Switchgrass (*Panicum virgatum*) possesses a divergent family of cinnamoyl CoA reductases with distinct biochemical properties

Luis L. Escamilla-Treviño¹,², Hui Shen¹,², Srinivasa Rao Uppalapati¹, Tui Ray¹, Yuhong Tang¹,², Timothy Hernandez¹,², Yanbin Yin²,³, Ying Xu²,³ and Richard A. Dixon¹,²

¹Plant Biology Division, Samuel Roberts Noble Foundation, 2510 Sam Noble Parkway, Ardmore, OK 73401, USA; ²US Department of Energy BioEnergy Science Center (BESC); ³Department of Biochemistry and Molecular Biology and Institute of Bioinformatics, University of Georgia, 120 Green Street, Athens, GA 30602, USA

Author for correspondence: Richard A. Dixon
Tel: +1 580 224 6601
Email: radixon@noble.org

Received: 12 June 2009
Accepted: 27 July 2009


Key words: allelic variation, bioenergy, bioethanol, cinnamoyl CoA reductase, gene family, switchgrass (*Panicum virgatum*).

Summary

- The down-regulation of enzymes of the monolignol pathway results in reduced recalcitrance of biomass for lignocellulosic ethanol production. Cinnamoyl CoA reductase (CCR) catalyzes the first step of the phenylpropanoid pathway specifically dedicated to monolignol biosynthesis. However, plants contain multiple CCR-like genes, complicating the selection of lignin-specific targets. This study was undertaken to understand the complexity of the CCR gene family in tetraploid switchgrass (*Panicum virgatum*) and to determine the biochemical properties of the encoded proteins.

- Four switchgrass cDNAs (most with multiple variants) encoding putative CCRs were identified by phylogenetic analysis, heterologously expressed in *Escherichia coli*, and the corresponding enzymes were characterized biochemically.

- Two cDNAs, *PvCCR1* and *PvCCR2*, encoded enzymes with CCR activity. They are phylogenetically distinct, differentially expressed, and the corresponding enzymes exhibited different biochemical properties with regard to substrate preference. *PvCCR1* has higher specific activity and prefers feruloyl CoA as substrate, whereas *PvCCR2* prefers caffeoyl and 4-coumaroyl CoAs. Allelic variants of each cDNA were detected, but the two most diverse variants of *PvCCR1* encoded enzymes with similar catalytic activity.

- Based on its properties and expression pattern, *PvCCR1* is probably associated with lignin biosynthesis during plant development (and is therefore a target for the engineering of improved biomass), whereas *PvCCR2* may function in defense.

Introduction

Globally, ethanol is the most widely produced liquid biofuel. Currently, more than 10 billion gallons of ethanol are produced per year from starch [from maize (*Zea mays*)] or sugar [from sugarcane (*Saccharum* species) and sugar beet (*Beta vulgaris*)] through mature industrial procedures for the hydrolysis of starch and fermentation of sugar (Goldemberg, 2007; Rass-Hansen et al., 2007). Maize grain and other cereals, such as sorghum (*Sorghum bicolor*), are the primary feedstocks for US ethanol production. However, the use of these as feedstocks for bioethanol has been criticized as diverting material away from the human food chain, with resulting food shortages and price rises (Runge & Senauer, 2007).

An alternative to starch or sugar ethanol (classified as first-generation biofuels) is ethanol produced from lignocellulosic biomass via saccharification and fermentation (a second-generation biofuel) (Smeets & Faaij, 2007; Schmer et al., 2008; Yuan et al., 2008). Dedicated perennial energy crops, such as switchgrass (*Panicum virgatum*), crop residues and forestry biomass are major cellulosic ethanol sources that could potentially displace 30% of our current petroleum consumption (Schmer et al., 2008).

Switchgrass has been proposed as a major perennial bioenergy feedstock in the USA because it is widely...
adapted, has high biomass production, high C-4 photosynthetic efficiency and efficient use of water and nitrogen (Yuan et al., 2008). Switchgrass grown and managed as a bioenergy crop produces five times more renewable energy than the energy consumed in its production as a source of bioethanol, and has significant environmental benefits related to soil conservation and low net greenhouse gas emissions (Schmer et al., 2008). Although higher energy ratios would be obtainable from switchgrass through direct combustion than through saccharification and fermentation to ethanol (Powlson et al., 2005), most research efforts in the USA are directed towards the improvement of the efficiency of switchgrass processing for the generation of liquid transportation fuels.

The most critical biological obstacle to be overcome in the transition from starch-based to lignocellulosic biofuels is the recalcitrance of cell walls to chemical and/or enzymatic breakdown. The processing of lignocellulosic biomass requires pretreatment (usually with hot acid), saccharification (hydrolysis) and fermentation (Ragasauskas et al., 2006). Reducing the recalcitrance of lignocellulosic biomass could eliminate or reduce pretreatment, reduce cellulase loading and, ultimately, facilitate consolidated bio-processing. These improvements would reduce production costs and increase the economic competitiveness of lignocellulosic bioethanol (Lynd et al., 2008).

Lignin is probably the major factor affecting the recalcitrance of lignocelluloses (Yuan et al., 2008). It is an aromatic polymer that fills the spaces in the secondary cell wall between the cellulose, hemicellulose and pectin components, especially in tracheids, sclereids and xylem. It is covalently linked to hemicellulose and thereby cross-links plant polysaccharides, conferring mechanical strength and hydrophobicity to the cell wall. Several studies have demonstrated the link between reduced lignin levels and decreased recalcitrance with improved saccharification efficiency (Grabber, 2005; Davison et al., 2006; Ralph et al., 2006; van der Rest et al., 2006; Talukder, 2006; Chen & Dixon, 2007; Jackson et al., 2008).

