



5-31-2016

## Transgenic Plant-Produced Hydrolytic Enzymes and the Potential of Insect Gut-Derived Hydrolases for Biofuels

Jonathan D. Willis  
*University of Tennessee, Knoxville*

Mitra Mazarei  
*University of Tennessee, Knoxville*

C. Neal Stewart Jr.  
*University of Tennessee, Knoxville, nealstewart@utk.edu*

Follow this and additional works at: [https://trace.tennessee.edu/utk\\_planpubs](https://trace.tennessee.edu/utk_planpubs)

 Part of the [Plant Sciences Commons](#)

---

### Recommended Citation

Willis JD, Mazarei M and Stewart C N Jr. (2016) Transgenic Plant-Produced Hydrolytic Enzymes and the Potential of Insect Gut-Derived Hydrolases for Biofuels. *Frontiers in Plant Science* 7. doi: 10.3389/fpls.2016.00675

This Article is brought to you for free and open access by the Plant Sciences at TRACE: Tennessee Research and Creative Exchange. It has been accepted for inclusion in Plant Sciences Publications and Other Works by an authorized administrator of TRACE: Tennessee Research and Creative Exchange. For more information, please contact [trace@utk.edu](mailto:trace@utk.edu).



# Transgenic Plant-Produced Hydrolytic Enzymes and the Potential of Insect Gut-Derived Hydrolases for Biofuels

Jonathan D. Willis<sup>1,2</sup>, Mitra Mazarei<sup>1,2</sup> and C. Neal Stewart Jr.<sup>1,2\*</sup>

<sup>1</sup> Department of Plant Sciences, University of Tennessee, Knoxville, TN, USA, <sup>2</sup> Oak Ridge National Laboratory, BioEnergy Science Center, Oak Ridge, TN, USA

## OPEN ACCESS

### Edited by:

Huanzhong Wang,  
University of Connecticut, USA

### Reviewed by:

Holger Klose,  
RWTH Aachen University, Germany  
Elizabeth E. Hood,  
Arkansas State University, USA

### \*Correspondence:

C. Neal Stewart Jr.  
nealstewart@utk.edu

### Specialty section:

This article was submitted to  
Plant Biotechnology,  
a section of the journal  
Frontiers in Plant Science

**Received:** 03 March 2016

**Accepted:** 02 May 2016

**Published:** 31 May 2016

### Citation:

Willis JD, Mazarei M and Stewart CN Jr. (2016) Transgenic Plant-Produced Hydrolytic Enzymes and the Potential of Insect Gut-Derived Hydrolases for Biofuels. *Front. Plant Sci.* 7:675. doi: 10.3389/fpls.2016.00675

Various perennial C4 grass species have tremendous potential for use as lignocellulosic biofuel feedstocks. Currently available grasses require costly pre-treatment and exogenous hydrolytic enzyme application to break down complex cell wall polymers into sugars that can then be fermented into ethanol. It has long been hypothesized that engineered feedstock production of cell wall degrading (CWD) enzymes would be an efficient production platform for exogenous hydrolytic enzymes. Most research has focused on plant overexpression of CWD enzyme-coding genes from free-living bacteria and fungi that naturally break down plant cell walls. Recently, it has been found that insect digestive tracts harbor novel sources of lignocellulolytic biocatalysts that might be exploited for biofuel production. These CWD enzyme genes can be located in the insect genomes or in symbiotic microbes. When CWD genes are transformed into plants, negative pleiotropic effects are possible such as unintended cell wall digestion. The use of codon optimization along with organelle and tissue specific targeting improves CWD enzyme yields. The literature teaches several important lessons on strategic deployment of CWD genes in transgenic plants, which is the focus of this review.

**Keywords:** lignocellulosic, insect cellulase, cell wall degrading (CWD) enzyme, biofuel, hydrolase, gene expression, enzyme targeting

## INTRODUCTION

The natural recalcitrance of plant cell walls is a major commercial hurdle for cellulosic biofuel production. Two economic barriers of biofuel production are pretreatment and hydrolytic enzymes (Wyman, 2007; Furtado et al., 2014). It has been hypothesized that using a transgenic plant vehicle for the production of hydrolytic enzymes would improve cellulosic biofuel economics. Thus far, plant species transformed to overproduce hydrolytic enzymes are not bioenergy feedstock crops, but rather, easily transformed species such as *Arabidopsis thaliana*, *Nicotiana tabacum* (tobacco), and *Oryza sativa* (rice) (Taylor et al., 2008). All of these models are annual plants, and therefore may not be entirely predictive of enzyme activity perennial crops forecasted to be the most useful for cellulosic bioenergy.

