Distinctively altered lignin biosynthesis by site-modification of *OsCAD2* for enhanced biomass saccharification in rice

Guifen Zhang[a,c], Lingqiang Wang[a,b*], Xukai Li[a], Shuming Bai[a], Yali Xue[a], Zihui Li[a], Shang-wen Tang[c], Yanting Wang[a, c], Youmei Wang[a], Zhen Hu[a], Ping Li [a] and Liangcai Peng[a,c*]

[a]Biomass & Bioenergy Research Centre, College of Plant Science & Technology, Huazhong Agricultural University, Wuhan 430070, China.
[b]State Key Laboratory for Conservation & Utilization of Subtropical Agro-Bioresources, College of Agriculture, Guangxi University, Nanning, China.
[c]Laboratory of Biomass Engineering & Nanomaterial Application in Automobiles, College of Food Science & Chemical Engineering, Hubei University of Arts & Science, Xiangyang, China.

*Correspondent author: Tel: +86-27-87281765; Fax: +86-27-87280016.
Email: lpeng@mail.hzau.edu.cn; lqwang@mail.hzau.edu.cn. Web: http://bbrc.hzau.edu.cn.

Running title: ***Modified OsCAD2 enhances biomass saccharification***

Abstract

Crop straws represent enormous biomass resource convertible for biofuels and bioproducts, but lignocellulose recalcitrance restricts its saccharification for commercial utility. Despite genetic modification of lignin biosynthesis has been attempted to reduce recalcitrance in bioenergy crops, it remains challenging to optimize lignin deposition without an unacceptable yield penalty. Based on gene expression analysis and phylogenetic tree profiling, this study selected a cinnamyl-alcohol dehydrogenase gene (*OsCAD2*) as the target for genetic engineering of lignin biosynthesis in rice. Using CRISPR/Cas9 technology, this study generated independent homozygous transgenic lines with precise site-mutation of *OsCAD2*, which showed slightly reduced lignin levels but markedly decreased *p*-hydroxyphenyl (H) contents in lignin by 34% and increased guaiacyl (G) contents by 16%, compared to the wild type. Under mild alkali pretreatment (1% NaOH, 50°C), the *OsCAD2* site-modified lines showed effective lignin extraction up to 70% (of total lignin) from mature rice straws, which caused either significantly increased biomass porosity and cellulose accessibility

or remarkably reduced cellulase adsorption to lignin in pretreated lignocellulose residues. These consequently led to almost complete biomass enzymatic saccharification with increased hexoses yields by 61%-72% in the modified lines, being much higher than those of the lignin-altered lines reported in previous studies. Hence, this study has demonstrated a novel genetic engineering strategy to reduce lignocellulose recalcitrance with minimized biomass loss for cost-effective biomass conversion to bioethanol in rice and bioenergy crops.

Keywords: CAD, Lignin, CRISPR/Cas9, Rice, Alkali pretreatment, Enzymatic saccharification, Bioethanol, Biomass porosity, Cellulose accessibility

Introduction

Crop feedstock provides large amounts of lignocellulose residues, however, its intrinsic recalcitrance requires high energy input for bioethanol production. Lignocellulose recalcitrance is fundamentally determined by major wall polymers (cellulose, hemicellulose, lignin) levels and features (Li et al., 2017). As a major contributor to the recalcitrance, lignin forms a physical barrier against cellulases accession by interlinking with hemicellulose and cellulose (Haarmeyer et al., 2017; Zeng et al., 2014; Li et al., 2014). Recently, lignin has been examined to adsorb cellulases, which affects enzymatic hydrolysis of lignocellulose (Jin et al., 2016; Deng et al., 2020). Harsh physical and chemical pretreatments are thus required for effective lignin extraction, demanding high energy input and leading to potential secondary chemical pollution into environment (Himmel et al., 2007).

Genetic modification of plant cell walls has been implemented as a promising solution to the lignocellulose recalcitrance (Hamelinck et al., 2005; Aden & Foust, 2009). In particular, the genetic engineering of lignin biosynthesis could enhance biomass enzymatic saccharification and bioethanol production in bioenergy crops (Ralph et al., 2004; Chen et al., 2003). However, due to complicated structures and diverse functions of plant cell walls, their modification to some degree could simply

affect plant growth and mechanic strength (Li et al., 2017, Wang et al., 2016). Hence, it becomes important to select desirable bioenergy crops using advanced genetic engineering approach.

As a characteristic component of secondary cell walls, lignin primarily consists of *p*-hydroxyphenyl (H), syringyl (S), and guaiacyl (G) units, which are polymerized by the radical coupling of three monolignols: *p*-coumaryl, sinapyl and coniferyl alcohols (Chen and Richard, 2007). Genetic engineering of the enzymes involved in the monolignol biosynthetic pathways has been conducted in various plants such as alfalfa (Fu et al, 2011a), switchgrass (Park et al., 2017; Zhou et al., 2015), poplar (Cass et al., 2015), *Brachypodium* (Wu et al., 2019) and tall fescue (Jackson et al., 2008). Cinnamyl alcohol dehydrogenase (CAD) is the enzyme involved in the last step of phenylpropanod pathway by catalyzing the NADPH-dependent reduction of hydroxy-cinnamaldehydes into the corresponding alcohols (Kim et al., 2002; Raes et al., 2003; Saballos et al., 2009). Downregulation of *CAD* genes has enhanced biomass enzymatic digestibility in several eudicot species (Baucher et al., 1996, 1999). Meanwhile, genetic manipulation of *CAD* genes is also implemented in grasses (Chen et al., 2003; Saathoff et al., 2011; Fornale et al., 2012; Bouvier et al., 2013), but little has yet been explored about how lignocellulose recalcitrance is effectively improved in the transgenic plants.

Rice is a major cereal crop worldwide, producing approximately 650-975 million tons of lignocellulose-based straw per year (Binod et al., 2010). Although *OsCAD2* has been identified in rice (Tobias & Chow, 2005; Zhang et al., 2006; Li et al., 2009; Hirano et al., 2012), much remains unknown about its impact on biomass properties. In this study, we identified totally twelve sequences that significantly matched to the Os*CAD* family by searching TIGR database. Then, we selected one of the members (*OsCAD2*) which preferential expressed in the lignocellulose-rich culm tissue, and generated the transgenic lines with precisely site-modification of *OsCAD2* using CRISPR/Cas9 technology. Hence, this study determined significantly reduced lignocellulose recalcitrance in the modified lines, leading to much enhanced biomass enzymatic saccharification and bioethanol production compared to the wild type. Finally, this work attempted to interpret how the lignocellulose recalcitrance is improved for

enhanced bioethanol production in the modified lines, thereby demonstrating a powerful strategy for lignin modification in bioenergy crops.

Materials and methods

Plant materials and samples collection

Rice (*Oryza sativa* L. *ssp*) transgenic plants and the wild type Nipponbare (NPB) were grown in the experimental fields of Huazhong Agricultural University, Wuhan, China. The rice mature straws were collected, dried at 50 °C to the constant weight, and stored in a dry container until use.

Phylogenetic analysis and motif identification

The multiple alignment analysis was performed using the Clustal X program (version 1.83) (Thompson et al., 1997) and MAFFT (Katoh et al., 2005). The unrooted phylogenetic trees were constructed with the MEGA7.0 program and the neighbor joining method with 1,000 bootstrap replicates (Kumar et al., 2004). Protein sequences were analyzed using the MEME program (http://meme.sdsc.edu/meme/cgi-bin/meme.cgi) for the confirmation of motifs. The MEME program (version7.0) was employed with the following parameters: number of repetitions; maximum number of motifs; optimum motif width set to > 6 and < 200. The motifs were annotated using the Inter Pro Scansearch program (http://www.ebi.ac.uk/Tools/InterProScan/).

OsCADs genes database search and expression profile in rice

The Hidden Markov Model (HMM) profile of the cellulose synthase and the cinnamyl alcohol dehydrogenase domain was downloaded from PFam (http://pfam.sanger.ac.uk/). A name search and the protein family ID were performed for the identification of genes from the rice genome. Information about the coding sequence (CDS), amino acid (AA) and full-length cDNA accessions was obtained from TIGR (http://www.tigr.organd KOME http://cdna01.dna.affrc.go.jp/cDNA). The gene

expression profile data in 33 rice tissue samples (Table S1) of Zhenshan 97 (ZS97) and Minghui 63 (MH63) were obtained from the CREP databasehttp://crep.ncpgr.cn. The expression values were log transformed, and cluster analyses were performed using a software cluster with Euclidean distances and the hierarchical cluster method of "complete linkage clustering". Clustering tree was constructed and viewed in JavaTreeview.

Plasmid vector construction and rice transformation

The plasmid vector pRGE32 was used to transiently express U3p:sgRNA along with UBIp:Cas9 in rice. The construct carrying the U3p:sgRNA and UBIp:Cas9 was transformed into the *Agrobacterium* strain (EHA105) as described by Xie et al. (2015). The transformation and regeneration of transgenic plants were performed as described by Fan et al. (2018). The positive transgenic lines were selected based on PCR and sequencing analysis. The sequences of sgRNA (*OsCAD2*) and the primers used in this study were listed in Table S2.

Hemicellulose and cellulose extraction and monosaccharide measurement

Cell wall fractionation procedure was performed as previously described (Peng, Hocart, & Williamson, 2000) with minor modification. The dry biomass powder (40 mesh) samples (0.1 g) were initially washed twice with 5.0 mL phosphate buffer and twice with 5.0 mL distilled water. The remaining pellet was stirred with 5.0 mL chloroform-methanol (1:1, v/v) for 1 h at 40 °C and washed twice with 5.0 mL methanol, followed by 5.0 mL acetone. The remaining pellet was added with 5.0-mL aliquot of DMSO-water (9:1, v/v), rocked gently on a shaker overnight. After centrifugation at 3000 *g*, the pellet was washed twice with 5.0 mL DMSO-water and then with 5.0 mL distilled water three times. The remaining pellet was defined as total crude cell walls. The crude cell walls were further suspended in 0.5% (w/v) ammonium oxalate (5.0 mL) and heated for 1 h in a boiling water bath and all supernatants were used as pectin fraction. The remaining pellet was suspended in 4 M KOH containing 1.0 mg/mL sodium borohydride (5.0 mL) and incubated for 1 h at 25 °C. Total hexoses and pentoses

released from the 4 M KOH extraction and the pentoses released from 67% (v/v) H_2SO_4 hydrolysis of the non-KOH-extractable pellet were summed as total hemicellulose content. Cellulose was estimated by calculating the hexoses from the non-KOH-extractable pellet. The anthrone/H_2SO_4 method (Fry, 1988) and orcinol/HCl method (Dische, 1962) were respectively applied for the hexose and pentose assays. D-glucose and D-xylose were prepared to plot standard curves, and the pentose reading at 660 nm was deducted for the final hexose calculation to eliminate the interference of pentoses on the hexose reading at 620 nm. The remaining pellet after pectin fraction was added with 2.5 mL TFA (2 M) in a sealed tube at 121°C in an autoclave (15 psi) for 1 h. The supernatant was used to determine monosaccharide composition of hemicellulose by GC-MS as previously described (Li. et al, 2017). For cellulose and hemicellulose contents analyses, all experiments were conducted at independent triplicates.