Lignin content and composition have been manipulated transgenically in a variety of plant species, mainly by targeting the monolignol biosynthesis pathway (Fig. S1, see Supporting Information) (Hoffmann et al., 2004; Chen et al., 2006; Ralph et al., 2006; Millar et al., 2007; Shadle et al., 2007; Wagner et al., 2007; Jackson et al., 2008). In the above studies, reduced lignin content and/or altered lignin composition were achieved by the down-regulation of the individual genes encoding phenylalanine ammonia-lyase, cinnamic acid 4-hydroxylase, 4-hydroxycinnaminate-CoA ligase, hydroxycinnamoyl-CoA shikimate hydroxycinnamoyl transferase (HCT), p-coumaroylshikimate 3-hydroxylase, caffeoyl CoA 3-O-methyltransferase, caffeic acid/5-hydroxyferulic acid 3-O-methyltransferase, cinnamoyl-CoA reductase (CCR) and cinnamyl alcohol dehydrogenase (CAD) (Fig. S1). However, when lignin levels become too low, plant growth and development can be affected, and this has been attributed to either the dysfunction of the vascular system or altered hormone transport (van der Rest et al., 2006; Besseau et al., 2007; Do et al., 2007; Leple et al., 2007; Shadle et al., 2007; Mir Derikvand et al., 2008).

In general, the down-regulation of the later enzymes in monolignol biosynthesis has less effect on plant growth and yield than the down-regulation of the earlier enzymes. This may result from the avoidance of pleiotropic growth effects as a result of perturbations in metabolites that feed into other pathways. CCR and CAD catalyze steps devoted specifically to monolignol biosynthesis, and could therefore be good targets for down-regulation in bioenergy crops to improve saccharification efficiency without compromising biomass yields. The reduction in lignin content through the down-regulation of CCR has been observed in tobacco (Dauwe et al., 2007), tomato (van der Rest et al., 2006), poplar (Leple et al., 2007) and other species. In a recent study (Jackson et al., 2008), CCR and CAD were down-regulated individually in alfalfa; a comparison between lines with similar reductions in lignin level revealed greater saccharification efficiency in the case of the CCR-down-regulated lines.

It is important to determine whether the improvements in saccharification efficiency recorded for dicot species, such as alfalfa, can be reproduced through targeting the monolignol pathway in switchgrass. However, little is known at present concerning the genes and corresponding enzymes involved in monolignol biosynthesis in this species. Here, we identify four cDNAs similar to CCRs in the tetraploid lowland switchgrass variety Alamo, and show that two of them (PvCCR1 and PvCCR2) encode enzymes with CCR activity. A combination of biochemical and gene expression analyses indicate that PvCCR1 is the enzyme involved in lignification, and therefore a target for down-regulation to improve switchgrass as a bioenergy crop. PvCCR2 appears to be involved in plant defense. We also describe potential allelic variants of switchgrass CCRs.

Materials and Methods

Plant material and reagents

A seedling from the Alamo cultivar of switchgrass (Panicum virgatum L.) was randomly selected, clonally propagated, and plants were grown under standard glasshouse conditions.

The CCR substrates 4-coumaroyl CoA, caffeoyl CoA, feruloyl CoA, 5-hydroxyferuloyl CoA and sinapoyl CoA were synthesized as described previously (Stockigt & Zenk, 1975). Aldehyde products of CCR activity for calibration curves were purchased from Sigma-Aldrich (St Louis, MO, USA) (coniferaldehyde and sinapaldehyde) or synthesized by the Schmer et al. group.
as described previously (Chen et al., 2001) (caffeoyl aldehyde and 5-hydroxyferuloyl aldehyde).

Cloning of CCR cDNAs

CCR sequences of Arabidopsis thaliana, Medicago truncatula and Zea mays were used as subjects for BLAST searches against the currently available switchgrass sequences in public databases. Full-length consensus sequences from multiple cDNA alignments were used for primer design. Total RNA was isolated from stem or leaf using Tri Reagent according to the manufacturer’s instructions (Molecular Research Center Inc., Cincinnati, OH, USA), and cDNA synthesis was performed using the SuperScript III First-Strand System for RT-PCR Kit (Invitrogen, Carlsbad, CA, USA) employing an oligo-dT primer. The full-length cDNA was amplified by PCR using KOD DNA polymerase (EMD, San Diego, CA, USA), and cloned into the Escherichia coli expression vectors pET28a (EMD) using regular cloning, or into pDEST17 using the Gateway technology (Invitrogen). The primers used for cloning were as follows:

\[
\text{PvCCR1} \\
\text{forward, 5'}-\text{AAGCATATGACCCTGTTGACGCGC-3'}; \\
\text{reverse, 5'}-\text{CCTGCAGCGCCGCACGATGCGAG-3'}.
\]

\[
\text{PvCCR2} \\
\text{forward, 5'}-\text{CACCATGCGGCATCCGCTGTCG-3'}; \\
\text{reverse, 5'}-\text{TCAAGGTTGAATCAATGTGCA-3'}.
\]

\[
\text{PvCCR-like1} \\
\text{forward, 5'}-\text{GAGCAGTGGGAGGGGAGGAA-3'}; \\
\text{reverse, 5'}-\text{GACTCGAGATTTAAGCCAATG-3'}.
\]

\[
\text{PvCCR-like2} \\
\text{forward, 5'}-\text{CACCATGCTCCGCGGCGGCGGAG-3'}; \\
\text{reverse, 5'}-\text{TCACAGATTGCTATAGTCAAAAC-3'}.
\]

Plasmids were isolated using the Wizard Plus Minipreps DNA Purification System (Promega, Madison, WI, USA), and their inserts were sequenced in both directions for confirmation.

Phylogenetic analysis

The DNA sequences of 16 previously characterized CCR proteins from several plant species (Pichon et al., 1998; Lauvergeat et al., 2001; Ruellan et al., 2003; Larsen, 2004a,b; Ma & Tian, 2005; Kawasaki et al., 2006; van der Rest et al., 2006; Leple et al., 2007; Ma, 2007) were collected. In order to perform phylogenetic reconstruction, a multiple sequence alignment was performed on the full-length sequences of the 16 proteins, together with the four proteins predicted from switchgrass cDNAs. MAFFT (Katoh et al., 2005) was used in the alignment by employing the highly accurate method L-INS-I (Aihola et al., 2006; Nuin et al., 2006).

