Biochemical and physiological flexibility accompanies reduced cellulose biosynthesis in Brachypodium *cesa1*^{S830N}

Chad Brabham[a,h*], Abhishek Singh[b*], Jozsef Stork[a], Ying Rong[c,d], Indrajit Kumar[c], Kazuhiro Kikuchi[c,e], Yaroslava G. Yingling[b], Thomas P. Brutnell[c,f], Jocelyn K.C. Rose[g], and Seth DeBolt[a,Ψ]

[a]Department of Horticulture, University of Kentucky, 1408 Veterans Drive, Lexington, KY 40546
[b]Department of Materials Science and Engineering, North Carolina State University, 3028B Engineering Building, Campus Box 7907, 911 Partners Way, Raleigh NC, 27695
[c]Donald Danforth Plant Science Center, 975 N. Warson Road, St. Louis, MO 63132
[d]KWS Gateway Research Center, 1005 N Warson Rd, St. Louis, MO 63132, USA
[e]Syngenta Japan K.K., 21st floor, Office Tower X, 1-8-10 Harumi, Chuo-ku, Tokyo 104-6201, Japan
[f]Current Address: Viridis Genomics, P.O. Box 430 Chesterfield, MO 63006-0430
[g] Plant Biology Section, School of Integrative Plant Science, Cornell University, Ithaca, NY 14853.
[h]Current Address: Department of Crop, Soil and Environmental Sciences, University of Arkansas, Fayetteville, Arkansas 72701

*Equal contribution

[Ψ] Corresponding author: sdebo2@email.uky.edu

Keywords: cellulose, molecular dynamics, class specific region, biomechanics, CESA, cellular expansion

Abstract

Here, we present a study into the mechanisms of primary cell wall cellulose formation in grasses, using the model cereal grass *Brachypodium distachyon*. The exon found adjacent to the Bd*CESA1* glycosyltransferase QXXRW motif was targeted using Targeting Induced Local Lesions in Genomes (TILLING) and Sequencing Candidate Amplicons in Multiple Parallel Reactions (SCAMPRing) leading to the identification of the Bd*cesa1*^{S830N} allele. Plants carrying this missense mutation exhibited a significant reduction in crystalline cellulose content in tissues that rely on the primary cell wall for biomechanical support. However, *Bdcesa1*^{S830N} plants failed to exhibit the predicted reduction in plant height. In a mechanism unavailable to eudicotyledons, *B. distachyon* plants homozygous for the *Bdcesa1*^{S830N} allele appear to overcome the loss of internode expansion anatomically by increasing the number of nodes along the stem. Stem biomechanics were resultantly compromised in Bd*cesa1*^{S830N}. The *Bdcesa1*^{S830N} missense mutation did not interfere with *BdCESA1* gene expression. However, molecular dynamic simulations of the CESA structure with modeled membrane interactions illustrated that Bd*cesa1*^{S830N} exhibited structural changes in the translated gene product responsible for reduced cellulose biosynthesis. Molecular dynamic simulations showed that substituting S830N resulted in a stabilizing shift in the flexibility of the CSR arm of the core catalytic domain of CESA, revealing the importance of this motion to protein function.

Introduction

Recent advances in molecular techniques have facilitated significant progress in the field of plant functional genomics. However, most such studies focus on model organisms, with the eudicotyledonous *Arabidopsis thaliana* (arabidopsis), leading the way. Indeed, great strides have been made in developing a mechanistic understanding of cellulose biosynthesis by studies of arabidopsis including the determination of genes involved in the heterotrimeric cellulose synthase complex (CSC)(Arioli et al., 1998; Schieble et al., 2001; Kurek et al., 2002; Desprez et al., 2002, 2007; Taylor et al., 2003), structure-function relationships of CSC components (Harris et al., 2012; Sethaphong et al., 2013; Slabaugh et al., 2014) and an active field of study into CSC accessory proteins (Endler et al., 2015; Lei et al., 2015). Furthermore, X-ray crystallography has resolved the bacterial cellulose synthase protein structure (Morgan et al., 2013), which is distinct from plants in many ways. In plants, the homomeric structure of a *Gossipium hirsutum* CESA has recently been modeled (Nixon et al., 2016) and assembled *in vitro* (Purushotham et al., 2016), revealing numerous unanswered structural questions.

The family Poaceae is the most economically important group of plants and includes crops such as cereals, forage grasses, biofuel feedstocks and a variety of weed species. The temperate, C3 annual grass, *Brachypodium distachyon (B. distachyon* or brachypodium)(Poales: Poaceae) has emerged as a model grass for molecular genetic studies (Draper et al., 2001; Vogel et al., 2010). Moreover, several studies have advanced brachypodium as a genetic model for grass cell wall development (Christensen et al., 2010), cereal-pathogen interactions (Fitzgerald et al., 2015), and grain development (Hands and Drea, 2012).

In Eudicots, the primary cell wall comprises approximately a 1:1:1 ratio of cellulose, hemicellulose (mainly xyloglucans) and an assortment of pectic polysaccharides. Cellulose is the major structural component and the biosynthetic machinery responsible for cellulose biosynthesis has been the subject of intense study, particularly in arabidopsis. There are 10 CELLULOSE SYNTHASE A (CESA) isoforms in arabidopsis (Carroll and Specht, 2011) and it has been shown that three different CESAs are required to form a functional cellulose synthase complex (CSC) in a 1:1:1 ratio (Gonneau et al., 2014). Genetic studies have revealed that CSCs in primary cell walls comprise a particular combination of CESAs (Desprez et al., 2007, Persson et al., 2007) that differs from the CSCs of secondary cell walls (Taylor et al., 2003).

The primary cell walls of grasses are also composed of a highly organized network of polysaccharides. However, the non-cellulosic fraction differs significantly between grasses and eudicots in terms of the relative abundance and type of polysaccharides (Carpita, 1996; Vogel, 2008). In grasses, such as brachypodium, cellulose represents a third of the primary cell wall, but the surrounding matrix glycans largely comprise glucuronoarabinoxylans, with some of the arabinosyl residues being esterified with ferulate, as well as mixed linkage glucans (1,3 1,4-β-glucans) but relatively little xyloglucan or pectin is present (Vogel, 2008).

Current understanding of cellulose biosynthesis in grasses is less detailed, but they have conserved CESA clusters, indicating commonalities in the mechanism of cellulose biosynthesis in eudicots and grasses (Handakumbura et al., 2015). However, it is notable that a highly potent cellulose biosynthesis inhibitor, isoxaben, has little effect on grasses (Brabham and DeBolt, 2013). This compound targets primary cell wall CESA

proteins in eudicots (Scheible et al., 2001; Desprez et al., 2002; Harris et al., 2012) and so the resistance observed in grasses suggests that structure-functional differences may exist in the cellulose biosynthetic machinery giving rise to primary cell wall formation.

