University of Nebraska - Lincoln

# DigitalCommons@University of Nebraska - Lincoln

Publications from USDA-ARS / UNL Faculty

U.S. Department of Agriculture: Agricultural Research Service, Lincoln, Nebraska

2015

# Characterization of novel *Brown midrib 6* mutations affecting lignin biosynthesis in sorghum

Erin D. Scully
*USDA-ARS*, erin.scully@ars.usda.gov

Tammy Gries
*USDA-ARS*, tgries2@unl.edu

Deanna L. Funnell-Harris
*USDA-ARS*, Deanna.Funnell-Harris@ars.usda.gov

Zhanguo Xin
*Plant Stress and Germplasm Development Unit, USDA-ARS, Lubbock, TX*

Frank A. Kovacs
*University of Nebraska-Kearney*, kovacsfa@unk.edu

*See next page for additional authors*

Follow this and additional works at: https://digitalcommons.unl.edu/usdaarsfacpub

## Authors

Erin D. Scully, Tammy Gries, Deanna L. Funnell-Harris, Zhanguo Xin, Frank A. Kovacs, Wilfred Vermerris, and Scott E. Sattler

**Research Article**

# Characterization of novel *Brown midrib 6* mutations affecting lignin biosynthesis in sorghum

Erin D. Scully[1,2], Tammy Gries[1], Deanna L. Funnell-Harris[1,3], Zhanguo Xin[4], Frank A. Kovacs[5], Wilfred Vermerris[6], and Scott E. Sattler[1,2*]

[1]Grain, Forage, and Bioenergy Research Unit, USDA-ARS, Lincoln, NE 68583, USA,

[2]Department of Agronomy and Horticulture, University of Nebraska-Lincoln, Lincoln, NE 68583, USA, [3]Department of Plant Pathology, University of Nebraska-Lincoln, Lincoln, NE 68583,USA, [4]Plant Stress and Germplasm Development Unit, USDA-ARS, Lubbock, TX 79414, USA, [5]Department of Chemistry, University of Nebraska-Kearney, Kearney, NE 68849, USA,

[6]Department of Microbiology & Cell Science and UF Genetics Institute, University of Florida, Gainesville, FL 32610, USA. *Correspondence: Scott.Sattler@ars.usda.gov

Running Title: Characterization of Novel Sorghum *Bmr6* Mutations

Edited by: Natalia Doudareva, Purdue University, USA

**Abstract**

The presence of lignin reduces the quality of lignocellulosic biomass for forage materials and feedstock for biofuels. In C4 grasses, the *brown midrib* phenotype has been linked to mutations to genes in the monolignol biosynthesis pathway. For example, the *Bmr6* gene in sorghum (*Sorghum bicolor*) has been previously shown to encode cinnamyl alcohol dehydrogenase (CAD), which catalyzes the final step of the monolignol biosynthesis pathway. Mutations in this gene have been shown to reduce the abundance of lignin, enhance digestibility, and improve saccharification efficiencies and ethanol yields. Nine sorghum lines harboring five different *bmr6* alleles were identified in an EMS-mutagenized TILLING population. DNA sequencing of *Bmr6* revealed that the majority of the mutations impacted evolutionarily conserved amino acids while three-dimensional structural modeling predicted that all of these alleles interfered with the enzyme's ability to bind with its NADPH cofactor. All of the new alleles reduced *in vitro* CAD activity levels and enhanced glucose yields following saccharification. Further, many of these lines were associated with higher reductions in acid detergent lignin compared to lines harboring the previously characterized *bmr6-ref* allele. These *bmr6* lines represent new breeding tools for manipulating biomass composition to enhance forage and feedstock quality.

**INTRODUCTION**

Plant cell walls represent a vast store of fixed carbon that is currently being utilized for biofuel production. Converting lignocellulosic biomass into biofuels requires liberation of sugars from plant cell wall polysaccharides. However, the presence of lignin, whose linkages are highly resistant to enzymatic and microbial degradation, impairs access to these polysaccharides. Lignin is a nonlinear, hydrophobic biopolymer that is localized to the secondary cell wall and imparts structural rigidity and support to the cell wall structure. Its presence within cell walls prevents vascular collapse and enhances water conductance (Ruel et al 2009) and also provides a protective barrier against insect pests and microbial pathogens (Siegrist et al 1994; Bonello et al 2003). Lignin is comprised of a variety of aromatic compounds linked predominantly by $\beta$-O-4 aryl ether linkages (Whetten and Sederoff 1995). Monolignols, which are the precursors of lignin, are derived from the amino acid phenylalanine through a pathway with at least eleven different enzymatic reactions (Vanholme et al 2013). The monolignols, *p*-coumaryl, coniferyl, and sinapyl alcohol, are transferred to the cell wall and polymerized into lignin through oxidative coupling reactions generated from enzymatically-generated free radicals (Boerjan et al 2003), which result in the formation of *p*-hydroxyphenyl (H), guaiacyl (G), and syringyl (S) lignin subunits, respectively. The presence of lignin reduces the yield of fermentable sugars from the plant cell wall and thus has negative impacts on agronomic and industrial uses of lignocellulosic materials (Alvira et al 2010). Therefore, efficient saccharification and conversion of plant biomass to biofuels requires pretreatments to disrupt the lignin barrier (Zeng et al 2014). Unfortunately, pretreatment often releases phenolic compounds that can be inhibitory to saccharification and fermentation processes (Klinke et al 2004; Ximenes et al 2010). The

presence of lignin also reduces the forage quality of sorghum by hindering access to digestible energy and nutrients (i.e., proteins and minerals) crosslinked to the cell wall matrix (Wilson and Kennedy 1996).

A potential solution to these problems lies in utilizing plants with lower abundances of lignin for biomass conversions and forages (Chen and Dixon 2007). The genes that encode enzymes of the monolignol biosynthetic pathway are targets for modifying lignin concentration and manipulating biomass composition (Sattler et al 2010b). In several plant species, mutations in genes in the monolignol biosynthetic pathway have been identified that result in reduced lignin phenotypes (Poovaiah et al 2014; Vermerris and Abril 2014). In C4 grasses, the *brown midrib* phenotype (*bmr* in sorghum and *bm* in maize) is associated with reduced lignin concentration and alterations in lignin subunit composition relative to wild-type counterparts (reviewed in Sattler et al 2010b). At least six naturally occurring *brown midrib* loci have been identified in maize (*Zea mays*) and are designated *bm1* through *bm6* (Ali et al 2010). In sorghum (*Sorghum bicolor*), *bmr* mutants have been identified in a chemically mutagenized population (Xin et al 2008; Xin et al 2009). At least seven independent *Bmr* loci have been identified in sorghum (Sattler et al 2014). Genes encoding three of these loci have been identified and characterized and include *bmr2*, *bmr6*, and *bmr12*, which encode enzymes in the monolignol biosynthetic pathway; 4-coumarate: coenzyme A ligase (4CL), cinnamyl alcohol dehydrogenase (CAD) and caffeic acid *O*-methyltransferase (COMT), respectively (Bout and Vermerris 2003; Sattler et al 2009; Saballos et al 2012). The sorghum genes encoding the loci *bmr19* and *bmr29* to *bmr32* have not been identified (Saballos et al 2008; Sattler et al 2014). In maize, *Bm1*, *Bm2* and *Bm3* have been cloned, and *Bm1* and *Bm3* encode orthologous genes to *Bmr6* and *Bmr12*, respectively (Vignols et al 1995; Chen et al 2012). The maize *Bm2* and *Bm4* genes were shown to encode a

methylenetetrahydrofolate reductase and a folylpolyglutamate synthase respectively, which are involved in *S*-adenosyl-l-methionine (SAM) biosynthesis (Tang et al 2014; Li et al 2015). In addition to reductions to lignin content and changes in lignin composition, *bmr* mutations are also associated with other compositional changes of the plant cell wall that could enhance biomass conversions and improve forage digestibility. For example, reduced incorporation of *p*-coumaric acid into cell walls was previously observed in *bmr6* and *bmr12* plants (Palmer et al 2008). Consequences of reduced abundances of lignin and altered cell wall composition include enhanced saccharification efficiencies and higher ethanol conversion rates from tissue collected from sorghum *bmr* mutants, making them promising sources of low lignin biomass (Dien et al 2009; Sattler et al 2010a; Sattler et al 2012).

*Bmr6* encodes CAD, a member of the alcohol dehydrogenase superfamily that catalyzes the terminal step of the monolignol biosynthetic pathway, the reduction of *p*-hydroxycinnamyl aldehydes into their corresponding alcohols using NADPH as a cofactor (Saballos et al 2009; Sattler et al 2009). The previously characterized *bmr6-ref* mutant encodes a nonsense mutation, which truncates the peptide prior to the NADPH-binding and C-terminal catalytic domains (Sattler et al 2009). Further characterization of *bmr6-ref* plants revealed that Bmr6 proteins were not detectable in stalk tissues based on immunodetection and CAD activity levels were significantly reduced relative to wild type (Sattler et al 2009). The *bmr6-ref* mutation also resulted in reductions in Klason lignin and acid detergent lignin (ADL) (Dien et al 2009), the incorporation of cinnamaldehydes instead of alcohols into the growing lignin polymer (Bucholtz et al 1980; Palmer et al 2008), altered S/G ratios (Sattler et al 2009), and an improved extractability of lignin (Dien et al 2009). As a consequence of these modifications to lignin, enhancements in saccharification efficiency and ethanol recovery were observed in tissues

collected from *bmr6* plants (Dien et al 2009). Sorghum carrying the *bmr6-ref* mutation has also

been shown to have improved forage quality and neutral detergent fiber (NDF) digestibility

(Oliver et al 2004), *in vitro* dry matter disappearance (IVDMD), and *in vitro* cell wall constituent

disappearance (IVCWCD) (Porter et al 1978), which significantly improved fiber digestion

(Aydin et al 1999) and lactational performance in cattle (Oliver et al 2004).

Here, we characterize a series of nine novel *bmr6* mutant sorghum lines derived from an

EMS-mutagenized TILLING population. Unlike *bmr6-ref*, which is likely a null allele, most of

these novel *bmr6* alleles contain missense mutations, which provide new tools to investigate how

hypomorphic alleles impact CAD activity levels, biomass composition and lignin concentration,

as well as digestibility and biomass conversion rates for forage and bioenergy-based breeding

programs.