Lignin level and monolignol detection

The biomass powder samples (300 mg) were extracted with 1% NaOH at 50 °C for 2 h. After centrifugation at 3000 *g*, the remaining residues were washed twice with distilled water and dried under vacuum. The dried residues were used for lignin analysis. Total lignin content of the raw materials and the residues after 1% NaOH pretreatment were determined by two-step acid hydrolysis method according to the Laboratory Analytical Procedure of the National Renewable Energy Laboratory (Sluiter et al., 2008). The acid-insoluble lignin was calculated gravimetrically as acid-insoluble residue after correction for ash, and the acid-soluble lignin was measured by UV spectroscopy. Monolignols were extracted by Nitrobenzene Oxidation (NBO) as described by Schultz (1986) with minor modifications as described by Li et al. (2014). Standard H-, G- and S-monolignols were purchased from Sinopharm Chemical Reagent Co., Ltd. A Kromat Universal C18 column (4.6 mm × 250 mm, 5 μm) was used for HPLC analysis with a SHIMADZU LC-20A machine with a UV-detector at 280 nm. CH_3OH: H_2O: HAc (16:63:1, v/v/v) was used as mobile phase (flow rate: 1.1 mL/min), the injection volume was 20 μL. For total lignin level assay, experiments were conducted at independent triplicates.

Characterization of metabolic products involved in lignin biosynthetic pathways

The second inverted internodes of rice stem tissues at heading stage were homogenized using a tissue grinder (Schwingműhle Tissue Lyser II, Germany) for 50 sec at 29 Hz. For each sample, 100 mg of dry powder was mixed with 1.0 ml of 70% methanol containing 0.1 mg/l acyclovir (internal standard), vortexed, and extracted for 10 h at 4°C. After centrifugation at 9500 g for 10 min, the resulting supernatant passed through a 0.22 μm organic filter (SCAA-104; ANPEL, China) and collected for analysis using a LC-ESI-MS/MS system (HPLC) (Shim-pack UFLC SHIMADZUCBM20A system, 5500 Q TRAP; Applied Biosystems, Framingham, MA, USA). A stepwise multiple ion monitoring-enhanced product ion was used to construct the MS2T library, as described by Chen et al. (2013). To facilitate the identification of the metabolites detected in this study, accurate m/z for each precursor ions was obtained using a time-of-flight mass spectrometry platform (HPLC, Shim-pack UFLCSHIMADZUCBM20A system, Triple TOF 5600; Applied Biosystems). The quantification of the metabolites was conducted using a scheduled multiple reaction monitoring (MRM) method as previously described by Chen et al. (2013). The scheduled MRM algorithm was used with an MRM detection window of 90 sec and a target scan time of 1.0 sec in Analyst 1.5 software. The metabolite data were $\log_2$ -transformed for statistical analysis to improve normality.

Biomass porosity and cellulose accessibility measurements

Simons' stain (SS) was applied to determine the overall biomass porosity as previously described (Chandra，Ewanick, &Saddler, 2008; Sun et al., 2017). Direct Blue 15 (DB: Phenamine Sky Blue A Conc) and Direct Yellow 11 (DY) were purchased from Pylam Products Co. Inc (Garden City, NY). The samples (0.10 g) were added with 1 mL of Alum electrolyte solution (5 mM $KAl(SO_4)_2$ + 1.5 mM NaCl) to 15-mL polypropylene centrifuge tubes. A set of tubes with a 1:1 solution of DB and DY were prepared by adding the dye solution (10 mg/mL) in a series of volumes (0.25, 0.50, 0.75, 1.0, 1.5 mL). Distilled water was added to a final volume of 10 mL, and the samples were incubated for 9 h at 70 ºC under shaken at 200 rpm. After centrifugation at 8000 *g*, the

absorbances of the supernatants were measured at 612.5 nm and 410.5 nm, which are the wave lengths of maximum absorbance for DB15 and DY11. The concentration of DB and DY dyes in the supernatants was calculated according to the Lambert–Beer law for binary solutions using equations 1 and 2. The maximum DB and DY dyes adsorbed to the biomasses were calculated using the Langmuir adsorption model as previously described by Chandra et al. (2008).

$$A_{410.5\ nm} = \varepsilon_{Y/410.5} L C_Y + \varepsilon_{B/410.5} L C_B \qquad (1)$$

$$A_{612.5\ nm} = \varepsilon_{Y/612.5} L C_Y + \varepsilon_{B/612.5} L C_B \qquad (2)$$

Where A is the absorbance of the supernatant from the dye mixture at 410.5 or 612.5 nm, Ɛ is the extinction coefficient of each dye at the respective wavelength, and L is the path length equal to the cuvette width (1 cm). The calculated extinction coefficient values used in this study were ƐY/410.5 = 33.50, ƐB/410.5= 3.56, ƐY/612.5 = 0.13, ƐB/612.5 = 24.66 L $g^{-1}cm^{-1}$.

Congo Red stain was applied to detect the cellulose accessibility as previously described (Wiman et al, 2012; Alam et al, 2019; Cheng et al, 2019). The biomass samples (100 mg) were incubated with dye solution at a series of concentrations (0.5, 1.0, 2.0, 3.0, 4.0, 5.0 g /L) in 0.3 M phosphate buffer at 60 ºC for 24 h. After centrifugation at 8000 *g*, the absorbance of the supernatant was measured at 498 nm. The cellulosic surface area was estimated from the maximum adsorption capacities by assuming that direct red adsorbs as dimer aggregates under the experimental conditions. The maximum dye adsorption capacity of the biomass was calculated using the monolayer Langmuir adsorption model. All measurements were conducted in independent triplicates.

Electron microscopy observation

The biomass residues were observed using scanning electron microscopy (SEM JSM-IT300, Akishima, Tokyo, Japan) as previously described by Xu et al. (2012). Well-mixed biomass residues were sputter-coated with gold in a JFC-1600 ion sputter (Mito City, Japan) and visualized for 5 to 8 times to acquire images.

Biomass enzymatic hydrolysis

Acid (H_2SO_4) and alkali (NaOH) pretreatment and enzymatic hydrolyses were respectively performed as described by Fan et al. (2017). For the alkali pretreatment, biomass powders were incubated with 6 mL of 1% NaOH (w/v) under shaking (150 rpm) at 50 °C for 2 h. For the acid pretreatment, biomass powder was added to 6 mL of 1% H_2SO_4 (v/v) and heated at 121 °C for 20 min in an autoclave. After centrifugation at 3000 *g* for 5 min, the biomass residues of alkali or acid pretreatment were incubated with the mixed-cellulase enzymes (10.60 FPU g^{-1} biomass and xylanase 6.72 U g^{-1} biomass), while 1% Tween was co-supplied. After enzymatic hydrolysis were performed under 150 rpm at 50 °C for 48 h, the supernatants were collected for yeast fermentation.

Yeast fermentation

Yeast fermentation was conducted with the *Saccharomyces cerevisiae* (purchased from Angel Yeast Co., Ltd., Yichang, China) strain using total hexoses obtained from saccharification as the carbon source at 3% solid loading. The fermentation liquid was distilled at 100 °C to collect ethanol liquor. Ethanol was measured using $K_2Cr_2O_7$ method as previously described by Alam et al. (2020). The fermentation liquid was distilled at 100 °C for 15 min to achieve ethanol liquor. Appropriate amount of ethanol sample in 2.00 mL 5% $K_2Cr_2O_7$ (5g $K_2Cr_2O_7$ dissolved in 90.00 mL distilled water and 10.00 mL 98% sulfuric acid) was heated for 10 min in a boiling water bath. After cooling, the absorbance was read at 600 nm. All experiments were conducted in independent triplicate.

Crude cellulose hydrolysis with β-1,4-exoglucanase (cellobiohydrolase-CBH I)

The mature straws were dried, ground into biomass powders through size 40 mesh, and extracted with 4 M KOH (containing 1.0 mg/mL sodium borohydride) at 25 °C for 2 h to remove hemicellulose. The remaining pellets were washed 5 times with ddH_2O and extracted with 8% (w/v) sodium chlorate (containing1.5% acetic acid, v/v) at 25 °C for 48 h to remove lignin. The remaining pellets were washed 6 to 8 times with ddH_2O,

dried under a vacuum and defined as crude celluloses. The crude cellulose samples (10 mg) were incubated with 5 U CBH I (E.C. 3.2.1.91; Megazyme, USA) PH 4.8 at 37 °C for 14 h. The reaction was performed in the 1.5 ml tube with 100 mM NaAC, pH 5.0. After centrifugation at 3000 *g*, the supernatants were collected and treated with 2 M TFA at 120 °C for 1 h, and *myo*-inositol (20 μg) was added as the internal standard. The glucose released from CHB I was determined by GC-MS (SHIMADZU GCMS-QP2010 Plus) as previously described by Huang et al. (2019).

Soluble enzyme detection and SDS-PAGE

The soluble cellulase enzymes were collected from the supernatants from biomass enzymatic hydrolysis described above, and total proteins were determined using Coomassie Brilliant Blue G250 assay (Sedmak et al., 1977). The Coomassie Brilliant Blue G250 dye was prepared in solution containing ethanol (w/v; 2:1) and phosphoric acid (w/v; 1:1) and filtered through a 0.22 μm filter. The absorbance of the protein-dye complex was reading at 595 nm using UV-vis spectrometer (V-1100D, Shanghai MAPADA Instruments Co., Ltd. Shanghai, China). Every 1.0 mL enzymatic hydrolysate was added into 10 mL ethanol for precipitating cellulase and reducing the Tween-80 interference. The precipitated protein was added into 1.0 mL distilled water and 3.0 mL Coomassie Brilliant Blue G250, and the absorbance was measured 10 min later. The protein samples were loaded into 12% SDS-PAGE gel as previously described by Sun et al. (2020). All protein content assays were performed in independent triplicate.

Data collection and statistical analysis

The SPSS statistics software was used for the statistical analysis. All data were initially analyzed by the Kolmogorov-Smirnov test to check for normal distribution. The Spearman correlation analysis was performed to calculate correlation coefficients, and the Student's *t*-test was used for comparison analysis.

Results and discussion

Selection of OsCAD2 as the target gene for rice genomic editing

By searching for TIGR database, this study initially identified twelve DNA sequences matched to the *OsCAD* gene family in rice, which were designated as *OsCAD1-D7, OsCAD8A, B, C*, *D,* and *OsCAD9* (Table S3). An unrooted phylogenetic tree was then generated from alignments of OsCADs proteins along with the orthologous AtCADs in *Arabidopsis* (Figure 1A). OsCAD2 was most close to the AtCAD4 and AtCAD5 in protein similarity (99% identity) and in motif constitution (Figure 1B). Furthermore, the *OsCADs* expression profiling was obtained from the CREP database (http://crep.ncpgr.cn; Wang et al., 2010) including the genome-wide expressions of 33 tissues almost throughout the entire life cycle of rice (Table S1). All *OsCAD* genes could be classified into three major groups. As a difference from other members, the *OsCAD2* was prominently expressed in the lignocellulose-rich tissues such as culm, mature sheath, panicle, hull and spikelet (Figure 1C), suggesting that it is an appropriate target for genomic editing in this study.

