.org/10.1007/s12155-012-9208-0

Huang CF, Lin TH, Guo GL, Hwang WS (2009) Enhanced ethanol production by fermentation of rice straw hydrolysate without detoxification using a newly adapted strain of Pichia stipitis. Bioresource Tech 100:3914–3920 http://doi.org/10.1016/j.biortech.2009.02.064

Jamaldheen SB, Sharma K, Rani A, Moholkar VS, Goyal A (2018) Comparative analysis of pretreatment methods on sorghum (*Sorghum durra*) stalk agrowaste for holocellulose content. Preparative Biochem Biotech 48:457-464 http://doi.org/10.1080/10826068.2018.1466148

Jang YS, Malaviya A, Cho C, Lee J, Lee SY (2012) Butanol production from renewable biomass by clostridia. Bioresource Technol 123:653–663 http://doi.org/10.1016/j.biortech.2012.07.104

Jiang W, Zhou H, Bi H, Fromm M, Yang B, Weeks DP (2013) Demonstration of CRISPR/Cas9/sgRNA-mediated targeted gene modification in Arabidopsis, tobacco, sorghum and rice. Nucleic Acids Res 41: e188

Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Sci 337:816–821

Jun SY, Walker AM, Kim H, Ralph J, Vermerris W, Sattler SE, Kang C (2017) The enzyme activity and substrate specificity of two major cinnamyl alcohol dehydrogenases in sorghum (*Sorghum bicolor*), SbCAD2 and SbCAD4. Plant Physiol 174:2128–2145

Karimi K, Emtiazi G, Taherzadeh MJ (2006) Ethanol production from dilute-acid pretreated rice straw by simultaneous saccharification and fermentation with *Mucor indicus*, *Rhizopus oryzae*, and *Saccharomyces cerevisiae*. Enzyme Microbial Technol 40: 138–144 http://doi.org/10.1016/j.enzmictec.2005.10.046

Kebrom TH, Burson BL, Finlayson SA (2006) Phytochrome B represses *Teosinte branched1* expression and induces sorghum axillary bud outgrowth in response to light signals. Plant Physiol 140:1109–1117 http://doi.org/10.1104/pp.105.074856

Klein RR, Mullet JE, Jordan DR, Miller FR, Rooney WL, Menz MA, Franks CD, Klein PE (2008) The effect of tropical sorghum conversion and inbred development on genome diversity as revealed by high-resolution genotyping. The Plant Genome 48:S12–S26 http://doi.org/10.2135/cropsci2007.06.0319tpg

Klein RR, Rodriguez-Herrera, R, Schlueter JA, Klein PE, Yu ZH, Rooney WL (2001) Identification of genomic regions that affect grain-mould incidence and other traits of agronomic importance in sorghum. Theor Appl Genetics 102:307–319 http://doi.org/10.1007/s001220051647

Lambin EF, Meyfroidt P (2011) Global land use change, economic globalization, and the looming land scarcity. Proced National Academy Sci USA. 108: 3465–3472 http://doi.org/10.1073/pnas.1100480108

Lan EI, Liao JC (2013) Microbial synthesis of *n*-butanol, *iso*-butanol, and other higher alcohols from diverse resources. Bioresource Technol 135:339–349 http://doi.org/10.1016/j.biortech.2012.09.104

Lan W, Morreel K, Lu F, Rencoret J, del Río JC, Voorend W, Vermerris W, Boerjan W, Ralph J (2016) Maize tricin-oligolignol metabolites and their implications for monocot lignification. Plant Physiol 171: 810-820 http://doi.org/10.1104/pp.16.02012

Li J, Nagpal P, Vitart V, McMorris TC, Chory J (1996) A role for brassinosteroids in light-dependent development of Arabidopsis. Sci 272:398–401 http://doi.org/10.1126/science.272.5260.398

Lin YR, Schertz KF, Paterson AH (1995) Comparative analysis of QTLs affecting plant height and maturity across the Poaceae, in reference to an interspecific sorghum population. Genetics,141:391–411