**RESULTS**

**Identification of novel *bmr6* alleles**

*bmr6* mutants derived from an EMS-mutagenized TILLING population were visually identified

by the presence of a tan to brown midrib (Figure 1), and complementation tests with *bmr6* tester

lines confirmed that the nine new *bmr* mutants were allelic to *bmr6* (Sattler et al. 2014). To

determine the mutations that were responsible for the *bmr6* phenotype, DNA from the coding

region of the *Bmr6* gene, which encodes CAD in sorghum, was PCR amplified and sequenced.

Five different transition mutations (C to T or G to A) within the *Bmr6* coding region were

identified, which are consistent with exposure to the mutagen EMS (Table 1). In the new *bmr6*

plants, the majority of the mutations identified were missense mutations that were predicted to

change evolutionarily conserved amino acids in CAD protein motifs involved in either substrate

binding, cofactor binding, or dimerization (Figure 2). The one exception was Q321stop, which

encodes a methionine residue in other plant species (e.g., *Linum usitatissimum*, *Prunus mume*, and *Medicago sativa*). This mutation was predicted to produce a C-terminally truncated protein missing the last 45 amino acids (Figure 2). Although these mutants were isolated independently from the M3 generation of the TILLING population, some of the lines identified in this study contained the same missense mutations in the *Bmr6* gene.

**Relative Expression of *bmr6* Alleles using qRT-PCR**

Quantitative real-time PCR (qRT-PCR) was used to determine the expression level of the different *bmr6* alleles (Figure 3). Similar expression levels of the *bmr6* alleles were observed relative to wild type. However, the *bmr6-ref* and *bmr6-971* alleles were expressed approximately 10-fold and 5-fold less than wild type. A reduction in expression was previously observed in *bmr6-ref*, and is likely due to nonsense-mediated mRNA decay. However, the reduced expression detected in *bmr6-971* plants was unexpected, because *bmr6-32* and *bmr6-741* also harbor P253L mutations and expressed *bmr6* at the approximately the same levels as wild type. Interestingly, *the bmr6-45* allele was expressed at the same level as wild type even though *bmr6-45* contains a nonsense mutation near the end of the coding region. This transcript may not be targeted for nonsense mediated decay because the nonsense mutation occurs near the end of the open reading frame of the transcript, unlike the *bmr6-ref* transcript where the nonsense mutation occurs in the middle of the transcript. Additionally, *bmr6-1277,* which contains the missense mutation A164V, had a 60% higher expression level compared to wild type.

**Detection of Bmr6 Protein via Western Blotting**

To examine the impact that these mutations had on CAD protein accumulation, protein extracts from *bmr6* and wild-type stalks were analyzed via western blotting using antibodies against Bmr6 and ascorbate peroxidase (loading control) (Figure 4). Bmr6 protein was nearly

undetectable in extracts from *bmr6-ref* and *bmr6-45* tissue, which both contained nonsense mutations. In addition, Bmr6 protein was nearly undetectable in *bmr6-23* and *bmr6-1103*, which both contain the same G184D missense mutation. However, no differences in mRNA expression levels were observed in *bmr6-23*, *bmr6-45*, and *bmr6-1103*. Additionally, the intensity of the Bmr6 band in *bmr6-1277* was slightly lower than wild type. Therefore, the increased mRNA expression levels observed in plants carrying the *bmr6-1277* allele may be a mechanism to compensate for the lower abundance of Bmr6 protein relative to WT in the stalk. Notably, the approximately 5-fold reduction in *bmr6-971* mRNA expression did not significantly impact Bmr6 protein accumulation. The remainder of the *bmr6* mutants accumulated protein at approximately the same levels as wild type, which indicates that these alleles did not have significant impacts on Bmr6 protein levels in stalk materials.

### *bmr6* mutations reduce CAD enzymatic activity

To ascertain the impact of these point mutations on CAD enzymatic activity, the corresponding *bmr6* cDNA sequences were cloned and expressed, and recombinant proteins were purified from *E. coli*. The recombinant wild-type protein purified from *E. coli* was slightly higher in molecular weight because it contained a leader sequence and an N-terminal His-tag (Figure 4). Recombinant proteins were successfully induced and purified for all *bmr6* mutations except G184D, which produced an insoluble protein. The inability to recover soluble protein from the construct containing this mutation is consistent with its absence on the western blot, which suggests that the protein is not being properly folded.

Recombinant proteins were assayed for CAD activity using two preferred substrates, coniferaldehyde and sinapaldehyde (Figure 5). Compared to wild type, all *bmr6* variants had substantially lower enzyme activities relative to wild type plants. A164V, G190S, and Q321stop

all displayed approximately 10-fold reductions in CAD activity relative to wild type proteins using either coniferaldehyde or sinapaldehyde as substrates. Interestingly, P253L had a 5-fold lower CAD activity level compared to wild type when coniferaldehyde was used as a substrate and a 65-fold lower activity using sinapaldehyde as a substrate, suggesting that this mutation impacts substrate preference. Even though Q321stop produces a truncated protein that is missing the last 45 amino acids, this recombinant protein still retained CAD activity at levels similar to the other *bmr6* point mutations.

**Impacts of *bmr6* point mutations on predicted protein structure**

All *bmr6* point mutations resulted in changes in amino acids in close proximity to the predicted NADPH cofactor binding site (Figure S1). To predict the impact that the missense and nonsense mutations had on the tertiary structure of CAD and to link these structural changes to the reductions in enzyme activity levels observed above, three-dimensional structures were modeled using I-TASSER. As expected, all of the *bmr6* point mutations were predicted to interfere with the ability to bind with the NADPH cofactor. For example, A164V (Figure S1B) replaces a relatively small residue with a much larger hydrophobic residue, potentially resulting in a steric clash in the NADPH binding pocket. The G184D mutation replaces an amino acid without a side-chain with a residue with a long, negatively charged side-chain in an environment with no obvious hydrogen bonding partner. Therefore, in order to accommodate this amino acid substitution, the backbone of the CAD protein would likely be significantly distorted, which could severely disrupt the structure of the protein. In both the G190S (Figure 6A) and P253L (Figure 6B) mutations, the substituted amino acids (serine and leucine) are predicted to extend into the NADPH binding pocket, creating steric interferences that likely interfere with binding of the substrate. In the case of G190S, the mutation occurs in the highly conserved

[188]GXGGV(L)[G193] motif (Saballos et al 2009) and replaces a residue lacking a side-chain with a polar side chain, would create a severe distortion in the predicted three dimensional structure that the *Arabidopsis* 2CF6 is not the top-scoring model. Finally, in the Q321stop mutation (Supplemental Figure 1D), the C-terminal region of the protein that shapes the NADPH binding pocket is missing; therefore, the interactions with NAPDH are predicted to be disrupted as the phosphate groups of NADPH would be exposed, creating an unfavorable interaction.

Additionally, these point mutations occur in close proximity to residues with key involvement in catalysis, zinc cofactor binding, and substrate binding and could disrupt these processes as well. For example, A164V is directly adjacent to Cys163, which forms the catalytic zinc binding site along with Cys47, His69, and Glu70. Thus, mutations in close proximity to these regions may interfere with zinc cofactor binding that are critical for enzyme functionality. G184D is distal to residues predicted to be for dimerization and mutations at or near these residues may interfere with CAD's ability to form a homodimer. Finally, the Q321stop mutant is missing the majority of the C terminal catalytic region, which could restrict its ability to reduce cinnamylaldehydes.

### *bmr6* mutants have reduced lignin concentration

To determine the impact that these novel *bmr6* alleles have on forage composition, fiber analysis was conducted to determine neutral detergent fiber (NDF), acid detergent fiber (ADF) and acid detergent lignin (ADL), which estimates the relative abundance of cellulose, hemicellulose and lignin; cellulose and lignin; and lignin, respectively (Table 2). All *bmr6* mutants had significantly lower ADL levels compared to wild type, which ranged from 59% (*bmr6-32*) to 87% (*bmr6-1277*) of the level observed in wild type. Additionally, many of these lines had significantly lower ADL levels than *bmr6-ref*, including *bmr6-971* (81%), *bmr6-32* (74%), and

*bmr6-45* (78%) of *bmr6-ref* level suggesting that a subset of these alleles might have a stronger impact on lignin biosynthesis or that they may harbor other mutations in the genome that influence the synthesis of lignin. Overall, modest impacts on ADL levels were observed in the *bmr6-23*, *bmr6-307*, *bmr6-1103*, and *bmr6-1277* lines, while the most significant reductions were observed in *bmr6-45* and *bmr6-32* plants, which had 62% and 59% of the wild-type ADL levels. Interestingly, several of the novel *bmr6* mutants also had slightly higher ADF levels relative to wild type, including *bmr6-741*. However, *bmr6-1277* had the most significant impact on NDF and ADF of any of the *bmr6* lines and was associated with a 115% and 108% of the levels detected in wild-type, respectively. Interestingly, *bmr6-32*, which harbors the same missense mutation as *bmr6-741* and *bmr6-971* (P253L), actually has a much lower ADF than wild type (86% of WT) and the other P253L mutants (*bmr6-741* and *bmr6-971*), suggesting that other mutations in *bmr6-32* may be impacting cell wall metabolism.

### *bmr6* mutations alter lignin subunit composition

The novel *bmr6* mutants also displayed reduced levels of guaiacyl (G-) and syringyl (S-) lignin relative to wild type as determined by thioacidolysis and GC-MS analysis of β-*O*-4 linked residues (Figure 7). At least 50% reductions in both S and G lignin were detected in all of the *bmr6* mutants, and the most significant impacts on S and G lignin levels were observed in *bmr6-ref* and *bmr6-307*, which displayed 13% and 20% of wild-type S lignin levels and 19% and 17% of wild-type G lignin levels, respectively. Levels of H-lignin were low in both wild-type and *bmr6* plants (Figure S2); however, the *bmr6* mutants tended to have slightly lower levels of H lignin relative to wild type plants. Overall, *bmr6-ref* had the lowest levels of H lignin. The slight impact of *bmr6* mutations on levels of H residues is consistent with the results of a

previous study in which *bmr6* plants in two different genetic backgrounds had marginally lower levels of H residues compared to wild-type plants (Palmer et al 2008).

***bmr6* biomass has increased glucose yields following enzymatic saccharification**

To assess the impact that these novel *bmr6* alleles had on the recovery of fermentable sugars from lignocellulosic materials, stover was enzymatically saccharified using cellulase and glucose yields were measured at 24 and 70 h.  Glucose yields from stover collected from *bmr6* plants were slightly elevated compared to wild type after 24 h, although yields from *bmr6-1103* were the most substantial (Table 3).  After 70 h, glucose yields were significantly higher compared to wild type from stover collected from all of the new *bmr6* mutants with the exceptions of *bmr6-23* (G184D), *bmr6-32* (P253L), and *bmr6-1277* (A164V).  The largest yields were observed from materials collected from *bmr6-741* and *bmr6-971* plants, which both harbored the P253L mutation.   High glucose yields were also obtained from *bmr6-45* (Q321stop), *bmr6-307* (G190S), and *bmr6-1103* (G184D).