Generation of homozygous OsCAD2 transgenic lines with site-mutation by CRISPR/Cas9

Using CRISPR/Cas9 editing technology, this study selected one specific region in exon 3 of the *OsCAD2*, with the 20-bp target site for the design of a sgRNA using CRISPR-P program (Figure 2A). The binary constructs carrying the sgRNA within target region with Cas9p driven by UBIp were generated (UBIp: Cas9-*OsCAD2*), and transformed into the *japonica* rice cultivar Nipponbare (NPB) via *agrobacterium*-mediated transformation, resulting in total 59 independent transgenic lines obtained in this study. A total of 43 T-DNA positive transgenic lines were identified by PCR, and 10 lines were further sequenced to verify gene editing occurrence.

We then screened out three independent homozygous rice transgenic *OsCAD2*-edited lines carrying single-site (1 bp) insertions in the target regions (Figure 2B). Compared to the wild type (WT)/NPB, those three homozygous *OsCAD2* site-mutated

lines exhibited normal plant growth with lower plant height (Figure 3A， Table S4). Furthermore, we found relatively lower lignin levels in the mature straws of three transgenic rice lines, and in particular, two lines showed significantly reduced lignin levels by 15% and 18%, compared to the WT (Figure 3B, Table 1). Notably, three transgenic rice lines showed significantly increased cellulose levels by 10%-12% and hemicellulose contents by 17%-21%. Hence, the increase of cellulose and hemicellulose contents should compensate for the reduction of lignin in the cell walls of the three transgenic lines, leading to normal plant growth and similar plant strength. In addition, this study showed similar monosaccharide compositions of hemicelluloses between modified lines and WT (Table S5), suggesting that the *OsCAD2* site-mutation has little effect on hemicellulose biosynthesis.

Enhanced hexoses and bioethanol yields from the biomass of OsCAD2 modified lines

To measure biomass saccharification, we calculated both hexoses and pentoses yields released from enzymatic hydrolysis after chemical pretreatments of mature rice straws. Using our previously-established approaches (Xu et al., 2012; Li et al., 2017), this study performed mild acid (1% H_2SO_4) and alkali (1% NaOH) pretreatments to enhance biomass saccharification for bioethanol fermentation in three modified lines (Figure 4; Figure S1). Compared to the WT, the modified lines were of significantly higher hexose yields against cellulose (% cellulose) or dry matter (% DM) after pretreatments (n=3) (Figure 4A-D). In comparison, the alkali pretreatments led to much more enhanced hexoses yields than those of the acid pretreatments. Importantly, when 1% Tween-80 was co-supplied into the enzymatic hydrolysis, the 1% NaOH pretreatment could cause an almost complete biomass saccharification with hexoses yields close to 100% (% cellulose) in two modified lines, whereas the WT only had the hexoses yield of 82% (Figure 4E). Meanwhile, this study also determined much more enhanced pentoses yields (% DM) released from enzymatic hydrolysis of pretreated biomass residues in all modified lines (Fig. S1), which should be due to relatively raised hemicellulose levels in the modified lines.

Furthermore, the modified lines showed much increased hexose yields by 60%,

compared to the WT, which was even higher than those of previously lignin-modified lines with raised rates of less than 46% (Table 2). Hence, the site-mutation of *OsCAD2* should be more effective for reducing lignocellulose recalcitrance. In addition, this study performed a yeast fermentation using total hexoses released from enzymatic hydrolysis, and determined significantly 25%-30% higher bioethanol yields in all modified lines than those of the WT at $p < 0.01$ (n=3) (Figure 4F).

Distinct alteration of H- and G-monomers in the OsCAD2 modified transgenic lines

With respect to the lignin reduction in the *OsCAD2* modified lines, this study determined three major monomer proportions of lignin in rice straws (Table 1). Compared to the WT, three modified lines showed reduced H-monomer and increased G-monomer levels with a similar S-monomer content. In detail, the average H-monomer of the three modified lines was reduced by 34% and the average G-monomer was increased by 16% relative to the WT, which resulted in three monomers ratios (H/S, S/G, H/G) reduced by 1.7-, 1.2-, 1.8-folds in the modified lines. Hence, the modified lines showed a distinct alteration of H- and G-monomers in the mature rice straws (Table 1). Furthermore, on the basis of dry weight per plant, three monomers (S, G, H) levels of lignin were all reduced in the transgenic lines, compared to the WT, indicating that the *OsCAD2* site-mutation could affect three monomers biosynthesis. The results also suggested that the site-mutation of *OsCAD2* should predominately affect H-monomer biosynthetic pathway, leading to a consequently enhanced G-monomer biosynthesis by redirecting the metabolic flux into the channel of coniferyl alcohol for G-monomer incorporation (Figure S2). However, it remains interesting to explore why the H-monomer biosynthesis was mostly inhibited in the transgenic lines in the future study. In addition, as many *CAD*-like putative genes have been identified with distinct catalytic activities for lignin biosynthesis in plants (Sibout et al., 2005; Youn et al., 2006; Park et al., 2018), site mutations of different *CAD* genes may lead to characteristically modified lignocellulose production in the transgenic lines.

Effective lignin extraction from alkali pretreatment in the OsCAD2 modified lines

As lignin deposition forms a barrier against biomass enzymatic hydrolysis, this study detected lignin extraction from alkali pretreatment with mature straws (Figure 5). By comparison, about 62%-70% lignin was extracted by the alkali pretreatment in the modified lines, but only 57% lignin removed in the WT (Figure 5A), indicating that the distinct alteration of H- and G-monomer should cause an effective lignin extraction in the modified lines. However, both modified lines and WT showed similar hemicellulose extraction rates of 11%-12% from the alkali pretreatments (Figure 5B), suggesting that the lignin modification may not affect its interaction with hemicellulose in the modified lines. Since correlation analysis has been well applied to account for wall polymer impacts on biomass enzymatic saccharification (Deng et al., 2020; Sun et al., 2017; Pei et al., 2016), this study found out that the lignin levels of biomass residues were negatively correlated with the hexoses yields released from enzymatic hydrolysis at $p < 0.01$ (n=24, R^2 =0.91) (Figure 5C), which confirmed that the lignin extraction should play a major role for enhanced biomass enzymatic saccharification in the modified lines.

Increased biomass porosity and accessibility in the OsCAD2 modified lines

Because the *OsCAD2* modified lines are of high biomass enzymatic saccharification, this study measured biomass porosity and cellulose accessibility in mature rice straws, which have been characterized as the direct parameters accounting for enzyme accession and loading capacities (Grethlein, 1985; Sun et al., 2017; Alam et al., 2019). Using our improved Simons' staining approach (Sun et al., 2017), we measured significantly increased yellow dye (DY) and total dye staining in raw biomass samples of modified lines at $p < 0.01$ level (n=3), but no significant difference in blue dye (DB) staining (Figure 6A-C). As DB and DY are accountable for small and large pore sizes of biomass residues respectively (Sun et al, 2017), the data revealed that the modification could increase the amounts of large-size pores in raw biomass materials. Interestingly, after the 1% NaOH pretreatments, the modified lines showed much more increased in both DB and DY staining than those of the WT (Figure 6D-F), consistent with the effective lignin extraction in the modified lines. On the other hands, the lignin

levels were negatively correlated with the total dyes at $p < 0.01$, whereas that both cellulose and hemicellulose levels showed the positive correlation with total dyes (Figure 6G), suggesting that the lignin extraction should play a specific role for increased biomass porosity.

Subsequently, the cellulosic surface area was estimated in the wet state using DR28 staining (Congo red). Since the DR28 molecule (7.4Å×4.3Å×26.2Å) is specific for smaller pores (Pelekani & Snoeyink, 2001), the cellulosic surface area was measured from the maximum adsorption capacities by assuming that direct red adsorbs as dimer aggregates under the experimental conditions. Hence, this study conducted Congo Red stains with pretreated biomass residues to detect the cellulose surface areas accountable for cellulose accessibility. By comparison, three modified lines showed significantly raised cellulose surface areas than those of the WT at $p < 0.05$ level (Figure 6H), suggesting that the modified lines should be of much more surface sites for cellulases loading and attacking. This result could be supported by the findings that the glucose yields released in the modified line are up to 1.5-2 folds higher than those of the WT from the cellobiohydrolase (CBH I) hydrolyses (Figure 6I). Because the CBHI specifically attacks on the reducing-ends of β-1, 4-glucan chains to produce cellobiose, the data could be directly accounting for increased cellulose accessibility in the modified lines (Huang et al., 2019).

In addition, this study observed surface images of biomass residues in mature rice straws under scan electronic microscopy (Figure S3). Without any pretreatment, three modified lines obviously exhibited coarse surfaces than those of the WT. However, although alkali pretreatments could cause smooth surfaces in all biomass samples, three modified lines remained much rougher faces than the WT did, consistent with the findings of effective lignin extraction and increased biomass porosity and accessibility in the modified lines.

Reduced cellulases adsorption for high biomass saccharification in the OsCAD2 modified lines

To test if the effective lignin extraction could also reduce mixed-cellulases adsorption,

this study further examined soluble protein contents in the supernatants of enzymatic hydrolyses with the pretreated biomass residues (Figure 7). Compared to the WT, three modified lines contained significantly increased soluble proteins at $p < 0.01$ (n=3), with the protein levels increased by 31%-37% (Figure 7A). SDS-PAGE profiling showed that the soluble proteins were derived from the commercial mix-cellulases incubated with the pretreated biomass residues for enzymatic hydrolysis (Figure 7B). In addition, the soluble proteins levels were positively correlated with the hexoses yields released from enzymatic hydrolysis at $p < 0.01$ (n=12, R=0.88) (Figure 7C), interpreting that the modified lines with increased biomass saccharification should be partially due to reduced mixed-cellulases adsorption with the pretreated biomass residues. Furthermore, the soluble proteins levels had a significantly negative correlation with lignin contents of the pretreated biomass residues at $p < 0.01$ (n=12, R=-0.81) (Figure 7D), but had no significant correlation with cellulose or hemicelluloses (data not shown), suggesting that lignin may be specific for enzyme adsorption.

Taken together, this study demonstrated that the effective lignin extraction could not only increase biomass porosity and cellulose accessibility for efficient cellulases access and loading, but may also reduce cellulases adsorption in the *OsCAD2* modification lines, leading to an integrated enhancement of its biomass enzymatic saccharification for high bioethanol production.