Prior studies of CESA clade members in arabidopsis revealed that some CESA gene mutations can be tolerated, whereas others are lethal, indicating where redundancy is present (e.g. Persson et al., 2007). In addition, point mutations in *AtCESA1* and *-3* have been useful for structure function predictions (Sethaphong et al., 2013). An array of functional genomic tools are now available for brachypodium (Vogel, 2010; Brutnell et al., 2015), creating opportunities to conduct investigations that were previously restricted to arabidopsis. An example of such an approach is reverse genetics via Targeting Induced Local Lesion IN Genomes (TILLING; Henikoff et al., 2004; McCallum et al., 2000), which enables the isolation of point mutations in a gene product of interest. Identifying TILLING mutants from within a mutagenized seed population can be accelerated by combining the polymerase chain reaction (PCR) amplification of a gene region of interest with next generation sequencing. This approach is referred to as SCAMPRing (sequencing candidate amplicons in multiple parallel reactions; Gilchrist et al., 2013).

The CESA1 gene in grasses is proposed to be an ortholog of *radially swollen1 (rsw1* or *AtCESA1*; Arioli et al., 1998; Handakumbura et al., 2015).), an arabidopsis gene that plays a crucial role in the production of cellulose in the primary cell wall. Here, we sought to gain functional insights into the role of CESA1 in grasses using both TILLING and SCAMPRing. We show that a single missense mutation in this gene results in a significant reduction in the cellulose content of brachypodium and we use molecular dynamic simulation to predict how this mutation alters the dynamics of the cellulose

synthase complex. We also discuss the significance of an apparent adaptive anatomical response in the stem of the mutant that occurs as a consequence of the cellulose deficiency.

Results

Identification of *Brachypodium* Primary Cell Wall CESAs

A combined phylogenetic and quantitative real time PCR approach was used to identify *BdCESA* genes involved in primary cell wall cellulose biosynthesis in root, shoot and leaf tissues (**Fig 1**). The brachypodium reference genome has 10 predicted *CESA* genes, although *BdCESA10* (Bradi1g36740) is not predicted to have catalytic residues required for glucosyltransferase activity (Morgan et al., 2013; Nixon et al., 2016; Purushotham et al., 2016) and so we did not consider it to be a candidate functional *CESA* gene. It is also worth noting that Bd*CESA5* (Bradi1g29060) does not have a predicted zinc finger domain, believed to be involved in CESA oligomerization (Kurek et al., 2002), but it was not excluded from these analyses. The brachypodium *CESA* naming system described by Handakumbura and authors (2015) was adopted. They classified *BdCESA* genes based upon their closest arabidopsis orthologs and our data supported their findings. To further validate the phylogenetic predictions, the relative gene expression profiles of *CESA* genes was measured in etiolated coleoptiles **(Fig 1**). Since the focus of the current study was primary cell wall *CESA* genes, we evaluated the relative fold change in *CESA* in actively growing tissues to seek *CESAs* that were expressed in all tissues. The relative expression profile of *CESA* genes showed that *BdCESA1* (Bradi2g34240), *BdCESA3* (Bradi1g54250) and *BdCESA6* (Bradi1g53207) were expressed in target tissues (**Fig 2**). *BdCESA9*

(Bradi1g36740) *BdCESA2* (Bradi1g04597) and *BdCESA5* were detectable but showed insufficient expression uniformity for further evaluation. *BdCESA4* (Bradi3g28350), *BdCESA7* (Bradi4g30540) and *BdCESA8* (Bradi2g49912) were the subject of an extensive prior study and are known to be required for cellulose biosynthesis in secondary cell walls (Handakumbura et al., 2015). As expected, their expression was significantly reduced in coleoptile and root tissue (**Fig 1**).

Based on these findings, and in accordance with Handakumbura et al. (2015), *BdCESA1, 3,* and *6* expression was consistent with their having a role in primary cell wall cellulose biosynthesis. Genetic evidence from arabidopsis suggests that AtCESA1 and 3 are required for cellulose biosynthesis (Arioli et al., 1998; Kurek et al., 2002; Persson et al., 2007) and we sought to understand the functional genetic significance of primary cell wall *CESAs* in grasses via TILLING for a mutation in *BdCESA1*.

Targeting and Identification of *BdCESA1* TILLING Mutant

Screens were conducted on an EMS-mutagenized population of brachypodium accession Bd21 (Methods). In brief, primers were designed to genomic regions with the highest probability for ethyl methane sulfonate (EMS)-induced missense and nonsense lesions in Bd*CESA1*, using the web-based tool CODDLE (Codons Optimized to Discover Deleterious Lesion; Henikoff et al., 2004; McCallum et al., 2000; Gilchrist et al., 2013). A 1,096 base pair region of the *BdCESA1* gene was selected for TILLING (see Methods), corresponding to approximately one sixth of the full-length genomic sequence. This DNA region encodes the last half of the glycosyltransferase domain to the sixth transmembrane domain (**Fig 2, black arrows indicate primer location**). To identify point mutations,

primers were used to amplify the region of interest, using pooled DNA samples from our TILLING population as a template. Next generation sequencing (NGS) of these samples revealed a point mutation, which was predicted to result in an asparagine instead of a serine at position 830 of the BdCESA1 protein. No further predicted or sequenced mutations were identified in *BdCESA1*.

The *Bdcesa1*^{S830N} is located in the cytosolic catalytic loop 10 amino acids downstream of the QXXRW motif and approximately 23 amino acids before the beginning of the third transmembrane alpha helix (**Fig 2, asterisk**). A sequence analysis of plant CESA1 proteins based on the 19 amino acids following the QXXRW motif revealed high conservation within this region (**Table 1**). A total of 88 plant CESA proteins were used to generate the sequence logo plot (**Fig 2**), where the height of a logo is proportional to the frequency at which it occurs at that position (Schneider et al., 1990). Analysis of the sequence logo plot revealed the *Bdcesa1*^{S830N} would change the conserved serine found in all identified CESAs to an asparagine.

Expression of *CESA1* in Bd*cesa1*^{S830N}

An important question related to the gene of interest was whether the S830N missense substitution influenced expression of the *BdCESA1 in planta*. The region of interest is near the catalytic domain, but the missense mutation was not predicted to result in a premature stop codon or alter *BdCESA1* gene expression. The expression of the native *BdCESA1* was evaluated in three biological replicates of the *Bdcesa1*^{S830N} compared with wild type. Using semi-quantitative polymerase chain reaction (25 cycles using GAPDH as a control gene) we found no change in gene expression in expanding shoot tissue (**Fig**

S1). These data support the prediction that the missense mutation will not interfere with *BdCESA1* gene expression, but might lead to structural differences in the translated gene product.

Relative changes in cellulose content in *cesa1*^{S830N}

Cellulose content was measured in leaf, sheath, stem and peduncle tissue of mature wild type (Bd21-3) and *Bdcesa1*^{S830N} plants, and a significant reduction was found in all mutant samples. On average, *Bdcesa1*^{S830N} mutants had 7% less cellulose in leaf and sheath tissue and 25% less cellulose in stem and peduncle tissues compared to wild-type (**Fig 3,** significance established via Tukeys mean separation P>0.05).