Plant cell walls are composed of three major types of polymers: cellulose, hemicelluloses, and lignin. Each of these polymers are cross-linked and/or intertwined with another type to provide stereochemical strength (Cosgrove, 2005). Plant cell walls have successfully been broken down

using mechanical and chemical means, however these processes can release metabolites that can inhibit fermentation. Furfural is one such metabolite that inhibits the growth of yeast (Boyer et al., 1992; Zhang K. et al., 2011). A mixture of mechanical, chemical, and enzymatic means might be required to reduce cell wall recalcitrance in lignocellulosic biomass processing. Enzymes represent one of the highest material costs in lignocellulosic processing (Taylor et al., 2008; Klein-Marcuschamer et al., 2012).

Synergistic action of multiple enzymes are required for complete cell wall degradation. Four-to-five classes of enzymes are required to break down cellulose into its smaller cellobiose sub-units (Bhat, 2000; Lopez-Casado et al., 2008), which is, in turn, composed of two glucose units connected by a  $\beta$ -1,4 linkage. The tightly-packed crystalline structure of cellulose prevents enzyme accessibility. Hemicelluloses have variable polymeric structures composed of pentose sugars and a variety of side chains, which necessitates a suite of enzymes for hydrolysis (Saha, 2003; Scheller and Ulvskov, 2010).

Before the advent of rapid DNA sequencing technologies, glycosyl hydrolase (GH) enzymes were categorized based on their substrate specificity. Substrate specificity categorization was inadequate with discovery of enzymes with multiple catalytic domains and/or multiple substrate adherence (Henrissat, 1991). The previous substrate specificity classification system was updated and GHs were re-grouped into families (Henrissat and Bairoch, 1993, 1996). The CAZy database (CAZy.org) provides tools for classification and defining novel cellulolytic enzymes (Lombard et al., 2014).

Enzymes used for industrial cell wall hydrolysis are prepared from mesophilic fungi and bacteria. Optimal enzyme activity in such cases is typically 50°C and a pH range of 5.0–7.0 (Bruce and Palfreyman, 1998; Bhalla et al., 2013). Enzymatic degradation of cellulose to glucose requires three enzyme classes: endo-glucanases, cellobiohydrolases, and  $\beta$ -glucosidases (**Box 1**). Endo-glucanases (EG) act randomly on internal portions of the cellulose polymer creating smaller chains. Cellobiohydrolases (CBH) (also reported as exo-glucanases) cleave these smaller units at the end portion into cellobiose units.  $\beta$ -glucosidases (BG) cleave cellobiose units into individual glucose monomers. For efficient degradation of cellulose, each of these enzymes are required (Kostylev and Wilson, 2012).

**Abbreviations:** AU, absorbance unit; AX, arabinoxylan; AZCL-xylan, azo-crosslinked-xylan; AZCL, azo-crosslinked; BG,  $\beta$ -glucosidase; BMCC, bacterial microcrystalline cellulose; *CaMV35S*, cauliflower mosaic virus 35S promoter; CBH, cellobiohydrolase; CBM, carbohydrate binding module; CMC, carboxymethylcellulose; *CMVar*, cucumber mosaic virus advanced replicating promoter; CWD, cell wall degrading enzyme; FU, fluorescence units; EG, endo-1,4- $\beta$ -glucanase; ER, endoplasmic reticulum; GH, glycosyl hydrolase; GHF, glycosyl hydrolase family; *GlutB-1*, rice glutelin endosperm promoter; HEPES, 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid; *Hor2-4*, barley B1 hordein promoter; *Mac*, promoter combination of the Ti plasmid mannopine synthetase and *CaMV 35S*; MUC, 4-methylumbelliferyl- $\beta$ -D-cellobioside; MUL, 4-methylumbelliferyl- $\beta$ -D-lactopyranoside; n.d., not determined; n.o., not optimized; n.r., not reported; pNPG, *p*-nitrophenyl  $\beta$ -D-glucopyranoside; RBB-xylan, remazol brilliant blue R-D-xylan; *RbcS-3C*, Rubisco small sub-unit promoter; *TRBO*, tobacco mosaic virus RNA-based overexpression vector; U, unit; Xyl, xylanase; *Zm-PEPc*, maize phosphoenolpyruvate carboxylase.