**DISCUSSION**

C4 grasses, including sorghum, represent promising sources of renewable lignocellulosic materials that are well suited for uses as forage for livestock and biofeedstock for chemical precursors and biofuels production (Sarath et al 2008).  However, poor digestibility and low yields represent two major hindrances to successfully utilizing lignocellulosic materials for forage and biofuels (Jung and Allen 1995; Abramson et al 2010).  Therefore, significant research has been devoted to enhancing the forage and biofeedstock qualities of C4 grasses through biotechnological approaches and concerted breeding efforts (Sattler et al 2010b; Byrt et al 2011). The low lignin *brown midrib* mutants of C4 grasses overcome many of these limitations by improving digestibility and enzymatic access to fermentable sugars stored in plant cell walls

(Cherney et al 1991; Sarath et al 2008). In sorghum, the *Bmr6* gene encodes the major cinnamyl alcohol dehydrogenase (CAD) involved in lignin biosynthesis. Previously, mutations in *Bmr6* were associated with overall reductions in lignin and changes in lignin subunit composition, which improved saccharification and fermentation efficiencies over wild-type plants (Saballos et al 2008; Palmer et al 2008; Dien et al 2009).

In this study, we identified and characterized nine sorghum lines derived from an EMS-mutagenized TILLING population(Xin et al 2008; Xin et al 2009) that carry novel missense and nonsense mutations in *Bmr6*, which are associated with significant reductions in CAD activity levels. These novel alleles also allow us to investigate how missense and nonsense mutations affect CAD activity levels and the abundance and composition of lignin. These novel alleles had impacted these phenotypes to varying degrees relative to the previously characterized *bmr6-ref* allele. Notably, several of the lines impacted lignin composition and its relative abundance more significantly than *bmr6-ref*, which suggest that a subset of these novel alleles also have strong impacts on lignin biosynthesis.

All of the new point mutations described in this study significantly reduced CAD activity levels relative to the wild type protein in recombinant enzyme assays. These reductions in CAD activity are consistent with previous findings for *bmr6-ref* where CAD activity was undetectable in internodes (Palmer et al 2008), and significantly lower CAD activity levels were detected in *bmr6* stalks than wild type (Pillonel et al 1991). The point mutation P253L appeared to impact the substrate preference of Bmr6 as a 65-fold reduction in the ability to reduce sinapaldehyde was observed. However, despite this apparent effect on Bmr6's ability to reduce sinapaldehyde to sinapyl alcohol (which is the monolignol precursor of syringyl residues), no major differences in the abundance of S-residues nor the S/G ratio were noted in the three lines harboring this

mutation (*bmr6-32*, *bmr6-741*, and *bmr6-971*) relative to the other *bmr6* lines assessed in this study. This observation indicates that there may be differences in enzyme activity *in vivo* that allows the P253L Bmr6 protein to reduce sinapaldehyde or, alternatively, that residual enzyme activity from other CAD-like enzymes may have catalyzed the reduction of sinapaldehyde in plants carrying the P253L mutation. There are approximately 13 CAD-like genes in the sorghum genome that share amino acid similarity to *Bmr6*, but only *Bmr6* is a member of the CAD2 clade, whose involvement lignification has been well-documented throughout vascular plants (Saballos et al 2009; Sattler et al 2009). CAD2 proteins have a specific set of key amino acid residues required for efficient activity against cinnamyl aldehydes. The CAD-like proteins in the sorghum genome lack these key residues, and CAD-like proteins from other plant displayed lower activities against cinnamyl aldehydes substrates than CAD2 proteins (Youn et al 2006; Saathoff et al 2012). For example, SbCAD4 can also bind and reduce sinapaldehyde, although at a much lower efficiency than Bmr6 (Sattler et al 2009).

In addition, all point mutations presented herein are predicted to disrupt the Bmr6 protein structure and its ability to interact with its NADPH cofactor. The previously characterized *bmr6-ref* mutant contains a nonsense mutation at position 132, and does not produce a detectable gene product via western blot, indicating that the resulting protein is severely truncated and not functional (Sattler et al 2009). Two of the novel point mutations presented in this study also may cause severe structural deformities to the Bmr6 protein. For example, G184D occurred in close proximity to residues involved in homodimerization, and was predicted to distort the backbone of Bmr6 significantly. Accordingly, the protein was undetectable in plant extracts that harbored this point mutation, and we were unable to express the recombinant protein in *E. coli*, which suggests that the protein was likely misfolded and degraded in both organisms. Likewise, the

Bmr6 protein was also not detectable in *bmr6-45* plant extracts, which carries a nonsense mutation near end of the open reading frame and is predicted to truncate significant portion of the C-terminal domain required for catalytic activity and protein folding.  In both cases, *bmr6* mRNA levels were not substantially different from wild type, which indicates that degradation of mis-folded proteins is likely responsible for the absence of the Bmr6 protein in plants harboring these mutations.  Despite the severity of the predicted impact of these point mutations on the Bmr6 protein structure and its functionality, plants harboring these mutations did not necessarily have the lowest levels of S- and G- lignin among the *bmr6* lines tested in this study.  For example, while S- and G- lignin levels were the lowest in *bmr6-307* (G184D) relative to the rest of the *bmr6* lines, with the exception of *bmr6-ref*, the levels of S- and G- lignin in *bmr6-23*, which also harbors the same G184D point mutation, were not substantially different than S- and G- lignin subunit levels of many of the other *bmr6* lines.  Furthermore, these lines were not necessarily associated with the most severe reductions in ADL or alterations in lignin subunit composition indicating that even mutations that result in minor structural changes to Bmr6 can have significant impacts on lignin composition and abundance.

Additionally, many of the new *bmr6* lines had lower ADL levels compared to both wild type and *bmr6-ref* plants suggesting that the new alleles presented in this study significantly reduce the abundance of lignin.  In particular, the P253L mutation found in the *bmr6-32, bmr6-741*, and *bmr6-971* lines was associated with the most significant reduction in ADL of all the mutations assessed, which ranged from 28% to 41% lower than ADL levels of wild type plants and 9% to 26% lower than ADL levels of *bmr6-ref* plants.  Accordingly, the highest glucose yields were observed from stover collected from *bmr6-741* and *bmr6-971* lines following enzymatic saccharification.  Interestingly, *bmr6-1277*, which is associated with an A164V

mutation, had lower ADL levels, but higher NDF and ADF levels than wild type, suggesting that in addition to lower levels of lignin, these plants may also have higher abundances of cell wall polysaccharides, such as cellulose and hemicellulose. Although the higher abundances of these carbohydrates directly increases the amount of fermentable and digestible energy localized to the cell wall, plant materials collected from *bmr6-1277* had the highest ADL levels of any of the novel *bmr6* mutants and the lowest glucose yields following enzymatic saccharification. Therefore, the A164V mutation does not appear to reduce lignin levels sufficiently to enhance extraction of these fermentable sugars from stover. Although increased cellulose and hemicellulose deposition has not been previously associated with the *bmr6-ref* mutation (Dien et al 2009), increased deposition of cellulose and arabinoxylans has previously been observed in RNAi *CAD*-downregulated maize (Fornalé et al 2012), suggesting that other secondary cell wall biosynthetic pathways can compensate for reduced lignin synthesis in *CAD*-downregulated lines. Alternatively, the *bmr6* lines harboring the A164V mutations could harbor mutations at other loci in the genome that influence cellulose and hemicellulose biosynthesis. This scenario is more likely as it has been previously estimated that EMS treatment introduces a mutation approximately every 516 kb in the genome (Xin et al 2008).

In conclusion, these nine novel *bmr6* lines expand the diversity of *bmr6* alleles available, which can be utilized in breeding programs to reduce lignin concentration and improve biomass for forage and bioenergy uses. The *bmr6-ref* allele has already shown great promise in improving digestibility and enhancing ethanol yield from lignocellulosic materials harvested from sorghum. Notably, the reductions in lignin caused by the *bmr6-ref* mutation were not associated with any serious reductions in fitness. For example, even though lignin provides a protective barrier against insect pests and pathogens (Moura et al 2010), previous studies have

shown that *bmr6* plants were not more susceptible to several common fungal pathogens and insect pests than wild type plants and, in some cases, were more resistant (Funnell-Harris et al 2010; Funnell-Harris et al 2014; Dowd and Sattler 2015). Additionally, *bmr6-ref* was not associated with significant negative impacts on several agronomic traits, including stand count, time to 50% anthesis, height, and yield (Sattler et al 2014). Taken together, these findings suggest that mutations at this locus have strong impacts on lignin biosynthesis without severely reducing fitness, which makes *bmr6* a useful tool for biomass manipulation. Although *bmr6-ref* has been previously deployed for germplasm development and is currently utilized in several major breeding programs, many of the *bmr6* lines described herein were associated with much higher reductions in ADL and, therefore, may be useful for lowering lignin content beyond what has been observed in *bmr6-ref* plants.

## MATERIALS AND METHODS

### Growth conditions

Near-isogenic *bmr6*-reference plants were obtained by crossing N121 (*bmr6-ref)* to the recurrent parent BTx623 as described previously (Sattler et al 2009). Several novel *bmr6* mutants were obtained from an EMS-mutagenized TILLING population (Xin et al 2008; Xin et al 2009). Mutants were identified by the presence of a brown midrib and allelism to *bmr6-ref* was confirmed through complementation testing with a *bmr6-ref* tester line (Sattler et al 2014). All mutants described in this manuscript failed to complement the *bmr6* leaf midrib phenotype. The newly identified *bmr6* mutants were self-pollinated for three generations. Seeds were planted in a soil mixture consisting of a 4:1:1:1:1 ratio of soil to peat moss to vermiculite to perlite to sand

and were allowed to germinate in 28 cm pots under greenhouse conditions. Growth conditions were augmented with high pressure sodium lights on a 16h daylight cycle. Approximately five mL of fertilizer containing 11:15:11 N:P:K (Ferti-lome, Gardener's Special) were applied to each pot every seven days. Temperatures in the greenhouse were maintained at 29 to 30 °C during the daylight hours and 26 to 27 °C at night. Stalk tissue was harvested from 5-6 week old plants and frozen immediately in liquid nitrogen. Frozen stalk tissue was ground using a freezer mill (SPEX Sample Prep, Metuchen, NJ), and ultimately used for qRT-PCR and western blotting.