Mutant phenotypes associated with reduced cellulose content

By TILLING in a mutagenized background we identified a single allele *Bdcesa1*^{S830N}. As noted by Liu et al. (2017), *CESA* mutations in grasses are lacking, which complicates comparisons to other alleles. Gene complementation is also made difficult by the fact that the mutant allele likely encodes a functional protein product that can interact with the wild-type product. Therefore, where appropriate, we compared phenotypes of *Bdcesa1*^{S830N} with those of the *dwf1-1* from *Sorghum bicolor* (Petti et al., 2015). The *dwf1-1* linked to an insertion mutation in *gibberellin20 (GA20)-oxidase* and *Bdcesa1*^{S830N} alleles displayed reduced stem cellulose content (35% and 28% respectively). and homozygous lethality. In eudicots, the loss of this vital component of the cell wall leads to stunted plants, in which root and shoot tissue expansion is severely reduced (Arioli et al., 1998). We observed a similar phenotype in the peduncles of

Bdcesa1^{S830N} mutants, with a 41% reduction in length compared to wild-type peduncles (P<0001, Student T-Test; **Fig 4A,B**). However, the stem lengths of *Bdcesa1*^{S830N} mutants were similar to their wild-type counterparts (**Fig 4A**), as were coleoptiles (wild type 3.7 cm ± 0.2; *Bdcesa1*^{S830N} 3.6 cm ± 0.2), which is atypical of the normal dwarfism seen in arabidopsis *rsw1-1* (Arioli et al., 1998). We therefore examined other stem characteristics that could be associated with lower cellulose content.

To determine whether the mutation affected the biomechanical properties of the stems, we measured the elastic modulus of isolated and dried lower, middle and upper stem sections, using an Instron materials testing instrument. These three regions were chosen for comparison because lower stems have secondary cell walls, whereas upper stems have predominately primary cell walls. There was no significant difference in the elastic modulus values for the lower and middle sections (n = 4, P>0.05, Tukeys HSD). However, we detected a highly significant reduction in the elastic modulus in the upper stem of *Bdcesa1*^{S830N} (1576 ± 386 s.d., n = 9) compared with controls (716 ± 269 s.d., n = 15, P = 0.009). Note that with Young's Modulus, the larger number equates to greater force needed to stretch a substrate (less elastic). In addition, the tensile strength of the lower and middle sections of the Bd*cesa1*^{S830N} stems was not significantly different to control (n = 3, P > 0.05) but the upper stems (60.6 ± 10.6 s.d.) were significantly weaker than the control at (86.9 ± 28.7 s.d., p = 0.003). Given that cellulose microfibrils act as the major load bearing component of the wall, these data indicate a functional association between the *Bdcesa1*^{S830N} allele and loss of structural integrity in the primary cell wall.

The stems of *Bdcesa1*^{S830N} were found to be both deficient in cellulose, as well as structurally weaker and less stiff, and yet the mutants grew to the same height as wild-

type plants. We therefore examined plant morphology and found that the stems of *Bdcesa1*^{S830N} mutants have a significantly higher mean density of nodes (10.3 ± 0.8)(**Fig 4C**) compared with wild type (6.6 ± 0.5)(two tailed T-test, P>0.001, n = 7). Correspondingly, significantly reduced internode lengths were observed but these differences were more pronounced in the upper stem internode distances. For instance, in wild type, the lower three internodes had a mean length of 9.3 cm ± 2.5 whereas *Bdcesa1*^{S830N} displayed average internode length of 7.0 cm ± 1.2. In contrast, the upper three internode lengths were significantly different (wild type 37.3 cm ± 0.96 compared to *Bdcesa1*^{S830N} 21.7 cm ± 2.2, n = 4). The reduced internode and peduncle length phenotypes were observed in plants growing in either growth chamber or greenhouse conditions (**Fig 4C**). To corroborate these data, we examined node frequency in the Sorghum *dwf1-1* allele (Petti et al., 2015). Here, node number was again significantly greater in *dwf1-1* compared to wild-type control plants (mean of 10.3 ± 0.6 for *dwf1-1* versus 7.8 for control). This was accompanied by severe reduction in internode expansion. It should be noted that *dwf1-1* is homozygous lethal and displays a greater reduction in cellulose than *Bdcesa1*^{S830N}, thus results could reflect the severity of this mutation. We carefully examined the morphology of the uppermost node and observed no gross anatomical difference between wild type and *Bdcesa1*^{S830N} (**Fig 4D**). Transverse sections were prepared from mature peduncles, which showed the greatest length reduction. No collapsed vascular xylem was observed (**Fig 4E**) in wild type or Bd*cesa1*^{S830N} but modest aberration in the integrity of the parenchyma layer and thickness of the cortex were distinguishable.

Non-Cellulosic Sugars in *Bdcesa1*^{S830N}

Of the sugars released from the non-cellulosic polysaccharides in the cell walls, negligible changes (1.2 fold increase or decrease) were observed in *Bdcesa1*^{S830N} compared to wild type (**Table 2**). The exception was galactose, which exhibited an approximately 30% greater relative abundance in stems and sheaths of the mutant plants. While a trend towards modest increases in the lignin content was observed in the *Bdcesa1*^{S830N} mutant stems (2.0 ± 0.2 % soluble and 20.4 ± 0.9 % insoluble), these differences were not significant compared with wild-type (1.91 ± 0.2 % soluble and 19.5 ± 0.9 % insoluble) (n = 4, P > 0.05, two tailed T-test).

Biochemical modeling of the *Bdcesa1*^{S830N} into a 3D atomistic model of cellulose synthase

The experimental determination of plant CESA structures presents a technical challenge due to the large size of the proteins, the fact that they contain eight membrane spanning domains and contain a relatively high proportion (~15%) of intrinsically disordered regions (Scavuzzo-Duggan et al., 2018). Part of the PCR region of OsCesA8 from rice (*Oryza sativa*) has been solved using X-ray crystallography (Rushton et al., 2017) and the configuration of AtCesA1 CatD region was determined using small angle X-ray scattering (SAXS)(Vandavasi et al., 2016). Additionally, the solution structure of bacterial cellulose synthase displays some biochemical dissimilarities to plant CESA(s) (Morgan et al., 2013). Therefore, standard homology modeling software, such as Phyre2 (Kelley et al., 2015), fails to produce robust models of plant CESA proteins (**Fig S2**). Here, we used the GhCESA model template and SWISS-MODEL (Biasini et al., 2014) to

build a 3D model of BdCESA1 to investigate possible structural changes occurring due to the amino acid substitution at S830. The complete model of BdCESA1 is shown in **Fig 5A**.