A maximum likelihood (ML) tree was built using PhyML (Guindon & Gascuel, 2003), and the Bayesian tree was built using MrBayes (Ronquist & Huelsenbeck, 2003). Specifically, PhyML analyses were conducted using the JTT model, 100 replicates of bootstrap analysis, the estimated proportion of invariable sites, four rate categories, the estimated gamma distribution parameter and the optimized starting BIONJ tree. MrBayes analyses were conducted with a mixed amino acid model estimated in the run, the estimated proportion of invariable sites, the estimated gamma distribution parameter and 1 000 000 generations.

DNA gel blot analysis

To isolate genomic DNA, 0.5 g of leaf tissue was ground under liquid nitrogen, and DNA was isolated using the cetyltrimethylammonium bromide (CTAB) procedure (Porebski et al., 1997) with the following modifications: the ground tissue was resuspended in 20 ml of CTAB, incubated at 65°C for 1 h and then extracted with chloroform–isoamyl alcohol. The genomic DNA was digested with EcoRI, HindIII, PstI or XhoI following the manufacturer’s instructions. Ten micrograms of digested DNA were loaded and resolved on a 1.0% agarose gel, and then transferred to a nylon membrane (Zeta-Probe GT Genomic; Bio-Rad, Hercules, CA, USA) using established protocols (Sambrook et al., 1989).

For copy number reconstruction, 10 µg of genomic DNA (digested with EcoRI, HindIII, PstI or XhoI) was loaded onto gels, together with amounts of PvCCR1 and PvCCR2 cDNAs equivalent to one, two, four or eight copies (per haploid genome) based on an estimated switchgrass genome size of 1.85 × 10⁹ bp (Hopkins et al., 1996).

Digoxigenin (DIG)-labeled probe was prepared by PCR, amplifying a region of 372 bp of PvCCR1 or 470 bp of PvCCR2; the membrane was hybridized with the probe overnight at 60°C. Probe preparation, hybridization, DIG-antibody reaction, detection and film development were performed following the DIG Application Manual for Filter Hybridization (https://www.roche-applied-science.com/prodinfo_fst.jsp?page=/PROD_INF/MANUALS/DIG_MAN/dig_toc.htm).

Expression of switchgrass CCRs in E. coli

Escherichia coli strain Rosetta cells harboring the PvCCR1 or PvCCR2 constructs were cultured at 37°C until the optical density at 600 nm reached 0.6–0.7, and protein expression was then induced by adding isopropyl 1-thio β-galactopyranoside at a final concentration of 0.5 mM, followed by incubation at 16°C for 18–20 h. Frozen cell
pellets from 25 ml of induced culture were thawed at room temperature and resuspended in 1.2 ml of extraction–washing buffer (10 mM imidazole, 50 mM Tris–HCl pH 8.0, 500 mM NaCl, 10% glycerol and 10 mM β-mercaptoethanol). The pellets were sonicated three times for 20 s and the supernatants were recovered after centrifugation at 16 000g, mixed with equilibrated Ni-NTA beads (Qiagen, Germantown, MD, USA) and incubated at 4°C for 30 min under constant inversion to allow the His-tagged proteins to bind to the beads. After washing the beads three times with 1 ml of extraction–washing buffer, the target proteins were eluted with 250 µl of elution solution (250 mM imidazole, 50 mM Tris–HCl buffer pH 8.0, 500 mM NaCl, 10% glycerol and 10 mM β-mercaptoethanol). The purity of the eluted target proteins was verified by sodium dodecyl-sulfate-polyacrylamide gel electrophoresis, and the protein concentrations were determined using the BioRad protein assay (BioRad, Hercules, CA, USA).

Enzyme activity assays and kinetics

Pure recombinant enzymes (1–5 µg) were incubated at 30°C (for 5, 10 or 30 min) with 100 mM sodium/potassium phosphate buffer pH 6.3, 250 mM NADPH and 5–100 µM substrate (4-coumaroyl CoA, caffeoyl CoA, feruloyl CoA, 5-hydroxyferuloyl CoA or sinapoyl CoA) in a final volume of 500 µl. The reactions were stopped by the addition of 70 µl of 24% w/v trichloroacetic acid, and the products were extracted with ethyl acetate (3×500 µl), dried under an N2 flow and resuspended in 100 µl of methanol. The reaction products were analyzed by reverse-phase high-performance liquid chromatography on a C18 column (Spherisorb 5µm ODS2; Waters, Milford, MA, USA) in a step gradient using 1% phosphoric acid in water as solvent A and acetonitrile as solvent B. Calibration curves were constructed with authentic standards of each product, except for 4-coumaroyl aldehyde, which was not available. As the curves for the different products are very similar, the calibration curve for caffeoyl aldehyde was employed to calculate the kinetic parameters using 4-coumaroyl CoA as substrate.

Real-time quantitative reverse transcription-polymerase chain reaction (qRT-PCR)

Total RNA was isolated from the internodes, leaves, nodes and leaf sheaths of E4 stage (elongation stage with four internodes) (Fig. S2, see Supporting Information) switchgrass plants (Moore et al., 1991) using Tri-reagent. RNA integrity was checked using an Agilent 2100 Bioanalyzer with 6000 Nano Chip (Agilent Technologies, Palo Alto, CA, USA). First-strand cDNA was synthesized from 3 µg of total RNA using oligo-dT20 primer (Invitrogen) and SuperScript III reverse transcriptase (Invitrogen) after an extensive Turbo-DNase (Ambion, Austin, TX, USA) treatment. DNA contamination was tested for each sample before reverse transcription.