Conclusion

Based on the precise editing of *OsCAD2* by CRISPR/Cas9 technology, this study generated the *OsCAD2* site-mutated lines showing a distinct lignin modification with predominately reduced H-monomer and increased G-monomers. Under mild alkali pretreatment, the *OsCAD2* modified lines exhibit an effective lignin extraction for either remarkably increased biomass porosity (cellulose accessibility) or reduced cellulase adsorption, leading to integrated enhancements for almost complete biomass enzymatic saccharification and much increased bioethanol yields, compared to the WT. Hence, this study has demonstrated a powerful strategy for genetic modification of plant

cell walls in bioenergy crops, and it has also provided insights into lignin biosynthesis in plants.

Acknowledgments

We thank Professor Kabin Xie for kindly providing CRISPR/cas9 constructs. This work was supported in part by grants from the Natural Science Foundation of Guangxi (2020GXNSFDA238027), National Natural Science Foundation of China (31771775; 31171524; 31670296; 31571721), National Key Research and Development Project of China (2018YFD0301303), the Project of Huazhong Agricultural University Independent Scientific & Technological Innovation Foundation (2014bs04), the National Key Research and Development Program (2016YFD0800804), the National 111 Project (BP0820035) and the Project of Hubei University of Arts and Science (XKQ2018006).

Conflicts of interest

The authors declare that they have no conflicts of interest.

References

Aden, A., & Foust, T. (2009). Technoeconomic analysis of the dilute sulfuric acid and enzymatic hydrolysis process for the conversion of corn stover to ethanol. *Cellulose*, 16, 535-545. doi: 10.1007/s10570-009-9327-8

Alam, A., Wang, Y., Liu, F., Kang, H., Tang, S., Wang, Y., … Peng, L. (2020). Modeling of optimal green liquor pretreatment for enhanced biomass saccharification and delignification by distinct alteration of wall polymer features and biomass porosity in *Miscanthus*. *Renewable Energy*, 159, 1128-1138. doi: org/10.1016/j. renene. 2020. 06. 013

Alam, A., Zhang, R., Liu, P., Huang, J., Wang, Y., Hu, Z., … Peng, L. (2019). A finalized determinant for complete lignocellulose enzymatic saccharification

potential to maximize bioethanol production in bioenergy *Miscanthus*. *Biotechnology for Biofuels*, 12. doi: org/10.1186/s13068-019-1437-4

Baucher, M., Bernard-Vailhe M., Chabbert, B., Besle, J., Opsomer, C., Van Montagu M., & Botterman, J. (1999). Down-regulation of cinnamyl alcohol dehydrogenase in transgenic alfalfa (Medicago sativa L.) and the effect on lignin composition and digestibility. *Plant Molecular Biology*, 39, 437-447. doi: 10.1023/A:1006182925584

Baucher, M., Chabbert, B., Pilate, G., Van Doorsselaere J., Tollier, M. T., Petit-Conil, M., … Boerjan, W. (1996). Red xylem and higher lignin extractability by down-regulating a cinnamyl alcohol dehydrogenase in poplar. *Plant Physiology*, 112,1479-1490. doi: 10.1104/pp.112.4.1479

Baxter, H. L., Poovaiah, C. R., Yee, K. L., Mazarei, M., Rodriguez, M., Thompson O.A., & Stewart, C.N. (2015). Field evaluation of transgenic switchgrass plants overexpressing *PvMYB4* for reduced biomass recalcitrance. *Bioenergy Research*, 8, 910-921. doi: 10.1007/s12155-014-9570-1

Binod, P., Sindhu, R., Singhania, R. R., Vikram, S., Devi, L., Nagalakshmi, S., …Pandey, A. (2010). Bioethanol production from rice straw: An overview. *Bioresource Technology*, 101, 4767-4774. doi: 10.1016/j.biortech.2009. 10. 079

Bouvier d'Yvoire, M., Bouchabke-Coussa, O., Voorend, W., Antelme, S., Cézard, L., Legée, F., … Sibout, R. (2013). Disrupting the cinnamyl alcohol dehydrogenase 1 gene (*BdCAD1*) leads to altered lignification and improved saccharification in *Brachypodium distachyon*. *The Plant Journal*, 73, 496-508. doi:10.1111/tpj.12053

Cass, C. L., Peraldi, A., Dowd, P. F., Mottiar, Y., Santoro, N., Karlen, S. D., … Sedbrook, J.C. (2015). Effects of phenylalanine ammonia lyase (PAL) knockdown on cell wall composition, biomass digestibility, and biotic and abiotic stress responses in *Brachypodium*. *Journal of Experimental Botany*, 66, 4317-4335. doi: 10.1093/jxb/erv269

Chandra, R., Ewanick, S., Hsieh, C., & Saddler, J. N. (2008). The characterization of pretreated lignocellulosic substrates prior to enzymatic hydrolysis, Part 1: A

modified Simons' staining technique. *Biotechnology Progress*, 24, 1178-1185. doi: 10.1002/btpr.33

Chen, F., & Dixon, R. A. (2007). Lignin modification improves fermentable sugar yields for biofuel production. *Nature Biotechnology*, *25*, 759. doi: 10.1038/nbt1316

Chen, L., Auh, C. K., Dowling, P., Bell, J., Chen, F., Hopkins, A., … Wang, Z. Y. (2003). Improved forage digestibility of tall fescue (*Festuca arundinacea*) by transgenic down-regulation of cinamyl alcohol dehydrogenase. *Plant Biotechnology Journal*, 1, 437-449. doi: 10.1046/j.1467-7652.2003.00040.x

Chen, W., Gong, L., Guo, Z., Wang, W., Zhang, H., Liu, X., … Luo J. (2013). A novel integrated method for large-scale detection, identification, and quantification of widely targeted metabolites: application in the study of rice metabolomics. *Molecular Plant*, 6, 1769-1780. doi: 10.1093/mp/sst080

Cheng, L., Wang, L., Wei, L., Wu, Y., Alam, A., Xu, C., … Xia, T. (2019). Combined mild chemical pretreatments for complete cadmium release and cellulosic ethanol co-production distinctive in wheat mutant straw. *Green Chemistry*, *21*, 3693-3700. doi: 10.1039/c9gc00686a

Deng, J., Zhu, X., Chen, P., He, B., Tang, S. W., Zhao, W., … Peng, L. (2020). Mechanism of lignocellulose modification and enzyme disadsorption for complete biomass saccharification to maximize bioethanol yield in rapeseed stalks. *Sustainable Energy & Fuels*. doi:10.1039/C9SE00906J.

Dische, Z. (1962). Color reactions of carbohydrates. In: RL Whistler, RL Wolfrom, eds. *Methods in Carbohydrate Chemistry*, Academic Press, New York. pp. 475-514.

Eudes, A., Sathitsuksanoh, N., Baidoo, E. E., George, A., Liang, Y., … Loque, D. (2015). Expression of a bacterial 3-dehydroshikimate dehydratase reduces lignin content and improves biomass saccharification efficiency. *Plant Biotechnology Journal*, 13, 1241-1250. doi: 10.1111/pbi.12310

Fan, C., Feng, S., Huang, J., Wang, Y., Wu, L., Li, X., … Peng, L. (2017). AtCesA8 - driven OsSUS3 expression leads to largely enhanced biomass saccharification and lodging resistance by distinctively altering lignocellulose features in rice.

Biotechnology for Biofuels, 10, 221. doi: 10.1186/s13068-017-0911-0

Fan, C., Li, Y., Hu, Z., Hu, H., Wang, G., Li, A., … Feng, S. (2018). Ectopic expression of a novel *OsExtensin-like* gene consistently enhances plant lodging resistance by regulating cell elongation and cell wall thickening in rice. *Plant Biotechnology Journal*, 16, 254-263. doi: 10.1111/pbi.12766

Fornale, S., Capellades, M., Encina, A., Wang, K., Irar, S., Lapierre, C., … Caparrós-Ruiz, D. (2012). Altered lignin biosynthesis improves cellulosic bioethanol production in transgenic maize plants down-regulated for cinnamyl alcohol dehydrogenase. *Molecular Plant*, 5, 817-830. doi: 10.1093/mp/ssr097

Fry, S. C. (1988). The growing plant cell wall: chemical and metabolic analysis. *Growing Plant Cell Wall Chemical & Metabolic Analysis*. doi:10.1016/0031-9422(88)83079-6

Fu, C., Mielenz, J. R., Xiao, X., Ge, Y., Hamilton, C. Y., Rodriguez, M. J., … Wang Z. Y. (2011a). Genetic manipulation of lignin reduces recalcitrance and improves ethanol production from switchgrass. *Proceedings of the National Academy of Sciences of the United States of America*, 108, 3803-3808. doi: 10.1073/pnas.1100310108

Fu, C., Xiao, X., Xi, Y., Ge, Y., Chen, F., Bouton, J., & Wang, Z. Y. (2011b). Downregulation of cinnamyl alcohol dehydrogenase (CAD) leads to improved saccharification efficiency in switchgrass. *Bioenergy Research*, 4, 153-164. doi: 10.1007/s12155-010-9109-z

Grethlein, H. E. (1985). The effect of pore size distribution on the rate of enzymatic hydrolysis of cellulosic substrates. *Nature Biotechnology*, 3, 155-160. doi: 10.1038/nbt0285-155

Haarmeyer, C. N., Smith, M. D., Chundawat, S., Sammond, D., & Whitehead, T. A. (2017). Insights into cellulase-lignin non-specific binding revealed by computational redesign of the surface of green fluorescent protein. *Biotechnology and Bioengineering*, 114(4), 740-750. doi: 10.1002/bit.26201

Hamelinck, C. N., Hooijdonk, G. V., & Faaij, A. P. C. (2005). Ethanol from lignocellulosic biomass: techno-economic performance in short-, middle- and

long-term. *Biomass & Bioenergy*, 28, 384-410. doi: 10.1016/j.biombioe.2004.09.002

Himmel, M. E., Ding, S., Johnson, D. K., Adney, W. S., Nimlos, M. R., Brady, J. W., & Foust, T. D. (2007). Biomass recalcitrance: engineering plants and enzymes for biofuels production. *Science*, 315, 804-807. doi: 10.1126/science.1137016

Hirano, K., Aya, K., Kondo, M., Okuno, A., Morinaka, Y., & Matsuoka, M. (2012). *OsCAD2* is the major *CAD* gene responsible for monolignol biosynthesis in rice culm. *Plant Cell Reports*, 31, 91-101. doi: 10.1007/s00299-011-1142-7

Huang, J., Xia, T., Li, G., Li, X., Li, Y., Wang, Y., … Wang, L. (2019). Overproduction of native endo-β-1,4-glucanases leads to largely enhanced biomass saccharification and bioethanol production by specific modification of cellulose features in transgenic rice. *Biotechnology for Biofuels*, 12, 11. doi: 10.1186/s13068-018-1351-1

Jackson, L. A., Shadle, G. L., Zhou, R., Nakashima, J., Chen, F., & Dixon, R. A. (2008). Improving saccharification efficiency of alfalfa stems through modification of the terminal stages of monolignol biosynthesis. *Bioenergy Research*, 1, 180-192. doi: 10.1007/s12155-008-9020-z

Jin, W., Chen, L., Hu, M., Sun, D, Li, A., Li, Y., … Peng, L. (2016). Tween-80 is effective for enhancing steam-exploded biomass enzymatic saccharification and ethanol production by specifically lessening cellulase absorption with lignin in common reed. *Applied Energy*, 175, 82-90. doi: 10.1016/j.apenergy.2016.04.104

Jung, J. H., Fouad, W. M., Vermerris, W., Gallo, M., & Altpeter, F. (2012). RNAi suppression of lignin biosynthesis in sugarcane reduces recalcitrance for biofuel production from lignocellulosic biomass. *Plant Biotechnology Journal*, 10, 1067-1076. doi: 10.1111/j.1467-7652.2012.00734.x

Katoh, K., Kuma, K., Toh, H., & Miyata, T. (2005). MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Research*, 33, 511-518. doi: 10.1093/nar/gki198

Kim, H., Ralph, J., Lu, F., Pilate, G., Leple, J. C., Pollet, B., & Lapierre, C. (2002).