Results suggest that Bd*cesa1*^{S830N} is on the solvent accessible side of CESA and that S830 is predicted to interact with the CSR. When incorporating the mutation into the model we see that the migration of the CSR domain to the membrane is inhibited (**Fig 5**). The 3D structure highlighted components of the catalytic core in orientations of PCR, CSR, QxxRW, and DCD relative to mutation site S830N (**Fig 5**). Simulation results predict that the mutation might cause structural alterations in the CSR region. Further, maps of the free energy landscape reveal the conformational space defined by the root mean square deviation of the CSR region and the distance of the CSR from the S830N mutation site (**Fig 5B,C)**. The energy landscapes include local metastable states separated by small energy barriers and had few 'lowest free energy' regions, seen in yellow. Computational results show that in contrast to wild-type CSR, the mutant CSR moves closer to the mutation site and QxxRW region. The CSR is largely unstructured and the prediction of the motion for migration indicates that it is possible (**Fig S3).** Using the functional amino acid Bd*cesa1*^{S830N} substitution as a target, we computationally evaluated a range of alternative substitutions (**Table 3**) at this site. It was computationally predicted that altering S830 to any alternative amino acid would have a deleterious influence on the CSR stability.

We looked further into the structural dynamics of the components of the CSR that may have biological significance related to the S830N point mutation (**Fig 6**). In the ancient bryophyte *Physcomitrella patens,* quantification of the hyper-variable nature of

CSR identified biologically significant molecular recognition of features (Morfs) that can undergo order disorder transition (Scavuzzo-Duggan et al., 2018). In **Fig 5c,d** the entire CSR domain in *Bdcesa1*^{S830N} is predicted to shift closer to the mutation site and QxxRW region. More specifically, the conserved and MoRF regions of mutant CeSA CSR were up to 10Å closer to the mutation site (**Fig 6**), while the proximity of MoRF2 to the mutation site in mutant CESA was comparable to that of wild type. **Fig S4** compares the solvent accessible surface area of components of CSR.

Predicted effect of mutation on the activity of BdCESA1

In **Fig 5**, we defined the free energy landscape based on the distance between DCD and QxxRW motifs and the root mean square deviations of QxxRW. The mutant CESA (**Fig 5f**) exhibited RMSD deviations with a tighter distribution. The S830N mutation was predicted to make the QxxRW motif more rigid than its counterpart in the wild type. In contrast, the mutant DCD motif displayed higher structural deviations (**Fig 5g,h**). The lowest energy basins show that DCD and QxxRW motifs were more distant from each other in the mutant CESA. We also compared the deleterious nature of the substitution of serine with other amino acids and found that S830N mutation would be mildly deleterious based on DUET prediction score (**Table 3**).

Discussion

Despite the majority of terrestrial biomass being produced by grasses, molecular genetic studies into cellulose biosynthesis are limited. It is worth noting that based on phylogenetic studies (Handakumbura et al., 2015), cellulose biosynthetic genes appear

quite conserved across grasses and eudicots. There are, however, some curious differences. For instance, it is not known why class-L herbicides, known as cellulose biosynthesis inhibitors are far more potent against broadleaf plants than grasses. One hypothesis is that differences in cell wall composition between grasses and eudicots allows for modest expansion during cellulose biosynthesis inhibitor stress (Brabham et al., 2018). In this study, we used TILLING and SCAMPRing targeting a region adjacent to the catalytic domain in *BdCESA1* because it is predicted to be orthologous to the severe *radially swollen1 (rsw1*)(Arioli et al., 1998; Persson et al., 2007). Unlike *rsw1,* which was severely dwarfed and null lethal, *Bdcesa1*^{S830N} internode expansion was greatly reduced, but normal plant height was reached (**Fig 4a**). Brachypodium plants carrying the *Bdcesa1*^{S830N} allele displayed an unexpected physiological change that correlated with the lack of expansion and reduced cellulose, which was to increase the number of nodes along the stem (**Fig 4c**). Hence, it is feasible that, similar to chemical disruptions of cellulose biosynthesis in grasses with certain Class L herbicides, the mutant with the genetic dysfunction can still grow due to having more nodes (Fig 4). Interpretation of complex phenotypes, however, must be tempered by the single allele derived from the EMS population used for TILLING experiments. Results could be linked to secondary mutations giving rise to the secondary phenotypes observed. In support of a possible correlation, the results observed were largely expansion driven, which is a consistent feature of cellulose deficit in expanding plant tissues (Desprez et al., 2007; Handakumbura et al., 2015). There are very few examples to cross reference our observations to. We did re-visit the phenotype of the sorghum *dwf1-1* mutation (Petti et al., 2015), which caused severe cellulose deficit linked to a mutation in GA20-oxidase.

Here, node density increased along the stem and internode length was shorter, but the *dwf1-1* plant failed to reach the height of the wild-type. This may reflect that *dwf1-1* contained a greater reduction in cellulose than Bd*cesa1*^{S830N}, which in turn influenced plant expansion akin to a dose effect. Alternatively, it could also be suggested that the *dwf1-1* mutation in GA20 oxidase, imparted a broader influence on expansion and development than targeting cellulose alone. Irrespective of mechanism, it was common to both grass taxa (Panicoideae and Poaceae) that cellulose deficit was accompanied by a increased node number along the stem axis and decreased internode expansion.

Consistent with a phenotype linked to the primary cell wall, the upper stems of Bd*cesa1*^{S830N} were biomechanically weaker, but the lower stem tissue exhibited no significant change in biomechanical properties. These data support the suggestion that the *Bdcesa1*^{S830N} phenotypes were primary cell wall linked, and that secondary cell wall thickening, which is more developed in the lower grass stem, helped compensate for any biomechanical deficiencies.

The limited number of homologous structures, the presence of intrinsically disordered regions, and the low number of solved transmembrane protein structures mean that an integrated modeling approach, using bioinformatics tools and molecular dynamic simulation can only generate hypothetical structural predictions, although this approach has been used to develop hypotheses and predictions (Sethaphong et al., 2013; Slabaugh et al., 2013; Lei et al., 2015; Nixon et al., 2016). While requiring further work and future biochemical validation, we hypothesize that structural changes in the active binding site affected CESA activity. When coupled with the reduced accessibility of MoRFs binding regions in the CSR domain these structural changes would have implications for the

assembly of the cellulose synthase complex. Future work is needed to understand the disordered protein region in CESA (Shoemaker et al., 2000, Scavuzzo-Duggan et al., 2018) as these regions have played diverse biochemical roles (Dunker et al., 2005, Mohan et al., 2006).

In conclusion, these data provide fundamental information about the nature of cellulose biosynthesis in grasses. Some practical applications may be foreseeable. For instance, stem lodging (breaking of the stem) of major cereal crops is tightly linked to nodal structure and alone accounts for 5-20% of annual losses in global grain production (Robertson et al., 2017). Our data supports the notion that primary cell-wall cellulose biosynthesis is one of many factors influencing the complex trait of stem morphology and biomechanics (Von Forell et al., 2015) and raises the possibility of strength gradients across the grass stem.