Mantilla Perez MB, Zhao J, Yin Y, Hu J, Salas Fernandez MG (2014) Association mapping of brassinosteroid candidate genes and plant architecture in a diverse panel of *Sorghum bicolor*. Theor App Genetic 127:2645–2662 http://doi.org/10.1007/s00122-014-2405-9

Matsuoka S, Kennedy AJ, Santos EGD dos, Tomazela AL, Rubio LCS (2014) Energy cane: its concept, development, characteristics, and prospects. Advan Bot 2014: Article ID 597275

McCollum T, Mccuistion K, Bean B (2005) *Brown midrib* and photoperiod-sensitive forage sorghums. In: Proceed 2005 Plains Nutrition Council Spring Conference. San Antonio, Texas, USA, p10

McIntosh S, Vancov T (2010) Enhanced enzyme saccharification of Sorghum bicolor straw using dilute alkali pretreatment. Bioresource Technol 101:6718–6727 http://doi.org/10.1016/j.biortech.2010.03.116

Meki MN, Ogoshi RM, Kiniry JR, Crow SE, Youkhana AH, Nakahata MH, Littlejohn K (2017) Performance evaluation of biomass sorghum in Hawaii and Texas. Industrial Crops Products 103:257–266. http://doi.org/10.1016/j.indcrop.2017.04.014

Mohan SM, Madhusudhana R, Mathur K, Howarth CJ, Srinivas G, Satish K, Reddy RN, Seetharama N (2009).Co-localization of quantitative trait loci for foliar disease resistance in sorghum. Plant Breed 128: 532–535 http://doi.org/10.1111/j.1439-0523.2008.01610.x

Morgan JLW, Strumillo J, Zimmer J (2013) Crystallographic snapshot of cellulose synthesis and membrane translocation. Nature 493:181–186 http://doi.org/10.1038/nature11744

Morris GP, Ramu P, Deshpande SP, Hash CT, Shah T, Upadhyaya HD, Riera-Lizarazu O, Brown PJ, Acharya CB, Mitchell SE, Harriman J, Glaubitz JC, Buckler ES, Kresovich S (2013) Population genomic and genome-wide association studies of agroclimatic traits in sorghum. Proceed National Acad Sci USA 110:453–458 http://doi.org/10.1073/pnas.1215985110

Motlhaodi T, Geleta M, Chite S, Fatih M, Ortiz R, Bryngelsson T (2017) Genetic diversity in sorghum [*Sorghum bicolor* (L .) Moench] germplasm from Southern Africa as revealed by microsatellite markers and agro-morphological traits. Genetic Resource Crop Evolution 64: 599–610 http://doi.org/10.1007/s10722-016-0388-x

Mueller SC, Brown RM (1980) Evidence for an intramembrance component associated with a cellulose microfibril-synthesizing complex in higher plants. J Cell Biol 84:315–326

Mulder MM, Van Der Hage ERE, Boon JJ (1992) Analytical in-source pyrolytic methylation electron impact mass spectrometry of phenolic acids in biological matrices. Phytochemical Analysis 3:165–172 http://doi.org/10.1002/pca.2800030405\

Mullet JE (2017) High-biomass C4 grasses — Filling the yield gap. Plant Sci. 261:10–17 http://doi.org/10.1016/j.plantsci.2017.05.003

Mullet JE, Rooney WL, Klein PE, Morishige D, Murphy R, Brady JA (2010) Discovery and utilization of sorghum genes (*Ma5*/*Ma6*). U.S. patent 8,309,793 B2

Multani DS, Briggs SP, Chamberlin MA, Blakeslee JJ, Murphy AS, Johal GS (2003) Loss of an MDR transporter in compact stalks of maize *br2* and sorghum *dw3* mutants. Sci 302: 81–84 http://doi.org/10.1126/science.1086072

Murphy RL, Klein RR, Morishige DT, Brady JA, Rooney WL, Miller FR, Dugas DV, Klein PE, Mullet J E (2011) Coincident light and clock regulation of pseudoresponse regulator protein 37 (PRR37) controls photoperiodic flowering in sorghum. Proceed National Acad Sci USA. 108:16469–16474 http://doi.org/10.1073/pnas.1106212108