Plants were also grown to full maturity in 2009 at the University of Nebraska Field Laboratory in Ithaca, NE (Sharpsburg silty clay loam; fine, smectitic, mesic Type Argiudoll) in two 7.6-m rows spaced 76 cm apart. The soil was supplemented with nitrogen fertilizer prior to planting at 157 kg ha$^{-1}$. Plots were seeded with a precision vacuum planter and 240 seeds per plot were planted. Materials were planted on 21 May 2009 and 3.8 cm of supplemental irrigation was applied on 6 and 29 August 2009. Three individual plants were harvested at random from the center row of each block on 7 October 2009. Panicles were removed at maturity and tissue was harvested using a silage cutter modified for small pot use. Materials were dried at 50 °C, ground in a Wiley mill fitted with a 2 mm mesh screen (Arthur H. Thomas Co., Philadelphia, PA), ground in a cyclone mill fitted with a 1 mm mesh screen (Udy Corp., Fort Collins, CO), and stored for analysis. Field harvested materials collected were used for thioacidolysis, fiber analysis and enzymatic saccharification.

**DNA Sequence Analysis**

Sorghum DNA was extracted from leaf tissue using a cetyl-trimethyl-ammonium bromide-based DNA extraction buffer (Rogers and Bendich 1985), and DNA from the three coding regions of the *Bmr6* locus was PCR amplified as described previously (Sattler et al 2009). Amplification products were sequenced bidirectionally (Operon, Huntsville, AL) and sequences were assembled and analyzed using MacVector (version 10.0). DNA sequences were translated into amino acid sequences and were aligned to CAD sequences from other plants using ClustalW (version 2.0).

## Gene Expression Analysis Using Quantitative Real-Time PCR (qRT-PCR)

Total RNA was extracted from five-week-old, greenhouse-grown plant stalks from three individual plants per line. In brief, RNA was extracted from approximately one gram of ground stalk material as described previously (Suzuki et al 2004; Sattler et al 2009). The RNA was treated with RNase-free DNase (Promega, Madison, WI) following the manufacturer's protocol. DNase was inactivated by adding EDTA to a final concentration of 200 mM and the RNA was purified using the EZNA Plant RNA Kit (Omega Bio-Tek, Norcross, GA). One microgram of DNase treated RNA was used for cDNA synthesis with the iScript Reverse Transcription Supermix for qRT-PCR (Bio-Rad, Hercules, CA). Twenty µL qPCR reactions were performed in triplicate with 0.6 µM F primer, 0.6 µM R primer, 1 µL 1:5 diluted cDNA, and 10 µL SsoAdvanced SYBR Green Supermix (Bio-Rad) using the Bio-Rad CFX Connect Real Time System. Primers for *Bmr*6 (Sb04g005950.1; *Bmr*6F: 5′-GAGGTGCTCCAGTTCTGCG-3′ and *Bmr*6R: 5′- CAGCGCCTCGTTCACGTAC-3′) and elongation initiation factor 4A-I (*elf4A1*) (Sb04g003390.1; *elf4A1*F: 5′-GGCGAAGGATGGTTTCTAAG-3′; *elf4A1*R: 5′-GCTTGATACCCCCACATCTT-3′) were designed as described previously (Sattler et al 2009).

Thermal cycling parameters were as follows:  initial denaturation for 30s at 95 °C, 40 cycles of denaturation at 95 °C for 5s and annealing at 62 °C for 10s.  Primer specificities were confirmed by dissociation curve analysis of the amplified products, consisting of denaturation at 95 °C for 5 s, cooling to 65 °C for 15 s, and gradual heating at 0.01 °C/s to a final temperature of 95 °C. Relative expression was computed using the $\Delta\Delta C_t$ method using *elf4A1* for normalization.  No-template and no-RT controls were analyzed to verify the absence of DNA contamination. Relative expression values were statistically analyzed using ANOVA.  Pairwise comparisons among genotypes were performed using Tukey's Honest Significant Difference (HSD) test with an experiment-wise error rate of $\alpha = 0.05$ using R (version 3.1.1 for Linux).

**Western blotting**

Proteins were isolated as described previously using an extraction buffer containing protease inhibitor (Sigma-Aldrich, St. Louis, MO) (Sattler et al 2009).  Protein extracts were separated on a 12% SDS Bio-Rad criterion XT precast gel and transferred to a nitrocellulose membrane in a solution containing 10 mM Tris, 100 mM glycine, and 8% v/v methanol using a wet blot system. The transfer was conducted at 65V for 45 minutes.  The membrane was treated with Ponceau S stain to confirm that the proteins were successfully transferred and the membrane was subsequently blocked with 3% nonfat dry milk in TBST (TBS + 0.5% Tween 20) for one hour. 1:5,000 dilutions of Bmr6 (Cocalico Biologicals, Reamstown, PA) (Sattler et al 2009) and ascorbate peroxidase (control) primary polyclonal antibodies obtained from rabbit antisera were added to the membrane.  After one hour incubation period at room temperature, the membrane was washed three times in TBST to remove unbound primary antibody.  A 1:10,000 dilution of a secondary antibody (goat antirabbit IgG + horseradish peroxidase; Sigma-Aldrich A-0545) was

added and the membrane was incubated for 1 hour.  Unbound secondary antibody was removed by washing the membrane twice in TBST for 5 minutes and once in TBST+0.5M NaCl for 5 minutes.  The secondary antibody was detected using chemiluminescence with Amersham ECL Western blotting reagent (GE Healthcare).  Each lane on the gel represents protein extract collected from one individual plant.  Immunoblotting was performed on three individual plants from each line.  The blot displayed are representative of all three replicates that were analyzed.

**Bmr6 recombinant protein**

The *Bmr6* coding region was PCR amplified from cDNA using the primers *Bmr6*F: 5'-gatctgggtaccatggggagcctggcgtccgagaggggga-3' and *Bmr6*R: 5'-cgcaagctttcagttgctcggcgcatcagcgg-3'and the product was ligated into a pET30a cloning vector (Novagen, Madison, WI) as a *Kpn*I-*Hin*dIII fragment (Sattler et al 2009).  The point mutations identified in the novel *bmr6* alleles were introduced into the coding region through site directed mutagenesis using the QuickChange Site Directed Mutagenesis Kit (Agilent, Santa Clara, CA), and validated through commercial DNA sequencing (Operon, Huntsville, AL).  Recombinant vectors were introduced into competent *E. coli* Rosetta cells (Novagen, EMD Millipore, Billerica, MA).  Cultures derived from a single colony were grown to log phase at 37 °C and protein expression was induced by exposure to 0.1 M isopropyl β-D-1-thiogalactopyranoside.  Induced cells were incubated for 18 hours at 20 °C.  Soluble proteins were extracted from transformed cells via sonication at 20 W for three minutes alternating between a 10 second pulse followed by a 30 second rest period. The expressed recombinant proteins containing N-terminal His tags were purified via affinity nickel column chromatography (GE Healthcare, Little Chalfont, United Kingdom) and were eluted with 0.5M imidazole.  Protein induction was confirmed by SDS-PAGE.

**Cinnamyl alcohol dehydrogenase (CAD) Enzyme Assays**

CAD enzyme assays were performed on recombinant proteins to avoid interference from residual CAD activity associated with other alcohol dehydrogenases found in whole-plant protein extracts. Cinnamyl alcohol dehydrogenase (CAD) enzyme activity was monitored for each *bmr6* mutant using two different CAD substrates, 200 mM coniferaldehyde or 200 mM sinapaldehyde dissolved in 100 mM Tris-HCl (pH=8.8). Ten μL of each recombinant Bmr6 protein were suspended in a buffer consisting of 100 mM Tris HCl (pH = 7.5), 5 mM dithiothreitol, and 5% ethylene glycol. Dilutions of recombinant proteins were prepared immediately prior to use from concentrated aliquots that were stored at -80 °C. 100 mM 2-(N -morpholino)ethanesulfonic acid (MES; pH 6.5) and 200 μM NADPH were added to the reaction mixture and the final volume of each reaction was adjusted to 200 μL with deionized water. CAD activity was measured indirectly by monitoring the gradual reduction in absorbance at 340 nm as the co-factor NADPH was converted to $NADP^+$. The reaction was carried out in triplicate for each mutant at room temperature and was measured every 20 seconds for three minutes using a 96-well plate reader (SpetraMax Plus 384, Molecular Devices). The amount (μg) of Bmr6 recombinant protein added to each reaction was used to calculate specific CAD activity. Data were analyzed using KaleidaGraph 4.0 (Synergy Software) and SAS for Windows 9.1 (SAS Institute). The method of linear least squares (PROC REG) was used to fit a simple two parameter model to the optical density data, which provided an estimate of the $NADP^+$ production rate.

**Protein structural modeling and analysis**

The three-dimensional structural predictions of CAD containing the various missense and nonsense mutations detected in the TILLING population were generated using the iterative

threading assembly refinement (I-TASSER) server (Roy et al 2010). I-TASSER identified *Arabidopsis thaliana* CAD with bound NADPH (PDB ID: 2CF6) as the best template to generate a structural model for our wild-type sequence (Youn et al 2006). I-TASSER was also used to generate models for all of our mutants allowing us to visualize the impact of these mutations on the predictive modeling. For visualization, the mutant models were superimposed with 2CF6 using Chimera (Petersen et al 2004).

**Lignin composition**

Thioacidolysis followed by GC-MS was used to determine the relative concentration of guaiacyl (G), syringyl (S), and H (*p*-hydroxyphenol) subunits in the lignin polymer in the novel *bmr6* mutants. The samples were washed, derivitized, and analyzed as described previously (Palmer et al 2008) and 200 µL 12.5 mg/mL docosane in dichloromethane was added to each sample as an internal standard. H, G, and S subunits were identified based on mass spectra and quantified using the peak area of the major ions (m/z) (H 239; G 269, and S 299) (Palmer et al 2008). The peak area of the internal standard and the weights of the dried material were used for normalization. Analysis was performed in duplicate on three individual plants per line.

**Fiber analysis**

One hundred milligrams of dried stover were used for acid detergent lignin (ADL) analysis. Relative abundance of neutral detergent fiber (NDF), acid detergent fiber (ADF), and acid detergent lignin (ADL) were estimated using an ANKOM 200 fiber analyzer (ANKOM Tech. Corp., Fairport, NY) (Vogel et al 1999). Detergent hemicellulose was estimated as the difference between ADF and NDF and detergent cellulose as the difference between ADF and NDL. Analysis was performed in duplicate on four individual plants per line.

