



REVIEW

Open Access

# Bioenergy grass feedstock: current options and prospects for trait improvement using emerging genetic, genomic, and systems biology toolkits

Frank Alex Feltus\* and Joshua P Vandenbrink

## Abstract

For lignocellulosic bioenergy to become a viable alternative to traditional energy production methods, rapid increases in conversion efficiency and biomass yield must be achieved. Increased productivity in bioenergy production can be achieved through concomitant gains in processing efficiency as well as genetic improvement of feedstock that have the potential for bioenergy production at an industrial scale. The purpose of this review is to explore the genetic and genomic resource landscape for the improvement of a specific bioenergy feedstock group, the *C4 bioenergy grasses*. First, bioenergy grass feedstock traits relevant to *biochemical conversion* are examined. Then we outline genetic resources available bioenergy grasses for mapping bioenergy traits to DNA markers and genes. This is followed by a discussion of genomic tools and how they can be applied to understanding bioenergy grass feedstock trait genetic mechanisms leading to further improvement opportunities.

**Keywords:** Bioenergy grass, Genetic improvement, Bioenergy traits

## Introduction

Paleobioenergy obtained from coal, natural gas and oil deposits has allowed mankind to implement unprecedented technological advances in the last 250 years. Clearly, fossil fuels will not go away any time soon, but they are a finite resource with a viable lifespan affected by rapid population expansion (7 billion+; [1]) and the threat of the further elevation of greenhouse gases on our ability to respond to unpredictable variations in climate [2,3]. While the urgency for renewable energy sources to supplant fossil fuels on a massive scale is debatable, the need for alternative energy sources is evident. Bioenergy obtained from renewable plant material is an excellent component to any alternative energy portfolio.

Bioenergy feedstock selection is dependent upon many economic factors including land use constraints [4] and impact on other non-energy commodities [5], both of which could be addressed through public policy. Other feedstock factors can be addressed via rational existing feedstock selection as well as improvement through plant breeding and genetic modification. These factors include

energy density [6] and yield, cultivation costs [6], transportation logistics [7], pre-processing requirements [7], and conversion process [8]. For example, the scale-up of fermentable corn biomass (grain) to ethanol production (1<sup>st</sup> generation biofuel) in the U.S. in recent years has been successful since the conversion technology and agricultural infrastructure has matured [9]. Similarly, decades of sugarcane production in Brazil made it possible to become a net energy exporting economy [10]. Conversely, the promise of converting biomass that is recalcitrant to fermentation (lignocellulose) into viable energy products (2<sup>nd</sup> generation biofuels) has yet to be realized primarily due to the lack of realistic conversion techniques [11]. Thus, there is no turn-key bioenergy lignocellulosic feedstock solution at this time, but extensive research into efficient conversion process engineering and favorable feedstock properties is well under way.

The purpose of this review is to explore the genetic and genomic resource landscape for the improvement a specific bioenergy feedstock group, the *bioenergy grasses*. We define bioenergy grasses as members of the grass family (Poaceae) that employ *C4* photosynthesis and are capable of producing high biomass yield in the form of lignocellulose, fermentable juice, or fermentable grain [12]. Given

\* Correspondence: [ffeltus@clemsun.edu](mailto:ffeltus@clemsun.edu)  
Department of Genetics & Biochemistry, Clemson University, 105 Collings Street, BRC #302C, Clemson, SC 29634, USA

their proven utility as feedstock in academic and industrial interests, we focus on resources available for five specific bioenergy grasses: *Zea mays* (maize), *Saccharum* spp. (sugarcane), *Sorghum bicolor* (sorghum), *Miscanthus* spp. (*Miscanthus*), and *Panicum virgatum* (switchgrass). First, we discuss which grass feedstock traits are relevant to bioenergy production with a focus on biochemical conversion. Next, we discuss genetic resources available for the five bioenergy grasses to map bioenergy traits to genes. Then, we discuss genomic tools and how they can be applied to understanding bioenergy grass feedstock trait genetic mechanisms leading to further improvement opportunities. Finally, we will make the case for how modern genetic, genomic, and systems biology approaches can be coupled with bioprocessing constraints (industrial phenotypes) to breed feedstock varieties tailored to an industrial application.

### Relevant bioenergy grass traits

There are many extant bioenergy grass feedstock varieties (genotypes), which are sufficient for select conversion processes. For example, specific maize and sugarcane genotypes have been successful bioenergy grass feedstocks since high-yielding genotypes (grain and juice, respectively) have been grown at large scale for decades, and the conversion process (yeast fermentation) is well understood at the industrial level. Recent attention has been given to the more difficult problem of 2<sup>nd</sup> generation lignocellulose biomass conversion into profitable bioenergy products, which has the potential for accessing the photosynthate locked into the plant cell wall for conversion into useful products. Clearly, 2<sup>nd</sup> generation genotypes that produce high dry weight yields are of paramount importance, which is the opposite direction of the Green Revolution which led to small plants with high grain yield [13]. However, the identification and improvement of bioenergy grass genotypes with high biomass that efficiently respond to a given conversion process is ideal.

While there is much potential for bioenergy grasses as feedstock into *thermal conversion* processes (e.g. combustion, torrefaction, pyrolysis, and gasification), in this section we explore traits relevant to lignocellulose *biochemical conversion* processes which convert biomass into fermentable products through enzymatic hydrolysis (saccharification) [11]. The bioenergy grass feedstock traits that underlie conversion efficiency are being elucidated opening the door to genetic enhancement from existing feedstock.

### Cellulase inhibition

Cellulase enzyme cost is estimated to be ~50% of the total cost of the commercial hydrolysis process [14]. In addition, the enzymatic hydrolysis of lignocellulosic material experiences a reduction in activity over time. This

reduction in activity has been attributed to hydrolysis inhibition (end product and other [15-18]), reduction in easily accessible cellulose (e.g. crystalline vs. amorphous cellulose [19]), and reduction in efficient enzyme adsorption. Increasing enzyme accessibility to cellulose has been shown to play a crucial role in improving enzymatic hydrolysis [20-24]. Finding efficient means to increase enzymatic hydrolysis is vital to the success of lignocellulosic bioenergy production.

Chemical inhibition of cellulase reduces the total amount of reducing sugar produced for fermentation. High concentrations of end-products have been known to cause a reduction in cellulase activity. For example, while cellobiose is often a product of cellulases, it has also been shown to be a significant inhibitor of the activity of some cellulase [25]. This inhibition has been shown to be reduced by supplementing  $\beta$ -glucosidase to cellulase solutions lacking sufficient  $\beta$ -glucosidase activity [26]. End-product inhibition by glucose has been shown to inhibit late stage hydrolysis rates [27-29]. In addition to cellobiose, glucose has been shown to inhibit cellulase activity in cellulases derived from *Trichoderma* species [30,31]. However, inhibitory effects of glucose do not appear to affect *Aspergillus* species to the same degree [32-35]. This often leads to *Trichoderma* cellulases being supplemented with *Aspergillus*  $\beta$ -glucosidase to increase saccharification efficiency on an industrial level [36,37]. Additionally, xylose and arabinose, which are produced during the hydrolysis of hemicellulose, have been shown to inhibit cellulase activity [18,38]. Substrate inhibition of cellulases has led to simultaneous saccharification and fermentation (SSF) systems becoming popular, alleviating end-product inhibition.

In addition to end-product inhibition, metal ions have been shown to be inhibitory to cellulase hydrolysis reactions. It is suggested that the Fe(II) and Cu(II) oxidize the reducing ends of cellulose, inhibiting the exo-cellulolytic activity of cellulase [39-43]. However, not all metal ions cause an inhibitory effect on hydrolysis. Kim *et al.* found that while  $Hg^{++}$ ,  $Cu^{++}$  and  $Pb^{++}$  caused decrease in the production of total reducing sugars, other metal ions ( $Mn^{++}$ ,  $Ba^{++}$ , and  $Ca^{++}$ ) caused an increase in the total production of reducing sugars, indicating a stimulating effect on hydrolysis [44]. Two of these ions ( $Hg^{++}$  and  $Mn^{++}$ ) were shown to play a direct role in enzyme adsorption. Additionally,  $Mg^{++}$  was shown to stimulate the activity of glucanase from *Bacillus cellulyticus* [45]. The activity of cellulase produced from *Chaetomium thermophilum* was shown to be increased by  $Na^+$ ,  $K^+$  and  $Ca^{++}$ , but inhibited by  $Hg^{++}$ ,  $Zn^{++}$ ,  $Ag^+$ ,  $Mn^{++}$ ,  $Ba^{++}$ ,  $Fe^{++}$ ,  $Cu^{++}$ , and  $Mg^{++}$  [46]. This indicates that metal ions play an important role in enzyme efficacy during hydrolysis, and that knowledge of the correct ratio of metal ions is essential to increasing hydrolysis activity.

Phenolic compounds are also known to inhibit cellulolytic enzymes. These phenolics are often found in lignin, and are released (as well as their derivatives) during pretreatment processes. The types of phenolics present depends largely on the composition of biomass in combination with the type of pretreatment method employed [47-49]. A variety of released phenolic compounds have been identified during chemical pretreatment of lignocellulosic biomass [50-52], which have been shown to inhibit conversion of carbohydrates into ethanol as well as to inhibit cellulase activity [38,53-56]. Cellulases, hemicellulases, and  $\beta$ -glucosidase enzymes have all been shown to be inhibited by these phenolic compounds [54,56-59]. The magnitude of inhibition may specific to enzyme source as *Aspergillus niger*  $\beta$ -glucosidase was shown to be more resilient to phenolic inhibition when compared to *Trichoderma reesei*  $\beta$ -glucosidase, requiring a 4x higher concentration for inhibition [38]. Introduction of tannic acid degrading enzymes (Tannases) has been shown to increase enzymatic hydrolysis, likely by reducing tannic acid's propensity to interact and inhibit cellulase [60]. Additionally, polyethylene glycol has been shown to reduce inhibition of cellulase by tannins [61] by breaking up tannin-protein complexes. Tween 80 and PEG-4000 have been shown to prevent inhibition of  $\beta$ -glucosidase by reducing the tannins ability to bind the cellulase protein [61,62]. Finding additional methods to reduce the role of inhibitors in enzymatic hydrolysis is an important factor in increasing hydrolysis efficiency and profitability. Reducing the process-specific release of cellulase inhibitors through tailored feedstock genotypes is an attractive approach to enhancing enzymatic hydrolysis.

### Cellulose accessibility

Lignocellulosic material is a complex matrix of cellulose, hemicellulose and lignin [63,64]. In un-pretreated lignocellulosic samples, only a fraction of the cellulose is accessible to enzymatic hydrolysis, while the rest of the exposed biomass is lignin and hemicellulose. In order to increase access to cellulose, pretreatment methods are employed that aim to remove the lignin and hemicellulose fraction and leave cellulose available for hydrolysis. In addition, phenolic compounds such as ferulate play an important role in crosslinking lignin within the cell wall (see reviews [65-70]) and have the potential to be genetically modified to aid in the removal of specific cell wall components. There are many grass-specific features of the cell wall which have the potential to be exploited for increased bioenergy production [71]. For example, the composition of grass lignin is composed of syringyl (S), guaiacyl (G) and p- hydroxyphenyl (H) subunits that when present in varying ratios may lead to increased digestibility [68]. However, debate remains involving the role of lignin subunits in conversion efficiency [72-75].

Removal of structural components such as hemicellulose via dilute sulfuric acid pretreatment has been shown to increase accessibility to cellulose for enzymatic hydrolysis [76]. Removal of hemicellulose has been reported to increase pore volume and surface area further increasing the accessibility of cellulase [21]. Drying lignocellulosic substrates after chemical pretreatment results in the collapse of the newly formed pores, resulting in a decrease in enzymatic hydrolysis rate through reduction in available cellulose for hydrolysis [24,77]. Another pretreatment strategy which uses ionic liquids on switchgrass was shown to increase the porosity by over 30 fold, greatly increasing the accessibility of cellulose to enzymatic digestion [78]. This indicates that pore size and volume may play a significant role in increasing the rate of enzymatic hydrolysis. The identification of bioenergy grass feedstock genotypes that respond favorably to chemical pretreatment can increase end-product yield.

Lignin has been shown to play a large role in enzymatic conversion efficiency [79]. In *Miscanthus sinensis*, Yoshida *et al.* showed that removal of lignin via sodium chlorite resulted in an increase in enzymatic hydrolysis rate [80]. Yoshida *et al.* further demonstrated that the addition of hemicellulases resulted in an increase in overall hydrolysis rate, indicating that hemicellulose is an additional inhibitor of cellulose hydrolysis rates [80]. Zhao *et al.* also reported an increase in the enzymatic hydrolysis rate of sugarcane bagasse after the removal of lignin with paracetic acid [81]. Dissolution of lignocellulosic material with ionic liquid has been shown to increase enzymatic hydrolysis rates in wheat straw [82], corn stover [83] and switchgrass [78]. Kimon *et al.* showed that dissolving lignocellulosic material in ionic liquid at temperatures  $>150^{\circ}\text{C}$  has a large effect on saccharification of sugarcane bagasse [84]. Additionally ionic liquid pretreatment of switchgrass was shown to increase hydrolysis kinetics by over 39 fold over untreated switchgrass [78]. Ionic liquid pretreatment has also been shown to break inter and intra-molecular hydrogen bonding between cellulose strands causing an increase in the removal of amorphous components (lignin, hemicellulose) as well as an increase in surface area for cellulase adsorption [85]. These methods were both shown to superiorly increase hydrolysis rates when compared to traditional methods (dilute acid and ammonium hydroxide, respectively). Singh *et al.* reported that ionic liquid caused disruption of the inter and intra-molecular hydrogen bonding between lignin and cellulose which initially causes swelling of the plant cell wall followed by complete dissolution [86]. Organosolv pretreatment of switchgrass was shown to preferentially remove both lignin and hemicelluloses, leaving a larger cellulose fraction which resulted in an increase in the enzymatic hydrolysis rate [87]. Rollin *et al.* showed that treating switchgrass with organosolv resulted in a similar increase in the surface area causing

increased cellulase adsorption [88]. It is important to note that the promising field of ionic liquid pretreatment is still in its infancy. The current high costs of ionic liquid pretreatment limits its application to industrial scale-up, and like enzyme costs, must be reduced in order to be economically feasible on a large scale.

In addition to chemical pretreatment, naturally occurring mutations found in grasses have been shown to increase the rate of enzymatic hydrolysis via reductions in lignin. Brown midrib (*bmr*) is a phenotype found in grasses (maize [89], sorghum [90] and pearl millet [91]) that is associated with a mutation in genes involved in monolignol biosynthesis. These mutations have been shown to lead to a reduction in the total lignin content of the plant [92,93]. The brown colored midrib of the leaf has been shown to associate with a mutation in cinnamyl-alcohol dehydrogenase (CAD), which causes incorporation of cinnamyl-aldehydes in place of cinnamyl-alcohol during lignin biosynthesis [72,94,95]. Additional *bmr* varieties have been shown to have mutation in caffeic acid O-methyltransferase (COMT) [96-98]. However, both CAD and COMT mutants only exhibit reduced monolignol biosynthesis as opposed to total cessation of monolignol biosynthesis, indicating that other CAD and COMT genes may individually override complete cessation of monolignol biosynthesis. Theerattananon *et al.* found that a *bmr* mutant sorghum variety had less total lignin than forage, grain, sweet and photoperiod sensitive sorghum varieties [99]. In addition to lower lignin contents, *bmr* varieties have been shown to have increased susceptibility to chemical pretreatments. In sorghum, it was found that *bmr* mutants were more susceptible to alkaline pretreatment than non-*bmr* varieties [100]. Corredor *et al.* demonstrated that *bmr* sorghum varieties had a 79% hexose yield after enzymatic hydrolysis, which was higher than two non-*bmr* varieties which yielded 43% and 48% [101]. Additionally, sorghum varieties that contain both the mutations in COMT and CAD have been shown to have lower lignin contents than either mutant individually [102]. It is possible that there are additional genes and alleles leading to lowered lignin or other traits associated with higher hydrolysis rates. The identification of new as well as known lignification genes could lead to novel breeding programs where stacking of genes could result in intrinsic increases in lignocellulosic digestibility.

It is important to note that some maize *bmr* varieties have been characterized as being susceptible to lodging [103]. However, these susceptibilities were not seen in other maize studies which may be attributed to differences in genetic background [104,105]. This suggests that selecting an optimal genotype for the *bmr* mutation may be important in creating a superior feedstock. In addition to lodging, *bmr* mutants have been labeled as

more susceptible to disease and pathogen attack due to reduction in the lignin barrier. However, accumulation of lignin precursors has been shown to prevent the production of virulence factors as well as limit fungal pathogens [106-108]. It has also been widely reported that *bmr* varieties experience a decrease in yield associated with reduced lignin content. This has been seen in maize [104,109,110] and sorghum [111,112] *bmr* varieties. However, sorghum *bmr* hybrid varieties have been created that experience yields similar to wild type [113], suggesting that the genetic background of the mutant variety is important in overcoming yield reduction.