Materials and Methods

Plant Material and Growth

Seeds were sterilized with 30% household bleach for 15 min and subsequently washed three times with sterile distilled water and kept at 4°C for 2 d or 3 weeks. The 3-week cold treatment sufficiently vernalized seeds to promote rapid flowering. For all measurement studies, plants were pre-germinated and seedlings with a protruding radicle < 1mm were selected for use. To measure coleoptile (dark grown) or root (light grown) length at 7 days after germination, seedlings were placed on agar (11 g L^{-1}) plates and grown vertically in growth chambers at 22°C with a 14-h photoperiod. Plates of dark grown plants were wrapped in aluminum foil. After 7 days, organ length was measured.

Seedlings were left in the growth chamber for an additional week and transferred to soil pots and growth was maintained under 24-hr supplemental lighting at room temperature. Plants were also grown in subsequent generations under greenhouse conditions (Lexington, Kentucky).

TILLING by Sequencing

The brachypodium line Bd21-3 was used in all experiments. Approximately 10,000 Bd21-3 seeds were mutagenized with exposure of 80 mM ethyl methane sulfonate in a fume hood at room temperature for 16 hrs. M1 seed were rinsed 5 times in sterile distilled water and sown in flats of Metromix 360. After an overnight treatment of dark imbibition at 4°C, flats were placed in a growth chamber at 24 °C/18 °C (day/night) with a 20/4 light/dark cycle. M2 seed from approximately 5000 M1 plants was harvested for the TILLING population. To identify mutations in the *BdCESA1* gene DNA was collected from 3840 M2 plants in 8 x pools resulting in 480 pools of DNA arrayed in 40 x 96 well PCR plates. Primers were designed to a conserved region of *CESA1* – *FOR*-AAACGCTTTGGCCAGTCTCCGATATTT and REV-CCACCAGGTTAATCACAAGCACAGTGG (see supplementary online data **Table S1**) using the web-based tool CODDLE (Codons Optimized to Discover Deleterious Lesion; Henikoff et al., 2004). PCR reactions were performed on each DNA pool (480 PCR reactions) to maintain low complexity of the pools and ensure amplification of DNA isolated from all individuals that were sampled for each pool. These PCR products were then pooled in row/column arrays (24 + 20) to create 44 superpools. The DNA from the superpools was purified and equal concentrations of DNA fragmented with NEBNext

dsDNA Fragmentase (M0348S). The fragments were end-repaired and ligated to adapters for PCR amplification. PCR amplification and multiplexing were performed with the Illumina universal primer and a unique indexing primer to enable deconvolution of pools following sequencing. Equal amounts of DNA from each of the 44 samples was added to a single pool that was sequenced on one lane of a Illumina HiSeq SE.

Bioinformatic analysis and pool deconvolution

Quality assessments of Illumina reads were performed (FASTQC), adapters trimmed (Trim Galore) and reads mapped to amplicon sequences. SNPs were detected independently in row and column pools to identify intersections that defined unique DNA sample pools comprised of DNA from 8 individuals. M3 seed corresponding to each of these M2 progenitors was then grown and DNA isolated from individual plants for Sanger sequencing.

Identification of brachypodium CESAs

The protein sequences of Arabidopsis and *Oryza sativa* (rice) CESAs were used to search the brachypodium predicted proteome (https://phytozome.jgi.doe.gov/pz/portal.html) using BLAST (Altschul et al., 1997) and putative *BdCESA* sequences were searched for CESA specific glycotransferase domains (Carroll and Specht, 2011). Handakumbura et al. (2015) named BdCESA genes after their closest Arabidopsis orthologs. We conducted a phylogentic analysis in Mesquite (100 bootstraps)(Maddison and Maddison, 2018) using the class specific protein region (D to QxxRW motif) from arabidopsis and brachypodium to confirm their results.

Sectioning

Tissue sections were produced using a vibratome as described by Zelko et al. (2012). Sections were stained with were treated with ammonium hydroxide and the fluorescence from ferulate and lignin in the cell walls was observed with an Olympus FV1000 laser scanning confocal microscope using a 10x N.A. objective.

Integrated modeling approach.

The homology model was refined using all atom molecular dynamics using AMBER 16 software suite (Case et al., 2018) with FF14 protein variant force field and TIP3P water model. Molecular dynamics simulations allow a high resolution of protein conformation states and relate to their energy landscapes. Generally, the atomistic models of large proteins are trapped in local minima, which could limit the scope of equilibrium dynamics. In this study, to enhance sampling, a hyper dynamics approach was implemented, where a harmonic boost is provided to the potential energy function (protein force field), thereby smoothing the potential energy surface (Miao et al., 2015). This allows accelerated transitions between low energy states, together with an accurate free energy profile. The simulation protocol included conventional molecular dynamics (cMD) stages of 1,000 step minimization using the conjugate gradient and the steepest descent solute constrained NPT simulations. The Gaussian accelerated molecular dynamics (GaMD) module involved 200,000 steps of conventional molecular dynamics (cMD) for equilibration followed by additional 1X 10^6 steps of cMD to obtain statistical information about potential energy, which is required for determining boost potential.

After adding boost potential, the system was equilibrated for 200,000 steps. Next 1000,000 steps were used to obtain Gaussian acceleration parameters, the threshold potential and the scaling factor. The upper limit of the standard deviation of the total potential boost and dihedral boost was set to the recommended value of 6.0 kcal/mol. The Particle Mesh Ewald (PME) (Darden et al., 1993) summation method was used to calculate the electrostatic potential under periodic boundary conditions applied in all directions. The non-bonded interactions were cut at 9 Å with 0.00001 tolerance of Ewald convergence. The temperature was maintained at 300 K using a Langeven thermostat. The simulations were run for 300 ns for each sample of CSR. The protein structure form with the lowest free energy basin was extracted and a DMPC lipid bilayer was constructed around the transmembrane region using CHARMM membrane builder GUI (Jo et al., 2010). An all atom conventional molecular dynamic simulation as performed using AMBER 16 software suite (Case et al., 2018) with FF14 protein variant force field and TIP3P water model and Lipid17 forcefield. 800 ns of production simulation runs up to was performed. The cpptraj module and in-house scripts were used to perform post processing of the simulation data.

Expression of Putative CESAs

For qRT-PCR, we followed the methods of Udvardi et al. (2008). During harvest, shoot tissue (coleoptile removed) was only harvested if the first leaf had not developed a collar, and for elongating coleoptile tissue the encapsulated shoot was removed. mRNA was extracted from each sample using an RNAeasy kit (Qiagen). Amplification cycles and primers are listed in **Table S1**. Data were logged transformed to meet basic ANOVA

assumptions. Mean values were separated at an alpha value of 0.01 using Tukeys honest significant difference (HSD)(Abdi and Williams, 2010) and back-transformed.

Instron biomechanical phenotyping

Sections of dried brachypodium stems were taken from lower, middle and upper internodes and analyzed using an Instron tabletop load frame (Instron series 3340, clamp model 2710-203: Norwood MA). The samples were stretched at a rate of 3 mm per minute until they snapped. Resistance to extension was measured with a 100 Newton force transducer (Instron model: 2519-103). Bluhill 2 software (Instron) was used to calculate the tensile stress and strain at maximum load, as well as the elastic modulus. Data were analyzed using R (version 3.3.1) and significance set at P 0.05 using Tukeys (Abdi and Williams, 2010).