**Enzymatic saccharification**

Alcohol insoluble residues (AIR) were prepared from stover samples derived from field-grown plants as described in Sattler et al. (2014). Saccharification was performed in duplicate on three individual plants per line. Enzymatic saccharification of 300 mg of AIR from individual plants was performed based on NREL Laboratory Analytical Protocol 009 (Brown and Torget 1996) in 15-mL polypropylene tubes containing 10 mL 50 mM citrate buffer pH 4.8, 6 FPU Cellic CTec 3 (Novozymes, Franklinton, NC, USA) and 20 µg mL$^{-1}$ tetracycline (to prevent microbial growth). The tubes were incubated at 50 °C while shaking at 150 rpm. No-stover controls were used to correct for any glucose present in the enzyme cocktail. Samples (200 uL) were collected after 24 and 70 hours, boiled for 5 min. to inactivate the cellulases, and kept at 4 °C until glucose analysis. Samples were diluted four-fold in 50 mM citrate buffer pH 4.8, and 10 uL of this dilution was added to 140 uL ddH$_2$O. A volume of 150 uL Trinder reagent (Sekisui Diagnostics, Charlottetown, PEI, Canada) was added. This assay relies on the peroxidase-catalyzed formation of a red dye, quinoneimine, using H$_2$O$_2$ generated by the glucose oxidase-mediated conversion of glucose to gluconic acid. Glucose standards of known concentrations in 50 mM citrate buffer pH 4.8 were included in the same 96-well plate as the stover hydrolysates and used to generate a calibration curve. Absorbance at 505 nm was measured using a BioTek Synergy HT plate reader, 15 minutes after addition of the Trinder reagent.

**Statistical Analysis**

Statistical analysis of CAD enzyme activities, thioacidolysis results, saccharification results, and fiber data was performed using SAS (Cary, NC) PROC MIXED v.9.2 (SAS, 2002-2008). Heterogeneous variances were detected using Levene's test for homogeneity and were accounted

for with PROC MIXED. Pairwise comparisons among genotypes were performed using F-protected least significant differences with an experiment-wise error rate of $\alpha = 0.05$ (SAS).

## ACKNOWLEDGEMENTS

The U.S. Department of Agriculture (USDA) prohibits discrimination in all its programs and activities on the basis of race, color, national origin, age, disability, and where applicable, sex, marital status, familial status, parental status, religion, sexual orientation, genetic information, political beliefs, reprisal, or because all or part of an individual's income is derived from any public assistance program. (Not all prohibited bases apply to all programs.) Persons with disabilities who require alternative means for communication of program information (Braille, large print, audiotape, etc.) should contact USDA's TARGET Center at (202) 720-2600 (voice and TDD). To file a complaint of discrimination, write to USDA, Director, Office of Civil Rights, 1400 Independence Avenue, S.W., Washington, D.C. 20250-9410, or call (800) 795-3272 (voice) or (202) 720-6382 (TDD). USDA is an equal opportunity provider and employer.

## FOOTNOTE

**While this article was in press, it was revealed that the EMS-induced mutation identified in *bmr6-1277* matched the spontaneous *bmr6-41* mutation in sorghum accession IS21549, as reported by Gorthy et al. 2013.**

References

Abramson M, Shoseyov O, Shani Z (2010) Plant cell wall reconstruction toward improved lignocellulosic production and processability. **Plant Sci** 178: 61–72

Ali F, Scott P, Bakht J, et al (2010) Identification of novel brown midrib genes in maize by tests of allelism. **Plant Breed** 129: 724–726

Alvira P, Tomás-Pejó E, Ballesteros M, Negro MJ (2010) Pretreatment technologies for an efficient bioethanol production process based on enzymatic hydrolysis: A review. **Bioresour Technol** 101: 4851–4861

Aydin G, Grant RJ, O'Rear J (1999) Brown midrib sorghum in diets for lactating dairy cows. **J Dairy Sci** 82: 2127–2135

Boerjan W, Ralph J, Baucher M (2003) Lignin biosynthesis. **Annu Rev Plant Biol** 54: 519–546

Bonello P, Storer AJ, Gordon TR, et al (2003) Systemic effects of *Heterobasidion annosum* on ferulic acid glucoside and lignin of presymptomatic ponderosa pine phloem, and Potential Effects on Bark-Beetle-Associated Fungi. **J Chem Ecol** 29: 1167–1182.

Bout S, Vermerris W (2003) A candidate-gene approach to clone the sorghum brown midrib gene encoding caffeic acid *O*-methyltransferase. **Mol Genet Genomics** 269: 205–214

Brown, L Torget R (1996) Enzymatic saccharification of lignocellulosic biomass. Golden, CO

Bucholtz DL, Cantrell RP, Axtell JD, Lechtenberg VL (1980) Lignin biochemistry of normal and brown midrib mutant sorghum. **J Agric Food Chem** 28: 1239–1241

Byrt CS, Grof CPL, Furbank RT (2011) C4 plants as biofuel feedstocks: optimising biomass production and feedstock quality from a lignocellulosic perspective. **J Integr Plant Biol** 53: 120–135

Chen F, Dixon RA (2007) Lignin modification improves fermentable sugar yields for biofuel production. **Nat Biotechnol** 25: 759–761. doi: 10.1038/nbt1316

Chen W, VanOpdorp N, Fitzl D, et al (2012) Transposon insertion in a cinnamyl alcohol dehydrogenase gene is responsible for a *brown midrib1* mutation in maize. **Plant Mol Biol** 80: 289–97

Cherney JH, Cherney DJR, Akin DE, Axtell JD (1991) Advances in Agronomy Volume 46. **Adv Agron** 46: 157–198. doi: 10.1016/S0065-2113(08)60580-5

Dien BS, Sarath G, Pedersen JF, et al (2009) Improved sugar conversion and ethanol yield for forage sorghum (*Sorghum bicolor* L. Moench) lines with reduced lignin contents. **Bioenergy Res** 2: 153–164

Dowd PF, Sattler SE (2015) *Helicoverpa zea* (Lepidoptera: Noctuidae) and *Spodoptera frugiperda* (Lepidoptera: Noctuidae) responses to *Sorghum bicolor* (Poales: Poaceae) rissues from lowered lignin lines. **J Insect Sci** 15: 2

Fornalé S, Capellades M, Encina A, et al (2012) Altered lignin biosynthesis improves cellulosic bioethanol production in transgenic maize plants down-regulated for cinnamyl alcohol dehydrogenase. **Mol Plant** 5: 817–830

Funnell-Harris DL, Pedersen JF, Sattler SE (2010) Alteration in lignin biosynthesis restricts growth of *Fusarium spp.* in brown midrib sorghum. **Phytopathology** 100: 671–681

Funnell-Harris DL, Sattler SE, Pedersen JF (2014) Response of *Fusarium thapsinum* to sorghum brown midrib lines and to phenolic metabolites. **Plant Dis** 98: 1300–1308

Gorthy S, Mayandi K, Faldu D, Dalal M (2013) Molecular characterization of allelic variation in spontaneous brown midrib mutants of sorghum (*Sorghum bicolor* (L.) Moench). Mol Breed 31:795-803.

Jung HG, Allen MS (1995) Characteristics of plant cell walls affecting intake and digestibility of forages by ruminants. **J Anim Sci** 73: 2774–2790

Klinke HB, Thomsen AB, Ahring BK (2004) Inhibition of ethanol-producing yeast and bacteria by degradation products produced during pre-treatment of biomass. **Appl Microbiol Biotechnol** 66: 10–26

Li L, Hill-Skinner S, Liu S, et al (2015) The maize *brown midrib4* ( *bm4*) gene encodes a functional folylpolyglutamate synthase. **Plant J** 81: 493–504

Moura JCMS, Bonine CAV, de Oliveira Fernandes Viana J, et al (2010) Abiotic and biotic stresses and changes in the lignin content and composition in plants. **J Integr Plant Biol** 52:360–76

Oliver AL, Grant RJ, Pedersen JF, O'Rear J (2004) Comparison of *brown midrib-6* and *-18* forage sorghum with conventional sorghum and corn silage in diets of lactating dairy cows. **J Dairy Sci** 87: 637–644. doi: 10.3168/jds.S0022-0302(04)73206-3

Palmer NA, Sattler SE, Saathoff AJ, et al (2008) Genetic background impacts soluble and cell wall-bound aromatics in brown midrib mutants of sorghum. **Planta** 229: 115–127

Pillonel C, Mulder MM, Boon JJ, et al (1991) Involvement of cinnamyl-alcohol dehydrogenase in the control of lignin formation in *Sorghum bicolor* L. Moench. **Planta** 185:538–544

Poovaiah CR, Nageswara-Rao M, Soneji JR, et al (2014) Altered lignin biosynthesis using biotechnology to improve lignocellulosic biofuel feedstocks. **Plant Biotechnol J** 12:1163–1173

Porter KS, Axtell JD, Lechtenberg VL, Colenbrander VF (1978) Phenotype, fiber composition, and *in vitro* dry matter disappearance of chemically induced *brown midrib* (*bmr*) mutants of sorghum. **Crop Sci** 18:  205

Rogers SO, Bendich AJ (1985) Extraction of DNA from milligram amounts of fresh, herbarium and mummified plant tissues. **Plant Mol Biol** 5: 69–76

Roy A, Kucukural A, Zhang Y (2010) I-TASSER: a unified platform for automated protein structure and function prediction. **Nat Protoc** 5: 725–738

Ruel K, Berrio-Sierra J, Derikvand MM, et al (2009) Impact of CCR1 silencing on the assembly of lignified secondary walls in *Arabidopsis thaliana*. **New Phytol** 184: 99–113

Saathoff AJ, Hargrove MS, Haas EJ, et al (2012) Switchgrass PviCAD1: understanding residues important for substrate preferences and activity. **Appl Biochem Biotechno**l 168:1086–100. doi: 10.1007/s12010-012-9843-0

Saballos A, Ejeta G, Sanchez E, et al (2009) A genomewide analysis of the cinnamyl alcohol dehydrogenase family in sorghum [*Sorghum bicolor* (L.) Moench] identifies *SbCAD2* as the *brown midrib6* gene. **Genetics** 181: 783–795

Saballos A, Sattler SE, Sanchez E, et al (2012) *Brown midrib2* (*Bmr2*) encodes the major 4-coumarate:coenzyme A ligase involved in lignin biosynthesis in sorghum (*Sorghum bicolor* (L.) Moench). **Plant J** 70: 818–830

Saballos A, Vermerris W, Rivera L, Ejeta G (2008) Allelic association, chemical characterization and saccharification properties of brown midrib mutants of sorghum (*Sorghum bicolor* (L.) Moench). **Bioenergy Res** 1: 193–204