Identification of the structure and origin of thioacidolysis marker compounds for cinnamyl alcohol dehydrogenase deficiency in angiosperms. *Journal of Biological Chemistry*, 277, 47412-47419. doi: 10.1074/jbc.M208860200

Kumar, S., Tamura, K., & Nei, M. (2004). MEGA3: Integrated software for molecular evolutionary genetics analysis and sequence alignment. *Briefings in Bioinformatics*, 5, 150-163. doi: 10.1093/bib/5.2.150

Lam, P. Y., Tobimatsu, Y., Takeda, Y., Suzuki, S., Yamamura, M., … Lo, C. (2017). Disrupting flavone synthase II alters lignin and improves biomass digestibility. *Plant Physiology*, 174, 972-985. doi: 10.1104/pp.16.01973

Li, F., Xie, G., Huang, J., Zhang, R., Li, Y., Zhang, M., … Peng, L. (2017). OsCESA9 conserved-site mutation leads to largely enhanced plant lodging resistance and biomass enzymatic saccharification by reducing cellulose DP and crystallinity in rice. *Plant Biotechnology Journal*, 15, 1093-1104. doi: 10.1111/pbi.12700

Li, M., Si, S., Hao, B., Zha, Y., Wan, C., Hong, S., … Peng, L. (2014). Mild alkali-pretreatment effectively extracts guaiacyl-rich lignin for high lignocellulose digestibility coupled with largely diminishing yeast fermentation inhibitors in *Miscanthus*. *Bioresource Technology*, 169, 447-454. doi: 10.1016/j.biortech.2014.07.017

Li, X., Yang, Y., Yao, J., Chen, G., Li, X., Zhang, Q., & Wu, C. (2009). Flexible culm 1 encoding a cinnamyl-alcohol dehydrogenase controls culm mechanical strength in rice. *Plant Molecular Biology*, 69, 685-697. doi: 10.1007/s11103-008-9448-8

Oyarce, P., De Meester, B., Fonseca, F., de Vries, L., Goeminne, G., Pallidis, A., & Boerjan, W. (2019). Introducing curcumin biosynthesis in *Arabidopsis* enhances lignocellulosic biomass processing. *Nature Plants*, 5, 225-237. doi: 10.1038/s41477-018-0350-3

Park, H. L., Kim, T. L., Bhoo, S. H., Lee, T. H., Lee, S. W., & Cho, M. H. (2018). Biochemical characterization of the rice cinnamyl alcohol dehydrogenase Gene Family. *Molecules, 23*(10). doi:10.3390/molecules23102659

Park, J. J., Yoo, C. G., Flanagan, A., Pu, Y., Debnath, S., Ge, Y., … Wang, Z. Y. (2017).

Defined tetra-allelic gene disruption of the 4-coumarate: coenzyme A ligase 1 (*Pv4CL1*) gene by CRISPR/Cas9 in switchgrass results in lignin reduction and improved sugar release. *Biotechnology for Biofuels*, 10, 284. doi: 10.1186/s13068-017-0972-0

Pei, Y., Li, Y., Zhang, Y., Yu, C., Fu, T., Zou, J., … Chen, P. (2016). G-lignin and hemicellulosic monosaccharides distinctively affect biomass digestibility in rapeseed. *Bioresource Technology*, 203, 325-333. doi: 10.1016/j.biortech.2015.12.072

Pelekani, C., & Snoeyink, V. L. (2001). A kinetic and equilibrium study of competitive adsorption between atrazine and Congo red dye on activated carbon: the importance of pore size distribution. *Carbon*, 66, 97-103. doi: 10.1016/S0008-6223(00)00078-6

Peng, L., Hocart, C. H., & Williamson, R. R. E. (2000). Fractionation of carbohydrates in *Arabidopsis* root cell walls shows that three radial swelling loci are specifically involved in cellulose production. *Planta*, 211, 406-414. doi: 10.1007/s004250000301

Raes, J., Rohde, A., Christensen, J. H., Van de Peer, Y., & Boerjan, W. (2003). Genome-wide characterization of the lignification toolbox in *Arabidopsis*. *Plant Physiol*, 133, 1051-1071. doi: 10.1104/pp.103.026484

Ralph, J., Lundquist, K., Brunow, G., Lu, F., Kim, H., Schatz, P. F., … Boerjan, W. (2004). Natural polymers from oxidative coupling of 4-hydroxyphenyl-propanoids. *Phytochemistry Reviews*, 3, 29-60. doi: 10.1023/b:phyt.0000047809.65444.a4

Saathoff, A. J., Sarath, G., Chow, E. K., Dien, B. S., & Tobias, C. M. (2011). Downregulation of cinnamyl-alcohol dehydrogenase in switchgrass by RNA silencing results in enhanced glucose release after cellulase treatment. *Plos One*, 6, e16416. doi: 10.1371/journal.pone.0016416

Saballos, A., Ejeta, G., Sanchez, E., Kang, C., & Vermerris, W. (2009). A genome wide analysis of the cinnamyl alcohol dehydrogenase family in sorghum [Sorghum bicolor (L.) Moench] identifies *SbCAD2* as the brown midrib 6 gene. *Genetics*,

181, 783-795. doi: org/10.1534/genetics.108.098996

Schultz, T.P., & Templeton, M.C. (1986). Proposed mechanism for the nitrobenzene oxidation of lignin. *Holzforschung-International Journal of the Biology, Chemistry, Physics and Technology of Wood*, 40, 93-97. doi: 10.1515/hfsg.1986.40.2.93

Sedmak, J. J., & Grossberg, S. E. (1977). A rapid, sensitive, and versatile assay for protein using Coomassie brilliant blue G250. *Analytical Biochemistry*, 79, 544-552. doi: 10.1016/0003-2697(77)90428-6

Sluiter, A., Hames, B., Ruiz, R., Scarlata, C., Sluiter, J., & Templeton, D. (2008). Determination of structural carbohydrates and lignin in biomass. *Technical Report*, NREL/TP-510-42618, National Renewable Energy Laboratory, Golden, Co.

Sibout, R., Eudes, A., Mouille, G., Pollet, B., Lapierre, C., Jouanin, L., & Seguin, A. (2005). Cinnamyl alcohol dehydrogenase-C and -D are the primary genes involved in lignin biosynthesis in the floral stem of *Arabidopsis*. *Plant Cell, 17*(7), 2059-2076. doi:10.1105/tpc.105.030767

Sun, D., Alam, A., Tu, Y., Zhou, S., Wang, Y., Xia, T., … Peng, L. (2017). Steam-exploded biomass saccharification is predominately affected by lignocellulose porosity and largely enhanced by Tween-80 in *Miscanthus*. *Bioresource Technology*, 239, 74-81. doi: 10.1016/j.biortech.2017.04.114

Sun, D., Yang, Q., Wang, Y., Gao, H., He, M., Lin, X., Lu, J., … Peng, L. (2020). Distinct mechanisms of enzymatic saccharification and bioethanol conversion enhancement by three surfactants under steam explosion and mild chemical pretreatments in bioenergy *Miscanthus*. *Industrial Crops and Products*, 153, 112559. doi.org/10.1016/j.indcrop.2020.112559

Takeda, Y., Tobimatsu, Y., Karlen, S. D., Koshiba, T., Suzuki, S., Yamamura, M., & Umezawa, T. (2018). Downregulation of p-coumaroyl ester 3-hydroxylase in rice leads to altered cell wall structures and improves biomass saccharification. *Plant Journal for Cell & Molecular Biology*. doi: 10.1111/tpj.13988

Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F., & Higgins, D. G. (1997).

The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. *Nucleic Acids Research*. doi: 10.1093/nar/25.24.4876

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

Wang, L., Guo, K., Li, Y., Tu, Y., Hu, H., Wang, B., … Peng, L. (2010). Expression profiling and integrative analysis of the CESA / CSL superfamily in rice. *BMC Plant Biology*, 10, 282. doi: 10.1186/1471-2229-10-282

Wang, Y., Fan, C., Hu, H., Li, Y., Sun, D., Wang, Y., & Peng, L. (2016). Genetic modification of plant cell walls to enhance biomass yield and biofuel production in bioenergy crops. *Biotechnology Advance*, 34, 997-1017. doi: 10.1016/j.biotechadv.2016.06.00

Wiman, M., Dienes, D., Hansen, M.A., van der Meulen, T., Zacchi, G., & Liden, G. (2012). Cellulose accessibility determines the rate of enzymatic hydrolysis of steam-pretreated spruce. *Bioresource Technology*, 126, 208-215. doi: 10.1016/j.biortech.2012.08.082

Wu, Z., Wang, N., Hisano, H., Cao, Y., Wu, F., Liu, W., … Fu, C. (2019). Simultaneous regulation of *F5H* in COMT-RNAi transgenic switchgrass alters effects of *COMT* suppression on syringyl lignin biosynthesis. *Plant Biotechnology Journal*, 17, 836-845. doi: 10.1111/pbi.13019

Xie, K., Minkenberg, B., & Yang, Y. (2015). Boosting CRISPR/Cas9 multiplex editing capability with the endogenous tRNA-processing system. *Proceedings of the National Academy of Sciences of the United States of America*, 112, 3570-3575. doi: 10.1073/pnas.1420294112

Xiong, W., Wu, Z., Liu, Y., Li, Y., Su, K., Bai, Z., & Fu, C. (2019). Mutation of 4-coumarate: coenzyme A ligase 1 gene affects lignin biosynthesis and increases the cell wall digestibility in maize brown midrib5 mutants. *Biotechnology for Biofuels*, 12, 82. doi: 10.1186/s13068-019-1421-z

Xu, N., Zhang, W., Ren, S., Liu, F., Zhao, C., Liao, H., … Peng, L.