Transgenic approaches have already shown potential to increase saccharification efficiency in grasses. Overexpression of miR156, which suppresses SQUAMOSA PROMOTER BINDING PROTEIN LIKE (SPL) genes, in switchgrass caused an increase in overall biomass accumulation coupled with an increase in conversion efficiency of 24.2% – 155.5% in non-pretreated lignocellulosic material and between 40.7%–72.3% increase in acid pretreated samples [114]. In addition, moderate overexpression of miR156 caused switchgrass plants not to flower, reducing the possibility of transgenic gene escape. However, it should be noted that overexpression of miR156 caused dwarfism in both rice [115] and maize [116], which greatly reduces the plants value as a bioenergy feedstock. In addition, overexpression of R3R3-MYB4 transcription factors has been shown to repress lignin biosynthesis in several species [117-120]. In switchgrass, overexpression of PvMYB4 resulted in a three-fold increase in hydrolysis efficiency [121]. However, like the overexpression of miR156, these plants experienced a smaller stature than control varieties, limiting the gains made from increased hydrolysis efficiency. Clearly, the identification of active small RNA regulatory genes that do not affect biomass yield using genomic approaches is an exciting avenue towards bioenergy grass improvement.

#### Crystallinity index

Crystallinity index (CI) is a parameter that is used to determine the relative amount of crystalline cellulose in lignocellulosic material. Increased crystallinity of cellulose causes reduction in cellulase binding to cellulose due to reduced surface area. Conversely, increased amorphous cellulose causes an increase in the surface area, causing an increase in hydrolysis rates. CI has been measured using x-ray diffraction [122], solid-state <sup>13</sup>C NMR [123], infrared spectroscopy (IR) [124-126] and Raman spectroscopy [127]. CI has been shown to be correlated with enzymatic hydrolysis of lignocellulosic material. In *Sorghum bicolor*, CI has been shown to be negatively correlated with hydrolysis rate in whole plant tissue [128]. It has also been shown in sorghum as well

as maize that stem has a higher crystalline content than leaf tissue [129]. Furthermore, sorghum *bmr* mutants as well as wild type varieties experience an increase in CI after pretreatment with 1M NaOH. This observation is attributed to the removal of the amorphous component of the lignocellulosic biomass, leaving a larger fraction of crystalline material. However, it was also observed that an increase in the concentration of NaOH to 5M showed a decrease in CI, which was attributed to the crystal structure change and cellulose amorphization [100]. A similar trend was seen in dilute acid pretreatment of five sorghum varieties. Dilute acid pretreatment of sorghum at 140°C resulted in an increase in CI, however increasing the temperature during pretreatment to 165°C resulted in a decrease in the CI of 4 of 5 sorghum varieties [99]. This change in cellulose composition after pretreatment has been previously demonstrated in various industrial cellulose samples pretreated with NaOH [130,131]. Sugarcane bagasse was also shown to experience an increase in crystallinity after pretreatment with peracetic acid, which was attributed to a decrease in the amorphous component of the plant biomass [81]. Corredor *et al.* demonstrated dilute acid pretreatment of *bmr* and non-*bmr* sorghum varieties were shown to increase CI after pretreatment [101]. In addition, hydrolysis of the same samples resulted in a reduction in CI. Liu *et al.* found that like sorghum, acid pretreatment of maize biomass causes an increase in CI. However, the harshest pretreatment conditions cause a decrease in crystallinity, likely due to disruption of the cellulose crystalline structure [132]. This trend was confirmed by Mittal *et al.*, who also demonstrated that crystallinity of corn stover depends on specific conditions of alkali pretreatment. Additionally, Barl *et al.* demonstrated that maize husks experienced an increase in CI after both acid (H<sub>2</sub>SO<sub>4</sub>) and alkali (NaOH) pretreatment processes [133]. It should be noted that previous studies have demonstrated that the cellulose binding domain of cellulases disrupt cellulose crystalline structure and causes a decrease in CI [134,135]. This suggests that cellulose binding plays a role in conjunction with a decrease in cellulose content in the reduction in crystallinity index during enzymatic hydrolysis. Therefore, finding favorable genetic variation in endogenous and pretreated CI is a logical approach to improve hydrolysis yield [128].

Not all pretreatment strategies lead to an increase in CI. Pretreatment strategies that are particularly harsh initially increase CI through removal of amorphous components, followed by subsequent dissolution of crystalline cellulose. For example, Kimon *et al.* demonstrated that dissolving sugarcane lignocellulosic material with ionic liquids at temperatures >150°C causes a reduction in the cellulose CI and a large increase in glucan saccharification, while temperatures <150°C has a small

effect on crystallinity, which was associated with a slower initial rate of glucan saccharification [84]. Therefore, a screen for bioenergy grass genotypes that respond to harsh pretreatments in a favorable way could identify better feedstocks.

CI has been shown to differ between plant species, as well as different varieties within a species. When compared to different sorghum varieties, maize has been shown to have a higher CI [99]. Vandenbrink *et al.* demonstrated that CI differed between 18 different varieties of *Sorghum bicolor*, and these differences in CI were associated with hydrolysis rate [128]. Harris *et al.* found that crystallinity index differed among a large variety of plants which included sweet sorghum, switchgrass, giant *Miscanthus*, sweet *Miscanthus*, flame *Miscanthus*, gamagrass, big bluestem and *Arabidopsis* [136]. However, it must be pointed out that many of these species were only tested on a small number of varieties, which may not give an accurate depiction of CI in a diverse population where one genotype is one data point. These studies provide evidence that due to differences in CI between species and variety, there may be a significant genetic component that is associated with the trait.

There is much debate about the changes in crystallinity experienced during enzymatic hydrolysis of lignocellulosic materials. Various studies have demonstrated that amorphous cellulose components are hydrolyzed preferentially to crystalline components, resulting in an increase in crystallinity as enzymatic hydrolysis occurs [80,137,138]. However, various other studies have demonstrated that hydrolysis results in little change to crystallinity over the course of enzymatic hydrolysis [139,140], which was attributed to the synergistic action of *endo* and *exo*-glucanase activities [87,141]. However, it should be noted that studies have shown that the cellulose binding domain of multiple cellulases disrupt the supermolecular structure of cellulose, resulting in a decrease in CI [134,135]. This creates a difficult task in measuring changes in CI during enzymatic hydrolysis.

#### Enzyme adsorption

Non-specific cellulase adsorption to biomass plays a crucial role in determining the effectiveness of enzymatic hydrolysis. Due to the high cost of enzymes for commercial scale hydrolysis, adsorption and desorption rates in specific genotypes should be pre-determined. After hydrolysis, enzymes can either remain adsorbed to the substrate or unbound in the hydrolysate [142]. Cellulase adsorption depends largely on the concentration of the protein, as well as cellulase concentration and available surface area [143]. Initial protein adsorption has been shown to correlate with the initial rate of cellulose hydrolysis [19,144]. Multiple studies have shown that total enzyme adsorption is directly related to hydrolysis rate

and yield [145-148]. Strong correlations between available surface area and rate of hydrolysis have also been observed [23,149,150]. This increase in hydrolysis rate can be attributed to increased adsorption. Nutor *et al.* found that initial protein adsorption occurs quickly, reaching a maximum in 30 minutes, followed by 55-75% desorption [151]. Increasing the amount of enzyme adsorbed onto cellulose substrate is a potential avenue to increase hydrolysis rates, and it remains untested if specific cellulases are better adsorbed in specific bioenergy grass feedstock varieties.

Cellulase adsorption to lignin reduces cellulase activity by sequestering the enzyme away from its substrate. After the completion of hydrolysis, non-specific binding to lignin that has been freed during hydrolysis has been shown to occur, where 30-60% remains bound to the lignin fraction [152,153]. This non-specific binding has been shown to be only partly reversible [154]. Adsorption of cellulases to isolated lignin has been reported, supporting claims that non-specific binding occurs to the lignin fraction during hydrolysis [155,156]. Any cellulase bound to lignin is not available to hydrolyze cellulose, limiting overall efficiency. Hydrolysis rates of cellulose has been shown to be correlated with the tightness and affinity of adsorption [157]. Removal of lignin does not only reduce the steric hindrance to the enzyme, but also reduces the lignin available for non-specific binding [158,159].

Protein adsorption interactions are usually non-covalent (hydrogen bonding, electrostatic or hydrophobic interactions [160]). Surface characteristics of lignocellulosic material are thought to play a major role in cellulase adsorption where the high surface area hydrophobicity results in increased adsorption. Cellulases have been shown to have hydrophobic amino acids exposed on the outside of the protein, which interact with the hydrophobic surface of cellulase [161]. The affinity of cellulase for hydrophobic substrates may explain non-specific binding to lignin which is highly hydrophobic. In addition to this, metal ions have been shown to increase (in the case of  $Mn^{++}$ ) and decrease (in the case of  $Hg^{++}$ ) the adsorption affinity and tightness of binding to the hydrophobic surface of cellulose [44].

In order to drive down the cost of enzymatic hydrolysis, strategies to recycle cellulases are being developed [141,162-165]. Enzymes can be recovered from either bound substrate or from the liquid hydrolysate that remains after the first round of hydrolysis. Recovery of the enzyme from bound substrate can be achieved through washing with surfactant (such as Tween 20 [166]) or through recovery of the solid substrate in which the cellulase remains bound [162]. Use of cellulase recovered from lignocellulosic residue for subsequent rounds of hydrolysis have been shown to

experience reduced activity, which has been attributed to accumulation of bound lignin after each successive round of hydrolysis [154,163]. Recovery of enzyme from the liquid hydrolysate has been traditionally been done through ultracentrifugation techniques [142,167,168]. While this method has been proven effective, it would be costly to scale up to industrial magnitudes. A more effective method may be to exploit cellulase affinity for cellulose, in which the addition of cellulose to cellulase-containing hydrolysate results in re-adsorption onto the fresh cellulose substrate [163,169,170]. Tu *et al.* found that addition of fresh substrate to hydrolysate recovered ~50% of cellulases [171]. Additionally, bound enzyme was shown to be able to be recovered by contacting the bound substrate with fresh substrate [172]. However, sequential hydrolysis with recovered enzyme results in decreasing hydrolysis rates due to non-specific binding. Additionally it must be noted that  $\beta$ -glucosidase does not bind to cellulose substrate, and must be added at the beginning of each round of hydrolysis in order to prevent the buildup of cellobiose and the resulting substrate inhibition [171]. It is therefore necessary to develop techniques that are able to efficiently desorb cellulase from bound substrate. Deshpande *et al.* found that 90% of cellulase was recoverable from steam-exploded wheat straw [152]. Jackson *et al.* found that using a surfactant such as Tween 80 resulted in a recover of 6 – 77%, depending on concentration of Tween 80 and pH of the solution [166]. Additionally, Jackson *et al.* revealed that the highest protein recovery does not necessarily dictate the highest activity recovery, and that alkali conditions may be responsible for deactivation of the enzyme. Otter *et al.* demonstrated that Tween 80 and Triton X were able to desorb 65-68% of bound cellulase under alkaline conditions [173]. Qi *et al.* demonstrated that enzyme recycling of alkali and dilute-acid wheat straw was comparable when using ultracentrifugation and additional substrate techniques [174]. However, the additional substrate technique requires addition of  $\beta$ -glucosidase after each round of hydrolysis, whereas ultracentrifugation does not. Finally, there was a noticeable difference in enzyme recovery between dilute-acid and alkali pretreated samples, where alkali pretreated samples were able to desorb a larger amount of cellulase. While this discussion is focused on the putative industrial processes, it may be that specific feedstock varieties naturally exhibit lower adsorption rates that would further enhance the engineering endeavors.

In order for bioenergy to become a sustainable alternative to traditional fossil-fuel based transportation fuels, significant improvements to current enzymatic hydrolysis methods must be made. Reduced enzyme activity has been shown to be related to end-product inhibition, production of phenolic compounds from

lignin, as well as metal ion inhibition. Additionally the reduction in easily accessible cellulose through steric hindrance and high crystalline to amorphous cellulose levels cause a reduction in cellulose available for enzymatic hydrolysis. Non-specific binding of cellulases to solubilized lignin has also been associated with reduced hydrolysis rates. Finally, adsorption has been shown to be correlated with the initial rate of hydrolysis, while enzyme desorption is essential for enzyme recycling and reducing the cost of enzymes in bioenergy production. While these process components are being examined at the engineering level, a simple screen of existing bioenergy grass varieties could identify genotypes with a favorable trait baseline making the process engineering task less difficult.

### **Bioenergy grass genetic mapping resources**

There are tens of thousands of bioenergy grass genotypes in seed banks that have yet to be screened for favorable bioenergy traits. In fact, many traits that have been shown to deeply impact bioconversion yields have only been tested in a handful of genotypes. Surely, there are a multitude of relevant traits yet to be discovered. Therefore, we believe that genetic improvement is often premature until all screening options have been exhausted. With this caveat, genetic improvement in bioenergy grass feedstock can be achieved through transgenic manipulation or plant breeding programs. For example, centuries of selection have led to crops that provide high grain yields ideal for food production [13,175]. Many “elite” cultivars are dwarf varieties that allocate photosynthate towards larger grain yields as opposed to high cellulosic biomass. In grasses, the trend towards reduced lignocellulosic biomass could be rapidly reversed as genetic loci for plant height are few and well characterized [176-178]. In addition, the bioenergy traits discussed above can be genetically mapped to genomes, DNA markers associated with the trait developed, and alleles sorted into elite and novel cultivars. Once relevant DNA markers are identified, these traits can be selected for in breeding programs using marker assisted selection (MAS; [179]) or genome selection (GS; [180]) techniques. If the causal gene is identified, it can be introduced transgenically [181] to create elite bioenergy feedstock varieties.

In this section, we discuss the extensive genetic tools available for mapping traits in the genomes of bioenergy grasses as well as examples of previously mapped bioenergy traits.

Genetic mapping techniques available for bioenergy grasses include mapping Quantitative Trait Loci (QTLs) through linkage mapping in biparental populations [182], association mapping in a genetically diverse population [183], and nested association mapping (NAM)

[184,185]. QTL mapping requires relatively sparse marker coverage but identifies broad chromosomal regions associated with a trait of interest [182]. Association mapping analysis often requires prior knowledge of genes of interest or a full genome scan with high marker coverage to be successful [186]. Nested association mapping (NAM) populations exploit the benefits of both QTL and association mapping approaches [184,185]. It should be noted that genetic population structure can cause confounding correlation between markers and phenotypes within subpopulations [187,188]. The existence of distinct subpopulations can cause bias in the estimation of allelic effects and errors in QTL detection [189]. Thus, it is critical to generate panels that are genetically diverse and where population structure is clarified and corrected prior to genotype-phenotype associations [190]. All three genetic resources exist for diploid maize and sorghum bioenergy grasses and have been successful in mapping traits for years (see examples below). These approaches are more difficult in complex polyploids such as switchgrass, *Miscanthus*, and sugarcane, but there has been success in QTL mapping for these species (see examples below).

### **Quantitative trait loci**

Genetically defined mapping populations are a useful resource for locating DNA markers and mapping genes associated with desirable bioenergy traits. In these populations, quantitative trait loci (QTLs), intervals in the genome where DNA markers show a non-random association with a quantitative trait, can be identified [191], and the causal gene can possibly be mapped, albeit with difficulty (but see below). DNA markers associated with bioenergy QTLs can be used to breed superior varieties without extensive phenotyping [179] that contain a collection of genes desirable in a bioenergy feedstock. A key advantage of QTL mapping is that polymorphic DNA markers can be easily developed without a reference genome and they do not need be at high density across the genome.

In the diploid species sorghum, QTLs have been identified for many potentially advantageous genes valuable to biofuel production. QTLs related to leaf size including leaf width and leaf length [192] as well as leaf yield and composition [193] have been identified. Stem morphological traits such as height [178,193-203], diameter [192] and tillering characteristics [191,193,195,202] as well as stem composition and sugar content [193,201] have been associated with QTLs in sorghum. In addition, QTLs for flowering time or maturity have been shown to increase overall biomass by increasing the period of plant growth [178,194-198,201-205]. QTLs have also been analyzed for kernel weight [191,194,195,199,200,206,207] as well as grain composition [200,206,208,209]. In addition, QTLs for

post-harvest regrowth (ratooning) [191,193] may have the potential to increase total biomass yield producing additional biomass post-harvest. A recent study has mapped bioenergy QTLs, including biomass and stem sugar content, in a cross between a grain and sweet sorghum, [210]. The DNA markers identified in these studies can be used in breeding programs and demonstrate that markers for novel bioenergy traits such as the traits described above can easily be generated in existing or novel QTL mapping populations.