If transgenic plants are to be used to overproduce GHs, there are three important aspects to be considered for the host plant: (1) as a biocatalyst production unit, (2) altered cell walls, and (3) auto-hydrolysis of biomass (Taylor et al., 2008; Furtado et al., 2014; Lambertz et al., 2014; Damm et al., 2016). Utilizing plants as enzyme factories is an intriguing concept. It may not be practically feasible in this case, so using plant cell cultures for enzyme overproduction has been proposed (Hellwig et al., 2004). The highest potential goal is to produce bioenergy feedstocks with altered cell walls for high saccharification and their own saccharification enzymes in tandem. An autohydrolytic plant system would require enzymes to be inactive, induced, or sequestered until harvest in order to not negatively affect plant growth. Research on GHs produced in plants is in its relative infancy.

In this review we will highlight successes and challenges of GH production in plants. Here, we compare past use of fungal and bacterial GHs produced in plants and focus on lessons that can inform on selection of appropriate expression systems. We focus especially on the types and properties of insect-derived GHs and assess their potential use in plants. Strategies and recommendations for conducting research and development of CWD enzymes in plants will be discussed with an eye toward improving bioenergy crops and integrated biofuel systems.

## ENDO-GLUCANASES IN TRANSGENIC PLANTS

Heterologous expression of lignocellulosic enzymes *in planta* for improved production of biofuels was proposed in the 1990s, but relatively few enzymes have been tested (Lopez-Casado et al., 2008; Furtado et al., 2014; Lambertz et al., 2014). Mostly, bacterial and fungal enzymes have been overproduced in plants to assess efficacy of the concept (Taylor et al., 2008; Lambertz et al., 2014). Endo-glucanases (EG) have, by far, been the most abundant class of CWD enzymes produced in plants (**Table 1**). The E1 cellulase from the bacterium *Acidothermus cellulolyticus* has been the most frequent EG produced in transgenic plants (15 papers; **Table 1**). One reason behind the attention given to E1 is that it is a hyperthermophilic enzyme, which would be inactive during temperatures favoring plant growth.

The first report on E1 production in plants was in tobacco in which E1 was targeted to the chloroplast. E1 was active with an optimal temperature of 81°C and pH of 5.25 (Dai Z. et al., 2000). The cauliflower mosaic virus 35S promoter (35S) was used to drive E1 gene expression in *A. thaliana* for apoplast targeting (Ziegler et al., 2000). In this study, E1 was accumulated up to 26% of total soluble protein in primary leaves, which is the highest reported for transgenic plant produced hydrolases (Ziegler et al., 2000; Taylor et al., 2008). Tobacco E1 expressing lines were developed for transport to the chloroplast and found that not all parts of the enzyme would translocate, however the catalytic domain was sufficient to maintain enzyme activity (Jin et al., 2003). In an attempt to optimize E1 expression in tobacco, the hybrid constitutive promoter *Mac* (a combination of the Ti plasmid mannopine

**BOX 1 | CLASSIFICATION OF CELLULASE ENZYMES.**

Endo-  $\beta$ -1,4-glucanase: Glycosyl hydrolase which has an open-ended catalytic side to randomly bind and cleave the internal  $\beta$ -1-4 linkages of cellulose polymer chains producing reduced ends. Commonly called "cellulose."

Exo-  $\beta$ -1,4-glucanase (aka Cellobiohydrolase): Similar to its Endo- counterpart, except cellulose polymers are threaded through the enzyme to reach the catalytic site.

$\beta$ -glucosidase (aka cellobiases): Glucosidase enzyme which hydrolyzes the  $\beta$ -1-4 linkage cellobiose molecules releasing two glucoses.