PRIMER EXPRESS® software (version 3.0; Applied Biosystems, Foster City, CA, USA) was used to design primer sets for PvCCR1 (forward, 5′-GCGTCGGGCTTCAACAAC-3′; reverse, 5′-TCGGGTCATCTGGGTTCCT-3′), PvCCR2 (forward, 5′-CGCCGATGTCACGACACAC-3′; reverse, 5′-TTGCCGGCCGATAATTTG-3′) and the reference gene polyubiquitin (forward, 5′-TTGGTGTTGGCCAGTAAG-3′; reverse, 5′-AGGCCAGAGCCGGCAGTA-CAG-3′) using default parameters according to the manufacturer’s recommendation. PCRs were performed in an ABI PRISM 7900HT Sequence Detection System (Applied Biosystems Inc.) in an optical 384-well plate. Power SYBR Green PCR Master Mix (Applied Biosystems Inc.) was used for the PCR in a final volume of 10 µl. Three technical replicates were included for each sample and the data were normalized against the reference gene. No-template control reactions were performed for each pair of primers. Gene-specific amplification was confirmed by a single dominant peak in the qRT-PCR dissociation curve analyses. PCR efficiency (E) was calculated by LinRegPCR software (Ramakers et al., 2003). The ACT values were used to quantify the relative expression ratio, as described previously (Czechowski et al., 2004).

Inoculation of switchgrass with rust fungus

The leaf rust pathogen Puccinia emaculata isolate NRF-1, collected from switchgrass research plots in Ardmore, Oklahoma, was maintained on susceptible switchgrass plants. The two-celled teliospores and internal transcribed spacer sequences were used to confirm the identity of the P. emaculata NRF-1 isolate (R. Uppalapati, Y. Ishiga, L. Szabo and M. Saha, unpublished). Urediniospores were harvested from infected leaves using a spore collector (designed by Dr Geoff Harms, University of Minnesota) attached to a vacuum-pump (GAST Manufacturing Inc., Benton Harbor, MI, USA), and were suspended in distilled water (containing 0.001% Tween 20) to a final concentration of 1×10^5 spores ml⁻¹. The spore suspension was spray inoculated onto the leaves of Panicum virgatum cv. Alamo, and distilled water (0.001% Tween 20) was used for mock-inoculation controls. The spray-inoculated plants were incubated in the dark for 18 h at 100% humidity in a dew chamber. The plants were subsequently moved to a growth chamber maintained at 70% relative humidity, 29°C day : night temperature, 16 h photoperiod, photon flux density 150–200 µmol m⁻² s⁻¹. Mock- and rust-inoculated leaf material was collected 2 and 10 d post-inoculation (dpi), and RNA was extracted using TRIzol Reagent (Invitrogen) according to the manufacturer’s instructions.
Results

Isolation and sequence analysis of switchgrass CCR cDNAs

cDNAs of \textit{PvCCR1}, \textit{PvCCR2}, \textit{PvCCR-like1} and \textit{PvCCR-like2} were isolated by RT-PCR using specific primers designed from full-length consensus sequences. Sequence analysis of several independent clones of \textit{PvCCR1} revealed five different cDNA sequences with small variations (Fig. 1). Based on the consensus sequence between the five clones, four \textit{PvCCR1}\textsubscript{s} (a-d) had between two to five different nucleotides, resulting in changes of between one and four amino acid residues. The other cDNA (\textit{PvCCR}\textsubscript{1}e) had 11 nucleotide differences and a three-nucleotide deletion when compared with the consensus sequence. The use of a proofreading polymerase makes it unlikely that the sequence differences were PCR artifacts. In tetraploid self-incompatible switchgrass, it is likely that the four most similar \textit{PvCCR1}\textsubscript{s} are allelic variants, whereas the most divergent sequence might be the product of a duplication event.

Similarly, we identified three variants for \textit{PvCCR2}, four for \textit{PvCCR-like2} (one, \textit{PvCCR-like2d}, incomplete), but only one for \textit{PvCCR-like1} (Fig. S3, see Supporting Information). These numbers will probably increase as additional sequence information for switchgrass becomes available.

The nucleotide and encoded amino acid sequences of the above \textit{PvCCR1} and \textit{PvCCR2} cDNAs have been deposited in the EMBL/GenBank/DDBJ Nucleotide Sequence Databases with the accession numbers GQ450296 (\textit{PvCCR1}\textsubscript{a}), GQ450297 (\textit{PvCCR1}\textsubscript{b}), GQ450298 (\textit{PvCCR1}\textsubscript{c}), GQ450299 (\textit{PvCCR1}\textsubscript{d}), GQ450300 (\textit{PvCCR1}\textsubscript{e}), GQ450301 (\textit{PvCCR2}\textsubscript{a}), GQ450302 (\textit{PvCCR2}\textsubscript{b}), GQ450303 (\textit{PvCCR2}\textsubscript{c}), GQ450304 (\textit{PvCCR-like1}), GQ450305 (\textit{PvCCR-like2a}), GQ450306 (\textit{PvCCR-like2b}) and GQ450307 (\textit{PvCCR-like2c}).

The \textit{PvCCR1} variant, which is most similar to the consensus sequence between the five clones (designated \textit{PvCCR1a}), has an open reading frame (ORF) of 1095 nucleotides encoding a 364-amino-acid protein with a predicted molecular mass of 39.04 kDa and a pI of 5.99. The \textit{PvCCR2} variant selected by the same criteria (designated \textit{PvCCR2a}) has an ORF of 1032 nucleotides encoding a protein of 343 amino acids (37.88 kDa) with a calculated pI of 6.61.

The protein sequences deduced from the \textit{PvCCR1}\textsubscript{a} and \textit{PvCCR2}\textsubscript{a} cDNAs show 54% similarity. \textit{PvCCR1} is the most closely related to previously characterized CCR proteins, sharing 69–89% similarity with \textit{AtCCR1}, \textit{TaCCR1} and \textit{ZmCCR1} (Table 1), whereas \textit{PvCCR2} shows lower similarity with the same protein sequences. \textit{PvCCR2} has a greater similarity to \textit{ZmCCR2} from corn (Table 1). The alignment between the switchgrass CCRs and other related proteins (Fig. S4, see Supporting Information) shows the conserved NADPH binding site and the signature motif NWYCY, which is thought to be a feature of the catalytic site of this enzyme (Lacombe \textit{et al.}, 1997).