In maize, extensive research into QTLs of agronomic traits has been conducted. QTLs for forage quality and biomass composition have been comprehensively studied [211-219] and may have the potential to increase conversion efficiency. Also, because corn is a major food crop, thorough investigation of mapping populations has been conducted leading to the identification of a multitude of grain yield QTLs [220-233] which may lead to larger starch-derived ethanol yields. Additionally, QTLs for biomass related traits including both plant height [177,234-242] and plant maturity/flowering time [234-240,243,244] have been characterized, which could lead to increases in overall biomass yield. Leaf biomass characteristic QTLs [236,245-247] have also been identified which can lead to increased biomass as well as increased crop density resulting in greater yields. As with sorghum QTL studies, the maize mapping populations used in these studies can be used to map additional bioenergy traits and these DNA markers can be used in selection programs.

Complex polyploids such as *Miscanthus sinensis*, switchgrass, and sugarcane have had substantially fewer QTLs identified relative to the diploid grasses: sorghum and maize. In *Miscanthus*, plant biomass including leaf yield, stem yield and total plant height have been identified [248,249] leading to potential increases in total biomass. Additionally, flowering time QTLs have been identified which may lead to increased biomass accumulation [250,251]. *Miscanthus* also has potential as an energy source for thermal conversion. This has led to the identification of QTLs that effect thermal conversion efficiency [252,253]. To date, there have not been QTLs identified for the composition of *Miscanthus* biomass or forage quality, but the extant mapping populations are an excellent resource for mapping these traits. In sugarcane QTLs for stem sugar content have been identified [254-257], but few other bioenergy QTLs have been identified. These representative studies demonstrate that QTL mapping is a realistic tool for mapping complex traits in polyploid species. Below we discuss how modern sequencing techniques can be used to sequence large DNA segments underlying the QTL that becomes a powerful resource for identifying candidate genes even in complex polyploids.

QTL mapping in polyploid bioenergy grasses should improve with the development of new genomic resources. Recently, a high density genetic map has been developed for switchgrass [258], and two high resolution linkage map were created for *Miscanthus sinensis* [259,260]. These high-density maps open the door mapping QTLs to other genome through comparative genomics. For example, the *Miscanthus* map studies found that that of the sequenced grass species, sorghum, has the closest syntenous relationship to *Miscanthus* and that *Miscanthus sinensis* is of tetraploid origin consisting of two sub-genomes. These genetic maps will allow researchers to translate genetic tools from sorghum such as QTL studies and a sequenced genome via synteny relationships, thereby expanding the toolkit available for *Miscanthus*. In addition, the high density linkage maps can be used for *Miscanthus* genome assembly as well as QTL studies. Known and as yet undetected QTLs are a valuable method to identify DNA markers, often in multiple genome positions, that can be used to select for improved feedstock varieties before a crop development cycle is complete.

Minimal progress has been made in the development of superior cultivars from the identification of QTL associated with bioenergy traits. This may be due to the limitations in the transferability of QTL information due to QTLs being specific to alleles from inbred mapping parents. It may be that robust QTLs detected under multiple genetic backgrounds will be required. However, MAS stacking of QTLs (pyramiding) has been successful in other plant species as an avenue of crop improvement. Zhang et al. used QTL pyramiding to increase downy mildew tolerance in wild lettuce (*Lactuca saligna*) [261]. In another example, rice yield [262] as well as grain size and shape [263] have been modified through QTL pyramiding strategies. This suggests that given the ideal genetic background, genetic improvement of bioenergy crops through QTL pyramiding may be a viable way to produce superior feedstocks.

The NAM method for mapping QTLs relies on selection of a genetically diverse founding population which is derived from a common parent to create a large population of related progeny (often in the form of Recombinant Inbred Lines or RILs). NAM has the benefit of providing high QTL mapping resolution without requiring high marker density within the population [264]. In maize, a NAM population was created consisting of one common parent crossed with 25 diverse parents to produce 5,000 genetically distinct offspring [264]. A sorghum NAM population is under development [265]. QTLs for leaf architecture (including leaf angle, leaf length and leaf width) have been identified using the maize NAM population [185]. In addition, NAM has been used to identify QTLs for complex traits such as

resistance to northern leaf blight in maize [266]. While NAM incorporates high resolution QTL mapping with low marker coverage and high heterogeneity, it also requires large population size and a structured population in order to be informative. This technique also requires the screening of a large number of individuals, which makes the identification of complex phenotypes potentially very labor intensive. However, the NAM and other advanced genetic approaches are a powerful approach to dissect the genetic architecture of complex bioenergy traits.

While QTL studies have potential for bioenergy gene discovery, they also have limitations. Due to genetic heterogeneity, QTLs may be overestimated or not detected. There are also a variety of problems that arise in QTL mapping of polyploid genomes such as sugarcane and *Miscanthus*. These include increases in the number of genotypes per marker or QTL due to the increased number of chromosomes in the homeologous set, the dosage of marker and QTL in the parents and progeny are not obvious or observable, additional copies of a marker can mask recombination events, and the pairing behavior of chromosomes during meiosis is usually unknown [267]. Furthermore, low density genetic maps make it difficult to locate genes within a QTL region, which can contain thousands of genes. Dense genetic maps based upon sequence tagged markers, as is the case for sorghum [268], are readily mapped to other genomes. In this way, bioenergy QTLs can be identified in diploid sorghum and mapped to complex genome bioenergy grasses for causal gene inference and validation.

#### **Association mapping (diversity) panels**

Association mapping is an alternative method for mapping QTLs that is based on linkage disequilibrium (LD) occurring from historical recombination events in genetically diverse populations [269,270]. Association mapping utilizes marker-phenotype associations to determine if certain DNA markers co-segregate with a phenotype of interest. Association mapping generally falls into one of two categories: i) candidate gene association mapping, which looks for markers and causal variation in a subset of genes that are of interest for polymorphisms and ii) genome wide scan association mapping (GWAS), which scans the whole genome using dense marker sets to find marker associations with complex traits. Association mapping offers multiple benefits over traditional QTL mapping populations. QTL mapping populations suffer from restrictions due to limited genetic heterogeneity in that a QTL that is mapped in one mapping population derived from two genetic backgrounds and may not be applicable to other populations with parents derived from different lineages [271,272]. Association mapping panels, however, benefit from having higher resolution of identified QTLs

than traditional QTL mapping methods [273]. While association mapping requires a large diverse germplasm (diversity panel) of individuals to map QTLs, it does not require generation of inbred or backcrossed populations.

Association mapping populations have been created for the bioenergy crops maize [274,275], sorghum [176,276] and sugarcane [277]. In sorghum, association mapping has led to the identification of markers for height, flowering time, tiller number and stem sugar [278,279]. In maize, association mapping has led to the identification of markers for flowering time [187,280,281], kernel composition [282] as well as starch accumulation [283]. Fewer studies have been conducted in sugarcane, which has a large complex genome with high ploidy levels ranging from 5x – 14x [284]. Wei *et al.* mapped disease resistance in 154 sugarcane cultivars [277]. A key drawback to association mapping is that the large population size required for successful identification of trait markers requires that phenotyping the plants be done in a high-throughput manner which requires a large labor force or robotics. Often, this reduces the scope of DNA markers that can be identified to traits where phenotyping is less intensive.

#### **Reverse genetics**

In addition to the forward identification of DNA markers (and genes) by mapping a bioenergy trait to a DNA polymorphism, reverse genetic tools exist for the identification of bioenergy genes from a panel of known mutants. If the mutants are created in a parent with a favorable bioenergy trait baseline, it is possible to map genes and improve feedstock at the same time. In the TILLING approach (Targeting Induced Local Lesions IN Genomes), point mutations are randomly created throughout the genome by treating seeds with a mutagen (e.g. ethyl-methanesulfonate (EMS)) [285-287]. These plants are selfed and screened for phenotypes of interest. The DNA sequences from plants with mutant phenotypes can be compared to the non-mutagenized parental DNA to determine the relevant mutation. For example, DNA can be purified in a high throughput manner [288] and sequenced using high-throughput techniques for the discovery of rare mutations [289]. If the founding parent of the TILLING population has a sequenced genome as a reference, sequencing of select mutant individuals in candidate genes or whole genome resequencing can be done to identify specific gene mutations that lead to phenotypes of interest (e.g. [290]). As proof of principle, a sorghum TILLING population has been effective in the discovery of mutations giving rise to the bioenergy-relevant brown mid-rib phenotype [291] and altered hydrogen cyanide potential [292]. Once the gene variant underlying a trait is identified, the gene can be sequenced (e.g. PCR amplicon sequencing),

and any DNA variants tested for association in additional genotypes from the source and related organisms.

TILLING populations have been created for the bioenergy crops maize [293] and sorghum [294]. TILLING has the potential to identify bioenergy traits such as flowering time, total biomass, grain yield, conversion efficiency, etc. TILLING as a strategy for biofuel improvement does have its limitations. Due to the mutations induced by EMS being distributed randomly throughout the genome, the TILLING strategy can require screening thousands of individual lines to identify mutants in a trait of interest. This requirement of high-throughput phenotyping techniques limits the throughput of mutant selection gene detection. Furthermore, polyploid genomes present problems associated with finding recessive mutants due to the number of gene copies present in the genome. In the case of bioenergy grasses, this is strong rationale for first identifying a causal genetic lesion in a diploid genome (e.g. sorghum) and then testing the effect of the mutation in more complex genomes through plant breeding or transgenics. In summary, advanced genetic and mutant populations are a powerful approach to create varieties and map genes relevant to bioenergy feedstock.

### **Bioenergy grass genomic resources**

The crop genetic studies reviewed above have identified DNA markers associated with some high priority bioenergy related traits such as total biomass and conversion efficiency. These biomarkers have immediate utility in bioenergy grass improvement, and it is certain that the future will reveal many more biomarkers linked to known and novel bioenergy traits. However, the DNA biomarker often merely tags DNA near the gene(s) causing the favorable phenotype. While effective in breeding, this level of information leaves the underlying casual biochemical pathways and mechanisms in the black box. If the molecular mechanisms (and specific genes) underlying a trait were to be deciphered, then the art of plant breeding could be enhanced by searching for gene variants in other genes in the same pathway(s) as the initially described causal gene. Fortunately, the genome blueprints for specific bioenergy crops have been deciphered in the last decade. Using a reference genome assembly as a guide, it is now possible to associate genetically mapped biomarkers with nearby candidate genes and their functional activities. This section surveys genomic resources available for bioenergy grasses and discusses their utility in a genetically mapped trait context.

While genome-wide measurements of gene output can be obtained and interpreted without a reference genome, a high-quality, annotated reference genome assembly provides a natural scaffold to organize and interpret genetic and genomic analyses. In the case of bioenergy grasses,

three key reference genomes have been sequenced and annotated: maize [295], sorghum [296], and switchgrass (<http://www.phytozome.org/panicumvirgatum.php>). Once a genome assembly is constructed, it is annotated for sequence features including gene models and copy number (gene duplications), regulatory features, heterologous genome alignments (synteny), and other dynamic features such as gene expression levels under different internal and external cues. An excellent genome assembly resource for many plants, including maize, sorghum and switchgrass, can be found at the DOE-JGI Phytozome website [297].

The genome assembly sequence is a stable coordinate system to associate genome-mapped genetic signals (e.g. QTL biomarkers, trait-associated SNPs) with functional genomics information such nearby genes, gene expression levels, and biochemical pathways. If the sequences of DNA biomarkers are known, one can often locate the approximate genome position of a genetic signal and find neighboring genes in a physical context. Through the genome browser, biomarker DNA sequences can be positioned using BLAT/BLAST alignment tools or possibly through keyword searches. In some cases, biomarker positions have been pre-computed such as maize genetic markers accessible at [298]. Neighboring gene models are often annotated for function, usually via homology mapping, and provide clues that a given gene could be involved in the expression of a bioenergy trait. Gene function annotations include conserved protein domains (e.g. Interpro [299], Gene Ontology (GO) terms [300], and biochemical pathways (e.g. KEGG; [301]) including well annotated metabolic enzymes (e.g. RiceCyc at Gramene [302]). These annotation terms provide clues into what a gene near the biomarker is doing including possible pathway involvement, an indicator of gene-gene interaction and complex trait mechanism. It should be noted that genome browsers are highly dynamic and are constantly being updated with new information relevant to basic biology and possible bioenergy trait mechanisms.

While a reference genome view of an individual organism is invaluable, there are a growing number of databases focused on genome comparison and mapping function between species. This *translational genomics* approach is very important for the bioenergy grasses as gene function information can be discovered in a well-studied diploid organism such as maize, rice, and sorghum for which the genome is easier to analyze relative to complex polyploids like switchgrass, sugarcane and *Miscanthus*. Translational genomics is possible between bioenergy grasses because grass genomes in general have maintained a similar structure analogous to mammalian genomes since they diverged from a common ancestor 50–70 million years ago [303]. Therefore, genomes of

non-bioenergy grasses including rice [304] and *Brachypodium* [305] are also useful reference blueprints for grass gene function discovery and genome comparison [306]. Through grass genome comparison, gene function can be inferred in a poorly studied genome by identifying orthologous chromosomal segments. For example, the VISTA comparative genome browser (<http://pipeline.lbl.gov>; [307]) visualizes pre-computed alignments between the genomes of maize and sorghum as well as many other plants. A rich resource for genetically mapped information and grass genome comparison is Gramene ([302,308]). Finally, the Comparative Saccharinae Genomics Resource (CSGR; [309]) is focused specifically on the grasses including and related the bioenergy grasses. For a deep study of these resources, the reader is directed to relevant chapters in [310]. The macroconservation of grass genome structure is critical for genomic translation between bioenergy grasses with complex genomes such as sugarcane, switchgrass and *Miscanthus*. It may be a long time before additional reliable assemblies of complex polyploid genomes are realized, and at this time, we suggest that sorghum is an ideal C4 bioenergy grass reference genome due a relatively small annotated genome and close evolutionary proximity to other C4 bioenergy grasses.

The genome assembly provides physical coordinates of known genes, and intergenome comparison explores the dynamic movement of genes over evolutionary time scales. A reference genome assembly is also a framework for organization dynamic gene output measurements. For example, bioenergy grass gene output at the RNA level has been measured for over a decade using 1<sup>st</sup> generation genomic tools including the conversion of tissue and treatment specific RNA samples into cDNA followed by tedious cloning and sequencing. These Expressed Sequence Tags (ESTs) have proven invaluable in gene identification and can be found in databases at the National Center for Biotechnology Information (NCBI EST database) as well as the genome databases mentioned above. Massively parallel measurements of the RNA transcriptome response under multiple treatments and conditions have been made for bioenergy grasses using DNA microarrays. These experiments are stored into raw and processed forms at the NCBI Gene Expression Omnibus (GEO) database and are an excellent functional genomic data mining resource for the bioenergy grasses. For example, differences in gene expression in a genetically defined population can be associated with traits as eQTLs [311]. In addition, thousands of gene co-expression interactions can be mined from these datasets and transformed into gene interaction networks (see examples below). These functional genomics resources have been effective in understanding the molecular function of many bioenergy grass genes.

In recent years, rapid advances in DNA sequencing technology coupled with a reference genome for mapping sequences have resulted in multiple powerful next generation genomic analytical tools [312]. New sequencing technologies are capable of sequencing  $10^5$ - $10^8$  DNA molecules in a single experiment. As opposed to measuring molecule levels through hybridization to microarrays, this depth of coverage allows for molecule counting such as RNA-derived cDNA (RNAseq) or genomic DNA (re-sequencing) fragments to such a degree that quantitative comparisons can be made between samples. Example applications include transcriptome profiling with RNAseq [313], *de novo* transcript assembly [314], single nucleotide polymorphism (SNP) discovery [315], is the discovery of rare mutations in mutagenized (e.g. TILLING) populations [289,290], genotyping by sequencing (GBS; [316]) followed by GWAS or GS [183]), as well as whole [317] or partial genome *de novo* genome assembly [318]. In short, emerging sequencing technologies provide a high resolution lens into the dynamic biology underlying organism development.