synthetase and 35S promoters) and the *RbcS-3C* (Rubisco small sub-unit) promoter were compared; the *RbcS-3C* was ranked higher in performance. However, the highest total soluble protein was reported as only 0.25%. Apoplast-targeted E1 transgenic potato plants produced 2.6% E1 relative to total soluble protein (Dai Z. Y. et al., 2000). Similar studies in maize, using an enhanced version of the 35S promoter driving the E1 gene resulted in 2.1% (Biswas et al., 2006) and 1.13% (Ransom et al., 2007) E1 accumulation relative to total soluble protein. Even though none of the above studies reported any negative effects to the host plants, E1 transgenic maize (Ransom et al., 2007), and tobacco cell walls developed differently than their respective non-transgenic parents. In each of these two studies increased saccharification was reported. Wildtype and E1 maize expressing lines were exposed to exogenous application of E1 and showed to be equivalent. The proposed reason for this is due to E1 is active *in vivo* during plant growth causing small nicks within the cellulose polymers. Addition of extra exogenous E1 application of enzyme was ineffectual due to the enzyme already hydrolyzing available cellulose recognition sites (Brunecky et al., 2011). In the same study, Brunecky et al. (2011) examined the transgenic E1 corn and tobacco made by Ziegelhoffer et al., (2001) via histological examination and postulated E1 was having an effect on cell wall growth during development (Ziegelhoffer et al., 2001; Brunecky et al., 2011). The presence of biologically active E1 during cell wall synthesis could play a role in reducing cellulose crystallinity or by creating gaps in the cell wall for increased enzyme accessibility (Brunecky et al., 2011). Aforementioned experiments demonstrate the utility of plant produced endoglucanases as a method for cell wall architecture manipulation.

One goal has been to use transgenic plants as biocatalyst producers in a commodity, such as E1 accumulation in maize seeds. The first 40 codons of E1 were codon-optimized in attempt to increase expression and endow tissue specificity (Hood et al., 2007). E1 in maize seed was further targeted to the cell wall, ER, and the vacuole. In the Hood et al. (2007) study, the ER appeared to be the best subcellular localization target, but none of the localizations led to off-effects in maize. Though active, the E1 enzyme was found to be truncated from a ~70 kDa size down to ~40 kDa. Furthermore, stable enzyme was observed after harvesting and drying (Hood et al., 2007). The researchers introgressed the transgenes into elite germplasm (Hood et al., 2012) for further field experiments (Garda et al., 2015). E1 accumulation increased up to seven-fold in the elite germplasm. Field produced E1 in maize seed appeared to be active, thus illustrating an economically-viable platform to overproduce and store E1 until needed.

In rice, E1 has been studied for production in vegetative biomass apoplasts (Oraby et al., 2007; Chou et al., 2011) and seed endosperm (Zhang Q. et al., 2012). Apoplastic targeting appeared to not affect plant growth (Oraby et al., 2007; Chou et al., 2011), whereas the endosperm production strategy led to dwarfing and early flowering (Zhang Q. et al., 2012). While introduction of E1 led to a suboptimal host plant phenotype, there are multiple variables which need to be addressed. In both cases, E1 was targeted to the apoplast, but used two different constitutive promoters and two different germplasm (Oraby et al., 2007; Chou et al., 2011). Chou et al. (2011) reported higher protein yield at 6.1% over Oraby et al. (2007) 4.9%. When E1 was targeted to the endosperm it was active in dried seeds, but seeds were smaller than wild type seed (Zhang Q. et al., 2012). The researchers speculated that lower seed weight is an effect of the transformation process as transgenic rice expressing non-cellulase proteins have a similar issue (Zhang Q. et al., 2012). No other phenotype was reported in the endosperm-targeted rice. Amongst the three cases, only Zhang Q. et al. (2012) performed an optimization experiment to demonstrate that the E1-produced rice maintained its high thermophilic optimum of 81°C.

E1 was transformed into duckweed (*Lemna minor* 8627), an aquatic plant, for accumulation in the cytosol in which extraction buffers were tested for efficacy. Three extraction buffers, sodium citrate (50 mM pH 4.8), sodium acetate buffer (50 mM pH 5), and HEPES (50 mM pH 8), were used and evaluated for their effects on enzyme yield and activity (Sun et al., 2007). The HEPES buffer extraction yielded more total soluble protein compared to other experimental buffers. When the proteins were extracted with citrate buffer and brief heating at 65°C enzyme activity was increased, although total recovered protein yield was lower than when the HEPES buffer extraction was used (Sun et al., 2007).

Transient expression of E1 was performed in sunflower to study the effects of various promoters: *CaMV35S*, *CMVar* (cucumber mosaic virus "advanced replicating"), and *TRBO* (tobacco mosaic virus RNA-based overexpression vector). Neither the *CMVar* nor *TRBO* led to E1 production. Utilizing the 35S promoter, it was shown that addition of methyl jasmonic acid led to a four-fold increase in production of active enzyme (Jung et al., 2014).