Cell Wall Analysis

Senesced plants were harvested and leaf, sheath and stem samples were dried for 1 week at 60°C. Tissue was either milled or sectioned (to approximately 3 mm sections) with a scalpel. To obtain alcohol insoluble residue (AIR), corresponding to crude cell wall, tissue samples we rewashed with 70% ethanol and placed in a 70°C water bath for 1 hr. This was repeated twice, except that the final ethanol wash was over night, followed by a brief acetone wash at room temperature. Sugars from non-cellulosic polysaccharides in the cell walls were quantified by the method of Foster et al. (2010) using AIR (3-5 mg). Neutral sugars (fucose, rhamnose, arabinose, galactose, glucose, mannose, xylose) were identified and quantified by pulsed electrochemical detection using a Dionex ED50

apparatus (Thermo Fisher Scientific). Sugars were separated using a CarboPAC-PA1 anion-exchange column as previously described (Endler et al., 2015). Cellulose was quantified colorimetrically using the anthrone-sulfuric acid method (Foster et al., 2010). Acid-soluble lignin, acid-insoluble lignin, and ash were measured using the laboratory analytical protocols NREL, LAP-004.

Online Supplementary Information

Table S1. Primers used in all experimental procedures

Figure S1. Semi quantitative PCR assessment of the *BdCESA1* transcript in the wild-type and mutant backgrounds

Figure S2. A computationally derived homology model utilized Phyre2.0.

Figure S3. Disorder prediction form ANCHOR probability scores.

Figure S4. Comparing solvent accessible surface area of conserved region, MoRF1, and MoRF2 in wild type and mutant CESA.

References

Abdi, H. & Williams L.J. (2010) Honestly significant difference (HSD) test. N.J. Salkind, D.M. Dougherty, B. Frey (Eds.), Encyclopedia of Research Design, Sage, Thousand Oaks, CA, USA pp. 583-585.

Altschul, S.F., et al. (1997) Gapped BLAST and PSI-BLAST: A New Generation of Protein Database Search Programs. Nucleic Acids Research, 25, 3389-3402.

Arioli T*, et al.* (1998) Molecular analysis of cellulose biosynthesis in Arabidopsis. *Science* 279:717-720.

Biasini M, *et al.* (2014) SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Research* 42, W252-W258

Brabham C & DeBolt S (2013) Chemical genetics to examine cellulose biosynthesis. *Front Plant Sci.* 3: 309.

Brabham C, Stork J, Barrett M & DeBolt S (2018) Grass Cell Walls Have a Role in the inherent Tolerance of Grasses to the Cellulose Biosynthesis Inhibitor Isoxaben. *Pest. Manag. Sci.* 74(4); 878-884.

Brutnell TP, Bennetzen JL, & Vogel JP (2015) Brachypodium distachyon and Setaria viridis: Model genetic systems for the grasses. *Annu. Rev. Plant Biol.* 66:465-485.

Carpita NC (1996) Structure and biogenesis of the cell walls of grasses. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* 47: 445-476.

Case, D.A., et al. (2018) AMBER 2018, University of California Press, San Francisco.

Carroll A & Specht CD. (2011) Understanding plant cellulose synthases through a comprehensive investigation of the cellulose synthase family sequences. *Front Plant Sci.* 2: 5.

Christensen U, Alonso-Simon A, Scheller HV, Willats WGT, & Harholt J (2010) Characterization of the primary cell walls of seedlings of Brachypodium distachyon- A potential model plant for temperature grasses. *Phytochem*. 71: 62-69.

Dalmais M, *et al*. (2013) A TILLING platform for function genomics in Brachypodium distachyon. *PLOS One* 8:e65503.

Darden T, York D, Pederson L (1993) Particle mesh Ewald: An N·log(N) method for Ewald sums in large systems**,** The Journal of Chemical Physics, 98, 10089-10092,

Desprez T, *et al.* (2002) Resistance against herbicide isoxaben and cellulose deficiency caused by distinct mutations in same cellulose synthase isoform *CesA6*. *Plant Physiol.* **128**, 482–490

Desprez T, *et al*. (2007) Organization of cellulose synthase Complexes involved in primary cell wall synthesis in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. USA*. 104: 15572-15577.

Draper J, Mur LAJ, Jenkins G, Ghosh-biswas GC, Bablak P, Hasterok R, & Routledge APM (2001) *Brachypodium distachyon*. A new model systems for functional genomics in grasses. *Plant Physiol.* 127: 1539-1555.

Dunker AK, Cortese MS, Romero P, Iakoucheva LM & Uversky VN. (2005) Flexible nets The roles of intrinsic disorder in protein interaction networks. *FEBS J*. 272, 5129-5148.

Endler A, Kesten C, Schneider R, Zhang Y, Ivakov A, Froehlich A, Funke N, Persson S (2015) A mechanism for sustained cellulose synthesis during salt stress. *Cell* 162, 1353–1364.

Fitzgerald TL, *et al.* (2015) Brachypodium as an emerging model for cereal-pathogen interactions. *Ann. Bot*. 115: 717-731.

Foster CE, Martin TM & Pauly M (2010) Comprehensive compositional analysis of plant cell walls (lignocellulosic biomass) Part II: Carbohydrates. *J. Vis Exp.* 37: 1837.

Gilchrist EJ, Sidebottom CHD, Koh CS, MacInnes T, Sharpe AG, & Haughn GW 2013. A mutant Brassica napus (Canola) population for the identification of new genetic diversity via TILLING and next generation sequencing. *PLOS One* 8: e84303.

Gonneau M, Desprez T, Guillot A, Vernhettes S, & Höfte H. (2014) Catalytic Subunit Stoichiometry within the Cellulose Synthase Complex. *Plant Physiol.* 166, 1709-1712.

Handakumbura PP, *et al*. (2015) Perturbation of Brachypodium distachyon. Cellulose synthase A4 or 7 results in abnormal cell walls. *BMC Plant Biol.* 13: 131-146.

Hands P & Drea S (2012) A comparative view of grain development in Brachypodium distachyon. *J. Cereal Sci.* 56: 2-8.

Harris DM, *et al*. (2012) Cellulose microfibril crystallinity is reduced by mutating C-terminal transmembrane region residues CESA1A-V903 and CESA3T-I942 *Proc. Natl. Acad. Sci. USA*. 109: 4098-103.

Henikoff S, Till BJ, & Comai L (2004) TILLING. Traditional mutagenesis meets functional genomics. *Plant Physiol.* 135: 630-636.

Kelley LA, Mezulis S, Yates CM, Wass MN & Sternberg MJE. (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nature Prot.* 10, 845-858.

Kurek I, Kawagoe Y, Jacob-Wilk D, Doblin M, & Delmer D (2002) Dimerization of cotton fiber cellulose synthase catalytic subunits occurs via oxidation of the zinc-binding domains. *Proc. Natl Acad. Sci. USA* **99**:11109–11114.