Murphy RL, Morishige DT, Brady JA, Rooney WL, Yang S, Klein PE, Mullet JE (2014) *Ghd7* (*Ma6*) represses sorghum flowering in long days: *Ghd7* alleles enhance biomass accumulation and grain production. The Plant Genome 7:1–10 http://doi.org/10.3835/plantgenome2013.11.0040

Murray SC, Rooney WL, Mitchell SE, Sharma A, Klein PE, Mullet JE, Kresovich SK (2008) Genetic improvement of sorghum as a biofuel feedstock: II. QTL for stem and leaf structural carbohydrates. Crop Sci 48:2180-2193

Narayanan S, Aiken R M, Vara Prasad PV, Xin Z, Yu J (2013) Water and radiation use efficiencies in sorghum. Agron J 105:649–656 http://doi.org/10.2134/agronj2012.0377

OECD/FAO (2015) OECD-FAO Agricultural Outlook 2015-2024 http://www.fao.org/3/a-i4738e.pdf

Okuno A, Hirano K, Asano K, Takase W, Masuda R, Morinaka Y, Ueguchi-Tanaka M, Kitano H, Matsuoka M (2014) New approach to increasing rice lodging resistance and biomass yield through the use of high gibberellin producing varieties. PloS One 9(2):e86870 http://doi.org/10.1371/journal.pone.0086870

Olek AT, Rayon C, Makowski L, Kim HR, Ciesielski P, Badger J, Paul LN, Ghosh S, Kihara D, Crowley M, Himmel ME, Bolim JT, Carpita NC (2014) The structure of the catalytic domain of a plant cellulose synthase and its assembly into dimers. The Plant Cell 26: 2996-3009 http://doi.org/10.1105/tpc.114.126862

Olson SN, Ritter KB, Rooney W, Kemanian A, McCarl BA, Zhang Y, Hall S, Packer D, Mullet J (2012) High biomass yield energy sorghum: developing a genetic model for C4 grass bioenergy. Biofuels

Bioproducts Biorefining 6:640–655. http://doi.org/10.1002/bbb.1357

Ordonio RL, Ito Y, Hatakeyama A, Ohmae-shinohara K, Kasuga S, Tokunaga T, Mizuno H, Kitano H, Matsuoka M, Sazuka T (2015) Gibberellin deficiency pleiotropically induces culm bending in sorghum: an insight into sorghum semi-dwarf breeding, Scientific Reports 4: 1–10 http://doi.org/10.1038/srep05287

Packer DJ, Rooney WL (2014) High-parent heterosis for biomass yield in photoperiod-sensitive sorghum hybrids. Field Crops Res 167:153–158 http://doi.org/10.1016/j.fcr.2014.07.015

Pao C, Morgan PW (1986) Genetic regulation of development in *Sorghum bicolor*. Plant Physiol 82:575–580

Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, et al. (2009) The *Sorghum bicolor* genome and the diversification of grasses. Nature 457:551-556 http://doi.org/10.1038/nature07723

Payne CE, Wolfrum EJ (2015) Rapid analysis of composition and reactivity in cellulosic biomass feedstocks with near-infrared spectroscopy. Biotech Biofuels 8:1–14. http://doi.org/10.1186/s13068-015-0222-2

Pereira MG, Lee M (1995) Identification of genomic regions affecting plant height in sorghum and maize. Theor App Genetics 90:380–388 http://doi.org/10.1007/BF00221980

Perumal R, Menz MA, Mehta PJ, Katilé S, Gutierrez-Rojas LA, Klein RR, Klein PE, Prom LK, Schlueter A, Rooney WL, Magill CW (2009) Molecular mapping of *Cg1*, a gene for resistance to anthracnose (*Colletotrichum sublineolum*) in sorghum. Euphytica 165:597–606 http://doi.org/10.1007/s10681-008-9791-5

Petrik DL, Karlen SD, Cass CL, Padmakshan D, Lu F, Liu S, Bris PL, Antelme S, Santoro N, Wilkerson C G, Sibout R, Lapierre C, Ralph J, Sedbrook JC (2014) *p*-Coumaroyl-CoA:monolignol transferase (PMT) acts specifically in the lignin biosynthetic pathway in *Brachypodium distachyon*. Plant J 77(5):713–726 http://doi.org/10.1111/tpj.12420