Sarath G, Mitchell RB, Sattler SE, et al (2008) Opportunities and roadblocks in utilizing forages and small grains for liquid fuels. **J Ind Microbiol Biotechnol** 35: 343–354

Sattler SE, Funnell-Harris DL, Pedersen JF (2010a) Efficacy of singular and stacked *brown midrib 6* and *12* in the modification of lignocellulose and grain chemistry. **J Agric Food Chem** 58: 3611–3616. doi: 10.1021/jf903784j

Sattler SE, Funnell-Harris DL, Pedersen JF (2010b) Brown midrib mutations and their importance to the utilization of maize, sorghum, and pearl millet lignocellulosic tissues. **Plant Sci** 178: 229–238

Sattler SE, Palmer NA, Saballos A, et al (2012) Identification and characterization of four missense mutations in *brown midrib 12* (*Bmr12*), the caffeic o-methyltranferase (COMT) of sorghum. **Bioenergy Res** 5: 855–865

Sattler SE, Saathoff AJ, Haas EJ, et al (2009) A nonsense mutation in a cinnamyl alcohol dehydrogenase gene is responsible for the sorghum brown midrib6 phenotype. **Plant Physiol** 150: 584–595

Sattler SE, Saballos A, Xin Z, et al (2014) Characterization of novel sorghum brown midrib mutants from an EMS-mutagenized population. **G3 (Bethesda)** 4: 2115-2124

Siegrist J, Jeblick W, Kauss H (1994) Defense responses in infected and elicited cucumber (*Cucumis sativus L.*) hypocotyl segments exhibiting acquired resistance. **Plant Physiol** 105: 1365–1374

Suzuki Y, Tetsu K, Hiroyuki K (2004) Isolation from siliques, dry seeds, and other tissues of *Arabidopsis thaliana*. **Biotechniques** 37: 542–544

Tang HM, Liu S, Hill-Skinner S, et al (2014) The maize *brown midrib2* (*bm2*) gene encodes a methylenetetrahydrofolate reductase that contributes to lignin accumulation. **Plant J** 77: 380–392

Vanholme R, Cesarino I, Rataj K, et al (2013) Caffeoyl shikimate esterase (CSE) is an enzyme in the lignin biosynthetic pathway in *Arabidopsis*. **Science** 341:1103–1106. doi: 10.1126/science.1241602

Vermerris W, Abril A (2015) Enhancing cellulose utilization for fuels and chemicals by genetic modification of plant cell wall architecture. **Curr Opin Biotechnol** 32C: 104–112

Vignols F, Rigau J, Torres MA, et al (1995) The *brown midrib3* (*bm3*) mutation in maize occurs in the gene encoding caffeic acid O-methyltransferase. **Plant Cell** 7: 407–416

Vogel KP, Pedersen JF, Masterson SD, Toy JJ (1999) Evaluation of a filter bag system for NDF, ADF, and IVDMD forage analysis. **Crop Sci** 39: 276

Whetten R, Sederoff R (1995) Lignin biosynthesis. **Plant Cell** 7: 1001–1013

Wilson J, Kennedy P (1996) Plant and animal constraints to voluntary feed intake associated with fibre characteristics and particle breakdown and passage in ruminants. **Aust J Agric Res** 47: 199

Ximenes E, Kim Y, Mosier N, et al (2010) Inhibition of cellulases by phenols. **Enzyme Microb Technol** 46: 170–176

Xin Z, Wang ML, Barkley NA, et al (2008) Applying genotyping (TILLING) and phenotyping analyses to elucidate gene function in a chemically induced sorghum mutant population. **BMC Plant Biol** 8:103

Xin Z, Wang ML, Burow G, Burke J (2009) An induced sorghum mutant population suitable for bioenergy research. **Bioenergy Res** 2:10–16

Youn B, Camacho R, Moinuddin SGA, et al (2006) Crystal structures and catalytic mechanism of the *Arabidopsis* cinnamyl alcohol dehydrogenases *AtCAD5* and *AtCAD4*. **Org Biomol Chem** 4:1687–1697

Zeng Y, Zhao S, Yang S, Ding S-Y (2014) Lignin plays a negative role in the biochemical process for producing lignocellulosic biofuels. **Curr Opin Biotechnol** 27: 38–45

**Tables**

**Table 1. Nonsense and missense mutations detected at the *Bmr6* gene**

Mutants with matching superscripts harbor the same point mutation at the *bmr6* locus. Nucleotide and amino acid changes referenced below correspond to positions in the *Sorghum bicolor* CAD2 gene (*Sorghum bicolor* v2.1 assembly: Sobic.004G071000; GenBank accession FJ554574). [*]*bmr6-ref* was previously identified in (Saballos et al 2009; Sattler et al 2009).

| *bmr6* mutant | Mutation (nucleotide position) | Amino acid change |
|---|---|---|
| *bmr6-ref*[*] | C3376T | Gln132stop |
| *bmr6-1277*[A] | C4195T | Ala164Val |
| *bmr6-23*[B] | G4255A | Gly184Asp |
| *bmr6-1103*[B] | | |
| *bmr6-31*[C] | G4275A | Gly190Ser |
| *bmr6-307*[C] | | |
| *bmr6-32*[D] | C4462T | Pro253Leu |
| *bmr6-741*[D] | | |
| *bmr6-971*[D] | | |
| *bmr6-45*[E] | C4665T | Gln321stop |

**Table 2.  Fiber analysis of the *bmr6* mutants**

NDF:  Neutral detergent fiber; ADF:  Acid detergent fiber; ADL:  acid detergent lignin.  Values with conflicting letters are significantly different via F-protected LSD test at an experiment-wise error rate of α=0.05.  Values in bold are significantly different from BTx623 (wild type) at α=0.05.  Mutants with matching superscripts harbor the same point mutation at the *bmr6* locus.

| Line | NDF (%) | | ADF (%) | | ADL (%) | |
|---|---|---|---|---|---|---|
| | Mean | Standard Error | Mean | Standard Error | Mean | Standard Error |
| BTx623 | 61.91 (c) | 0.81 | 37.02 (bc) | 0.65 | 5.68 (a) | 0.12 |
| *bmr6-ref* | **65.46 (b)** | 0.96 | 38.43 (ab) | 0.96 | **4.53 (bc)** | 0.22 |
| *bmr6-1277*[A] | **71.03 (a)** | 0.94 | **39.95 (a)** | 0.49 | **4.92 (b)** | 0.21 |
| *bmr6-23*[B] | 62.23 (c) | 1.34 | **33.62 (ef)** | 0.98 | **4.49 (bc)** | 0.28 |
| *bmr6-1103*[B] | 62.39 (c) | 0.79 | 35.83 (cd) | 0.72 | **4.38 (bc)** | 0.31 |
| *bmr6-31*[C] | 61.39 (c) | 1.13 | **33.44 (ef)** | 0.71 | **4.18 (cd)** | 0.15 |
| *bmr6-307*[C] | 59.62(cd) | 0.70 | **31.44 (g)** | 0.62 | **4.68 (bc)** | 0.12 |
| *bmr6-32*[D] | **58.32 (d)** | 0.36 | **31.78 (fg)** | 0.34 | **3.34 (f)** | 0.11 |
| *bmr6-741*[D] | 62.04(c) | 0.55 | **34.58 (de)** | 0.44 | **4.11 (cde)** | 0.28 |
| *bmr6-971*[D] | 61.95 (c) | 1.06 | **34.70 (de)** | 0.69 | **3.68 (def)** | 0.15 |
| *bmr6-45*[E] | **56.92 (d)** | 1.49 | **31.49 (g)** | 0.99 | **3.53 (ef)** | 0.34 |

**Table 3. Enzymatic saccharification of *bmr6* biomass.** Glucose yields (mg/g stover) after enzymatic saccharification at 50 °C at 60 FPU/g cellulose of unpretreated stover after 24 and 72 hours. Means in the same column with different letters are significantly different based on a Fisher's F-protected LSD with an experiment-wise error rate of α=0.05. Mutants with matching superscripts harbor the same point mutation at the *bmr6* locus.

| Glucose yield (mg/g stover) | | | | |
|---|---|---|---|---|
| | 24 hours | | 70 hours | |
| Line | Mean | Standard Error | Mean | Standard Error |
| BTx623 | 78.7 (ab) | 12.0 | 95.1 (ab) | 15.6 |
| *bmr6-ref* | 91.2 (bc) | 7.7 | 119.3 (c) | 10.5 |
| *bmr6-1277*[A] | 83.8 (ab) | 14.0 | 112.0 (bc) | 17.5 |
| *bmr6-23*[B] | 86.0 (bc) | 7.8 | 112.1 (bc) | 19.8 |
| *bmr6-1103*[B] | 99.1 (c) | 11.3 | 119.8 (c) | 8.8 |
| *bmr6-31*[C] | 87.6 (bc) | 11.9 | 115.1 (c) | 11.0 |
| *bmr6-307*[C] | 89.4 (bc) | 7.4 | 120.7 (c) | 6.0 |
| *bmr6-32*[D] | 95.7 (abc) | 20.5 | 111.4 (abc) | 25.7 |
| *bmr6-741*[D] | 90.8 (bc) | 6.8 | 124.8 (c) | 9.6 |
| *bmr6-971*[D] | 94.4 (bc) | 18.9 | 127.5 (c) | 7.6 |
| *bmr6-45*[E] | 84.7 (b) | 5.6 | 120.0 (c) | 5.8 |

## Figure Legends

### Figure 1.  Images of midribs of 6-week old wild type and *bmr6* plants

Nine lines carrying mutations at the *bmr6* locus were identified from this population and allelism to *brm6* was confirmed through complementation testing.

### Figure 2.  Amino acid sequence alignment of predicted cinnamyl alcohol dehydrogenease proteins surrounding the *bmr6* mutations

The predicted amino acid sequence of Bmr6 was aligned to CAD proteins from a variety of other plant species using ClustalW (version 2.0).  Bmr6 *Sorghum bicolor*:  ACL80888, SoCAD *Saccharum officinarum*:  CAA13177, Bm1 *Zea mays*:  AFW70530, SiCAD *Setaria italic*: XP_004951629, PvCAD *Panicum virgatum*:  XP_004951629, LuCAD *Linum usitatissimum*: ABF48495, ObCAD *Oryza brachyantha*:  XP_006647011, FaCAD *Festuca arundinacea*: AAK97809, LpCAD *Lolium perenne*:  AAB70908, BdCAD *Brachypodium distachyon*: XP_003570974, PmCAD *Prunus mume*:  XP_003570974, EpCAD *Eucalyptus pilularis*: BAM05553, VvCAD *Vitis vinifera*:  XP_002285406, GmCAD *Glycine max*:  XP_003555961, MsCAD *Medicago sativa*:  CAA79625, NtCAD *Nicotiana tabacum*:  CAA79625, PaCAD *Picea abies*:  CAI30877, and PtCAD *Populus tomentosa*:  AGU43755.  Consensus amino acid sequence is indicated below the alignment and point mutations analyzed in this study are denoted with an asterisk.  Dark gray shading indicates conserved amino acids while light grey shading denotes similar amino acids.  Residues with predicted involvement in substrate binding or catalysis are underlined.