Lei L, Singh A, Bashline L, Li S, Yingling YG, Gu Y (2015) The Arabidopsis cellulose synthase interactive 1 is required for a fast recycling of cellulose synthase complex to the plasma membrane. *Plant Cell* 27(10): 2926-2940.

Jo, S, Kim, T, Iyer, VG, Im W (2008). CHARMM-GUI: A Web-based Graphical User Interface for CHARMM. *J. Comput. Chem.* 29:1859-1865

Liu, D. *et al*. (2017). Imaging cellulose synthase motility during primary cell wall synthesis in the grass Brachypodium distachyon. *Sci Rep***7**, 15111

Maddison, W. P. and D.R. Maddison. 2018. Mesquite: a modular system for evolutionary analysis. Version 3.51 http://www.mesquiteproject.org

McCallum CM, Comai L, Greene EA, & Henikoff S (2000) Targeting induced local lesions in genomes (TILLING) for plant functional genomics. *Plant Physiol.* 123: 439-422.

Miao YL, Feher VA & McCammon JA (2015) Gaussian Accelerated Molecular Dynamics: Unconstrained Enhanced Sampling and Free Energy Calculation. *J. Chem. Theory Comp.***11**, 3584-3595.

Mohan A, Oldfield CJ, Radivojac P, Vacic V, Cortese MS, Dunker AK & Uversky VN (2006) Analysis of molecular recognition features (MoRFs). *J. Mol. Biol*. 362, 1043-1059.

Morgan JLW, Strumillo J, & Zimmer J (2013) Crystallographic snapshot of cellulose synthesis and membrane translocation. *Nature* 493: 181-186.

Nixon TB *et al.,* (2016) Comparative structural and computational analysis supports eighteen cellulose synthases in the plant cellulose synthase complex. *Sci Rep* **6,** 28696.

Nixon TB, *et al.* (2016) Comparative structural and computational analysis supports eighteen

cellulose synthases in the plant cellulose synthesis complex. *Sci. Rep.* 6: 28696.

Persson S*, et al.* (2007) Genetic evidence for three unique components in primary cell-wall cellulose synthase complexes in Arabidopsis. *Proc. Natl. Acad. Sci. USA* 104:15566-15571.

Petti, C., Hirano, K., Stork, J., and DeBolt, S. (2015). Mapping of a cellulose-deficient mutant named dwarf1-1 in Sorghum bicolor to the green revolution gene gibberellin20-oxidase reveals a positive regulatory association between gibberellin and cellulose biosynthesis. *Plant Physiol.* 169, 705–716.

Rushton, PS, Olek, AT, Makowski, L, Badger, J, Steussy, NC, Carpita NC, Stauffacher, CV (2017) Rice Cellulose SynthaseA8 Plant-Conserved Region Is a Coiled-Coil at the Catalytic Core Entrance *Plant Physiol.* 173(1): 482–494.

Purushotham P, Cho SH, Diaz-Moreno S, Kumar M, Nixon BT, Bulone V, & Zimmer J (2016) A single heterologously expressed plant cellulose synthase isoforms is sufficient for cellulose microfibril formation in vitro. *Proc. Natl. Acad. Sci. USA*. 113: 11360-11365.

Robertson DJ, Julias M, Lee SY & Cook DD (2017) Maize Stalk Lodging: Morphological Determinants of Stalk Strength. *Crop Sci.* 57, 926-934.

Scavuzzo-Duggan ET, Chaves AM, Singh A, Sethaphong L, Slabaugh E, Yingling YG, Haigler CH, Roberts AW (2018) Cellulose synthase 'class specific region' are intrinsically disordered and functionally undifferentiated. *J. Int. Plant Biol.* DOI: 10.1111/jipb.12637.

Scheible WR, Eshed R, Richmond T, Delmer D, & Somerville CR (2001) Modifications of cellulose synthase confer resistance to isoxaben and thiazolidinone herbicides in Arabidopsis *Ixr1* mutants. *Proc. Natl Acad. Sci. USA* 98:10079-10084

Schneider TD (1990) Sequence Logos: A new way to display consensus sequences. *Nucleic Acids Res.* 18: 6097-6100.

Sethaphong L, *et al.* (2013) Tertiary model of a plant cellulose synthase. Proc Natl Acad. Sci USA 110(18):7512–7517.

Shoemaker BA, Portman JJ &Wolynes PG. (2000) Speeding molecular recognition by using the folding funnel: The fly-casting mechanism. *Proc Natl Acad. Sci USA* 97, 8868.

Slabaugh E, Sethaphong L, Xiao C, Amick J, Anderson C T, Haigler CH, & Yingling YG. (2014) Computational and genetic evidence that different structural conformations of a non-catalytic region affect the function of plant cellulose synthase. *J Exp. Bot.* 65, 6645–6653.

Sugimoto K, Williamson, RE & Wasteneys, GO (2001) Wall architecture in the cellulose-deficient rsw1 mutant of Arabidopsis thaliana: microfibrils but not microtubules lose their transverse alignment before microfibrils become unrecognizable in the mitotic and elongation zones of roots. *Protoplasma*, 215, 172–183.

Taylor NG, Howells RM, Huttly AK, Vivkers K & Turner SR (2003). Interactions among three distinct CesA proteins essential for cellulose synthase. Proc Natl Acad. Sci USA 100: 1450-1455.

Udvardi MK, Czechowshi T, & Scheible, WR (2008) Eleven golden rules of quantitative RT-PCR. *Plant Cell* 20:1736-1737.

Vandavasi, V, *et al.* (2016). A Structural Study of CESA1 Catalytic Domain of Arabidopsis Cellulose Synthesis Complex: Evidence for CESA Trimers. *Plant Physiol*, 170, 123–135

Vogel J (2008) Unique aspects of the grass cell wall. *Curr. Opin Plant Biol.* 11: 201-207.

Vogel JP and the International Brachypodium Initiative. (2010). Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* 463: 763-768.

Von Forell G, Robertson D, Lee SY & Cook DD (2015) Preventing lodging in bioenergy crops: a biomechanical analysis of maize stalks suggests a new approach. *J. Exp. Bot.* 66; 4367-4371.

Zelko I, Lux A, Sterckeman T, Martinka M, & Kollarova K (2012) An easy method for cutting and fluorescent staining of thin roots, *Ann. Bot.* 110. 475-478.