Petti C, Harman-Ware AE, Tateno M, Kushwaha R, Shearer A, Downie AB, Crocker M, Debolt S (2013) Sorghum mutant *RG* displays antithetic leaf shoot lignin accumulation resulting in improved stem saccharification properties. Biotech Biofuels 6:146 http://doi.org/10.1186/1754-6834-6-146

Porter KS, Axtell JD, Lechtenberg VL, Colenbrander VF (1978) Phenotype, fiber composition, and in vitro dry matter disappearance of chemically induced brown midrib (bmr) mutants of sorghum. Crop Sci 18: 205–208 http://doi.org/10.2135/cropsci1978.0011183X001800020002x

Pu Y, Hu F, Huang F, Ragauskas AJ (2015) Lignin structural alterations in thermochemical pretreatments with limited delignification. BioEnergy Res. 8:992–1003 http://doi.org/10.1007/s12155-015-9655-5

Quinby J (1974) Sorghum Improvement and the Genetics of Growth. College Station, Texas: Texas A&M University Press.

Quinby JR (1966) Fourth maturity gene locus in sorghum. Crop Sci. 6:516–518 http://doi.org/10.2135/cropsci1966.0011183X000600060005x

Quinby, J. R. (1967). The maturity genes of sorghum. Advances in Agronomy, 19, 267–305. http://doi.org/10.1016/S0065-2113(08)60737-3

Quinby JR (1975) The genetics of sorghum improvement. J Heredity 66: 56–62 http://doi.org/10.1093/oxfordjournals.jhered.a108582

Quinby JR, Karper RE (1954) Inheritance of height in sorghum. Agron J 46(5):211–216 http://doi.org/10.2134/agronj1954.00021962004600050007x

Ragauskas AJ, Williams CK, Davison BH, Britovsek G, Cairney J. Eckert CA, Frederick Jr WJ, Hallett J P, Leak DJ, Liotta CL, Mielenz JR, Murphy R, Templer R, Tschaplinski T (2006) The path forward for biofuels and biomaterials. Sci 311:484–489

Ralph J, Hatfield RD (1991) Pyrolysis-GC-MS characterization of forage materials. J Agri Food Chem 39:

1426–1437 http://doi.org/10.1021/jf00008a014

Rooney WL, Aydin, S (1999). Genetic control of a photoperiod-sensitive response in *Sorghum bicolor* (L.) Moench. Crop Sci 39:397–400 https://doi.org/10.2135/cropsci1999.0011183X0039000200016x

Rooney WL, Blumenthal J, Bean B, Mullet JE (2007) Designing sorghum as a dedicated bioenergy feedstock. Biofuels, Bioproducts Biorefining 1:147–157 http://doi.org/10.1002/bbb

Rosegrant MW, Msangi S (2014) Consensus and contention in the food-versus-fuel debate. Ann Rev Environ Resour 39:271–294 http://doi.org/10.1146/annurev-environ-031813-132233

Saballos A, Ejeta G, Sanchez E, Kang C, Vermerris W (2009) A genomewide analysis of the cinnamyl alcohol dehydrogenase family in sorghum [*Sorghum bicolor* (L.) Moench] identifies *SbCAD2* as the *brown midrib6* gene. Genetics 181(2):783-795 http://doi.org/10.1534/genetics.108.098996

Saballos A, Sattler SE, Sanchez E, Foster TP, Xin Z, Kang C, Pedersen JF, Vermerris W (2012) *Brown midrib2* (*Bmr2*) encodes the major 4-coumarate:coenzyme A ligase involved in lignin biosynthesis in sorghum (*Sorghum bicolor* (L.) Moench). The Plant J 70:818–830 http://doi.org/10.1111/j.1365-313X.2012.04933.x

Saballos A, Vermerris W, Rivera L, Ejeta G (2008) Allelic association, chemical characterization and saccharification properties of *brown midrib* mutants of sorghum (*Sorghum bicolor* (L.) Moench). BioEnergy Res 1:193–204 http://doi.org/10.1007/s12155-008-9025-7