\textit{PvCCR-like1} and \textit{PvCCR-like2} protein sequences show less similarity to the previously characterized CCRs (Table 1), and the NWYCY signature motif, which is 100% conserved in all CCR enzymes with proven activity, is not conserved in these proteins (Fig. S4). The signature motif is replaced by QWYY in \textit{PvCCR-like1} and by DWNY in \textit{PvCCR-like2}.

A phylogenetic tree was constructed to show the relationships of the \textit{PvCCR1}, \textit{PvCCR2}, \textit{PvCCR-like1} and \textit{PvCCR-like2} protein sequences to those of previously characterized CCRs from several dicots and monocots. The

\begin{itemize}
  \item \textbf{PvCCR1\textsubscript{c}}: MTV...PKNAHLKALDGA...TEYVISAAAEEAGTVRR...AAWEAARQ...VGPL\textsubscript{QQP}
  \item \textbf{PvCCR1\textsubscript{d}}: MTV...PKNAHLKALDGA...TEYVISAAAEEAGTVRR...AAWEAARQ...VGPL\textsubscript{QQP}
  \item \textbf{PvCCR1\textsubscript{b}}: MTV...PKNAHLKALDGA...TEYVISAAAEEAGTVRR...AAWEAARQ...VGPLL\textsubscript{QQP}
  \item \textbf{PvCCR1\textsubscript{a}}: MTV...PKNAHLKALDGA...TEYVISAAAEEAGTVRR...AAWEAARQ...VGPLL\textsubscript{QQP}
  \item \textbf{PvCCR1\textsubscript{e}}: MT...PKNAHLKALDGA...TEYVISAAAEEAGTVRR...AAWEAARQ...VGPLL\textsubscript{QQP}
  \item \textbf{PvCCR2\textsubscript{c}}: AQHLRVESPAASGRHLCAERVlH...AYKFTNQK...QEKKHLPVL\textsubscript{LDEQA}...IRA
  \item \textbf{PvCCR2\textsubscript{d}}: AQHLRVESPAASGRHLCAERVlH...AYKFTNQK...QEKKHLPVL\textsubscript{LDEQA}...IRA
  \item \textbf{PvCCR2\textsubscript{b}}: AQHLRVESPAASGRHLCAERVlH...AYKFTNQK...QEKKHLPVL\textsubscript{LDEQA}...IRA
  \item \textbf{PvCCR2\textsubscript{a}}: AQHLRVESPAASGRHLCAERVlH...AYKFTNQK...QEKKHLPVL\textsubscript{LDEQA}...IRA
  \item \textbf{PvCCR2\textsubscript{e}}: AQHLRVESPAASGRHLCAERVlH...AYKFTNQK...QEKKHLPVL\textsubscript{LDEQA}...IRA
\end{itemize}

\textbf{Fig. 1} Aligned partial protein sequences of \textit{PvCCR1} variants. Differences are highlighted.
phylogenies of 20 CCR proteins built using both ML and Bayesian analyses were very consistent with each other. In Fig. 2, only the topology from the ML analysis is shown, but statistical confidence values of both ML and Bayesian analyses are indicated. This reveals that there are three major branches of the CCR gene family (Fig. 2). Eight dicot CCRs cluster together, whereas the 12 monocot CCRs are arranged into two clusters. PfCCR1 belongs to one of the monocot clusters and is orthologous to ZmCCR1 and SoCCR, whereas the other three switchgrass sequences, together with ZmCCR2, TaCCR2 and OsCCR, belong to another monocot cluster with long branches, suggesting that these genes have evolved relatively more rapidly (Fig. 2). In this cluster, PfCCR2 is orthologous to ZmCCR2 and TaCCR2, whereas the other two CCR-like sequences form a separate subcluster with even longer branches, suggesting a distant evolutionary relationship between the two sequences and the true CCR proteins (bootstrap values are lower than 50% and thus not shown). The addition of more CCR homologs from moss (Physcomitrella patens spp. patens) and spike moss (Selaginella moellendorfii) into the phylogenetic analyses revealed that the two switchgrass CCR-like proteins form a cluster with mosses (data not shown), indicating that they are descendents of an ancient CCR clade.

Genomic organization of CCRs in switchgrass

To estimate the complexity of the CCR gene family in switchgrass, DNA gel blot analysis was carried out using genomic DNA digested with EcoRI, HindIII, PvuII and XhoI. A 372 bp fragment of PfCCR1 and a 470 bp fragment of PfCCR2 were used separately as probes. On the blot probed with PfCCR1, two strong bands and one weak band were detected in the EcoRI digest, one strong band and two weak bands in the HindIII digest and one very high molecular weight band and one or two weaker, lower molecular weight bands in the PvuII and XhoI digests (Fig. 3a). The PfCCR2 blot showed three or four bands (some possibly doublets) in the EcoRI, HindIII and PvuII lanes. The XhoI lane showed the expected band of c. 270 bp allowing for the possible presence of a small intron, resulting from the presence of XhoI sites within the probe region. These results indicate that the switchgrass genome contains more than one copy of both genes. As we do not yet have information on the presence of possible intron sequences, and sequence polymorphisms would not be detected by the above analysis, the CCR gene number was further addressed by copy number reconstruction analysis (Fig. 3b). The intensity of the genomic DNA band in the HindIII digest was comparable with eight copies for PfCCR1, whereas, using the PfCCR2 probe, the intensity of the band on XhoI digestion appeared to be equivalent to four copies; it should be noted that there was no cross-hybridization between PfCCR1 and PfCCR2 probes and plasmids (Fig. 3b).