Ongoing and historical genetic studies of bioenergy traits can be the immediate beneficiaries of these new sequencing technologies in that known gene regions can be sequenced and validated. For example, given the correct mix of resources, candidate genes and QTLs can now be cloned in a cost effective manner. In one scenario, a QTL for a relevant trait is mapped even at low marker resolution without a reference genome. Then, marker probes proximal to the QTL are used to screen a BAC library to identify nearby BACs. Once candidate BACs are identified, they can be pooled and cheaply sequenced as has been performed for melon (57 BACs; [319]), the complex genome of barley (91 BACs; [320]), and cacao (27 BACs; [318]). BAC pool assemblies can be annotated for candidate genes, used to design probes for additional BAC selection, and act as a reference sequence for resequencing applications. Of course, the process of BAC selection is enhanced if a physical map exists that can be used to identify a BAC minimum tiling path (e.g. [318]). In the case of switchgrass, a physical map might resolve the polyploidy issue in BAC selection [321], so individual genomes can be separately pooled thereby reducing the probability of intergenome misassembly.

Many bioenergy traits including those outlined above are complex in that they are controlled by multiple genes. By looking at a bioenergy trait (like those discussed above) as a systems biology problem, it may be possible to identify multiple markers or causal alleles that can be mixed in an appropriate genetic background to achieve the desired effect on yield. A near complete set of genes is known for a growing number of grasses (e.g. sorghum, maize, rice), but how these genes function in concert is poorly understood. Fortunately, modern genomic tools allow for the detection of gene dependencies

in the context of a relevant biochemical pathway or mapped trait that can be woven into gene interaction networks [322]. For example, gene interaction networks can be constructed that represent the non-random co-expression of transcripts between genes [323,324] or the physical interaction of gene products at the level of protein:protein interaction (PPI; [325,326]). Integrated gene sub-networks can be parsed from the overall network and non-randomly coupled with known biochemical pathways (e.g. fermentable sugar metabolism) or genetic signals (e.g. biomass yield) through a reference genome using systems biology techniques [323,327,328]. For example, gene co-expression networks have been constructed for many plants including rice [329,330] and maize [323]. Co-expressed gene modules have been identified in these networks, and some of the networks are enriched in genes that when mutated give rise to specific phenotypes that can be translated to the maize genome via conserved sub-graphs [323]. Gene regulatory networks can also be mapped to co-expressed gene modules [331]. It is possible to construct additional co-expression networks from other bioenergy grasses using RNAseq input (e.g. potato network [332]).

A systems genetics approach allows for both the prediction of complex polygenic genotype-phenotype interactions and also the ability to translate this information from diploid to polyploid genomes, a key asset in bioenergy grass improvement. We believe that gene interaction networks will significantly reduce the candidate gene list underlying a bioenergy trait if the requirement is made that interacting genetic signal genomic positions (e.g. a QTL set, multiple LD blocks from a GWAS study, or genes mapped in mutant lines that result in the same phenotype) must overlap with tightly interacting genes from the network (e.g. [323]). It is at the intersection of genetics and genomics that complex bioenergy traits, which by definition are polygenic, can be tested as a genetic sub-system as opposed to breaking the system into individual genetic components such as a single large-effect QTL.

## Conclusions

Given the uncertainties involved with long term fossil fuel production and increased carbon emissions affecting global climate, the pursuit of sustainable fuels from lignocellulosic biomass is important. We conclude that a deeper understanding of feedstock traits affecting bioconversion such as enzyme inhibition, cellulose accessibility, and enzyme adsorption will ameliorate hurdles to bioenergy production so that it is competitive with current fossil fuel based transportation fuels. While these factors limit the efficiency of enzymatic bioconversion, they also provide a myriad of opportunities for end-product yield improvement through feedstock

genetics coupled with process engineering. Breeding programs that have historically focused on increased grain yields can be shifted to focus on traits yielding high-biomass, hydrolysis-efficient bioenergy crop varieties. It should be noted, however, that vast bioenergy grass seed stocks still need to be screened for high yield baselines prior to breeding new varieties. For example, future or extant varieties that contain low lignin (such as *bmr* maize, sorghum and millet) may help to reduce steric hindrance to hydrolytic enzymes as well as reduce non-specific binding and increased enzyme recovery. Additionally, reduced lignin content has potential to reduce the amount of phenolic compounds released during pretreatment and hydrolysis, which reduces inhibition to cellulase. Through the coupling of DNA biomarkers to these traits, better crops can be developed through marker-assisted selection, and rapid advances in genomic and systems biology techniques should reveal novel biochemical mechanisms that can be engineered into current feedstock varieties. It is our belief that close collaboration between the plant breeder, systems biologist, and process engineer will result in accelerated development of bioenergy grass feedstock tailored to a specific conversion process thereby increasing bioenergy viability through *industrial genetics*.

## Abbreviations

CAD: Cinnamyl-alcohol dehydrogenase; CI: Crystallinity index; COMT: Caffeic acid O-ethyltransferase; DNA: Deoxyribonucleic acid; EMS: Ethyl-methanesulfonate; GS: Genome selection; GWAS: Genome wide scan association mapping; LD: Linkage disequilibrium; MAS: Marker assisted selection; NAM: Nested association mapping; QTL: Quantitative trait loci; RIL: Recombinant inbred line; RNA: Ribonucleic acid; SNP: Single nucleotide polymorphism; TILLING: Targeting Induced Local Lesions IN Genomes.

## Competing interests

The authors declare that they have no competing interests.

## Authors' contributions

FAF and JPV researched and wrote the manuscript. All authors read and approved the final manuscript.

## Acknowledgements

This research was supported in part by the Clemson Experiment Station project #SC-1700381 to FAF under technical contribution #6038.

Received: 11 May 2012 Accepted: 5 October 2012

Published: 2 November 2012

## References

1. Central Intelligence Agency World Factbook. [www.cia.gov](http://www.cia.gov).
2. Hansen J, Sato M: **Greenhouse gas growth rates**. *Proc Natl Acad Sci U S A* 2004, **101**:16109–16114.
3. Pachauri RK, Reisinger A: *IPCC Climate Change 2007*. Geneva, Switzerland: Synthesis Report; 2007.
4. Fargione J, Hill J, Tilman D, Polasky S, Hawthorne P: **Land clearing and the biofuel carbon debt**. *Science* 2008, **319**:1235–1238.
5. Elobeid A, Tokgoz S, Hayes DJ, Babcock BA, Hart CE: **The Long-Run Impact of Corn-Based Ethanol on the Grain, Oilseed, and Livestock Sectors: A Preliminary Assessment**. In *Book The Long-Run Impact of Corn-Based Ethanol on the Grain, Oilseed, and Livestock Sectors: A Preliminary Assessment*. City: Iowa State University; 2006.

6. Graham RL, Lichtenberg E, Roningen VO, Shapouri H, Walsh ME: *The Economics of Biomass Production in the United States*. <http://bioenergy.ornl.gov/papers/bioam95/graham3.html>.
7. Vidal BC Jr, Dien BS, Ting KC, Singh V: **Influence of feedstock particle size on lignocellulose conversion—a review**. *Appl Biochem Biotechnol* 2011, **164**:1405–1421.
8. Schell C, Riley C, Petersen GR: **Pathways for development of a biorenewables industry**. *Bioresour Technol* 2008, **99**:5160–5164.
9. Renewable Fuels Association. [www.ethanolrfa.org](http://www.ethanolrfa.org).
10. Valdes C: **Brazil's Ethanol Production: Looking Forward**. <http://www.ers.usda.gov/media/126865/bio02.pdf>.
11. Sims RE, Mabee W, Saddler JN, Taylor M: **An overview of second generation biofuel technologies**. *Bioresour Technol* 2010, **101**:1570–1580.
12. Vermerris W (Ed): *Genetic Improvement of Bioenergy Crops*. Gainesville, FL: Springer; 2008.
13. Evenson RE, Gollin D: **Assessing the impact of the green revolution, 1960 to 2000**. *Science* 2003, **300**:758–762.
14. Galbe M, Zacchi G: **A review of the production of ethanol from softwood**. *Appl Microbiol Biotechnol* 2002, **59**:618–628.
15. Gruno M, Våljamäe P, Pettersson G, Johansson G: **Inhibition of the *Trichoderma reesei* cellulases by cellobiose is strongly dependent on the nature of the substrate**. *Biotechnol Bioeng* 2004, **86**:503–511.
16. Holtzapfel M, Cognata M, Shu Y, Hendrickson C: **Inhibition of *Trichoderma reesei* cellulase by sugars and solvents**. *Biotechnol Bioeng* 1990, **36**:275–287.
17. Johnson EA, Reese ET, Demain AL: **Inhibition of *Clostridium thermocellum* Cellulase by End Products of Cellulolysis**. *J Appl Biochem* 1982, **4**:64–71.
18. Xiao Z, Zhang X, Gregg D, Saddler J: **Effects of sugar inhibition on cellulases and  $\beta$ -glucosidase during enzymatic hydrolysis of softwood substrates**. *Appl Biochem Biotechnol* 2004, **115**:1115–1126.
19. Wald S, Wilke CR, Blanch HW: **Kinetics of the enzymatic hydrolysis of cellulose**. *Biotechnol Bioeng* 1984, **26**:221–230.
20. Stone JE, Scallan AM, Donefer E, Ahlgren E: **Digestibility as a Simple Function of a Molecule of Similar Size to a Cellulase Enzyme**. In *Cellulases and Their Applications*, Advances in Chemistry, Volume 95. Washington, D.C.: American Chemical Society; 1969:219–241.
21. Grethlein HE: **The Effect of Pore Size Distribution on the Rate of Enzymatic Hydrolysis of Cellulosic Substrates**. *Nat Biotechnol* 1985, **3**:155–160.
22. Burns D, Ooshima H, Converse A: **Surface area of pretreated lignocellulosics as a function of the extent of enzymatic hydrolysis**. *Appl Biochem Biotechnol* 1989, **20**–21:79–94.
23. Thompson DN, Chen H-C, Grethlein HE: **Comparison of pretreatment methods on the basis of available surface area**. *Bioresour Technol* 1992, **39**:155–163.
24. Esteghlalian AR, Bilodeau M, Mansfield SD, Saddler JN: **Do Enzymatic Hydrolyzability and Simons' Stain Reflect the Changes in the Accessibility of Lignocellulosic Substrates to Cellulase Enzymes?** *Biotechnol Prog* 2001, **17**:1049–1054.
25. Holtzapfel MT, Caram HS, Humphrey AE: **Determining the inhibition constants in the HCH-1 model of cellulose hydrolysis**. *Biotechnol Bioeng* 1984, **26**:753–757.
26. Bommarius AS, Katona A, Cheben SE, Patel AS, Ragauskas AJ, Knudson K, Pu Y: **Cellulase kinetics as a function of cellulose pretreatment**. *Metab Eng* 2008, **10**:370–381.
27. Caminal G, López-Santín J, Solà C: **Kinetic modeling of the enzymatic hydrolysis of pretreated cellulose**. *Biotechnol Bioeng* 1985, **27**:1282–1290.
28. Gusakov AV, Sinitsyn AP, Klyosov AA: **Kinetics of the enzymatic hydrolysis of cellulose: 2. A mathematical model for the process in a plug-flow column reactor**. *Enzym Microb Technol* 1985, **7**:383–388.
29. Lee Y-H, Fan LT: **Kinetic studies of enzymatic hydrolysis of insoluble cellulose: (II). Analysis of extended hydrolysis times**. *Biotechnol Bioeng* 1983, **25**:939–966.
30. Chen H, Hayn M, Esterbauer H: **Purification and characterization of two extracellular  $\beta$ -glucosidases from *Trichoderma reesei***. *Biochimica et Biophysica Acta (BBA) - Protein Structure and Molecular Enzymology* 1992, **1121**:54–60.
31. Gong C-S, Ladisch MR, Tsao GT: **Cellobiase from *Trichoderma viride*: Purification, properties, kinetics, and mechanism**. *Biotechnol Bioeng* 1977, **19**:959–981.
32. Decker CH, Visser J, Schreier P:  **$\beta$ -Glucosidases from Five Black *Aspergillus* Species: Study of Their Physico-Chemical and Biocatalytic Properties**. *J Agric Food Chem* 2000, **48**:4929–4936.
33. Günata Z, Vallier M-J: **Production of a highly glucose-tolerant extracellular  $\beta$ -glucosidase by three *Aspergillus* strains**. *Biotechnol Lett* 1999, **21**:219–223.
34. Watanabe T, Sato T, Yoshioka S, Koshijima T, Kuwahara M: **Purification and properties of *Aspergillus niger* beta-glucosidase**. *Eur J Biochem* 1992, **209**:651–659.
35. Yan TR, Lin CL: **Purification and characterization of a glucose-tolerant  $\beta$ -glucosidase from *Aspergillus niger***. *Biosci Biotechnol Biochem* 1997, **61**:965–970.
36. Reczey KA, Brumbauer A, Bollók M, Szengyel Z, Zacchi G: **Use of hemicellulose hydrolysate for  $\beta$ -glucosidase fermentation**. *Appl Biochem Biotechnol* 1998, **70–72**:225–235.
37. Sternberg D, Vijayakumar P, Reese ET:  **$\beta$ -Glucosidase: microbial production and effect on enzymatic hydrolysis of cellulose**. *Can J Microbiol* 1977, **23**:139–147.
38. Ximenes E, Kim Y, Mosier N, Dien B, Ladisch M: **Inhibition of cellulases by phenols**. *Enzym Microb Technol* 2010, **46**:170–176.
39. Tejirian A, Xu F: **Inhibition of Cellulase-Catalyzed Lignocellulosic Hydrolysis by Iron and Oxidative Metal Ions and Complexes**. *Appl Environ Microbiol* 2010, **76**:7673–7682.
40. Ferchak JD, Pye EK: **Effect of cellobiose, glucose, ethanol, and metal ions on the cellulase enzyme complex of *Thermomonospora fusca***. *Biotechnol Bioeng* 1983, **25**:2865–2872.
41. Gardner RM, Doerner KC, White BA: **Purification and characterization of an exo-beta-1,4-glucanase from *Ruminococcus flavefaciens* FD-1**. *J Bacteriol* 1987, **169**:4581–4588.
42. Harada K, Tanaka K, Fukuda Y, Hashimoto W, Murata K: **Degradation of rice bran hemicellulose by *Paenibacillus* sp. strain HC1: gene cloning, characterization and function of  $\beta$ -D-glucosidase as an enzyme involved in degradation**. *Arch Microbiol* 2005, **184**:215–224.
43. Heredia A, Fernandez-Bolaños J, Guillen R: **Inhibitors of cellulolytic activity in olive fruits (*Olea europaea*, Hojiblanca var.)**. *Zeitschrift für Lebensmitteluntersuchung und -Forschung A* 1989, **189**:216–218.
44. Kim DW, Jang YH, Kim CS, Lee NS: **Effect of Metal Ions on the Degradation and Adsorption of Two Cellobiohydrolases on Microcrystalline Cellulose**. *Bulletin of the Korean Chemical Society* 2001, **22**:716–720.
45. Lee TK, Kim CH: **Molecular cloning and expression of an endo-beta-1,4-D-glucanase I (Avicelase I) gene from *Bacillus cellulolyticus* K-12 and characterization of the recombinant enzyme**. *Appl Biochem Biotechnol* 1999, **80**:121–140.
46. Li D-C, Lu M, Li Y-L, Lu J: **Purification and characterization of an endocellulase from the thermophilic fungus *Chaetomium thermophilum* CT2**. *Enzym Microb Technol* 2003, **33**:932–937.
47. Dekker RFH: **Inhibitors of *Trichoderma reesei*  $\beta$ -glucosidase activity derived from autohydrolysis-exploded *Eucalyptus regnans***. *Appl Microbiol Biotechnol* 1988, **29**:593–598.
48. Excoffier G, Toussaint B, Vignon MR: **Saccharification of steam-exploded poplar wood**. *Biotechnol Bioeng* 1991, **38**:1308–1317.
49. Panagiotou G, Olsson L: **Effect of compounds released during pretreatment of wheat straw on microbial growth and enzymatic hydrolysis rates**. *Biotechnol Bioeng* 2007, **96**:250–258.
50. Klinke HB, Ahring BK, Schmidt AS, Thomsen AB: **Characterization of degradation products from alkaline wet oxidation of wheat straw**. *Bioresour Technol* 2002, **82**:15–26.
51. Du B, Sharma LN, Becker C, Chen S-F, Mowery RA, van Walsum GP, Chambliss CK: **Effect of varying feedstock-pretreatment chemistry combinations on the formation and accumulation of potentially inhibitory degradation products in biomass hydrolysates**. *Biotechnol Bioeng* 2010, **107**:430–440.
52. Fenske JJ, Griffin DA, Penner MH: **Comparison of aromatic monomers in lignocellulosic biomass prehydrolysates**. *J Ind Microbiol Biotechnol* 1998, **20**:364–368.
53. Ximenes E, Kim Y, Mosier N, Dien B, Ladisch M: **Deactivation of cellulases by phenols**. *Enzym Microb Technol* 2011, **48**:54–60.
54. Pan X: **Role of Functional Groups in Lignin Inhibition of Enzymatic Hydrolysis of Cellulose to Glucose**. *J Biobased Materials and Bioenergy* 2008, **2**:25–32.
55. Mandels M, Reese ET: **Inhibition of Cellulases**. *Annu Rev Phytopathol* 1965, **3**:85–102.