(2012). Hemicelluloses negatively affect lignocellulose crystallinity for high biomass digestibility under NaOH and H_2SO_4 pretreatments in *Miscanthus*. *Biotechnology for Biofuels*, 5, 58-70. doi: 10.1186/1754-6834-5-58

Yang, H., Zhang, X., Luo, H., Liu, B., Shiga, T. M., Li, X., & Meilan, R. (2019). Overcoming cellulose recalcitrance in woody biomass for the lignin-first biorefinery. *Biotechnology for Biofuels, 12*, 171. doi: 10.1186/s13068-019-1503-y

Youn, B., Camacho, R., Moinuddin, S. G., Lee, C., Davin, L. B., Lewis, N. G., & Kang, C. (2006). Crystal structures and catalytic mechanism of the *Arabidopsis* cinnamyl alcohol dehydrogenases AtCAD5 and AtCAD4. *Org Biomol Chem, 4*(9), 1687-1697. doi:10.1039/b601672c

Zeng, Y., Zhao, S., Yang, S., & Ding, S. (2014). Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. *Current Opinion in Biotechnology*, 27, 38-45. doi: 10.1016/j.copbio.2013.09.008

Zhang, K., Qian, Q., Huang, Z., Wang, Y., Li, M., Hong, L., … Cheng, Z. (2006). Gold hull and internode 2 encodes a primarily multifunctional cinnamyl-alcohol dehydrogenase in rice. *Plant Physiology*, 140, 72-83. doi: 10.1104/pp.105.073007

Zhou, X., Jacobs, T. B., Xue, J., Harding, S. A., & Tsai, C. J. (2015). Exploiting SNPs for biallelic CRISPR mutations in the outcrossing woody perennial Populus reveals 4-coumarate: CoA ligase specificity and redundancy. *New Phytologist*, 208, 298-301. doi: 10.1111/nph.13470

Figure Legends

Figure 1. Structural analysis and co-expression profiling of OsCAD family in rice.

(A) Unrooted tree analysis of the OsCAD protein family. (B) Comparison of exons and introns of the *OsCAD* genes and motif constitution of the OsCAD protein family. (C) Genes co-expression profiling involved in lignin biosynthesis in two rice varieties (Z, Zhenshan 97 and M, Minghui 63) throughout the life cycles. The color scale represents the relative signal values: green refers to low expression; black refers to medium expression and red refers to high expression).

Motif constitution in (B): first blue color-codes : LSPFS FSRRE TGEDD VTIKV LYCGI CHTD LHTIK NEWGN SMYPL VPGHE I； gray color-codes: SSDAE ZMKAA AGTMD GIIDT VSATH PLAPL LSLLK PNGKL VLVGA PEKPL； yellow color-codes: VDYDG TVTQG GYSDH IVVHZ RFVVR IPDGL PLDSA APLLC A；red color-codes: PGKHL GVVGL GGLGH VAVKF AKAFG LKVTV ISTSP GKRE EA； second blue color-codes: GVVTE VGANV TKFKA GDRVG VGCIV GSCRS CESCN QGLEN YCPKM VF.

The X-axis indicates the tissues at the developmental stages in (C): 1, Calli (15 days after subculture); 2, Calli (5 days after regeneration); 3, Calli (Screening stage); 4, Calli (15 days after induction T2); 5, Calli (15 days after induction T3); 6, Seed imbibition; 7, Seed germination; 8, Plumule (48 h after emergence, Dark); 9, Plumule (48 h after emergence, Light); 10, Radicle (48 h after emergence, Dark); 11, Radicle(48 h after emergence, Light); 12, Seedling; 13, Young shoot; 14, Young root; 15, Mature leaf; 16, Old leaf; 17, Mature sheath; 18, Old sheath; 19, Young flag leaf; 20, Old flag leaf; 21, Young panicle stages 3 (secondary branch primordium differentiation stage); 22, Young panicle stages 4 (pistil/stamen primordium differentiation stage); 23, Young panicle stages 5 (pollen-mother cell formation stage); 24, Young panicle; 25, Old panicle; 26, Young stem; 27, Old stem; 28, Hull; 29, Spikelet; 30, Stamen; 31, Endosperm (7 days after pollination); 32, Endosperm (14 days after pollination); 33, Endosperm (21 days after pollination). The Y-axis represents the genes relative expression levels obtained from microarray analysis.

Figure 2. *OsCAD2* gene models with target sequences.

(A) Construct for editing of target genomic sites in rice. (B) PCR detection of target sequences in three independent lines. The *OsCAD2* gene model includes four exons separated by three introns (represented by the solid line) and Red mark as mutation site.

Figure 3. Phenotype observation of *OsCAD2* modified lines.

(A) Plant images of *OsCAD2* modified lines and WT at heading stage (Scale bar as 15 cm). (B) Contents of three major cell wall polymers (% dry matter) of mature rice straws. Student's *t*-test between the modified line and WT at *$p < 0.05$ and **$p < 0.01$; increased or decreased percentage obtained by subtraction between the results of transgenic line and WT divided by the result of WT.

Figure 4. Biomass enzymatic saccharification in *OsCAD2* modified lines compared with wild type.

(A, C, E) Hexoses yields (% cellulose) released from enzymatic hydrolysis after 1.0% H_2SO_4 pretreatment or 1.0 % NaOH pretreatment or 1.0 % NaOH pretreatment plus 1% Tween-80. respectively. (B, D, F) Hexoses yields (% dry matter/DM) released from enzymatic hydrolysis with 1.0% H_2SO_4 pretreatment or 1.0 % NaOH pretreatment or 1.0 % NaOH pretreatment plus 1.0 % Tween-80. Student's *t*-test between the modified line and WT at *$p < 0.05$ and **$p < 0.01$ increased or decreased percentage obtained by subtraction between the results of transgenic line and WT divided by the result of WT.

Figure 5. Lignin and hemicellulose extraction from 1%NaOH pretreatment in *OsCAD2* modified lines and wild type. (A, B) Lignin and hemicellulose levels in raw materials and pretreated biomass residues, respectively; (C) Correlation analysis between the lignin levels and hexoses yields released from enzymatic hydrolyses of total 24 raw materials and pretreated residues (n=24). The percentages as wall polymer extraction rates of the pretreated biomass residues against the raw materials. ** As significant correlation at $p < 0.01$ level (n=24).

Figure 6. Characterization of biomass porosity and cellulose accessibility in *OsCAD2* modified lines and wild type. (A, B, C) Biomass porosity of raw materials detected by Simons' stain; Direct Yellow (DY) and Direct Blue (DB) as large pores and small pores, respectively. (D, E, F) Biomass

porosity of 1% NaOH pretreated residues. (G) Correlation analysis between wall polymers levels (% dry matter/DM) and total dye values (DY+DB); * and ** As significant correlations at $p < 0.05$ and 0.01 level (n=12). (H) Cellulose surface area accountable for accessibility by Congo Red stain in 1% NaOH pretreated residues. (I) Glucose yield released from CBHI hydrolysis of the crude cellulose substrates. Student's *t*-test between the modified line and WT at $*p < 0.05$ and $**p < 0.01$. The increased percentage calculated by subtraction between the values of transgenic line and WT divided by the value of WT.

Figure 7. Characterization of the mixed-cellulases added into the biomass enzymatic hydrolysis of *OsCAD2* modified lines and wild type after 1% NaOH pretreatment plus 1% Tween-80.

(A) Total soluble enzyme levels (μg/mL) of the supernatant collected from enzymatic hydrolysis; Student's *t*-test between the modified line and WT at $**p < 0.01$; increased percentage obtained by subtraction between the results of transgenic line and WT divided by the result of WT. (B) SDS-PAGE profiling of the soluble enzymes; EG: endo-β-1,4-glucanase; CBH: cellobiohydrolase; XYNI: endo-1,4-β-xylanase 1. (C) Correlation analysis between the soluble enzymes levels and hexoses yields (% DM) released from enzymatic hydrolyses of lignocellulose substrates examined. (D) Correlation analysis between the soluble enzyme levels and lignin contents in the 1.0 % NaOH pretreated biomass residues. ** As significant correlation at $p < 0.01$ level (n=12).

Table captions

Table 1. Lignin composition of three rice transgenic lines relative to wide type (Nipponbare)

* and ** indicating the significant differences between transgenic lines and WT by *t*-test at $P < 0.05$ and 0.01, respectively. H, S and G represented *p*-hydroxyphenyl, syringyl, and guaiacyl.

Table 2. Comparison of biomass saccharification between the *CAD2* rice transgenic lines and previously-reported engineered plants

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version at doi: http://

Supplementary materials

Figure S1. Pentoses yields released from enzymatic hydrolysis (% cellulose) of mature straws in *OsCAD2* modified lines and wild type under different pretreatments (1% H_2SO_4, 1% NaOH, 1% NaOH plus 1% Tween-80). Student's *t*-test between the modified line and WT at $*p < 0.05$ and $**p < 0.01$ with the increased percentage and data as means ± SD (n=3).

Figure S2. A hypothetic model interpreting how the *OsCAD2* site-modification may lead to a dynamical regulation on lignin biosynthetic pathways and its derivate metabolism.

Decreased compounds marked by dashed boxes with arrows in red, increased compounds by dashed boxes with arrows in green. Solid arrows represent abbreviated biosynthetic steps. PTAL, Phe and Tyr ammonia lyase; TAL, Tyr ammonia lyase; PAL, Phe ammonia lyase; C4H, cinnamate 4-hydroxylase; 4CL, 4-coumarate CoA ligase; HCT, p-hydroxycinnamoyl-CoA:quinate/ shikimate p-hydroxycinnamoyl transferase; C3H, p-coumaroyl ester 3-hydroxylase; CSE, caffeoyl shikimate esterase; CCoAOMT, caffeoyl-CoA O-methyltransferase; COMT, 5-hydroxyconiferaldehyde O-methyltransferase; CCR, cinnamoyl-CoA reductase; CAD, cinnamyl alcohol dehydrogenase; PMT, p-coumaroyl-CoA: monolignol transferase; LAC, laccase; PRX, peroxidase. S, G and H: syringyl, guaiacyl and p-hydroxyphenyl units of lignin. CHS, chalcone synthase; CHI, chalcone isomerase; FNSII, flavone synthase II; F3'H, flavonoid 3'-hydroxylase; OMT, O-methyltransferases; FOMT, flavonoid O-methyltransferase; FLS, flavonol synthase; ANS, anthocyanidin synthase; RT, rhamnosyl transferase UFGT; UDP Flavonoid glucosyl transferase; F3'5'H, flavonoid 3', 5'-hydroxylase; DFR; dihydroflavonol 4-reductase.

Figure S3. Scanning electron microscopic observations of raw materials and pretreated residues in *OsCAD2* modified lines and wild type. Arrows marked rough sites of lignocellulose surfaces.