Biochemical characterization of switchgrass CCRs

For biochemical characterization of the CCR enzymes, E. coli-expressed His-tagged fusion proteins were purified to homogeneity, as determined by sodium dodecylsulfate-polyacrylamide gel electrophoresis (Fig. 4a). Under denaturing electrophoresis conditions, these proteins exhibited a molecular mass of between 40 and 43 kDa, in agreement with the calculated sizes plus the tag which includes the 6-His motif. The enzymatic activity of the purified proteins was initially screened with the five possible substrates (4-coumaroyl CoA, caffeoyl CoA, feruloyl CoA, 5-hydroxyferuloyl CoA and sinapoyl CoA). PfCCR1 (variants a and e) and PfCCR2 showed activity with every substrate, whereas PfCCR-like1 and PfCCR-like2 showed no activity. Chromatograms and curves of reaction velocity vs substrate concentration for three substrates are shown in Fig. 4b,c. These confirm the identities of the expected aldehyde products, and show classic saturation curves over the range of substrate concentrations chosen.

Kinetic parameters were determined for PfCCR1a, PfCCR1e and PfCCR2a (Table 2). The three enzymes exhibited substrate inhibition at high concentrations (> 60 μM) of feruloyl CoA and 5-hydroxyferuloyl CoA (data not shown), but not with the other substrates. The
two PvCCR1 variants were overall more active than PvCCR2a based on the specific activity, and exhibited a very strong preference for feruloyl CoA, the $K_m$ value for this substrate being the lowest and the turnover rate ($K_{cat}$) being significantly higher than that for any other substrate (c. 10-fold higher than for 4-coumaroyl CoA). PvCCR2 preferred caffeoyl CoA and 4-coumaroyl CoA, but the catalytic efficiencies were very low in comparison with the efficiency of PvCCR1 towards feruloyl CoA (Table 2).

To determine whether the variants of $CCR1$ from Alamo encode proteins with significantly different biochemical properties, we also performed kinetic analysis of purified recombinant CCR1e, the most divergent of the five forms detected to date. Generally, the kinetic parameters of CCR1a and CCR1e were quite similar, particularly for the major substrate feruloyl CoA, although CCR1e was somewhat less active than CCR1a with caffeoyl and 4-coumaroyl CoAs (Table 2).

Fig. 2 Phylogenetic tree of switchgrass (Panicum virgatum) cinnamoyl CoA reductase (CCR) and CCR-like sequences, and 16 characterized CCRs from dicots (circles) and monocots (triangles). The tree was built using PhyML, and the neighbor-joining tree was built using MEGA4 (see Materials and Methods section). Only the topology from the maximum likelihood (ML) analysis is shown, but statistical confidence values for both ML and Bayesian analyses are indicated. Protein IDs or accession numbers are: Arabidopsis thaliana CCR1, AAG46037.1; Arabidopsis thaliana CCR2, AAG53687.1; Medicago truncatula CCR1, TC106830; Medicago truncatula CCR2, TC100678; Triticum aestivum CCR1, ABE01883.1; Triticum aestivum CCR1, AY771357; Zea mays CCR1, CAA74071.1; Zea mays CCR2, NP_001105715.1; Oryza sativa CCR1, BAI9248.1; Populus trichocarpa CCR, CAA12276.1; Lycopersicon esculentum CCR1, AAY41879.1; Lycopersicon esculentum CCR2, AAY41880.1; Eucalyptus gunnii CCR, CAA56103.1; Lolium perenne CCR, AAG09817.1; Hordeum vulgare CCR, AAN71760.1; Saccharum officinarum CCR, CAA13176.1.
Developmental expression of *PvCCR1* and *PvCCR2*

To investigate the expression patterns of *CCR* genes, qRT-PCR analysis was carried out using total RNA from plant tissues in the E4 stage (Figs S2,5a). In general, *PvCCR1* exhibited higher expression than *PvCCR2*, and was preferentially expressed in the stem, with highest expression in internode 1 (Fig. 5b). In all plant parts except the leaves, expression of *PvCCR1* was much higher than that of *PvCCR2*, and the largest difference was in internode 1, where the expression of *PvCCR1* was over 200 times that of *PvCCR2* (Fig. 5). *PvCCR2* was expressed at a higher level than *PvCCR1* in leaves, except for the youngest (leaf 5).

Expression of *PvCCR1* and *PvCCR2* in rust-infected switchgrass

qRT-PCR was used to determine the expression of *PvCCR1* and *PvCCR2* in leaves and internodes at 0, 2 and 10 dpi with the rust fungus *P. emaculata*. It is important to note that Alamo is moderately susceptible to rust. At 10 dpi, the expression of *PvCCR2* in leaves was almost 3.5-fold higher than that in the 0 dpi control (Fig. 6), although the gene was not induced at 2 dpi and *PvCCR1* transcript levels did not increase in infected leaves. *PvCCR2* expression was not significantly induced above control levels in internodes at 2 or 10 dpi.

Discussion

Allelic variation and genome complexity

DNA gel blot analysis with copy number reconstruction of *PvCCR1* and *PvCCR2* suggests the presence of multiple copies of each of these genes in the switchgrass genome. *CCR1* appears to be represented by more copies than *CCR2*. The values of four copies of *CCR2* and eight copies of *CCR1* per haploid genome are probably only approximate, as they are based on a switchgrass genome size determined only by flow cytometry (Hopkins et al., 1996). However, they do appear to be consistent with a high level of sequence polymorphism observed between individual PCR-amplified cDNA clones for the switchgrass *CCR* or *CCR*-like genes. This can be explained, at least in part, by allelic variation in the highly heterozygous, self-incompatible tetraploid species (Casler, 2005), and we have seen the same phenomenon with other lignin biosynthesis enzyme genes in switchgrass, e.g. *HCT* and *CAD* (L.L. Escamilla-Trevino, H. Shen, Y. Tang, T. Hernandez and R.A. Dixon, unpublished).