**Figure 3.  Relative RNA expression of *Bmr6* in stalks collected from 6-week old sorghum plants**

Total RNA was isolated from wild type and *bmr6* stalks and qRT-PCR was used to measure *Bmr6* RNA expression.  Expression levels were determined using the ΔΔCt method and *Bmr6* gene expression was relativized against WT levels.  Means with conflicting letters are different at p=0.05 using a Tukey HSD test (n=3).  Error bars depict standard error.

**Figure 4.  Immunoblot detection of Bmr6 from stalks collected from 6-week old sorghum plants**

Protein extracts obtained from wild type (WT) and plants containing a mutation at the *Bmr6* locus were separated by SDS-PAGE and were probed with rabbit polyclonal antibodies that target the Bmr6 protein (top band).   Polyclonal antibodies against ascorbate peroxidase were used as a loading and transfer control (bottom band).   RP:  recombinant CAD protein.  Asterisks denote the absence of Bmr6 protein.

**Figure 5.  Cinnamyl alcohol dehydrogenase (CAD) activities of wild-type and mutated Bmr6 proteins**

CAD activity levels were assayed using 200 mM of coniferaldehyde and sinapaldehyde substrates.  The enzyme velocity was normalized to the amount (mg) of recombinant protein added to the reaction.  A164V represents *bmr6-1277*, G184D represents *bmr6-23* and *bmr6-1103*, G190S represents *bmr6-31* and *bmr6-307*, P253L represents *bmr6-32*, *bmr6-741*, and *bmr6-971*, and Q321stop represents *bmr6-45*.  Recombinant protein containing G184D was not

assayed for CAD activity because *E. coli* clones did not produce protein.  Means with conflicting letters are different at p=0.05 using an F-protected LSD test.  Error bars depict standard error.

**Figure 6.  Disruptions to Bmr6 protein structure caused by G190S (A) and P253L (B) point mutations**

The impact of the point mutations on Bmr6 structure was modeled using iTASSER.  In panel A, the protein structure of the G190S mutant protein (orange) is overlaid over the wild type Bmr6 protein (blue).  The much larger serine residue represented as pink ball and stick protrudes into the binding pocket and likely inhibits NADPH binding.  In panel B, the mutant P253L Bmr6 protein is represented as a yellow ribbon overlaid on the wild type protein (blue ribbon).  In the P253L mutant, the leucine residue represented by pink ball and stick replaces the proline (green ball and stick), extends into the NADPH binding pocket and likely interferes with cofactor binding.

**Figure 7.  Abundance of syringyl and guaiacyl lignin in wild type and *bmr6* plants**

The abundance of S and G lignin from *bmr6* and wild type plants was determined via thioacidolysis and GC-MS analysis of β-*O*-4 linked residues (S lignin: black bars; G lignin:  grey bars).  The ratio of S/G lignin for each sample is displayed beneath the bar graph.  Statistical comparisons were performed using an F-protected LSD test.  Error bars represent standard error (n=3).  Mutants denoted a and b differ at α=0.05 from wild type and *bmr6-ref*, respectively.  See Materials and Methods for more information.

**SUPPORTING INFORMATION**

**Figure S1.  Predicted disruptions in Bmr6 protein structures caused by point mutations**

In panel A, the Q132stop mutant (orange) is overlaid on the blue wild type Bmr6 protein.  The mutant is completely missing the NADPH binding domain and interactions with its cofactor are lost. In panel B, the A164V mutant (yellow) is overlaid on the blue wild type protein.  Unlike the alanine in green, the valine in pink extends into the NADPH binding pocket, potentially inhibiting interactions with this cofactor.  In panel C, the G184D mutation places the much larger aspartic acid (pink) in the middle of a hydrophobic pocket where steric clashes occur. The adjacent large nonpolar residues are shown in blue ball and stick.  For reference, the NADPH binding pocket is shown on the right. In panel D, the Q321stop is represented by a green ribbon. In this mutant, the majority of the C terminal domain is severely truncated and leaves the phosphate groups of the NADPH cofactor exposed.  The model indicates that favorable polar interactions are lost between Q340 and H48 (blue) of the full length protein and NADPH. Also, a steric clash with H48 (pink) of the truncated protein and NADPH would be predicted by the model.