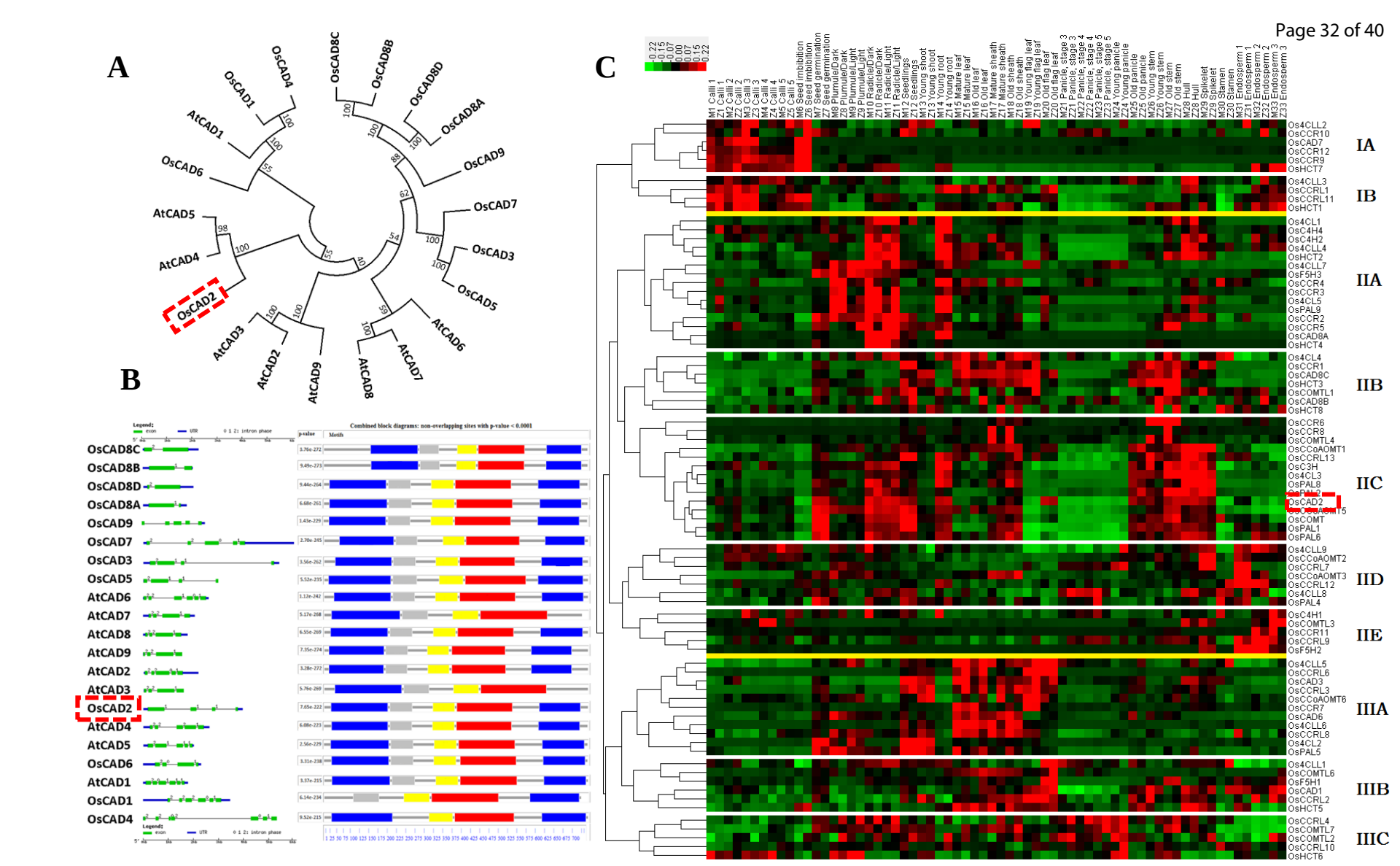
Table S1. The gene expression profile data in 33 rice tissue samples.

Table S2. List of PCR primers used in this study.

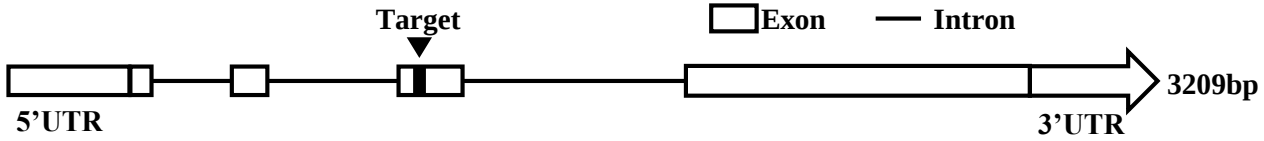
Table S3. List of the 12 *OsCAD* genes identified in rice.

Table S4. Major agronomic traits of *OsCAD2* transgenic rice lines and wild type.

Table S5. Monosaccharide composition (% of total) of hemicellulose in *OsCAD2* transgenic rice lines and wild type.


A
OsCAD4
OsCAD1
AtCAD1
OsCAD6
AtCAD5
AtCAD4
OsCAD2
AtCAD3
AtCAD2
AtCAD9
AtCAD8
AtCAD7
AtCAD6
OsCAD7
OsCAD3
OsCAD5
OsCAD9
OsCAD8A
OsCAD8D
OsCAD8B
OsCAD8C
B
Combined block diagrams: non-overlapping sites with p-value < 0.0001
OsCAD8C
OsCAD8B
OsCAD8D
OsCAD8A
OsCAD9
OsCAD7
OsCAD3
OsCAD5
AtCAD6
AtCAD7
AtCAD8
AtCAD9
AtCAD2
AtCAD3
OsCAD2
AtCAD4
AtCAD5
OsCAD6
AtCAD1
OsCAD1
OsCAD4
C
IA
IB
IIA
IIB
IIC
IID
IIE
IIIA
IIIB
IIIC