Heterosis (hybrid vigor) has recently been studied in switchgrass (Martinez-Reyna & Vogel, 2008; Vogel & Mitchell, 2008) with a view to improving the plant’s characteristics as a bioenergy crop. Allelic variation in a polyploid genome has been suggested as being responsible for heterosis in maize (Springer & Stupar, 2007). It is possible that allelic variation in lignin biosynthetic genes could contribute to heterosis. Equally likely is the possibility that it might have an impact on recalcitrance to saccharification; in this respect, association mapping studies are ongoing in switchgrass through the activities of the DOE Bioenergy Sciences Center, with high-throughput analysis of recalcitrance phenotypes, and it should soon be possible to determine whether variation in lignin pathway gene alleles is linked to recalcitrance phenotypes. In the present work, the catalytic efficiencies of the two most extreme variants of *CCR1* appeared to be quite similar. However, the other variants were not tested, and it is possible that other properties of a protein, such as its...
ability to associate with other proteins, may have an impact on its overall biological activities.

PvCCR vs PvCCR-like proteins

Most plant species with extensive sequence information appear to contain multiple genes annotated as encoding CCR. All previously characterized CCR enzymes, which are included in the phylogenetic tree shown in Fig. 2, contain the motif NWYCY. In the present work, the recombinant enzymes that have the NWYCY motive (PvCCR1 and PvCCR2) showed CCR activity, whereas enzymes that do not have the signature motif (PvCCRLike-1 has QWYVY and PvCCRLike-2 has DWYNY) do not show any CCR.

Fig. 4 Functional characterization of switchgrass (Panicum virgatum) cinnamoyl CoA reductases (CCRs). (a) Protein gel blot analysis of recombinant switchgrass CCRs expressed in Escherichia coli; M, molecular weight markers; 1, total proteins from noninduced cells; 2, total proteins from induced cells; 3, soluble proteins from induced cells; 4, pure recombinant protein. (b) High-performance liquid chromatograms and UV spectra (insets) of products. (c) Curves of reaction velocity vs substrate concentration for PvCCR1a (diamonds) and PvCCR2a (triangles) towards feruloyl CoA, caffeoyl CoA and 4-coumaroyl CoA.
activity. These results confirm that the NWYCY motif is indispensable for CCR activity, and therefore diagnostic for true CCRs, as suggested previously (Lacombe et al., 1997).

It is also possible to distinguish CCR from CCR-like on the basis of the overall sequence identity. PvCCR-like proteins show less than 46% similarity to functionally characterized CCRs at the amino acid level, whereas PvCCR1 has between 63% and 89% and PvCCR2 has between 53% and 60% similarity to functionally identified CCRs.

PvCCR1 and PvCCR2 have different biochemical properties

PvCCR1 and PvCCR2 encode functionally proven CCR enzymes, but the two have very different phylogenetic, biochemical and gene expression characteristics, and share only 53% similarity at the amino acid sequence level. In the phylogenetic tree, PvCCR1 is very closely related to ZmCCR1 (89% similarity) and seems to be its ortholog, whereas the closest CCR to PvCCR2 is ZmCCR2 with 60% similarity. In spite of their differences, PvCCR1 and PvCCR2 have almost identical NADPH binding motifs and a fully conserved NWYCY CCR signature motif, which is thought to be involved in the catalytic site of the enzyme (Lacombe et al., 1997). Recombinant PvCCR1 is active towards the five possible substrates, but clearly prefers feruloyl CoA, whereas recombinant PvCCR2 prefers 4-coumaroyl CoA and caffeoyl CoA, although the catalytic efficiencies are very low in comparison with the efficiency of PvCCR1 towards feruloyl CoA. Even the $K_{cat}/K_m$ values for PvCCR1 towards 4-coumaroyl CoA or caffeoyl CoA are between two- and three-fold higher than those observed for PvCCR2. In spite of its substrate preference, the kinetic data and transcript expression patterns suggest that PvCCR2 is unlikely to be involved in the formation of hydroxyphenyl (H) lignin, derived from coumaroyl alcohol, during switchgrass

<table>
<thead>
<tr>
<th>Substrate</th>
<th>$K_m$ (μM)</th>
<th>$K_{cat}$ (s$^{-1}$)</th>
<th>$K_{cat}/K_m$ (s$^{-1}$ μM$^{-1}$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Feruloyl CoA</td>
<td>25.4</td>
<td>2.01</td>
<td>0.079</td>
</tr>
<tr>
<td>4-Coumaroyl CoA</td>
<td>31.6</td>
<td>0.32</td>
<td>0.010</td>
</tr>
<tr>
<td>Caffeoyl CoA</td>
<td>29.6</td>
<td>0.23</td>
<td>0.008</td>
</tr>
<tr>
<td>Sinapoyl CoA</td>
<td>56.5</td>
<td>0.20</td>
<td>0.004</td>
</tr>
</tbody>
</table>

ND, not determined.
development. Furthermore, as both enzymes exhibit very low efficiencies towards sinapoyl CoA, it is likely that the pathway to syringyl (S) lignin in switchgrass does not operate via the reduction of sinapoyl CoA.

It is interesting to compare the catalytic activities of switchgrass CCRs 1 and 2 with pairs of CCR enzymes reported from other species. The $K_{\text{cat}}/K_m$ value of PvCCR1a for feruloyl CoA is 8–10 times higher than that towards 4-coumaroyl and caffeoyl CoAs, and more than 20 times higher than that towards sinapoyl CoA. These results are similar to those reported for TaCCR1 (Ma, 2007). However, the activity ($K_{\text{cat}}$) of PvCCR1a towards feruloyl CoA is c. 30 times higher than the activity of PvCCR2a, whereas the difference in activity between TaCCR1 and TaCCR2 is less than a factor of two (Ma & Tian, 2005). TaCCR2 has high activity towards caffeoyl CoA and feruloyl CoA, but very low activity towards 4-coumaroyl CoA, whereas PvCCR2 prefers caffeoyl and 4-coumaroyl CoAs and has low activity towards feruloyl CoA. Thus, although TaCCR2 is phylogenetically related to PvCCR2, its overall activity and substrate preference are distinct from those of the switchgrass enzyme, and are more consistent with an involvement in lignin biosynthesis during development.