*Thermobifida fusca* (formerly named *Thermomonospora fusca*), a thermophilic bacterium, produces Cel6A (formerly E2), an EG that was transformed into alfalfa, tobacco, and potato for cytosolic production. Active Cel6A showed no negative phenotype on whole plants. Cel6A was evaluated for thermostability and was found to be still active from 60–65°C, but became inactive at higher temperatures in the absence of substrate which was comparable to commercially produced

TABLE 1 | Endo- $\beta$ -1,4-glucanase transgenic plant studies in which various host plant species and subcellular localization were targeted for enzyme overproduction.

Enzyme Name	Source organism	Host plant	Organelle targeting	GHF	Temp. optimum	pH optimum	Substrate	Activity	Notes	Reference
E1	<i>Acidothermus cellulolyticus</i>	<i>Arabidopsis thaliana</i>	Apoplast	5	65°C n.o.	5.5 n.o.	CMC# MUC	26% TSP		Ziegler et al., 2000
E1	<i>Acidothermus cellulolyticus</i>	<i>Lemna minor</i> 8627	Cytosol	5	80°C	5.0	CMC	0.24 U/g fresh tissue		Sun et al., 2007
E1	<i>Acidothermus cellulolyticus</i>	<i>Nicotiana tabacum</i>	Chloroplast	5	81°C	5.25	MUC	2756.3 pmol MU mg <sup>-1</sup> TSP min <sup>-1</sup>		Dai Z. et al., 2000
E1	<i>Acidothermus cellulolyticus</i>	<i>Nicotiana tabacum</i>	Apoplast	5	81°C	5.25	MUC	18056 pmol MU mg <sup>-1</sup> TSP min <sup>-1</sup>		Dai et al., 2005
E1	<i>Acidothermus cellulolyticus</i>	<i>Nicotiana tabacum</i>	Chloroplast	5	65°C n.o.	5.5 n.o.	MUC	2824 FU		Jin et al., 2003
E1	<i>Acidothermus cellulolyticus</i>	<i>Nicotiana tabacum</i>	Apoplast Chloroplast	5	65°C n.o.	5.5 n.o.	CMC	1.6% TSP	Catalytic domain	Ziegelhoffer et al., 2001
E1	<i>Acidothermus cellulolyticus</i>	<i>Nicotiana tabacum</i> , <i>Zea mays</i>	Cytosol Apoplast	5	50°C n.o.	5.0 n.o.	MUC MUC	0.3 ng/mg biomass	Tobacco provided by (Ziegelhoffer) corn provided by Biswall	Brunecky et al., 2011
E1	<i>Acidothermus cellulolyticus</i>	<i>Oryza sativa</i>	Apoplast	5	65°C n.o.	5.5 n.o.	MUC	25000 pmol MU/mg protein/min		Chou et al., 2011
E1	<i>Acidothermus cellulolyticus</i>	<i>Oryza sativa</i>	Apoplast	5	65°C n.o.	5.5 n.o.	MUC	4.9% TSP		Oraby et al., 2007
E1	<i>Acidothermus cellulolyticus</i>	<i>Oryza sativa</i>	Endosperm	5	80°C	5.0	CMC Avicel CMC	0.6 g/L glucose released 0.25 g/L glucose released 830 U/g seed		Zhang Q. et al., 2012
E1	<i>Acidothermus cellulolyticus</i>	<i>Zea mays</i>	Endosperm cell wall	5	50°C n.o.	5.0 n.o.	MUC	16% TSP	First 40 codons optimized	Hood et al., 2007
E1	<i>Acidothermus cellulolyticus</i>	<i>Helianthus annuus</i>	Apoplast	5	65°C n.o.	5.5 n.o.	MUC	2.5 mg/kg fresh tissue	Plant codon optimized	Jung et al., 2014
E1	<i>Acidothermus cellulolyticus</i>	<i>Solanum tuberosum</i>	Chloroplast	5	55°C n.o.	5.5 n.o.	MUC	61000 pmol mg <sup>-1</sup> min <sup>-1</sup>		Dai Z. Y. et al., 2000
E1	<i>Acidothermus cellulolyticus</i>	<i>Zea mays</i>	Apoplast	5	65°C n.o.	5.5 n.o.	MUC	0.845 nmol/ $\mu$ g/min	Catalytic domain	Biswas et al., 2006
E1	<i>Acidothermus cellulolyticus</i>	<i>Zea mays</i>	Apoplast	5	50°C n.o.	4.8 n.o.	MUC	1.13% TSP	Catalytic domain	Ransom et al., 2007
Cell6A	<i>Thermobifida fusca</i>	<i>Medicago truncatula</i> , <i>Nicotianatabacum</i> , <i>Solanum tuberosum</i>	Cytosol	6	55°C n.o.	5.5 n.o.	Avicel CMC Corn stover CMC Corn stover	0.37 g/L 0.47 g/L 0.18 g/L 0.47 g/L 0.18 g/L		Ziegelhoffer et al., 1999