Saltonstall K, Lambert A, Meyerson LA, Saltonstall K, Lambert A, Meyerson LA (2010) Genetics and reproduction of common (*Phragmites australis*) and giant reed (Arundo donax). Invasive Plant Sci Manag 3:495–505 http://doi.org/10.1614/IPSM-09-053.1

Santoro N, Cantu SL, Tornqvist CE, Falbel TG, Bolivar JL, Patterson SE, Pauly M, Walton JD (2010) A high-throughput platform for screening milligram quantities of plant biomass for lignocellulose digestibility. BioEnergy Res 3:93–102 http://doi.org/10.1007/s12155-009-9074-6

Sattler SE, Palmer NA, Saballos A, Greene AM, Xin Z, Sarath G, Vermerris W, Pedersen JF (2012) Identification and characterization of four missense mutations in *Brown midrib 12* (*Bmr12*), the caffeic acid *O*-methyltranferase (COMT) of sorghum. BioEnergy Res 5: 855-865 http://doi.org/10.1007/s12155-012-9197-z

Sattler SE, Saathoff AJ, Haas EJ, Palmer NA, Funnell-Harris DL, Sarath G, Pedersen JF (2009) A nonsense mutation in a cinnamyl alcohol dehydrogenase gene is responsible for the sorghum brown midrib6 phenotype. Plant Physiol 150:584–595 http://doi.org/10.1104/pp.109.136408

Sattler SE, Saballos A, Xin Z, Funnell-Harris DL, Vermerris W, Pedersen JF (2014) Characterization of novel sorghum *brown midrib* mutants from an EMS-mutagenized population. G3 [Genes Genomes Genetics] 4:2115–24 http://doi.org/10.1534/g3.114.014001

Saxena IM, Brown Jr, RM (2005) Cellulose biosynthesis: current views and evolving concepts. Ann Bot 96:9–21 http://doi.org/10.1093/aob/mci155

Schmer MR, Vogel KP, Mitchell RB, Perrin RK (2008). Net energy of cellulosic ethanol from switchgrass. Proceed National Acad Sci USA, 105:464–469

Schnepf R, Yacobucci BD (2013) Renewable fuel standard (RFS): overview and issues. Congressional Reserach Service Report for Congress. http://doi.org/10.1109/TNS.1981.4331499

Scully ED, Gries T, Sarath G, Palmer NA, Baird L, Serapiglia MJ, Dien BS, Boateng AA, Ge Z, Funnel-Harris DL, Twigg P, Clemente TE, Sattler SE (2016) Overexpression of SbMyb60 impacts phenylpropanoid biosynthesis and alters secondary cell wall composition in *Sorghum bicolor*. The Plant J 85:378–395 http://doi.org/10.1111/tpj.13112

Searchinger T, Heimlich R, Houghton RA, Dong F, Elobeid A, Fabiosa J, Tokgoz S, Hayes D, Yu T (2008) Use of U.S. croplands for biofuels increases greenhouse gases through emissions from land-use change. Sci 319(5867):1238–1240 http://doi.org/10.1126/science.1151861

Selig MJ, Tucker MP, Law C, Doeppke C, Himmel ME, Decker SR (2011) High throughput determination

of glucan and xylan fractions in lignocelluloses. Biotechnology Letters 33: 961–967 http://doi.org/10.1007/s10529-011-0526-7

Selig MJ, Tucker MP, Sykes RW, Reichel KL, Brunecky R, Himmel ME, Davis MF, Decker SR (2010) Lignocellulose recalcitrance screening by integrated high-throughput hydrothermal pretreatment and enzymatic saccharification. Industrial Biotechnology 6:104–111 http://doi.org/10.1089/ind.2010.0009

Selig MJ, Viamajala S, Decker SR, Tucker MP, Himmel ME, Vinzant TB (2007) Deposition of lignin droplets produced during dilute acid pretreatment of maize stems retards enzymatic hydrolysis of cellulose. Biotech Progress 23:1333–1339