56. Siniero J, Dominguez H, Nunez MJ, Lema JM: **Inhibition of cellulase activity by sunflower polyphenols.** *Biotechnol Lett* 1997, **19**:521–524.
57. Berlin A, Balakshin M, Gilkes N, Kadla J, Maximenko V, Kubo S, Saddler J: **Inhibition of cellulase, xylanase and  $\beta$ -glucosidase activities by softwood lignin preparations.** *J Biotechnol* 2006, **125**:198–209.
58. Akin D, Rigsby L: **Corn Fiber: Structure, Composition, and Response to Enzymes for Fermentable Sugars and Coproducts.** *Appl Biochem Biotechnol* 2008, **144**:59–68.
59. Ximenes E, Brandon S, Doran-Peterson J: **Evaluation of a *Hypocrea jecorina* Enzyme Preparation for Hydrolysis of Tifton 85 Bermudagrass.** *Appl Biochem Biotechnol* 2008, **146**:89–100.
60. Tejirian A, Xu F: **Inhibition of enzymatic cellulolysis by phenolic compounds.** *Enzym Microb Technol* 2011, **48**:239–247.
61. Kumar R: ***Prosopis cineraria* leaf tannins: their inhibitory effect upon ruminal cellulase and the recovery of inhibition by polyethylene glycol-4000.** *Basic Life Science* 1992, **59**:699–704.
62. Zanolini A, Vanni P, Firenzuoli A: **Effect of tween 80 on protein-tannic acid complex.** *Cellular and Molecular Life Sciences* 1967, **23**:1015–1016.
63. Barriere Y, Guillet C, Goffner D, Pichon M: **Genetic variation and breeding strategies for improved cell-wall digestibility in annual forage crops.** *Animal Res* 2003, **52**:193–228.
64. Barriere Y, Ralph J, Mechin V, Guillaumie S, Grabber JH, Argillier O, Chabbert B, Lapiere C: **Genetic and molecular basis of grass cell-wall biosynthesis and degradability. II. Lessons from brown-rib mutants.** *C R Biol* 2004, **327**:847–860.
65. Buanafina MM d O: **Feruloylation in Grasses: Current and Future Perspectives.** *Molecular Plant* 2009, **2**:861–872.
66. Hatfield RD, Ralph J, Grabber JH: **Cell wall cross-linking by ferulates and diferulates in grasses.** *J Sci Food Agric* 1999, **79**:403–407.
67. Anterola AM, Lewis NG: **Trends in lignin modification: a comprehensive analysis of the effects of genetic manipulations/mutations on lignification and vascular integrity.** *Phytochemistry* 2002, **61**:221–294.
68. Grabber JH, Ralph J, Lapiere C, Barriere Y: **Genetic and molecular basis of grass cell-wall degradability. I. Lignin–cell wall matrix interactions.** *Comptes Rendus Biologies* 2004, **327**:455–465.
69. Barriere Y, Ralph J, Méchin V, Guillaumie S, Grabber JH, Argillier O, Chabbert B, Lapiere C: **Genetic and molecular basis of grass cell wall biosynthesis and degradability. II. Lessons from brown-midrib mutants.** *Comptes Rendus Biologies* 2004, **327**:847–860.
70. Ralph J, Hatfield Ronald D, Grabber John H, Jung Hans-Joachim G, Quideau S, Helm Richard F: **Cell Wall Cross-Linking in Grasses by Ferulates and Diferulates.** In *Lignin and Lignan Biosynthesis*, ACS Symposium Series, Volume 697. Washington, D.C.: Am Chemical Soc; 1998:209–236.
71. Vogel J: **Unique aspects of the grass cell wall.** *Current Opinion in Plant Biology* 2008, **11**:301–307.
72. Baucher M, Bernard-vailhé MA, Chabbert B, Besle J-M, Opsomer C, Van Montagu M, Botterman J: **Down-regulation of cinnamyl alcohol dehydrogenase in transgenic alfalfa (*Medicago sativa* L.) and the effect on lignin composition and digestibility.** *Plant Mol Biol* 1999, **39**:437–447.
73. Jung HG, Casler MD: **Relationship of lignin and esterified phenolics to fermentation of smooth bromegrass fibre.** *Anim Feed Sci Technol* 1991, **32**:63–68.
74. Jung H-JG, Buxton DR: **Forage quality variation among maize inbreds: Relationships of cell-wall composition and in-vitro degradability for stem internodes.** *J Sci Food Agric* 1994, **66**:313–322.
75. Jung H-JG, Ni W, Chapple CCS, Meyer K: **Impact of lignin composition on cell-wall degradability in an Arabidopsis mutant.** *J Sci Food Agric* 1999, **79**:922–928.
76. Grohmann K, Torget R, Himmel M: **Optimization of dilute acid pretreatment of biomass.** In *Biotechnology and Bioengineering Symposium*: Wiley; 1986.
77. Wong KKY, Deverell KF, Mackie KL, Clark TA, Donaldson LA: **The relationship between fiber-porosity and cellulose digestibility in steam-exploded *Pinus radiata*.** *Biotechnol Bioeng* 1988, **31**:447–456.
78. Arora R, Manisseri C, Li C, Ong MD, Scheller HV, Vogel K, Simmons BA, Singh S: **Monitoring and Analyzing Process Streams Towards Understanding Ionic Liquid Pretreatment of Switchgrass (*Panicum virgatum* L.).** *Bioenergy Research* 2010, **3**:134–145.
79. Mooney CA, Mansfield SD, Touhy MG, Saddler JN: **The effect of initial pore volume and lignin content on the enzymatic hydrolysis of softwoods.** *Bioresour Technol* 1998, **64**:113–119.
80. Yoshida M, Liu Y, Uchida S, Kawarada K, Ukagami Y, Ichinose H, Kaneko S, Fukuda K: **Effects of cellulose crystallinity, hemicellulose, and lignin on the enzymatic hydrolysis of *Miscanthus sinensis* to monosaccharides.** *Biosci Biotechnol Biochem* 2008, **72**:805–810.
81. Zhao X-B, Wang L, Liu D-H: **Peracetic acid pretreatment of sugarcane bagasse for enzymatic hydrolysis: a continued work.** *J Chem Technol Biotechnol* 2008, **83**:950–956.
82. Fu D, Mazza G, Tamaki Y: **Lignin Extraction from Straw by Ionic Liquids and Enzymatic Hydrolysis of the Cellulosic Residues.** *J Agric Food Chem* 2010, **58**:2915–2922.
83. Varanasi S, Schall CA, Dadi AP, Anderson J, Rao K: **Biomass Pretreatment.** 2008. <http://www.faqs.org/patents/app/20080227162#b>.
84. Kimon KS, Leslie Alan E, William Orlando Sinclair D: **Enhanced saccharification kinetics of sugarcane bagasse pretreated in 1-butyl-3-methylimidazolium chloride at high temperature and without complete dissolution.** *Bioresour Technol* 2011, **102**:9325–9329.
85. Li C, Knerim B, Manisseri C, Arora R, Scheller HV, Auer M, Vogel KP, Simmons BA, Singh S: **Comparison of dilute acid and ionic liquid pretreatment of switchgrass: Biomass recalcitrance, delignification and enzymatic saccharification.** *Bioresour Technol* 2010, **101**:4900–4906.
86. Singh S, Simmons BA, Vogel KP: **Visualization of biomass solubilization and cellulose regeneration during ionic liquid pretreatment of switchgrass.** *Biotechnol Bioeng* 2009, **104**:68–75.
87. Cateto C, Hu G, Ragauskas A: **Enzymatic hydrolysis of organosolv Kanlow switchgrass and its impact on cellulose crystallinity and degree of polymerization.** *Energy & Environmental Sci* 2011, **4**:1516–1521.
88. Rollin JA, Zhu Z, Sathitsuksanoh N, Zhang YHP: **Increasing cellulose accessibility is more important than removing lignin: A comparison of cellulose solvent-based lignocellulose fractionation and soaking in aqueous ammonia.** *Biotechnol Bioeng* 2011, **108**:22–30.
89. Jorgenson LR: **Brown midrib in maize and its linkage relations.** *J Am Soc Agronomy* 1931, **23**:549–557.
90. Porter KS, Axtell JD, Lechtenberg VL, Colenbrander VF: **Phenotype, Fiber Composition, and in vitro Dry Matter Disappearance of Chemically Induced Brown Midrib (bmr) Mutants of Sorghum.** *Crop Sci* 1977, **18**:205–208.
91. Cherney JH, Axtell JD, Hassen MM, Anliker KS: **Forage quality characterization of a chemically-induced brown-midrib mutant in pearl-mille.** *Crop Sci* 1988, **28**:783–787.
92. Kuc J, Nelson OE: **The abnormal lignins produced by the brown-midrib mutants of maize: I. The brown-midrib-1 mutant.** *Archives of Biochemistry and Biophysics* 1964, **105**:103–113.
93. Gee MS, Nelson OE, Kuć J: **Abnormal lignins produced by the brown-midrib mutants of maize: II. Comparative studies on normal and brown-midrib-1 dimethylformamide lignins.** *Arch Biochem Biophys* 1968, **123**:403–408.
94. Baucher M, Chabbert B, Pilate G, Van Doorslaere J, Tollier MT, Petit-Conil M, Cornu D, Monties B, Van Montagu M, Inze D, et al: **Red Xylem and Higher Lignin Extractability by Down-Regulating a Cinnamyl Alcohol Dehydrogenase in Poplar.** *Plant Physiol* 1996, **112**:1479–1490.
95. Sibout R, Eudes A, Mouille G, Pollet B, Lapiere L, Jouanin L, Seguin A: **Cinnamyl Alcohol Dehydrogenase-C and -D Are the Primary Genes Involved in Lignin Biosynthesis in the Floral Stem of Arabidopsis.** *Plant Cell* 2005, **17**:2059–2076.
96. Vincent D, Lapiere C, Pollet B, Cornic G, Negroni L, Zivy M: **Water Deficits Affect Caffeate O-methyltransferase, Lignification and Related Enzymes in Maize Leaves. A Proteomic Investigation.** *Plant Physiology* 2005, **137**:949–960.
97. Bout S, Vermerris W: **A candidate-gene approach to clone the sorghum Brown midrib gene encoding caffeic acid O-methyltransferase.** *Mole Genetics and Genomics* 2003, **269**:205–214.
98. Vignols F, Rigau J, Torres MA, Capellades M, Puigdomenech P: **The brown midrib3 (bm3) Mutation in Maize Occurs in the Gene Encoding Caffeic Acid O-Methyltransferase.** *Plant Cell* 1995, **7**:407–416.
99. Theeraratnanon K, Wu X, Staggrenborg S, Propher R, Madl R, Wang D: **Evaluation and characterization of sorghum biomass as feedstock for sugar production.** *Trans ASABE* 2010, **53**:509–525.