A

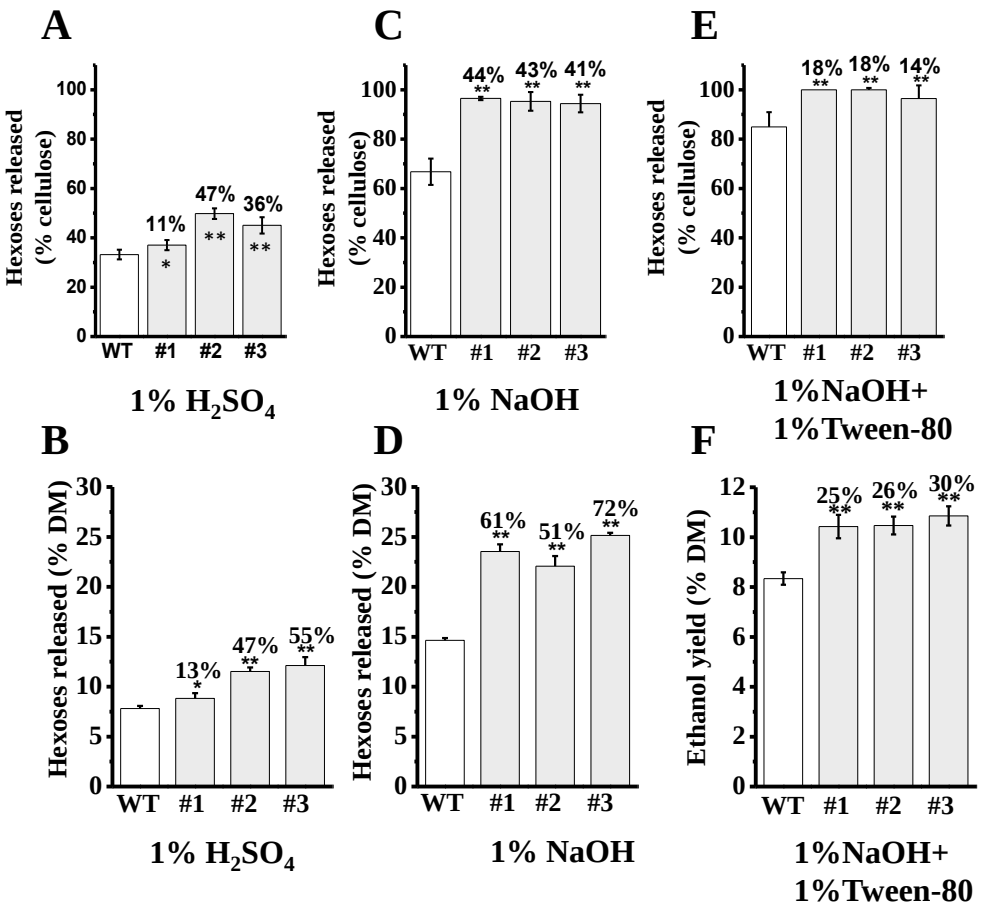


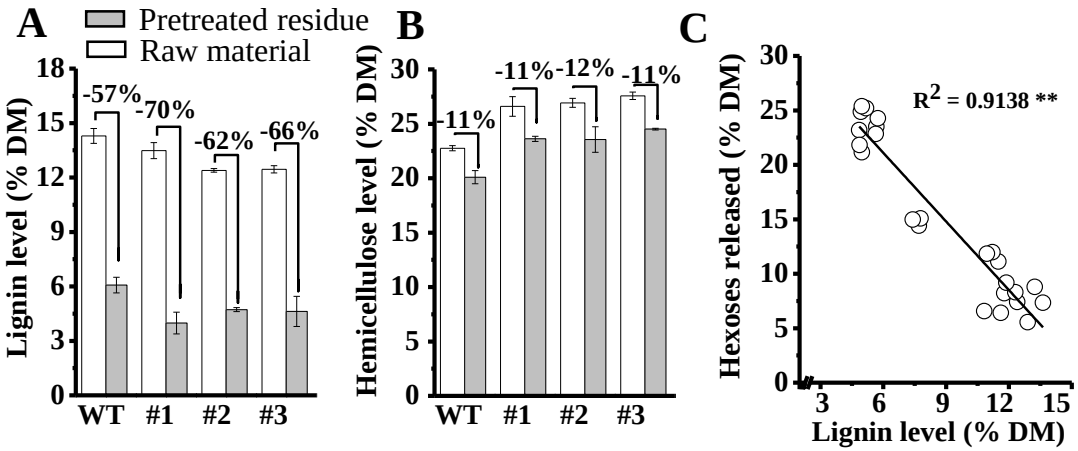
B

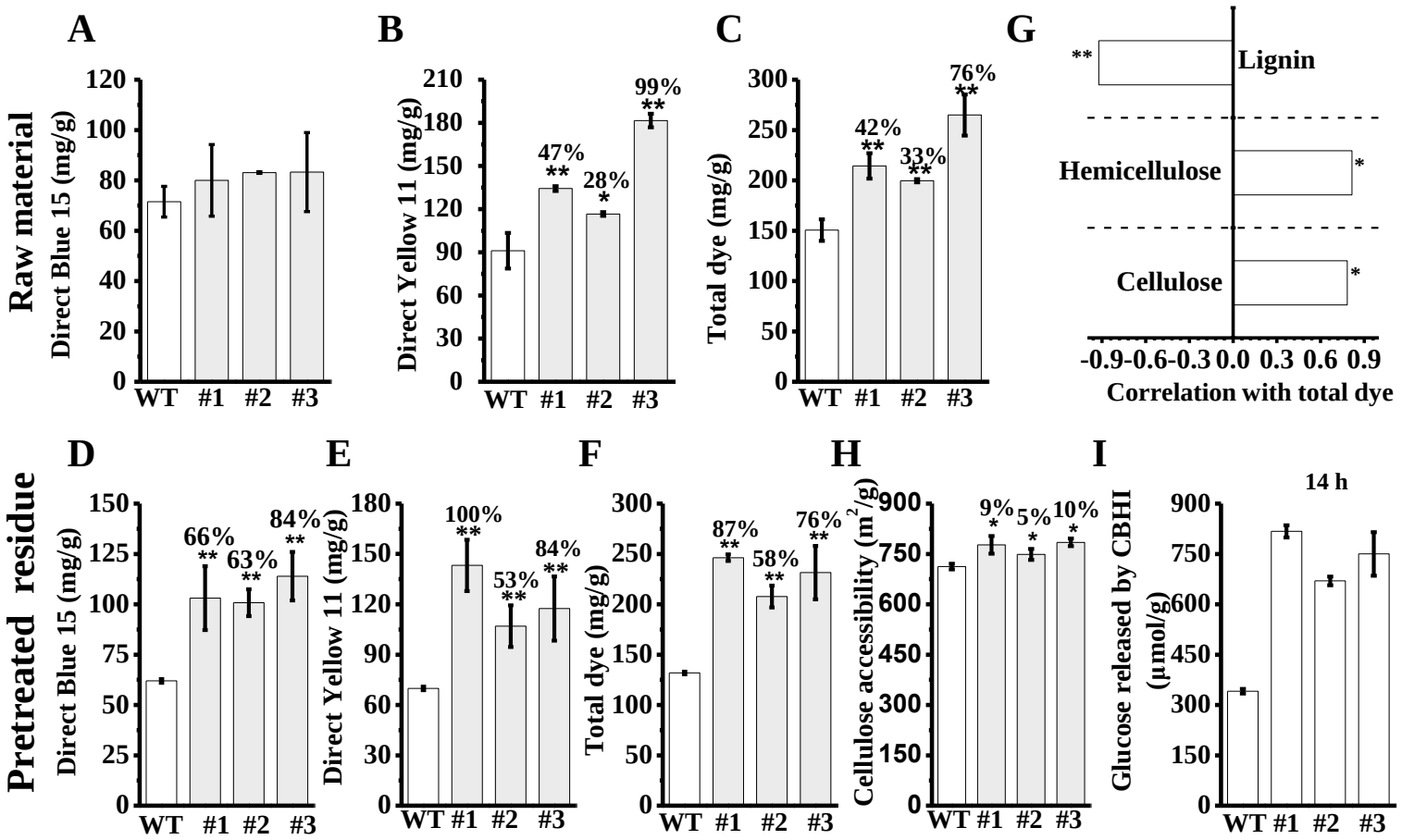
Target WT: **GCCAATGTGGAGCAGTACTGCAA**

Transgenic lines	**#1:**	**GCC**AATGGTGGAGCAGTACTGCAA	+1bp
	#2:	**GCC**AATGATGGAGCAGTACTGCAA	+1bp
	#3:	**GCC**AATGTTGGAGCAGTACTGCAA	+1bp


A
WT
#1
#2
#3
15cm
B
Level (% dry matter)
30
25
20
15
10
5
0
10%
*
10%
*
12%
**
17%
*
18%
**
21%
**
9%
18%
*
15%
*
WT #1 #2 #3
WT #1 #2 #3
WT #1 #2 #3
Cellulose
Hemicellulose
Lignin


A
Hexoses released (% cellulose)
11%
47%
36%
WT #1 #2 #3
1% H2SO4
C
44%
43%
41%
1% NaOH
E
18%
18%
14%
1%NaOH+ 1%Tween-80
B
Hexoses released (% DM)
13%
47%
55%
D
61%
51%
72%
F
Ethanol yield (% DM)
25%
26%
30%


A
Pretreated residue
Raw material
Lignin level (% DM)
18
15
12
9
6
3
0
-57%
-70%
-62%
-66%
WT
#1
#2
#3
B
Hemicellulose level (% DM)
30
25
20
15
10
5
0
-11%
-11%
-12%
-11%
WT
#1
#2
#3
C
Hexoses released (% DM)
30
25
20
15
10
5
0
R2 = 0.9138 **
3
6
9
12
15
Lignin level (% DM)


A
B
C
G
Raw material
Direct Blue 15 (mg/g)
Direct Yellow 11 (mg/g)
Total dye (mg/g)
WT #1 #2 #3
47% 28% 99%
42% 33% 76%
Lignin
Hemicellulose
Cellulose
-0.9 -0.6 -0.3 0.0 0.3 0.6 0.9
Correlation with total dye
Pretreated residue
D
E
F
H
I
66% 63% 84%
100% 53% 84%
87% 58% 76%
9% 5% 10%
Cellulose accessibility (m2/g)
Glucose released by CBHI (μmol/g)
14 h

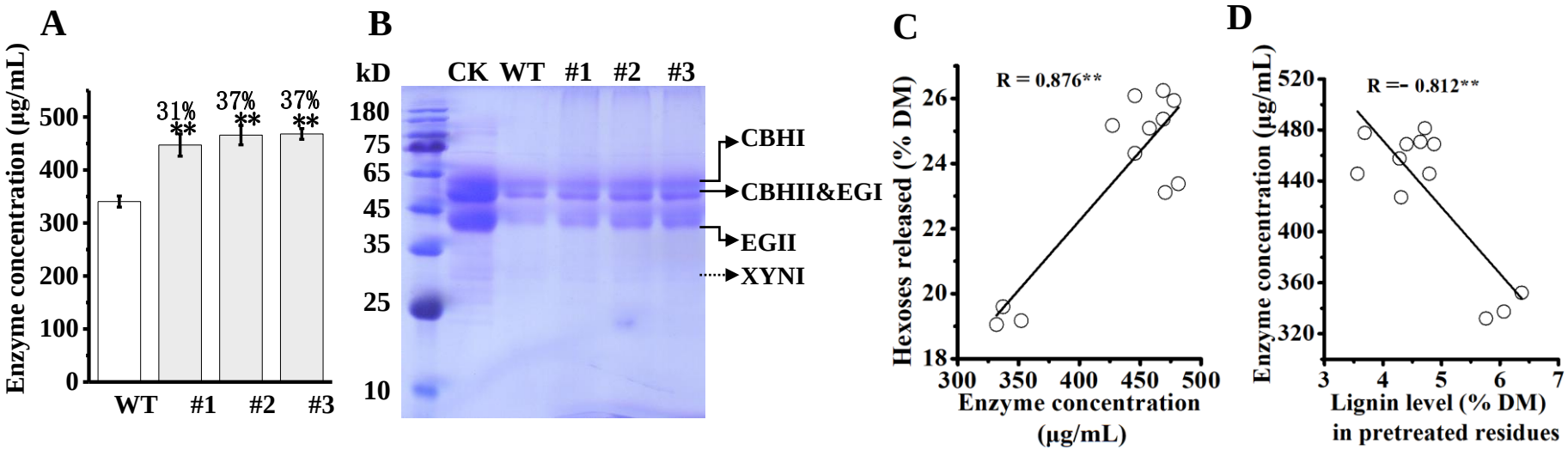

A
Enzyme concentration (μg/mL)
0
100
200
300
400
500
31%
37%
37%
**
**
**
WT
#1
#2
#3
B
kD
CK
WT
#1
#2
#3
180
75
65
45
35
25
10
CBHI
CBHII&EGI
EGII
XYNI
C
Hexoses released (% DM)
18
20
22
24
26
R = 0.876**
300
350
400
450
500
Enzyme concentration
(μg/mL)
D
Enzyme concentration (μg/mL)
320
360
400
440
480
520
R =- 0.812**
3
4
5
6
7
Lignin level (% DM)
in pretreated residues

Table 1. Lignin composition of three rice transgenic lines relative to wide type (Nipponbare)

	WT	#1	#2	#3	Relative to WT
Cellulose	24.4±0.3	26.6±0.9*	26.7±0.6*	27.2±0.3**	+10%
Hemicellulose	22.8±0.2	26.6±0.9*	26.9±0.4**	27.6±0.3**	+19%
Lignin	14.6±0.6	13.3±0.4	12.0±0.6*	12.5±0.2*	-14%
H-monomer of total (%)	21.0±4.2	13.3±2.0*	13.1±1.1*	15.4±3.6	-33.7%
S-monomer of total (%)	32.9±2.7	33.6±1.5	33.7±2.5	31.2±1.9	/
G-monomer of total (%)	46.1±2.0	53.1±0.5**	53.3±1.4**	53.4±1.6**	+15.5%
H/S	0.7	0.4	0.4	0.4	-1.7 folds
S/G	0.7	0.6	0.6	0.6	-1.2 folds
H/G	0.5	0.3	0.2	0.3	-1.8 folds

* and ** indicating the significant differences between transgenic lines and WT by *t*-test at $P < 0.05$ and 0.01, respectively. H, S and G represented *p*-hydroxyphenyl, syringyl, and guaiacyl.

Table 2. Comparison of biomass saccharification between the *CAD2* rice transgenic lines and previously-reported engineered plants

Samples	Genetic engineering	Altered lignin composition	Pretreatment	Hexoses yield increased	Reference
Rice	***OsCAD2*-CRISPR/Cas9**	**H reduced; G increased**	**1% NaOH, 50℃, 2h**	**61~72%**	**This study**
Arabidopsis	pCesA4:DCS_CURS2 expression	Curcumin introduced	0.25% NaOH, 48h	14.3~24.2%,	Oyarce et al, 2019
Maize	*4CL* mutant	G reduced; H increased	1.5% H_2SO_4, 121 ℃ 40min	17.6%	Xiong et al, 2019
Switchgrass	RNAi -*COMT*	S, G reduced	0.5 % H_2SO_4 18h	16.5~21.5%	Fu et al, 2011a
Poplar	*AtF5H1* expression	S increased	TFA (1:20, W/V) -20 ℃ 15h	< 35%	Yang et al, 2019
Switchgrass	RNAi -*CAD*	S, G reduced	1.5% H_2SO_4, 121 ℃ 40 min	19~43%	Fu et al, 2011b
Sugarcane	RNAi -*COMT*	S reduced	1.3% H_2SO_4, 121 ℃ 40 min	20~34%	Jung et al, 2012
Arabidopsis	*QsuB* expression	H, S increased; G reduced	1.2% H_2SO_4, 120 ℃ 1h	26~37%	Eudes et al, 2015
Switchgrass	*PvMYB4* overexpressing	S/G decreased	Condensing steam at 180 ℃ for 17.5 min	28.9 %	Eudes et al, 2015
Rice	*OsFNSII*-knockout mutant	S/G decreased	Without pretreatment	25~40%	Lam et al, 2017
Rice	*OsC3H*-CRISPR/Cas9	H, *p*-coumarate &Tricin increased	Without pretreatment	24~46%	Takeda et al, 2018