**Figure S2.  Abundance of *p*-hydroxyphenol (H) lignin in wild type and *bmr6* plants**

The abundance of H lignin from wild type and bmr6 plants was determined via thioacidolysis and GC-MS analysis of β-*O*-4 linked residues.  The levels of H lignin were low in all *bmr6* plants, but did not deviate significantly from wild type levels.  No major differences between samples were detected using an F-protected LSD test.  Error bars represent standard error (n=3).
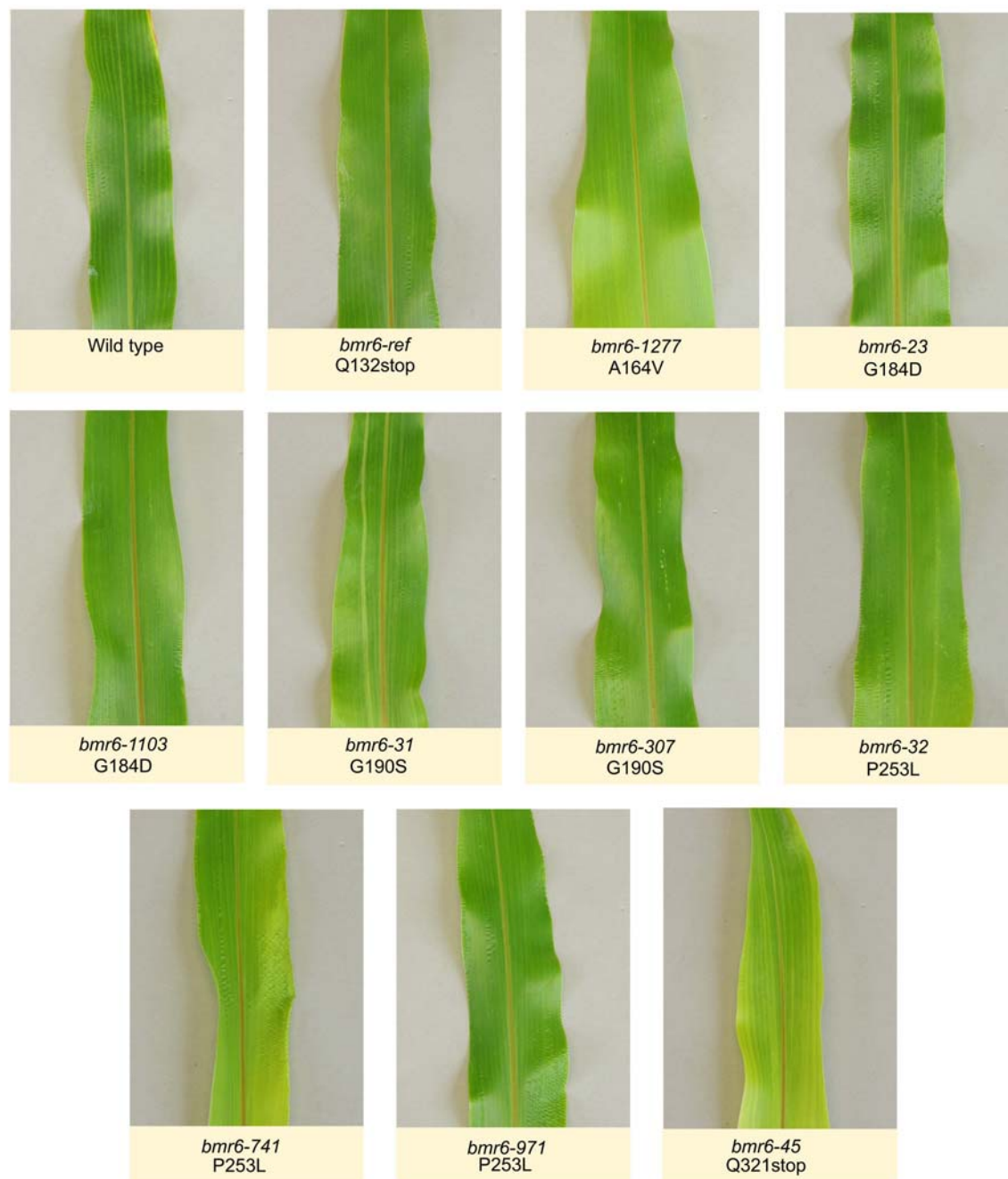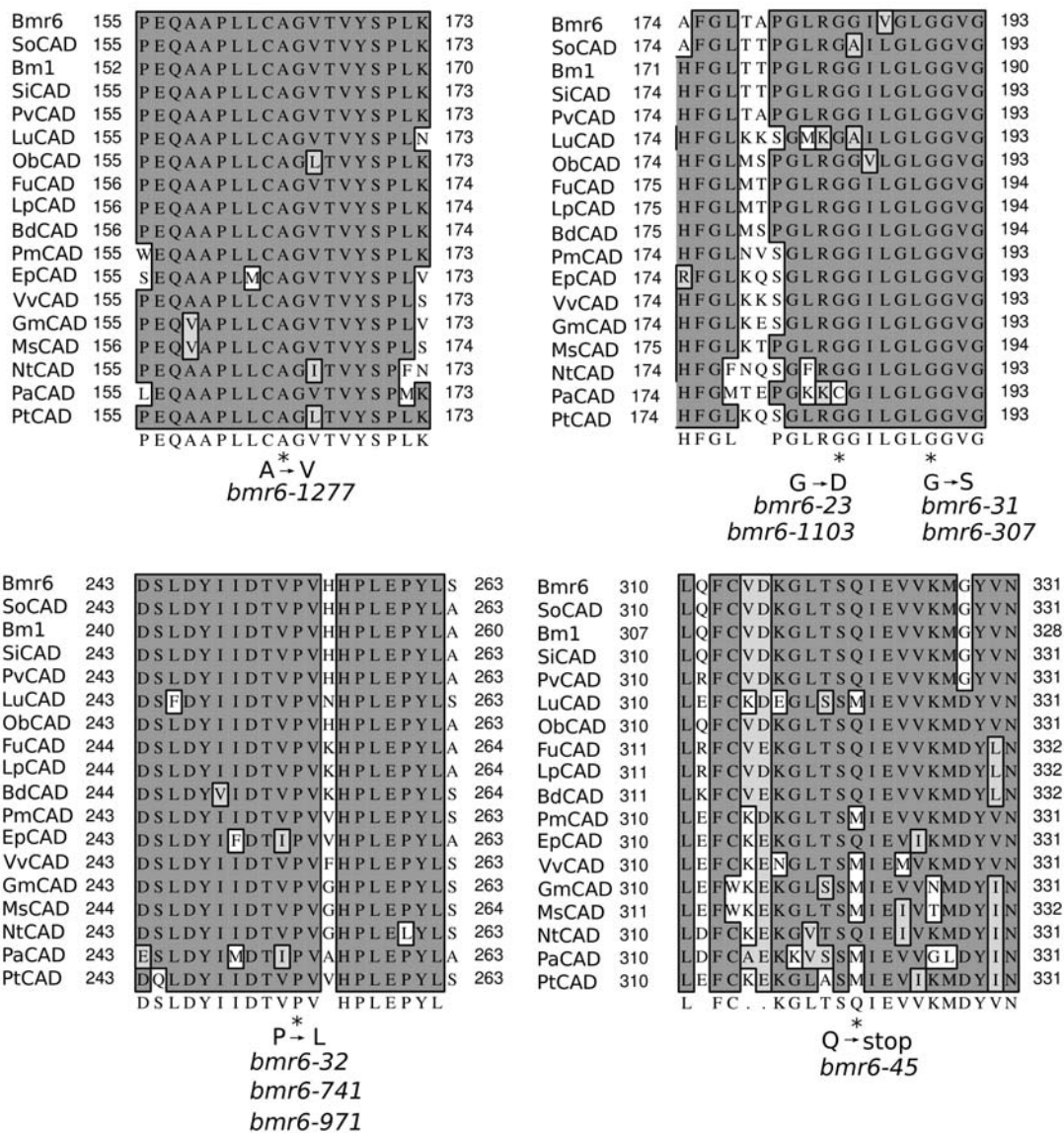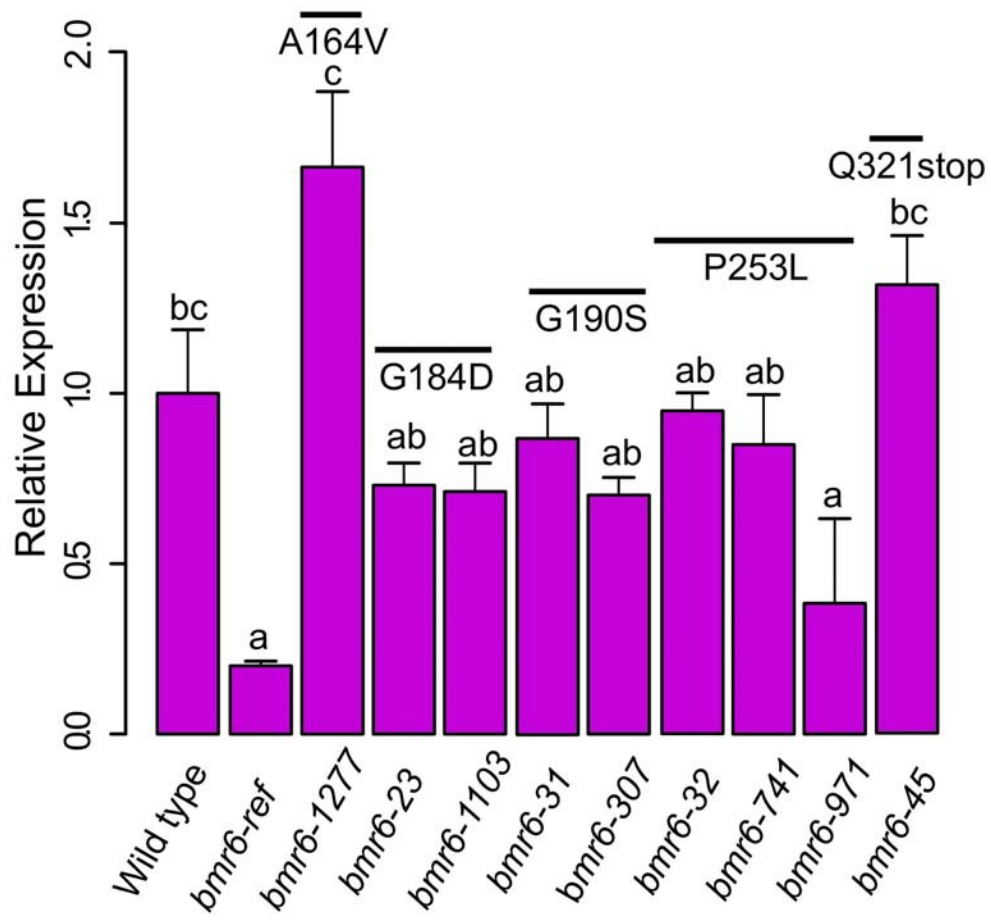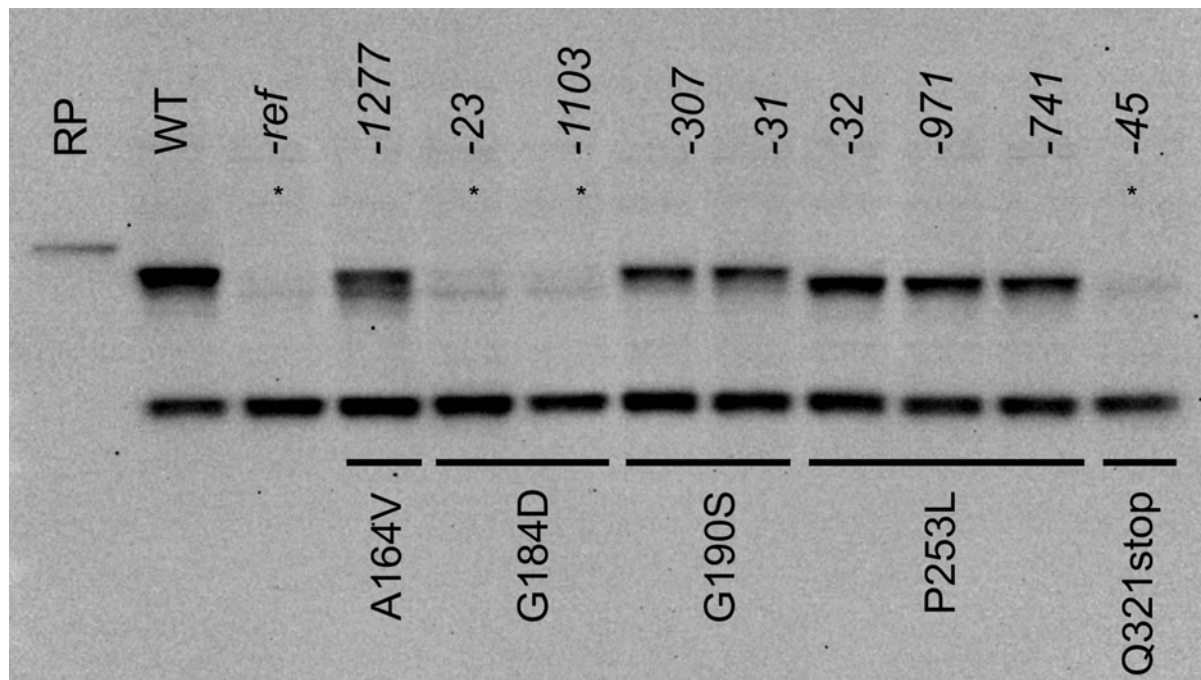
**Figure 1**

Figure 2

**Figure 3**

**Figure 4**
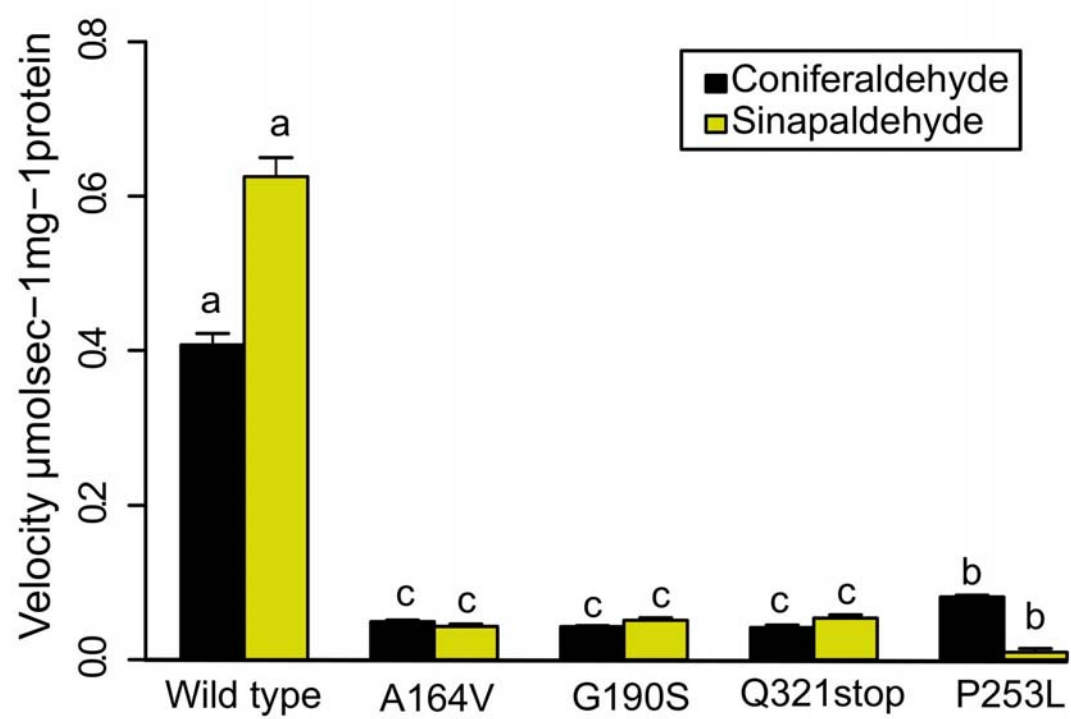
**Figure 5**

Figure 6

**Figure 7**