100. Wu L, Arakane M, Ike M, Wada M, Takai T, Gau M, Tokuyasu K: **Low temperature alkali pretreatment for improving enzymatic digestibility of sweet sorghum bagasse for ethanol production.** *Bioresour Technol* 2011, **102**:4793–4799.
101. Corredor DY, Salazar JM, Hohn KL, Bean S, Bean B, Wang D: **Evaluation and Characterization of Forage Sorghum as Feedstock for Fermentable Sugar Production.** *Appl Biochem Biotechnol* 2009, **158**:164–179.
102. Sattler SE, Funnell-Harris DL, Pedersen JF: **Efficacy of Singular and Stacked brown midrib 6 and 12 in the Modification of Lignocellulose and Grain Chemistry.** *J Agric Food Chem* 2010, **58**:3611–3613.
103. Miller JE, Geadelmann JL, Marten GC: **Effect Of The Brown Midrib-allele On Maize Silage Quality And Yield.** *Crop Sci* 1983, **23**:493–496.
104. Inoue N, Kasuga S: **Agronomic traits and nutritive value of stover in brown midrib-3 maize hybrids.** *J Japanese Soc Grassland Sci* 1989, **35**:220–227.
105. Weller RF, Phipps RH, Cooper A: **The effect of the brown midrib-3 gene on the maturity and yield of forage maize.** *Grass and Forage Sci* 1985, **40**:335–339.
106. Hua SST, Grosjean OK, Baker JL: **Inhibition of aflatoxin biosynthesis by phenolic compounds.** *Lett Appl Microbiol* 1999, **29**:289–291.
107. McKeenhen JD, Busch RH, Fulcher RG: **Evaluation of Wheat (*Triticum aestivum* L.) Phenolic Acids during Grain Development and Their Contribution to Fusarium Resistance.** *J Agric Food Chem* 1999, **47**:1476–1482.
108. Dowd PF, Duvick JP, Root T: **Comparative toxicity of allelochemicals and their enzymatic oxidation products to maize fungal pathogens, emphasizing Fusarium graminearum.** *Natural Toxins* 1997, **5**:180–185.
109. Lee MH, Brewbacker JL: **Effects of brown midrib-3 on yields and yield components of maize.** *Crop Sci* 1984, **24**:105–108.
110. Gentinetta E, Bertolini M, Rossi I, Lorenzoni C, Motto M: **Effect of brown midrib-3 mutant on forage quality and yield in maize.** *J Genet Breeding* 1990, **44**:21–26.
111. Bean B, McCollum T, McCuiston K, Robinson J, Villeareal B, VanMeter R, Pietsch D: **Texas panhandle Forage Sorghum Silage Trial.** In *Book Texas panhandle Forage Sorghum Silage Trial*. City:exas Cooperative Extension and Texas Agricultural Experiment Station; 2006.
112. Bean B, McCollum T, Villeareal B, Blumenthal J, Robinson J, Brandon R, Buttrey E, VanMeter R, Pietsch D: **Texas panhandle Forage Sorghum Silage Trial.** In *Book Texas panhandle Forage Sorghum Silage Trial*. City: Texas Cooperative Extension and Texas Agricultural Experiment Station; 2008.
113. Oliver AL, Pedersen JF, Grant RJ, Klopfenstein TJ, Jose HD: **Comparative Effects of the Sorghum *bmr-6* and *bmr-12* Genes: II. Grain Yield, Stover Yield, and Stover Quality in Grain Sorghum.** *Crop Sci* 2005, **45**:2240–2245.
114. Fu C, Sunkar R, Zhou C, Shen H, Zhang J-Y, Matts J, Wolf J, Mann DGJ, Stewart CN, Tang Y, Wang Z-Y: **Overexpression of miR156 in switchgrass (*Panicum virgatum* L.) results in various morphological alterations and leads to improved biomass production.** *Plant Biotechnology J* 2012, **10**:443–452.
115. Xie K, Wu C, Xiong L: **Genomic Organization, Differential Expression, and Interaction of SQUAMOSA Promoter-Binding-Like Transcription Factors and microRNA156 in Rice.** *Plant Physiol* 2006, **142**:280–293.
116. Chuck G, Cigan AM, Saetern K, Hake S: **The heterochronic maize mutant *Corngrass1* results from overexpression of a tandem microRNA.** *Nat Genet* 2007, **39**:544–549.
117. Tamagnone L, Merida A, Parr A, Mackay S, Culianez-Macia FA, Roberts K, Martin C: **The AmMYB308 and AmMYB330 Transcription Factors from *Antirrhinum* Regulate Phenylpropanoid and Lignin Biosynthesis in Transgenic Tobacco.** *Plant Cell* 1998, **10**:135–154.
118. Jin H, Cominelli E, Bailey P, Parr A, Mehrtens F, Jones J, Tonelli C, Weisshaar B, Martin C: **Transcriptional repression by AtMYB4 controls production of UV-protecting sunscreens in *Arabidopsis*.** *EMBO J* 2000, **19**:6150–6161.
119. Fornalé S, Shi X, Chai C, Encina A, Irar S, Capellades M, Fuguet E, Torres J-L, Rovira P, Puigdomènech P, et al: **ZmMYB31 directly represses maize lignin genes and redirects the phenylpropanoid metabolic flux.** *The Plant J* 2010, **64**:633–644.
120. Fornalé S, Sonbol F-M, Maes T, Capellades M, Puigdomènech P, Rigau J, Caparrós-Ruiz D: **Down-regulation of the maize and *Arabidopsis thaliana* caffeic acid O-methyl-transferase genes by two new maize R2R3-MYB transcription factors.** *Plant Mol Biol* 2006, **62**:809–823.
121. Shen H, He X, Poovaiyah CR, Wuddineh WA, Ma J, Mann DGJ, Wang H, Jackson L, Tang Y, Neal Stewart C, et al: **Functional characterization of the switchgrass (*Panicum virgatum*) R2R3-MYB transcription factor PvMYB4 for improvement of lignocellulosic feedstocks.** *New Phytol* 2012, **193**:121–136.
122. Segal L, Creely JJ, Martin AE Jr, Conrad CM: **An Empirical Method for Estimating the Degree of Crystallinity of Native Cellulose Using the X-Ray Diffractometer.** *Text Res J* 1959, **29**:786–794.
123. Newman RH: **Homogeneity in cellulose crystallinity between samples of *Pinus radiata* wood.** *Holzforchung* 2004, **58**:91–96.
124. Åkerholm M, Hinterstoisser B, Salmén L: **Characterization of the crystalline structure of cellulose using static and dynamic FT-IR spectroscopy.** *Carbohydr Res* 2004, **339**:569–578.
125. Evans R, Newman RH, Roick UC: **Changes in cellulose crystallinity during kraft pulping. Comparison of infrared, x-ray diffraction and solid state NMR results.** *Holzforchung* 1995, **49**:498–504.
126. Kataoka Y, Kondo T: **FT-IR Microscopic Analysis of Changing Cellulose Crystalline Structure during Wood Cell Wall Formation.** *Macromolecules* 1998, **31**:760–764.
127. Schenzel K, Fischer S, Brendler E: **New Method for Determining the Degree of Cellulose I Crystallinity by Means of FT Raman Spectroscopy.** *Cellulose* 2005, **12**:223–231.
128. Vandenbrink J, Hilten R, Das K, Paterson A, Feltus F: **Analysis of Crystallinity Index and Hydrolysis Rates in the Bioenergy Crop *Sorghum bicolor*.** *BioEnergy Res* 2011, **5**:1–11.
129. Reddy N, Yang YQ: **Structure and properties of natural cellulose fibers obtained from sorghum leaves and stems.** *J Agric Food Chem* 2007, **55**:5569–5574.
130. Isogai A, Atalla RH: **Dissolution of Cellulose in Aqueous NaOH Solutions.** *Cellulose* 1998, **5**:309–319.
131. Mansikkamäki P, Lahtinen M, Rissanen K: **Structural Changes of Cellulose Crystallites Induced by Mercerisation in Different Solvent Systems; Determined by Powder X-ray Diffraction Method.** *Cellulose* 2005, **12**:233–242.
132. Liu L, Sun J, Li M, Wang S, Pei H, Zhang J: **Enhanced enzymatic hydrolysis and structural features of corn stover by FeCl<sub>3</sub> pretreatment.** *Bioresour Technol* 2009, **100**:5853–5858.
133. Barl B, Billaderis CG, Murray ED: **Effect of chemical pretreatments on the thermal degradation of corn husk lignocellulosics.** *J Agric Food Chem* 1986, **34**:1019–1024.
134. Xiao Z, Gao P, Qu Y, Wang T: **Cellulose-binding domain of endoglucanase III from *Trichoderma reesei* disrupting the structure of cellulose.** *Biotechnol Lett* 2001, **23**:711–715.
135. Ciolacu D, Kovac J, Kokol V: **The effect of the cellulose-binding domain from *Clostridium cellulovorans* on the supramolecular structure of cellulose fibers.** *Carbohydr Res* 2010, **345**:621–630.
136. Harris D, DeBolt S: **Relative Crystallinity of Plant Biomass: Studies on Assembly, Adaptation and Acclimation.** *PLoS One* 2008, **3**:e2897.
137. Hallac BB, Pu Y, Ragauskas AJ: **Chemical Transformations of *Buddleja davidii* Lignin during Ethanol Organosolv Pretreatment.** *Energy Fuel* 2010, **24**:2723–2732.
138. Fan LT, Lee YH, Beardmore DH: **Mechanism of the enzymatic-hydrolysis of cellulose - effects of major structural features of cellulose on enzymatic-hydrolysis.** *Biotechnol Bioeng* 1980, **22**:177–199.
139. Penttilä PA, Várnai A, Leppänen K, Peura M, Kallonen A, Jääskeläinen P, Lucenius J, Ruokolainen J, Siika-aho M, Viikari L, Serimaa R: **Changes in Submicrometer Structure of Enzymatically Hydrolyzed Microcrystalline Cellulose.** *Biomacromolecules* 2010, **11**:1111–1117.
140. Hall M, Bansal P, Lee JH, Realf MJ, Bommarius AS: **Cellulose crystallinity - a key predictor of the enzymatic hydrolysis rate.** *FEBS J* 2010, **277**:1571–1582.
141. Ramos LP, Nazhad MM, Saddler JN: **Effect of enzymatic hydrolysis on the morphology and fine structure of pretreated cellulosic residues.** *Enzym Microb Technol* 1993, **15**:821–831.
142. Lu Y, Yang B, Gregg D, Saddler JN, Mansfield SD: **Cellulase adsorption and an evaluation of enzyme recycle during hydrolysis of steam-exploded softwood residues.** *Appl Biochem Biotechnol* 2002, **98–100**:641–654.
143. Lee Y-H, Fan LT: **Kinetic studies of enzymatic hydrolysis of insoluble cellulose: Analysis of the initial rates.** *Biotechnol Bioeng* 1982, **24**:2383–2406.
144. Converse A, Ooshima H, Burns D: **Kinetics of enzymatic hydrolysis of lignocellulosic materials based on surface area of cellulose accessible to enzyme and enzyme adsorption on lignin and cellulose.** *Appl Biochem Biotechnol* 1990, **24–25**:67–73.

145. Jeoh T, Ishizawa CI, Davis MF, Himmel ME, Adney WS, Johnson DK: Cellulose digestibility of pretreated biomass is limited by cellulose accessibility. *Biotechnol Bioeng* 2007, **98**:112–122.
146. Karlsson J, Medve J, Tjerneld F: Hydrolysis of steam-pretreated lignocellulose. *Appl Biochem Biotechnol* 1999, **82**:243–258.
147. Kumar R, Wyman CE: Access of cellulase to cellulose and lignin for poplar solids produced by leading pretreatment technologies. *Biotechnol Prog* 2009, **25**:807–819.
148. Kumar R, Wyman CE: Cellulase adsorption and relationship to features of corn stover solids produced by leading pretreatments. *Biotechnol Bioeng* 2009, **103**:252–267.
149. Gharpuray MM, Lee Y-H, Fan LT: Structural modification of lignocellulosics by pretreatments to enhance enzymatic hydrolysis. *Biotechnol Bioeng* 1983, **25**:157–172.
150. Sinitsyn A, Gusakov A, Vlasenko E: Effect of structural and physico-chemical features of cellulosic substrates on the efficiency of enzymatic hydrolysis. *Appl Biochem Biotechnol* 1991, **30**:43–59.
151. Nutor J, Converse A: The effect of enzyme and substrate levels on the specific hydrolysis rate of pretreated poplar wood. *Appl Biochem Biotechnol* 1991, **28–29**:757–772.
152. Deshpande MV, Eriksson KE: Reutilization of enzymes for saccharification of lignocellulosic materials. *Enzym Microb Technol* 1984, **6**:338–340.
153. Ooshima H, Sakata M, Harano Y: Adsorption of cellulase from *Trichoderma viride* on cellulose. *Biotechnol Bioeng* 1983, **25**:3103–3114.
154. Girard D, Converse A: Recovery of cellulase from lignaceous hydrolysis residue. *Appl Biochem Biotechnol* 1993, **39–40**:521–533.
155. Chernaglazov VM, Ermolova OV, Klyozov AA: Adsorption of high purity endo-1-4- $\beta$ -glucanases from *Trichoderma reesei* on components of lignocellulosic materials: cellulose, lignin and xylan. *Enzym Microb Technol* 1988, **10**:503–507.
156. Palonen H, Tjerneld F, Zacchi G, Tenkanen M: Adsorption of *Trichoderma reesei* CBH I and EG II and their catalytic domains on steam pretreated softwood and isolated lignin. *J Biotechnol* 2004, **107**:65–72.
157. Kim DW, Kim TS, Jeong YK, Lee JK: Adsorption kinetics and behaviors of cellulase components on microcrystalline cellulose. *J Ferment Bioeng* 1992, **73**:461–466.
158. Eriksson T, Karlsson J, Tjerneld F: A model explaining declining rate in hydrolysis of lignocellulose substrates with cellobiohydrolase I (Cel7A) and endoglucanase I (Cel7B) of *Trichoderma reesei*. *Appl Biochem Biotechnol* 2002, **101**:41–60.
159. Yang B, Wyman CE: BSA treatment to enhance enzymatic hydrolysis of cellulose in lignin containing substrates. *Biotechnol Bioeng* 2006, **94**:611–617.
160. Brash John L, Horbett Thomas A: Proteins at Interfaces. In *Proteins at Interfaces II*, ACS Symposium Series, Volume 602. Washington, D.C.: Am Chemical Soc; 1995:1–23.
161. Reinikainen T, Teleman O, Teeri TT: Effects of pH and high ionic strength on the adsorption and activity of native and mutated cellobiohydrolase I from *Trichoderma reesei*. *Proteins: Structure, Function, and Bioinformatics* 1995, **22**:392–403.
162. Ramos L, Saddler J: Enzyme recycling during fed-batch hydrolysis of cellulose derived from steam-exploded *Eucalyptus viminalis*. *Appl Biochem Biotechnol* 1994, **45–46**:193–207.
163. Lee D, Yu AHC, Saddler JN: Evaluation of cellulase recycling strategies for the hydrolysis of lignocellulosic substrates. *Biotechnol Bioeng* 1995, **45**:328–336.
164. Lee D, Yu A, Wong K, Saddler J: Evaluation of the enzymatic susceptibility of cellulosic substrates using specific hydrolysis rates and enzyme adsorption. *Appl Biochem Biotechnol* 1994, **45–46**:407–415.
165. Mes-Hartree M, Hogan CM, Saddler JN: Recycle of enzymes and substrate following enzymatic hydrolysis of steam-pretreated aspenwood. *Biotechnol Bioeng* 1987, **30**:558–564.
166. Jackson LS, Joyce TW, Heitmann JA, Giesbrecht FG: Enzyme activity recovery from secondary fiber treated with cellulase and xylanase. *J Biotechnol* 1996, **45**:33–44.
167. Tan LUL, Yu EKC, Campbell N, Saddler JN: Column cellulose hydrolysis reactor: An efficient cellulose hydrolysis reactor with continuous cellulase recycling. *Appl Microbiol Biotechnol* 1986, **25**:250–255.
168. Mores W, Knutsen J, Davis R: Cellulase recovery via membrane filtration. *Appl Biochem Biotechnol* 2001, **91–93**:297–309.
169. Ramos LP, Breuil C, Saddler JN: The use of enzyme recycling and the influence of sugar accumulation on cellulose hydrolysis by *Trichoderma* cellulases. *Enzym Microb Technol* 1993, **15**:19–25.
170. Castanon M, Wilke CR: Adsorption and recovery of cellulases during hydrolysis of newspaper. *Biotechnol Bioeng* 1980, **22**:1037–1053.
171. Tu M, Chandra RP, Saddler JN: Evaluating the Distribution of Cellulases and the Recycling of Free Cellulases during the Hydrolysis of Lignocellulosic Substrates. *Biotechnol Prog* 2007, **23**:398–406.
172. Vallander L, Eriksson K-E: Enzyme recirculation in saccharification of lignocellulosic materials. *Enzym Microb Technol* 1987, **9**:714–720.
173. Otter DE, Munro PA, Scott GK, Geddes R: Desorption of *Trichoderma reesei* cellulase from cellulose by a range of desorbents. *Biotechnol Bioeng* 1989, **34**:291–298.
174. Qi B, Chen X, Su Y, Wan Y: Enzyme adsorption and recycling during hydrolysis of wheat straw lignocellulose. *Bioresour Technol* 2011, **102**:2881–2889.
175. Paterson AH, Lin Y, Li Z, Schertz KF, Doebley JF, Pinson SRM, S L, JW S, Irvine JE: Convergent Domestication of Cereal Crops by Independent Mutations at Corresponding Genetic Loci. *Science* 1995, **269**:1714–1718.
176. Brown PJ, Rooney WL, Franks C, Kresovich S: Efficient mapping of plant height quantitative trait loci in a sorghum association population with introgressed dwarfing genes. *Genetics* 2008, **180**:629–637.
177. Ji-hua T, Wen-tao T, Jian-bing Y, Xi-qing M, Yi-jiang M, Jin-rui D, Jian-Sheng L: Genetic dissection of plant height by molecular markers using a population of recombinant inbred lines in maize. *Euphytica* 2007, **155**:117–124.
178. Lin YR, Schertz KF, Paterson AH: Comparative Analysis of QTLs Affecting Plant Height and Maturity across the Poaceae, in Reference to an Interspecific Sorghum Population. *Genetics* 1995, **141**:391–411.
179. Schnell RJ, Kuhn DN, Brown JS, Olano CT, Phillips-Mora W, Amores FM, Motamayor JC: Development of a marker assisted selection program for cacao. *Phytopath* 2007, **97**:1664–1669.
180. Jannink JL, Lorenz AJ, Iwata H: Genomic selection in plant breeding: from theory to practice. *Brief Funct Genomics* 2010, **9**:166–177.
181. Luo H, Lyznik LA, Gidoni D, Hodges TK: FLP-mediated recombination for use in hybrid plant production. *The Plant J* 2000, **23**:423–430.
182. Lander E, Schork N: Genetic dissection of complex traits. *Science* 1994, **265**:2037–2048.
183. Hamblin MT, Buckler ES, Jannink JL: Population genetics of genomics-based crop improvement methods. *Trends Genet* 2011, **27**:98–106.
184. McMullen MD, Kresovich S, Villeda HS, Bradbury P, Li HH, Sun Q, Flint-Garcia S, Thornsberry J, Acharya C, Bottoms C, et al: Genetic Properties of the Maize Nested Association Mapping Population. *Science* 2009, **325**:737–740.
185. Tian F, Bradbury PJ, Brown PJ, Hung H, Sun Q, Flint-Garcia S, Rocheford TR, McMullen MD, Holland JB, Buckler ES: Genome-wide association study of leaf architecture in the maize nested association mapping population. *Nat Genet* 2011, **43**:159–162.
186. Hirschhorn JN, Daly MJ: Genome-wide association studies for common diseases and complex traits. *Nat Rev Genet* 2005, **6**:95–108.
187. Thornsberry JM, Goodman MM, Doebley J, Kresovich S, Nielsen D, Buckler ES: Dwarf8 polymorphisms associate with variation in flowering time. *Nat Genet* 2001, **28**:286–289.
188. Pritchard JK, Stephens M, Donnelly P: Inference of Population Structure Using Multilocus Genotype Data. *Genetics* 2000, **155**:945–959.
189. Kennedy BW, Quinton M, van Arendonk JA: Estimation of effects of single genes on quantitative traits. *J Anim Sci* 1992, **70**:2000–2012.
190. Pritchard JK, Stephens M, Rosenberg NA, Donnelly P: Association Mapping in Structured Populations. *Am J Hum Genet* 2000, **67**:170–181.
191. Paterson AH, Schertz KF, Lin Y-R, Liu S-C, Chang Y-L: The Weediness of Wild Plants: Molecular Analysis of Genes Influencing Dispersal and Persistence of Johnsongrass, *Sorghum halepense* (L.) Pers. *Proc Natl Acad Sci U S A* 1995, **92**:6127–6131.
192. Xiao-ping L, Jin-feng Y, Cui-ping G, Acharya S: Quantitative trait loci analysis of economically important traits in *Sorghum bicolor*  $\times$  *S. sudanense* hybrid. *Canadian J Plant Sci* 2011, **91**:81–90.
193. Murray SC, Rooney WL, Mitchell SE, Sharma A, Klein PE, Mullet JE, Kresovich S: Genetic Improvement of Sorghum as a Biofuel Feedstock: II. QTL for Stem and Leaf Structural Carbohydrates. *Crop Sci* 2008, **48**:2180–2193.
194. Brown P, Klein P, Bortiri E, Acharya C, Rooney W, Kresovich S: Inheritance of inflorescence architecture in sorghum. *TAG Theor Appl Genet* 2006, **113**:931–942.