In the case of Arabidopsis, both AtCCR1 and AtCCR2 prefer feruloyl CoA, although AtCCR1 performs better than AtCCR2 (Lauvergeat et al., 2001). The $K_{\text{cat}}$ value of CCR from *Eucalyptus gunnii* is three to four times higher for feruloyl CoA than for sinapoyl CoA, 4-coumaroyl CoA or caffeoyl CoA (Goffner et al., 1994), but is c. 10 times lower than the $K_{\text{cat}}$ value of PvCCR1a for feruloyl CoA.

Another interesting biochemical property of switchgrass CCRs is their substrate inhibition by feruloyl CoA and 5-hydroxyferuloyl CoA at concentrations above 60 $\mu$M. A similar phenomenon was recently observed in the model legume *Medicago truncatula*, where one CCR form was substrate-inhibited by feruloyl and sinapoyl CoAs, whereas the other was not (R. Zhou & R. A. Dixon, unpublished). The physiological significance of these observations currently remains unclear, but they do suggest the possibility for the fine control of flux into different branches of the monolignol pathway.

Different functions for PvCCR1 and PvCCR2

The strongest expression of *PvCCR1* was in stem tissues, particularly in internode 1 of switchgrass plants at the E4 stage, which correlates with the high lignin content at this stage (H. Shen, Y. Tang and R. A. Dixon, unpublished). Its expression level was many times higher (223 times higher in internode 1) than that of *PvCCR2* in the stem, suggesting that *PvCCR1* is the enzyme involved in the reduction of hydroxycinnamoyl CoAs destined for lignin formation in switchgrass stems. The expression pattern of *PvCCR1* is similar to that of *ZmCCR1* in maize (Pichon et al., 1998), *TaCCR1* in wheat (Ma, 2007) and *LpCCR* in ryegrass (Larsen, 2004a). *PvCCR1* has a strong preference for feruloyl CoA as substrate, consistent with the currently accepted position of CCR in the monolignol pathway (Nakashima et al., 2008).

The constitutive expression of *PvCCR2* was barely detectable in growing tissues, suggesting that this gene may be associated with inducible physiological processes. *ZmCCR2* and *AtCCR2* are also expressed at very low levels in growing tissue. Recently, it has been shown that *ZmCCR2* is highly induced by water deficiency in the root elongation zone, suggesting that it may facilitate root acclimation to drying environments (Fan et al., 2006), whereas *AtCCR2* was induced in response to pathogenesis (Lauvergeat et al., 2001). Rust disease is common in switchgrass (Parrish & Fike, 2005), and *PvCCR2* was induced 3.5-fold at 10 dpi in leaves, whereas *PvCCR1* was not induced at all, suggesting that *PvCCR2* could have a specific role in defense.

Concluding remarks

Switchgrass has been proposed as the major perennial feedstock for lignocellulosic ethanol production in the USA (Schmer et al., 2008). A reduction in recalcitrance is a major target to make lignocellulosic ethanol production more competitive. According to the present phylogenetic, biochemical and gene expression studies, *PvCCR1* is probably the best target among the switchgrass genes annotated as encoding potential CCR enzymes. Three previous studies have reported the impacts on lignin content and composition of the down-regulation of *dicot CCR* genes possessing 69–72% amino acid similarity to *PvCCR1*. In tomato, the lignin content was significantly reduced by the down-regulation of a gene with 70% identity to *PvCCR1* (van der Rest et al., 2006); in alfalfa, the reduction in lignin levels in *CCR* down-regulated transgenic plants was shown to be associated with improved cell wall saccharification efficiency (Jackson et al., 2008); in poplar, *CCR* down-regulation led to reduced levels of lignin and hemicelluloses with a corresponding increase in cellulose levels (Leple et al., 2007). Future studies will address the impacts of down-regulation of *PvCCR1* in switchgrass, and also investigate further whether potential allelic variation is likely to be a complicating factor in lignin modification in this highly heterozygous tetraploid species. This work is currently in progress; however, although several studies have reported genetic transformation of switchgrass by either biolistic or *Agrobacterium*-mediated procedures (see Wang & Ge, 2006), the procedures are still both time-consuming and of low efficiency. As an alternative approach, the DOE Bioenergy Sciences Center has been developing virus-induced gene silencing (Unver & Budak, 2009) in the very closely related species foxtail millet (*Setaria italica*), and this might provide a first-
pass approach to determining the suitability of specific switchgrass CCR genes as targets for the improvement of cell wall saccharification. The present studies could also find application for the improvement of lignocellulose processing in species such as maize, sugarcane and sorghum, which possess CCR gene family members with 89–90% amino acid sequence similarity to PvCCR1.

Acknowledgements

We thank Dr Fang Chen for assistance with the synthesis of caffeoyl aldehyde and 5-hydroxyferuloyl aldehyde, Dr Zengu Wang for providing us with the cDNA of PvCCR-like1, and Drs Huanzhong Wang and Malay Saha for critical reading of the manuscript. This work was supported by the US Department of Energy Bioenergy Research Centers, through the Office of Biological and Environmental Research in the DOE Office of Science.

References


Supporting Information

Additional supporting information may be found in the online version of this article.

Fig. S1 The monolignol biosynthesis pathway depicted as a metabolic grid.

Fig. S2 Developmental stages of switchgrass.

Fig. S3 Nucleotide sequences of allelic variants of PcvCCR1, PcvCCR2 and PcvCCR-like2.

Fig. S4 Multiple sequence alignments of PcvCCR1 (consensus), PcvCCR2, PcvCCR-like1 and PcvCCR-like1 and 2 with cinnamoyl-CoA reductases from other plant species.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the New Phytologist Central Office.