Sené CFB, McCann MC, Wilson RH, Crinter R (1994) Fourier-transform Raman and Fourier-transform infrared spectroscopy. Plant Physiol 106:1623–1631 http://doi.org/10.1104/pp.106.4.1623

Sethaphong L, Haigler CH, Kubicki JD, Zimmer J, Bonetta D, DeBolt S, Yingling YG (2013). Tertiary model of a plant cellulose synthase. Proceed National Acad Sci USA, 110:7512–7517 http://doi.org/10.1073/pnas.1301027110

Singh M, Chaudhary K, Singal HR, Magill CW, Boora KS (2006) Identification and characterization of RAPD and SCAR markers linked to anthracnose resistance gene in sorghum [*Sorghum bicolor* (L.) Moench]. Euphytica 149:179–187 http://doi.org/10.1007/s10681-006-9306-1

Singh V, van Oosterom EJ, Jordan DR, Messina CD, Cooper M, Hammer GL (2010) Morphological and architectural development of root systems in sorghum and maize. Plant Soil 333:287–299 http://doi.org/10.1007/s11104-010-0343-0

Slewinski TL (2012) Non-structural carbohydrate partitioning in grass stems: a target to increase yield stability, stress tolerance, and biofuel production. J Exper Bot 63: 4547–4570 http://doi.org/10.1093/jxb/err313

Smith CW, & Frederiksen RA (2000) History of cultivar development in the United States. In: Smith CW, Frederiksen RA (eds) Sorghum: Origin, History, Technology, and Production. John Wiley & Sons, New York, pp 191-223

Snider JL, Raper RL, Schwab EB (2012) The effect of row spacing and seeding rate on biomass production and plant stand characteristics of non-irrigated photoperiod-sensitive sorghum (*Sorghum bicolor* (L.) Moench). Industrial Crops Products 37:527–535 http://doi.org/10.1016/j.indcrop.2011.07.032

Somerville C (2006) Cellulose synthesis in higher plants. Ann Rev Cell Develop Bio 22: 53–78 http://doi.org/10.1146/annurev.cellbio.22.022206.160206

Studer MH, DeMartini JD, Brethauer S, McKenzie HL, Wyman CE (2010) Engineering of a high-throughput screening system to identify cellulosic biomass, pretreatments, and enzyme formulations that enhance sugar release. Biotech Bioeng 105: 231–238 http://doi.org/10.1002/bit.22527

Tanabe S, Ashikari M, Fujioka S, Takatsuto S, Yoshida S, Yano M (2005) A novel cytochrome P450 is implicated in brassinosteroid biosynthesis via the characterization of a rice dwarf mutant, *dwarf11*, with reduced seed length. The Plant Cell 17:776–790 http://doi.org/10.1105/tpc.104.024950

Thomas MD, Sissoko I, Sacko M (1995) Development of leaf anthracnose and its effect on yield and grain weight of sorghum in West Africa. Plant Disease 79(2):151-153 http://doi.org/80:151-153

Thurber CS, Ma JM, Higgins RH, Brown PJ (2013). Retrospective genomic analysis of sorghum adaptation to temperate-zone grain production. Genome Bio 14:R68. http://doi.org/10.1186/gb-2013-14-6-r68

Upadhyaya HD, Wang YH, Sharma R, Sharma S (2013) Identification of genetic markers linked to anthracnose resistance in sorghum using association analysis. Theor App Genetics 126:1649–1657

Valdivia M, Galan JL, Laffarga J, Ramos JL (2016) Biofuels 2020: Biorefineries based on lignocellulosic materials. Microbial Biotech 9(5):585–594 http://doi.org/10.1111/1751-7915.12387

Van Rijn R, Nieves IU, Shanmugam KT, Ingram LO, Vermerris W.(2018) Techno-economic evaluation of cellulosic ethanol production based on pilot biorefinery data: a case study of sweet sorghum bagasse processed via L+SScF. BioEnergy Res 11:414–425 https://doi.org/10.1007/s12155-018-9906-3