195. Feltus F, Hart G, Schertz K, Casa A, Kresovich S, Abraham S, Klein P, Brown P, Paterson A: **Alignment of genetic maps and QTLs between inter- and intra-specific sorghum populations.** *TAG Theor Appl Genet* 2006, **112**:1295–1305.
196. Hart GE, Schertz KF, Peng Y, Syed NH: **Genetic mapping of *Sorghum bicolor* (L.) Moench QTLs that control variation in tillering and other morphological characters.** *TAG Theor Appl Genet* 2001, **103**:1232–1242.
197. Kebede H, Subudhi PK, Rosenow DT, Nguyen HT: **Quantitative trait loci influencing drought tolerance in grain sorghum (*Sorghum bicolor* L. Moench).** *TAG Theor Appl Genet* 2001, **103**:266–276.
198. Parh D, Jordan D, Aitken E, Mace E, Jun-ai P, McIntyre C, Godwin I: **QTL analysis of ergot resistance in sorghum.** *TAG Theor Appl Genet* 2008, **117**:369–382.
199. Pereira MG, Lee M: **Identification of genomic regions affecting plant height in sorghum and maize.** *TAG Theor Appl Genet* 1995, **90**:380–388.
200. Rami JF, Dufour P, Trouche G, Fliedel G, Mestres C, Davrieux F, Blanchard P, Hamon P: **Quantitative trait loci for grain quality, productivity, morphological and agronomical traits in sorghum (*Sorghum bicolor* L. Moench).** *TAG Theor Appl Genet* 1998, **97**:605–616.
201. Ritter K, Jordan D, Chapman S, Godwin I, Mace E, Lynne McIntyre C: **Identification of QTL for sugar-related traits in a sweet x grain sorghum (*Sorghum bicolor* L. Moench) recombinant inbred population.** *Mol Breed* 2008, **22**:367–384.
202. Shiringani A, Frisch M, Friedt W: **Genetic mapping of QTLs for sugar-related traits in a RIL population of *Sorghum bicolor* L. Moench.** *TAG Theoretical and Applied Genet* 2010, **121**:323–336.
203. Srinivas G, Satish K, Madhusudhana R, Nagaraja Reddy R, Murali Mohan S, Seetharama N: **Identification of quantitative trait loci for agronomically important traits and their association with genic-microsatellite markers in sorghum.** *TAG Theor Appl Genet* 2009, **118**:1439–1454.
204. Chantreau J, Trouche G, Rami JF, Deu M, Barro C, Grivet L: **RFLP mapping of QTLs for photoperiod response in tropical sorghum.** *Euphytica* 2001, **120**:183–194.
205. Crasta OR, Xu WW, Rosenow DT, Mullet J, Nguyen HT: **Mapping of post-flowering drought resistance traits in grain sorghum: association between QTLs influencing premature senescence and maturity.** *Mol Gen Genet* 1999, **262**:579–588.
206. Murray SC, Sharma A, Rooney WL, Klein PE, Mullet JE, Mitchell SE, Kresovich S: **Genetic Improvement of Sorghum as a Biofuel Feedstock: I. QTL for Stem Sugar and Grain Nonstructural Carbohydrates.** *Crop Sci* 2008, **48**:2165–2179.
207. Tuinstra MR, Grote EM, Goldsbrough PB, Ejeta G: **Genetic analysis of post-flowering drought tolerance and components of grain development in *Sorghum bicolor* (L.) Moench.** *Mol Breed* 1997, **3**:439–448.
208. Fernandez MGS, Hamblin MT, Li L, Rooney WL, Tuinstra MR, Kresovich S: **Quantitative Trait Loci Analysis of Endosperm Color and Carotenoid Content in Sorghum Grain.** *Crop Sci* 2008, **48**:1732–1743.
209. Winn JA, Mason RE, Robbins AL, Rooney WL, Hays DB: **QTL Mapping of a High Protein Digestibility Trait in Sorghum bicolor.** *Int J Plant Genomics* 2009, **2009**:1–6.
210. Felderhoff TJ, Murray SC, Klein PE, Sharma A, Hamblin MT, Kresovich S, Vermerris W, Rooney WL: **QTLs for Energy-related Traits in a Sweet x Grain Sorghum [*Sorghum bicolor* (L.) Moench] Mapping Population.** *Crop Sci* 2012, **52**:2040–2049.
211. Méchin V, Argillier O, Hébert Y, Guingo E, Moreau L, Charcosset A, Barrière Y: **Genetic Analysis and QTL Mapping of Cell Wall Digestibility and Lignification in Silage Maize.** *Crop Sci* 2001, **41**:690–697.
212. Roussel V, Gibelin C, Fontaine AS, Barrière Y: **Genetic analysis in recombinant inbred lines of early dent forage maize. II - QTL mapping for cell wall constituents and cell wall digestibility from per se value and top cross experiments.** *Maydica* 2002, **47**:9–20.
213. Cardinal A, Lee M, Moore K: **Genetic mapping and analysis of quantitative trait loci affecting fiber and lignin content in maize.** *TAG Theor Appl Genet* 2003, **106**:866–874.
214. Krakowsky M, Lee M, Coors J: **Quantitative trait loci for cell wall components in recombinant inbred lines of maize (*Zea mays* L.) I. stalk tissue.** *TAG Theor Appl Genet* 2005, **111**:337–346.
215. Krakowsky M, Lee M, Coors J: **Quantitative trait loci for cell wall components in recombinant inbred lines of maize (*Zea mays* L.) II. leaf sheath tissue.** *TAG Theor Appl Genet* 2006, **112**:717–726.
216. Ni W, Phillips RL, Jung HG: **Quantitative trait loci for cell-wall traits in maize.** In *8th international cell wall meeting; Sept 1–5*. Norwich, UK: John Innes Centre; 1998:1–38.
217. Fontaine A-S, Bout S, Barrière Y, Vermerris W: **Variation in Cell Wall Composition among Forage Maize (*Zea mays* L.) Inbred Lines and Its Impact on Digestibility: Analysis of Neutral Detergent Fiber Composition by Pyrolysis-Gas chromatography–mass spectrometry.** *J Agric Food Chem* 2003, **51**:8080–8087.
218. Fontaine AS, Briand M, Barrière Y: **Genetic variation and QTL mapping of para-coumaric and ferulic acid contents in maize stover at silage harvest.** *Maydica* 2003, **48**:75–84.
219. Krakowsky MD, Lee M, Holland JB: **Genotypic Correlation and Multivariate QTL Analyses for Cell Wall Components and Resistance to Stalk Tunneling by the European Corn Borer in Maize.** *Crop Sci* 2007, **47**:485–488.
220. Ribaut JM, Jiang C, Gonzalez-de-Leon D, Edmeades GO, Hoisington DA: **Identification of quantitative trait loci under drought conditions in tropical maize. 2. Yield components and marker-assisted selection strategies.** *TAG Theor Appl Genet* 1997, **94**:887–896.
221. Ribaut J-M, Fracheboud Y, Monneveux P, Banziger M, Vargas M, Jiang C: **Quantitative trait loci for yield and correlated traits under high and low soil nitrogen conditions in tropical maize.** *Mol Breed* 2007, **20**:15–29.
222. Ajmone Marsan P, Monfredini G, Ludwig WF, Melchinger AE, Franceschini P, Pagnotto G, Motto M: **In an elite cross of maize a major quantitative trait locus controls 1/4 of the genetic variation for grain-yield.** *TAG Theor Appl Genet* 1995, **90**:415–424.
223. Frova C, Krajewski P, di Fonzo N, Villa M, Sari-Gorla M: **Genetic analysis of drought tolerance in maize by molecular markers I. Yield components.** *TAG Theoretical and Applied Genet* 1999, **99**:280–288.
224. Frascaroli E, Cane MA, Landi P, Pea G, Gianfranceschi L, Villa M, Morgante M, Enrico Pe M: **Classical Genetic and Quantitative Trait Loci Analyses of Heterosis in a Maize Hybrid Between Two Elite Inbred Lines.** *Genetics* 2007, **176**:625–644.
225. Stuber CW, Lincoln SE, Wolff DW, Helentjaris T, Lander ES: **Identification of genetic factors contributing to heterosis in a hybrid from two elite maize inbred lines using molecular markers.** *Genetics* 1992, **132**:823–839.
226. Beavis WD, Smith OS, Grant D, Fincher R: **Identification of Quantitative Trait Loci Using a Small Sample of Topcrossed and F4 Progeny from Maize.** *Crop Sci* 1994, **34**:882–896.
227. Xiao Y, Li X, George M, Li M, Zhang S, Zheng Y: **Quantitative trait locus analysis of drought tolerance and yield in Maize in China.** *Plant Mol Biol Rep* 2005, **23**:155–165.
228. Li Y, Niu S, Dong Y, Cui D, Wang Y, Liu Y, Wei M: **Identification of trait-improving quantitative trait loci for grain yield components from a dent corn inbred line in an advanced backcross BC2F2 population and comparison with its F2:3 population in popcorn.** *TAG Theor Appl Genet* 2007, **115**:129–140.
229. Austin DF, Lee M: **Comparative mapping in F2:3 and F6:7 generations of quantitative trait loci for grain yield and yield components in maize.** *TAG Theor Appl Genet* 1996, **92**:817–826.
230. Austin DF, Lee M, Veldboom LR, Hallauer AR: **Genetic Mapping in Maize with Hybrid Progeny Across Testers and Generations: Grain Yield and Grain Moisture.** *Crop Sci* 2000, **40**:30–39.
231. Veldboom LR, Lee M: **Genetic Mapping of Quantitative Trait Loci in Maize in Stress and Nonstress Environments: I. Grain Yield and Yield Components.** *Crop Sci* 1996, **36**:1310–1319.
232. Lu H, Romero-Severson J, Bernardo R: **Genetic basis of heterosis explored by simple sequence repeat markers in a random-mated maize population.** *TAG Theor Appl Genet* 2003, **107**:494–502.
233. Agrama HAS, Moussa ME: **Mapping QTLs in breeding for drought tolerance in maize (*Zea mays* L.).** *Euphytica* 1996, **91**:89–97.
234. Austin DF, Lee M: **Genetic resolution and verification of quantitative trait loci for flowering and plant height with recombinant inbred lines of maize.** *Genome* 1996, **39**:957–968.
235. Khairallah MM, Bohn M, Jiang C, Deutsch JA, Jewell DC, Mihm JA, Melchinger AE, González-De-León D, Hoisington DA: **Molecular mapping of QTL for southwestern corn borer resistance, plant height and flowering in tropical maize.** *Plant Breeding* 1998, **117**:309–318.
236. Koester RP, Sisco PH, Stuber CW: **Identification of Quantitative Trait Loci Controlling Days to Flowering and Plant Height in Two Near Isogenic Lines of Maize.** *Crop Sci* 1993, **33**:1209–1216.

237. Veldboom LR, Lee M: Genetic Mapping of Quantitative Trait Loci in Maize in Stress and Nonstress Environments: II. Plant Height and Flowering. *Crop Sci* 1996, **36**:1320–1327.
238. Veldboom LR, Lee M, Woodman WL: Molecular marker-facilitated studies in an elite maize population: I. Linkage analysis and determination of QTL for morphological traits. *TAG Theor Appl Genet* 1994, **88**:7–16.
239. Berke TG, Rocheford TR: Quantitative Trait Loci for Flowering, Plant and Ear Height, and Kernel Traits in Maize. *Crop Sci* 1995, **35**:1542–1549.
240. Sari-Gorla M, Krajewski P, Di Fonzo N, Villa M, Frova C: Genetic analysis of drought tolerance in maize by molecular markers. II. Plant height and flowering. *TAG Theoretical and Applied Genet* 1999, **99**:289–295.
241. Bohn M, Khairallah MM, González-de-León D, Hoisington DA, Utz HF, Deutsch JA, Jewell DC, Mihm JA, Melchinger AE: QTL Mapping in Tropical Maize: I. Genomic Regions Affecting Leaf Feeding Resistance to Sugarcane Borer and Other Traits. *Crop Science* 1996, **36**:1352–1361.
242. Yan J, Tang H, Huang Y, Shi Y, Li J, Zheng Y: Dynamic analysis of QTL for plant height at different developmental stages in maize (*Zea mays* L.). *Chin Sci Bull* 2003, **48**:2601–2607.
243. Wisser RJ, Balint-Kurti PJ, Nelson RJ: The Genetic Architecture of Disease Resistance in Maize: A Synthesis of Published Studies. *Phytopathology* 2006, **96**:120–129.
244. Jines M, Balint-Kurti P, Robertson-Hoyt L, Molnar T, Holland J, Goodman M: Mapping resistance to Southern rust in a tropical by temperate maize recombinant inbred topcross population. *TAG Theor Appl Genet* 2007, **114**:659–667.
245. Raymond M, Muller B, Leonardi A, Charcosset A, Tardieu F: Combining Quantitative Trait Loci Analysis and an Ecophysiological Model to Analyze the Genetic Variability of the Responses of Maize Leaf Growth to Temperature and Water Deficit. *Plant Physiol* 2003, **131**:664–675.
246. Raymond M, Muller B, Tardieu F: Dealing with the genotype × environment interaction via a modelling approach: a comparison of QTLs of maize leaf length or width with QTLs of model parameters. *J Exp Bot* 2004, **55**:2461–2472.
247. Sadok W, Naudin P, Bousuge B, Muller B, Welcker C, Tardieu F: Leaf growth rate per unit thermal time follows QTL-dependent daily patterns in hundreds of maize lines under naturally fluctuating conditions. *Plant Cell Environ* 2007, **30**:135–146.
248. Atienza SG, Satovic Z, Petersen KK, Dolstra O, Martín A: Identification of QTLs influencing agronomic traits in *Miscanthus sinensis* Anders. I. Total height, flag-leaf height and stem diameter. *TAG Theor Appl Genet* 2003, **107**:123–129.
249. Atienza SG, Satovic Z, Petersen KK, Dolstra O, Martín A: Identification of QTLs associated with yield and its components in *Miscanthus sinensis* Anders. *Euphytica* 2003, **132**:353–361.
250. Jensen EF, Jones ST, Farrar K, Clifton-Brown JC, Donnison IS: Unravelling the genetic control of flowering time in the bioenergy grass *Miscanthus*. *Comparative Biochemistry and Physiology* 2008, **150**:5181.
251. Atienza SG, Ramirez MC, Martín A: Mapping QTLs controlling flowering date in *Miscanthus sinensis* Anders. *Cereal Res Communications* 2003, **31**:265–271.
252. Atienza SG, Satovic Z, Petersen KK, Dolstra O, Martín A: Influencing combustion quality in *Miscanthus sinensis* Anders.: Identification of QTLs for calcium, phosphorus and sulphur content. *Plant Breeding* 2003, **122**:141–145.
253. Atienza SG, Satovic Z, Petersen KK, Dolstra O, Martín A: Identification of QTLs influencing combustion quality in *Miscanthus sinensis* Anders. II. Chlorine and potassium content. *TAG Theor Appl Genet* 2003, **107**:857–863.
254. Ming RM, Wang YW, Draye XD, Moore PM, Irvine JI, Paterson AP: Molecular dissection of complex traits in autopolyploids: mapping QTLs affecting sugar yield and related traits in sugarcane. *TAG Theor Appl Genet* 2002, **105**:332–345.
255. Hoarau JY, Offmann B, D'Hont A, Risterucci AM, Roques D, Glaszmann JC, Grivet L: Genetic dissection of a modern sugarcane cultivar (*Saccharum* spp.). I. Genome mapping with AFLP markers. *TAG Theor Appl Genet* 2001, **103**:84–97.
256. Ming R, Liu S-C, Moore PH, Irvine JE, Paterson AH: QTL Analysis in a Complex Autopolyploid: Genetic Control of Sugar Content in Sugarcane. *Genome Res* 2001, **11**:2075–2084.
257. Aitken K, Jackson P, McIntyre C: Quantitative trait loci identified for sugar related traits in a sugarcane (*Saccharum* spp.) cultivar × *Saccharum officinarum* population. *TAG Theor Appl Genet* 2006, **112**:1306–1317.
258. Liu L, Wu Y, Wang Y, Samuels T: A high-density simple sequence repeat-based genetic linkage map of switchgrass. *G3 (Bethesda)* 2012, **2**:357–370.
259. Ma X-F, Jensen E, Alexandrov N, Troukhan M, Zhang L, Thomas-Jones S, Farrar K, Clifton-Brown J, Donnison I, Swaller T, Flavell R: High Resolution Genetic Mapping by Genome Sequencing Reveals Genome Duplication and Tetraploid Genetic Structure of the Diploid *Miscanthus sinensis*. *PLoS One* 2012, **7**:e33821.
260. Swaminathan K, Chae W, Mitros T, Varala K, Xie L, Barling A, Glowacka K, Hall M, Jezowski S, Ming R, et al: A framework genetic map for *Miscanthus sinensis* from RNAseq-based markers shows recent tetraploidy. *BMC Genomics* 2012, **13**:142.
261. Zhang NW, Pelgrom K, Niks RE, Visser RG, Jeuken MJ: Three combined quantitative trait loci from nonhost *Lactuca saligna* are sufficient to provide complete resistance of lettuce against *Bremia lactucae*. *Mol Plant Microbe Interact* 2009, **22**:1160–1168.
262. Wang P, Xing Y, Li Z, Yu S: Improving rice yield and quality by QTL pyramiding. *Mol Breed* 2012, **29**:903–913.
263. Wang S, Wu K, Yuan Q, Liu X, Liu Z, Lin X, Zeng R, Zhu H, Dong G, Qian Q, et al: Control of grain size, shape and quality by OsSPL16 in rice. *Nat Genet* 2012, **44**:950–954.
264. Yu J, Holland JB, McMullen MD, Buckler ES: Genetic Design and Statistical Power of Nested Association Mapping in Maize. *Genetics* 2008, **178**:539–551.
265. Kresovich S, Paterson AH, Feltus FA: Genomic and Breeding Foundations for Bioenergy Sorghum Hybrids. [http://genomicscience.energy.gov/research/DOEUSDA/abstracts/2011Kresovich\\_abstract.shtml](http://genomicscience.energy.gov/research/DOEUSDA/abstracts/2011Kresovich_abstract.shtml).
266. Poland JA, Bradbury PJ, Buckler ES, Nelson RJ: Genome-wide nested association mapping of quantitative resistance to northern leaf blight in maize. *Proc Natl Acad Sci U S A* 2011, **108**:6893–6898.
267. Doerge RW, Craig BA: Model selection for quantitative trait locus analysis in polyploids. *Proc Natl Acad Sci* 2000, **97**:7951–7956.
268. Bowers JE, Abbey C, Anderson S, Chang C, Draye X, Hoppe AH, Jessup R, Lemke C, Lenington J, Li Z, et al: A high-density genetic recombination map of sequence-tagged sites for sorghum, as a framework for comparative structural and evolutionary genomics of tropical grains and grasses. *Genetics* 2003, **165**:367–386.
269. Casa AM, Pressoir G, Brown PJ, Mitchell SE, Rooney WL, Tuinstra MR, Franks CD, Kresovich S: Community Resources and Strategies for Association Mapping in Sorghum. *Crop Sci* 2008, **48**:30–40.
270. Ersoz ES, Yu J, Buckler ES: Applications of Linkage Disequilibrium and Association Mapping in Crop Plants. In *Genomics-Assisted Crop Improvement. Volume 1*. Edited by Varshney RK, Tuberosa R. Netherlands: Springer; 2007:97–119.
271. Buckler IV ES, Thornsberry JM: Plant molecular diversity and applications to genomics. *Current Opinion in Plant Biology* 2002, **5**:107–111.
272. Flint-Garcia SA, Thornsberry JM, Buckler ES: Structure of linkage disequilibrium in plants. *Annual Rev Plant Biol* 2003, **54**:357–374.
273. Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, Kresovich S, Goodman MM, Buckler ES: Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc Natl Acad Sci U S A* 2001, **98**:11479–11484.
274. Flint-Garcia SA, Thuitet A-C, Yu J, Pressoir G, Romero SM, Mitchell SE, Doebley J, Kresovich S, Goodman MM, Buckler ES: Maize association population: a high-resolution platform for quantitative trait locus dissection. *The Plant J* 2005, **44**:1054–1064.
275. Yang X, Yan J, Shah T, Warburton M, Li Q, Li L, Gao Y, Chai Y, Fu Z, Zhou Y, et al: Genetic analysis and characterization of a new maize association mapping panel for quantitative trait loci dissection. *TAG Theor Appl Genet* 2010, **121**:417–431.
276. Murray SC, Rooney WL, Hamblin MT, Mitchell SE, Kresovich S: Sweet Sorghum Genetic Diversity and Association Mapping for Brix and Height. *Plant Gen* 2009, **2**:48–62.
277. Wei X, Jackson P, McIntyre C, Aitken K, Croft B: Associations between DNA markers and resistance to diseases in sugarcane and effects of population substructure. *TAG Theor Appl Genet* 2006, **114**:155–164.
278. Wang Y-H, Bible P, Loganathanaraj R, Upadhyaya H: Identification of SSR markers associated with height using pool-based genome-wide association mapping in sorghum. *Mol Breed* 2011, **30**:1–12.