Tables

Table 1. Sequence alignment of various isoforms showing conserved region E(K/R)xFGxS

```
BdCESA1      --------RMMKRTESSAPIFNMEDIEEGIE--GYEDERSMLMSQKRLEKRFGQSPIFTA
BdCESA3      --------KSNKHVDSSVPVFNLEDIEEGVEGAGFDDEKSLLMSQMSLEKRFGQSAAFVA
BdCESA6      --------LFFKRAENQSPAYALGEIEEGIPGA--ENDKAGIVNQEKLEKKFGQSSVFAA
BdCESA9      ----------LRRTMSVVPLLESEEDEEGIAEGGR--RRRLRSYSAALERHFGQSPLFIA
BdCESA4      KDKLGGAPKKGGSYRKQQRGFELEEIEEGIEGYD-ELERSSLMSQKNFEKRFGQSPVFIA
BdCESA7      -----GLP---------------ESVGDGMDG-----DKEMLMSQMNFEKRFGQSAAFVT
BdCESA8      --------RDSRREDLESAIFNLREIDNY-----DEYERSMLISQMSFEKSFGQSSVFIE
                                      .  :     :     .  :*: ****  *
```

Table 2. Quantification of non-cellulosic TFA soluble sugars in the stem, sheath, and leaf of wild type and TILLING mutants.

	Rhamnose[b]	Arabinose	Galactose	Glucose	Xylose
	mol (%)				
Stem [a]					
Wild Type	0.8 ± 0.1[c]	16 ± 0.4	3.2 ± 0.1	20 ± 0.7	60 ± 0.9
cesa1^{S830N}	0.9 ± 0.1*	16 ± 0.5	4.3 ± 0.2*	21± 0.6	58 ± 0.7*
Sheath					
Wild Type	1.1 ± 0.1	20 ± 1.0	5.0 ± 0.2	18 ± 0.4	56 ± 1.0
cesa1^{S830N}	1.2 ± 0.1	20 ± 0.6	6.4 ± 0.3*	17 ± 0.6	55 ± 1.3
Leaf					
Wild Type	2.3 ± 0.1	19 ± 0.6	7.6 ± 0.4	27 ± 1.7	44 ± 2.0
cesa1^{S830N}	1.8 ± 0.1	21 ± 0.5*	7.5 ± 0.2	23 ± 0.8*	47 ± 1.0

[a] Tissue from 6 biological reps was measured in triplicate for each genotype for neutral sugars

[b] Fucose and mannose values were either less than 0.7% or not detectable in tissue and not shown.

[c] A two-tailed t-test was used to compare means (± 1 standard error) of sugars within a tissue type and an asterisk indicates a significantly difference at an alpha value of 0.05. All values are rounded to the nearest one or tenth.

Table 3. Effect of point mutation predictions with other amino acids

Variant	PROVEAN Score	PROVEAN Prediction (cutoff= -2.5)	Duet ΔΔG	DUET prediction
S830A	-2.853	Deleterious	-0.636	Mildly Deleterious
S830V	-5.704	Deleterious	-0.074	Mildly Deleterious
S830L	-5.701	Deleterious	0.35	Stabilizing
S830G	-3.808	Deleterious	-1.025	Mildly Deleterious
S830W	-6.659	Deleterious	-0.924	Mildly Deleterious
S830T	-2.853	Deleterious	-0.445	Mildly Deleterious
S830Q	-3.805	Deleterious	-0.428	Mildly Deleterious
S830E	-3.807	Deleterious	-0.339	Mildly Deleterious
S830C	-4.755	Deleterious	-0.245	Mildly Deleterious
S830R	-4.756	Deleterious	-0.336	Mildly Deleterious
S830P	-4.758	Deleterious	-0.764	Mildly Deleterious
S830D	-3.808	Deleterious	-0.54	Mildly Deleterious
S830F	-5.705	Deleterious	-0.689	Mildly Deleterious
S830I	-5.703	Deleterious	0.327	Stabilizing
S830H	-4.758	Deleterious	-1.604	Mildly Deleterious
S830N	**-2.855**	**Deleterious**	**-0.619**	**Mildly Deleterious**
S830M	-4.752	Deleterious	0.248	Stabilizing
S830Y	-5.706	Deleterious	-0.446	Mildly Deleterious
S830K	-3.805	Deleterious	-0.211	Mildly Deleterious

Figure Legends

Fig 1. Characterizing relative transcript abundance of *BdCESA* genes in 3-4 day old roots, shoots, and coleoptiles to determine primary cell wall CesA. Fold change values were determined by comparing against gene expression in 3 week-old stem tissue. Means followed by a different letter within a tissue type are considered significantly different at alpha 0.05 using Tukeys test.

Fig 2. Gene structure, protein topology, TILLING region, and conserved sequence logo of point mutation in *BdCESA1*. Sequence logo plot of 88 eukaryotic *CESAs* following the QXXRW motif. The height of a logo is proportional to the conservation frequency at that position. The asterisk is the residue change in the *cesa1*^{S830N} mutant sequence. Gene structure and protein topology of plant *CESAs*. The black line and boxes are introns and exons, respectively. Colored boxes or lines within a box represent unique *CESA* protein domains: zinc finger (green box), class specific region (orange boxes), black lines (transmembrane domains), catalytic domains D, D, D, QxxRW (blue lines). The black arrows indicate the location of TILLING forward and reverse primers. The scale represents the length of *CESA* gene in kilobase pairs.

Fig 3. Cellulose content of stem, peduncle, leaf and sheath, from mature and senesced wild type (black) and *Bdcesa1*^{S830N} mutants. A two-tailed t-test was used to compare means (± 1 standard error) within a tissue type and an asterisk indicates a significantly difference at an alpha value of 0.05.

Fig 4. Morphological characteristics of wild type and the Bd*cesa1*^{S830N}. (A) Wild type and the *Bdcesa1*^{S830N} plants were propagated side by side and plant height and the image was captured during seed fill growth stages (P, peduncle). (B) Representative sample of stem tissue (right) and peduncle (left) demonstrates nodes frequency and peduncle length differences (N node; P penduncle and F floret). (C) The morphology of the uppermost node on the stem reveals little change between wild type and Bd*cesa1*^{S830N} (N node; LS leaf sheath; and S stem), note that the leaf blade was excised from the LS. Measurement of coleoptile (dark grown) and roots (light grown) length after 7 days and peduncle length for each genotype. A two-tailed t-test was used to compare means (± 1 standard error) of growth characteristics within a tissue type and an asterisk indicates a significantly difference at an alpha value of 0.05. Scale bars (A) 2.54 cm, (B, C) 1 cm, D 200 μm, E 40 μm.

Fig 5. The location of S830 in the BdCESA1 model and the BscA model. The mutation site is colored magenta, DCD motif is green and QxxRW is orange. In BdCESA1 model the PCR is red and CSR is blue. BscA does not have PCR and CSR-(c),(d) Conformational sampling of the CSR region as function of distance from mutation site and root mean square deviation for wild type and mutant respectively. -(e),(f) Conformational sampling of QxxRW region as a function of distance from the DCD site and root mean square deviation for wild type and mutant respectively. (g),(h) Conformational sampling of the DCD region as function of distance from the QxxRW and root mean square deviation for wild type and mutant respectively. Conformational spaces (c)-(h) are colored based on free energy with yellow being lowest energy state.

Fig 6. Time evolution of the interactions measured as a proximity distance between S830N site and the elements of the CSR region (a) conserved motif E(K/R)xFGxS. (b) MoRF1, (c) MoRF2 located in the wild type and the mutant CESA.