Vandenbrink JP, Delgado MP, Frederick JR, Feltus FA (2010) A sorghum diversity panel biofuel feedstock screen for genotypes with high hydrolysis yield potential. Industrial Crops Products 31:444–448

Vandenbrink JP, Hilten RN, Das KC, Paterson AH, Feltus FA (2012) Analysis of crystallinity index and hydrolysis rates in the bioenergy crop *Sorghum bicolor*. BioEnergy Res 5:387-397

Venuto B, Kindiger B (2008) Forage and biomass feedstock production from hybrid forage sorghum and sorghum-sudangrass hybrids. Grassland Sci 54: 189–196 http://doi.org/10.1111/j.1744-697X.2008.00123.x

Vermerris W, Abril AA (2015) Enhancing cellulose utilization for fuels and chemicals by genetic modification of plant cell wall architecture. Curr Opinion Biotech 32:104-112

Vermerris W, Boon JJ (2001) Tissue-specific patterns of lignification are disturbed in the *brown midrib2* mutant of maize (*Zea mays* L.). J Agri Food Chem 49:721–728 http://doi.org/10.1021/jf000740r

Vermerris W, Saballos A, Ejeta G, Mosier NS, Ladisch MR, Carpita NC (2007) Molecular breeding to enhance ethanol production from corn and sorghum stover. Crop Sci 47(S3): S142–S153 http://doi.org/10.2135/cropsci2007.04.0013IPBS

Vermerris W, Thompson KJ, McIntyre LM (2002. The maize *Brown midrib1* locus affects cell wall composition and plant development in a dose-dependent manner. Heredity 88: 450–457 http://doi.org/10.1038/sj.hdy.6800078

Wang ML, Zhu C, Barkley NA, Chen Z, Erpelding JE, Murray SC, Tuinstra MR, Tesso T, Pederson GA, Yu J (2009) Genetic diversity and population structure analysis of accessions in the US historic sweet sorghum collection. Theor Appl Genetics 120:13–23 http://doi.org/10.1007/s00122-009-1155-6

Wolabu TW, Zhang F, Niu L, Kalve S, Bhatnagar-Mathur P, Muszynski MG, Tadege M (2016) Three *Flowering Locus T-like* genes function as potential florigens and mediate photoperiod response in sorghum. New Phytol 210:946–959 http://doi.org/10.1111/nph.13834

Xin Z, Wang ML, Barkley NA, Burow G, Franks C, Pederson G, Burke J (2008) Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. BMC Plant Bio 8:103 http://doi.org/10.1186/1471-2229-8-103

Yang S, Murphy RL, Morishige DT, Klein PE, Rooney WL, Mullet JE (2014) Sorghum phytochrome B inhibits flowering in long days by activating expression of SbPRR37 and SbGHD7, repressors of SbEHD1, SbCN8 and SbCN12. PloS One 9:e105352 http://doi.org/10.1371/journal.pone.0105352

Yong W, Link B, O'Malley R, Tewari J, Hunter CT, Lu CA, Li X, Bleecker AB, Koch KE, McCann MC, McCarty DR, Patterson SE, Reiter WD, Staiger C, Thomas SR, Vermerris W, Carpita NC (2005) Genomics of plant cell wall biogenesis. Planta 221:747–751 http://doi.org/10.1007/s00425-005-1563-z

Yu J, Zhang X, Tan T (2007) An novel immobilization method of *Saccharomyces cerevisiae* to sorghum bagasse for ethanol production. J Biotech 129:415–420 http://doi.org/10.1016/j.jbiotec.2007.01.039

Zeng M, Ximenes E, Ladisch MR, Mosier NS, Vermerris W, Huang CP, Sherman DM (2012) Tissue-specific biomass recalcitrance in corn stover pretreated with liquid hot-water: Enzymatic hydrolysis (part 1). Biotech Bioeng 109(2):390-397 http://doi.org/10.1002/bit.23337

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

Zhang B, Zhao T, Yu W, Kuang B, Yao Y, Liu T, Chen X, Zhang W, Wu AM (2014) Functional conservation of the glycosyltransferase gene *GT47A* in the monocot rice. J Plant Res 127:423–432 http://doi.org/10.1007/s10265-014-0631-5