279. Shehzad T, Iwata H, Okuno K: **Genome-wide association mapping of quantitative traits in sorghum (*Sorghum bicolor* (L.) Moench) by using multiple models.** *Breed Sci* 2009, **59**:217–227.
280. Andersen JR, Schrag T, Melchinger AE, Zein I, Lübberstedt T: **Validation of *Dwarf8* polymorphisms associated with flowering time in elite European inbred lines of maize (*Zea mays* L.).** *TAG Theor Appl Genet* 2005, **111**:206–217.
281. Camus-Kulandaivelu L, Veyrieras J-B, Madur D, Combes V, Fourmann M, Barraud S, Dubreuil P, Gouesnard B, Manicacci D, Charcosset A: **Maize Adaptation to Temperate Climate: Relationship Between Population Structure and Polymorphism in the *Dwarf8* Gene.** *Genetics* 2006, **172**:2449–2463.
282. Wilson LM, Whitt SR, Ibáñez AM, Rocheford TR, Goodman MM, Buckler ES: **Dissection of Maize Kernel Composition and Starch Production by Candidate Gene Association.** *Plant Cell* 2004, **16**:2719–2733.
283. Tracy WF, Whitt SR, Buckler ES: **Recurrent Mutation and Genome Evolution: Example of and the Origin of Sweet Maize.** *Crop Sci* 2006, **46**:S-49–S-54.
284. Burner DM, Legendre BL: **Cytogenetic and fertility characteristics of elite sugarcane clones.** *Sugar Cane* 1994, **1**:6–10.
285. Comai L, Henikoff S: **TILLING: practical single-nucleotide mutation discovery.** *The Plant J* 2006, **45**:684–694.
286. Slade AJ, Knauf VC: **TILLING moves beyond functional genomics into crop improvement.** *Transgenic Res* 2005, **14**:109–115.
287. Gilchrist EJ, Haughn GW: **TILLING without a plough: a new method with applications for reverse genetics.** *Current Opinion in Plant Biology* 2005, **8**:211–215.
288. Xin Z, Chen J: **A high throughput DNA extraction method with high yield and quality.** *Plant Methods* 2012, **8**:26.
289. Tsai H, Howell T, Nitcher R, Missirlian V, Watson B, Ngo KJ, Lieberman M, Fass J, Uauy C, Tran RK, et al: **Discovery of rare mutations in populations: TILLING by sequencing.** *Plant Physiol* 2011, **156**:1257–1268.
290. Austin RS, Vidaurre D, Stamatiou G, Breit R, Provart NJ, Bonetta D, Zhang J, Fung P, Gong Y, Wang PW, et al: **Next-generation mapping of Arabidopsis genes.** *The Plant Journal* 2011, **67**:715–725.
291. Xin Z, Wang M, Barkley N, Burrow G, Franks C, Pederson G, Burke J: **Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population.** *BMC Plant Biology* 2008, **8**:103.
292. Blomstedt CK, Gleadow RM, O'Donnell N, Naur P, Jensen K, Laursen T, Olsen CE, Stuart P, Hamill JD, Moller BL, Neale AD: **A combined biochemical screen and TILLING approach identifies mutations in *Sorghum bicolor* L. Moench resulting in acyanogenic forage production.** *Plant Biotechnol J* 2012, **10**:54–66.
293. Till B, Reynolds S, Weil C, Springer N, Burtner C, Young K, Bowers E, Codomo C, Enns L, Odden A, et al: **Discovery of induced point mutations in maize genes by TILLING.** *BMC Plant Biology* 2004, **4**:1–8.
294. Xin Z, Cross M, Burrow G, Burke J: **Sorghum TILLING Population - A Community Resource for Sorghum Improvement.** [abstract]. In *American Society of Plant Biologists Annual Meeting 2011*. 2011. Paper No. 15020.
295. Schnable PS, Ware D, Fulton RS, Stein JC, Wei F, Pasternak S, Liang C, Zhang J, Fulton L, Graves TA, et al: **The B73 maize genome: complexity, diversity, and dynamics.** *Science* 2009, **326**:1112–1115.
296. Paterson AH, Bowers JE, Bruggmann R, Dubchak I, Grimwood J, Gundlach H, Haberer G, Hellsten U, Mitros T, Poliakov A, et al: **The *Sorghum bicolor* genome and the diversification of grasses.** *Nature* 2009, **457**:551–556.
297. Phytozome. <http://www.phytozome.net>.
298. Maize Genome Database. <http://www.maizegenome.org>.
299. Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, Bork P, Das U, Daugherty L, Duquenne L, et al: **InterPro: the integrative protein signature database.** *Nucleic Acids Res* 2009, **37**:D211–D215.
300. Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT, et al: **Gene ontology: tool for the unification of biology.** The Gene Ontology Consortium. *Nat Genetics* 2000, **25**:25–29.
301. Kanehisa M, Araki M, Goto S, Hattori M, Hirakawa M, Itoh M, Katayama T, Kawashima S, Okuda S, Tokimatsu Y, Yamanishi Y: **KEGG for linking genomes to life and the environment.** *Nucleic Acids Res* 2008, **36**:D480–D484.
302. Jaiswal P, Ni J, Yap I, Ware D, Spooner W, Youens-Clark K, Ren L, Liang C, Zhao W, Ratnapu K, et al: **Gramene: a bird's eye view of cereal genomes.** *Nucleic Acids Res* 2006, **34**:D717–D723.
303. Paterson AH, Bowers JE, Peterson DG, Estill JC, Chapman BA: **Structure and evolution of cereal genomes.** *Curr Opin Genet Dev* 2003, **13**:644–650.
304. IRGSP: **The map-based sequence of the rice genome.** *Nature* 2005, **436**:793–800.
305. InternationalBrachypodiumInitiative: **Genome sequencing and analysis of the model grass *Brachypodium distachyon*.** *Nature* 2010, **463**:763–768.
306. Paterson AH, Bowers JE, Feltus FA, Tang H, Lin L, Wang X: **Comparative genomics of grasses promises a bountiful harvest.** *Plant Physiol* 2009, **149**:125–131.
307. Frazer KA, Pachter L, Poliakov A, Rubin EM, Dubchak I: **VISTA: computational tools for comparative genomics.** *Nucleic Acids Res* 2004, **32**:W273–W279.
308. Gramene: A Resource for Comparative Plant Genomics. <http://www.gramene.org>.
309. Comparative Saccharinae Genomics Resource. <http://csgr.pgml.uga.edu>.
310. Paterson AH: *Genomics of the Saccharinae*. New York, NY: Springer; 2012.
311. Kliebenstein D: **Quantitative genomics: analyzing intraspecific variation using global gene expression polymorphisms or eQTLs.** *Annu Rev Plant Biol* 2009, **60**:93–114.
312. Morrell PL, Buckler ES, Ross-Ibarra J: **Crop genomics: advances and applications.** *Nat Rev Genet* 2011, **13**:85–96.
313. Cloonan N, Grimmond SM: **Transcriptome content and dynamics at single-nucleotide resolution.** *Genome Biol* 2008, **9**:234.
314. Wang Y, Zeng X, Iyer NJ, Bryant DW, Mockler TC, Mahalingam R: **Exploring the switchgrass transcriptome using second-generation sequencing technology.** *PLoS One* 2012, **7**:e34225.
315. Xu X, Liu X, Ge S, Jensen JD, Hu F, Li X, Dong Y, Gutenkunst RN, Fang L, Huang L, et al: **Resequencing 50 accessions of cultivated and wild rice yields markers for identifying agronomically important genes.** *Nat Biotechnol* 2012, **30**:105–111.
316. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, Mitchell SE: **A Robust, Simple Genotyping By Sequencing (GBS) Approach For High Diversity Species.** *PLoS One* 2011, **6**:e19379.
317. Earl D, Bradnam K, St John J, Darling A, Lin D, Fass J, Yu HO, Buffalo V, Zerbino DR, Diekhans M, et al: **Assemblathon 1: a competitive assessment of de novo short read assembly methods.** *Genome Res* 2011, **21**:2224–2241.
318. Feltus FA, Saski CA, Mockaitis K, Haiminen N, Parida L, Smith Z, Ford J, Staton ME, Ficklin SP, Blackmon BP, et al: **Sequencing of a QTL-rich region of the *Theobroma cacao* genome using pooled BACs and the identification of trait specific candidate genes.** *BMC Genomics* 2011, **12**:379.
319. Gonzalez VM, Benjak A, Henaff EM, Mir G, Casacuberta JM, Garcia-Mas J, Puigdomenech P: **Sequencing of 6.7 Mb of the melon genome using a BAC pooling strategy.** *BMC Plant Biol* 2010, **10**:246.
320. Steuernagel B, Taudien S, Gundlach H, Seidel M, Ariyadasa R, Schulte D, Petzold A, Felder M, Graner A, Scholz U, et al: **De novo 454 sequencing of barcoded BAC pools for comprehensive gene survey and genome analysis in the complex genome of barley.** *BMC Genomics* 2009, **10**:547.
321. Saski CA, Li Z, Feltus FA, Luo H: **New genomic resources for switchgrass: a BAC library and comparative analysis of homoeologous genomic regions harboring bioenergy traits.** *BMC Genomics* 2011, **12**:369.
322. Barabasi AL, Oltvai ZN: **Network biology: understanding the cell's functional organization.** *Nat Rev Genet* 2004, **5**:101–113.
323. Ficklin SP, Feltus FA: **Gene coexpression network alignment and conservation of gene modules between two grass species: maize and rice.** *Plant Physiol* 2011, **156**:1244–1256.
324. Butte AJ, Tamayo P, Slonim D, Golub TR, Kohane IS: **Discovering functional relationships between RNA expression and chemotherapeutic susceptibility using relevance networks.** *Proc Natl Acad Sci U S A* 2000, **97**:12182–12186.
325. Gu H, Zhu P, Jiao Y, Meng Y, Chen M: **PRIN: a predicted rice interactome network.** *BMC Bioinforma* 2011, **12**:161.
326. Seo YS, Chern M, Bartley LE, Han M, Jung KH, Lee I, Walia H, Richter T, Xu X, Cao P, et al: **Towards establishment of a rice stress response interactome.** *PLoS Genet* 2011, **7**:e1002020.
327. Nadeau JH, Dudley AM: **Genetics. Systems genetics.** *Science* 2011, **331**:1015–1016.
328. Plaisier CL, Horvath S, Huertas-Vazquez A, Cruz-Bautista I, Herrera MF, Tusie-Luna T, Aguilar-Salinas C, Pajukanta P: **A systems genetics approach implicates USF1, FADS3, and other causal candidate genes for familial hyperlipidemia.** *PLoS Genet* 2009, **5**:e1000642.

329. Lee TH, Kim YK, Pham TT, Song SI, Kim JK, Kang KY, An G, Jung KH, Galbraith DW, Kim M, *et al*: RiceArrayNet: a database for correlating gene expression from transcriptome profiling, and its application to the analysis of coexpressed genes in rice. *Plant Physiol* 2009, **151**:16–33.
330. Ficklin SP, Luo F, Feltus FA: The Association of Multiple Interacting Genes with Specific Phenotypes In Rice (*Oryza sativa*) Using Gene Co-Expression Networks. *Plant Physiol* 2010, **154**:13–24.
331. Spangler JB, Ficklin SP, Luo F, Freeling M, Feltus FA: Conserved Non-Coding Regulatory Signatures in Arabidopsis Co-expressed Gene Modules. *PLoS One* 2012, **7**:e45041.
332. Massa AN, Childs KL, Lin H, Bryan GJ, Giuliano G, Buell CR: The transcriptome of the reference potato genome *Solanum tuberosum* Group Phureja clone DM1-3 516R44. *PLoS One* 2011, **6**:e26801.

doi:10.1186/1754-6834-5-80

**Cite this article as:** Feltus and Vandenbrink: Bioenergy grass feedstock: current options and prospects for trait improvement using emerging genetic, genomic, and systems biology toolkits. *Biotechnology for Biofuels* 2012 **5**:80.

**Submit your next manuscript to BioMed Central  
and take full advantage of:**

- Convenient online submission
- Thorough peer review
- No space constraints or color figure charges
- Immediate publication on acceptance
- Inclusion in PubMed, CAS, Scopus and Google Scholar
- Research which is freely available for redistribution

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