(Continued)

TABLE 1 | Continued

Enzyme Name	Source organism	Host plant	Organelle targeting	GHF	Temp. optimum	pH optimum	Substrate	Activity	Notes	Reference
EG	Bacteria-like synthetic	<i>Saccharum officinarum</i>	Chloroplast ER	5	40°C n.o.	4.75 n.o.	CMC	23 nmol glucose/min/mg protein $\text{€}$		Harrison et al., 2011
ENG1	<i>Oryza sativa</i>	<i>Oryza sativa</i>	Vacuole Cytosol	9	n.d.	n.d.	MUC	n.d.	No transgenic plants regenerated	Nigorikawa et al., 2012
t-EG1	<i>Ruminococcus albus</i>	<i>Nicotiana tabacum</i>	Cytosol	5	30°C n.o.	5.6 n.o.	CMC	1.7 $\mu\text{mol/mg protein}$	Truncated enzyme	Kawazu et al., 1999
Cel6A	<i>Thermobifida fusca</i>	<i>Nicotiana tabacum</i>	Chloroplast	6	50°C n.o.	5.5	CMC BMCC	0.65 $\mu\text{M}$ glucose released $\text{€}$		Yu et al., 2007
ssol354	<i>Sulfolobus solfataricus</i>	<i>Nicotiana tabacum</i>	Apoplast ER	12	90°C	4.5	CMC $\neq$ , Azo-CMCE MUC	488 nmol 4MU min $^{-1}$ mg $^{-1}$ 0.1 g/L		Klose et al., 2012
TrCel5A	<i>Trichoderma reesei</i>	<i>Nicotiana tabacum</i>	Apoplast	5	55°C	4.8	Avicel CMC $\neq$ Azo-CMCE	5.5 mg L $^{-1}$ h $^{-1}$ 27 nmol 4MU min $^{-1}$ mg $^{-1}$	Ethanol-inducible	Klose et al., 2013
TrCel5A	<i>Trichoderma reesei</i>	<i>Nicotiana tabacum</i>	Apoplast ER	5	55°C n.o.	4.8 n.o.	MUC Azo-CMC	2.1 U mg $^{-1}$ 46 nmol 4MU min $^{-1}$ mg $^{-1}$		Klose et al., 2015

GHF represents glycosyl hydrolase family as defined by CAZY.org. The temperature and pH optimum are listed. Substrate and other abbreviations used in the table are listed below. ER, endoplasmic reticulum; CMC, carboxymethylcellulose; MUC, 4-methylumbelliferyl- $\beta$ -D-cellobioside; BMCC, bacterial microcrystalline cellulose; n.o., not optimized, used when reference reported single point activity instead of a range; TSP, total soluble protein; U, one unit is the release of 1  $\mu\text{mol}$  substrate per min; FU, Fluorescence units;  $\neq$ , substrate was used only for qualitative stain for activity;  $\text{€}$ , estimated values from published figure when exact numbers not provided;  $\text{€}$ , used to measure % relative to a control within the experiment.

Cel6A (Ziegelhoffer et al., 1999). In a subsequent study, Cel6A was targeted to tobacco chloroplasts and no morphological changes were observed in the host plant (Yu et al., 2007). Cel6A production resulted in an improved total soluble protein yield (0.6–4%).

Tobacco cells have been used to express two other EGs from bacteria. The t-EGI from *Ruminococcus albus* was targeted to the cytosol of tobacco cells. The t-EGI was active at 30°C. The t-EGI tobacco plants successfully autohydrolysed their cell walls (Kawazu et al., 1999). The EG SSO1354 from *Sulfolobus solfataricus* was successfully expressed in tobacco apoplasts and endoplasmic reticulum. SSO1354 produced in tobacco had highest activity at 90°C and pH 4.5 and showed no growth differences relative to non-transgenic plants. SSO1354 tobacco was tested with ionic liquid pre-treatment solutions and found to still be active (Klose et al., 2012).

The TrCel5A EG from the fungus *Trichoderma reesei* was produced in tobacco under two construct parameters (Klose et al., 2013). The TrCel5A was active under both construct versions at 55°C and pH 4.8 (Klose et al., 2013). Under the control of the 35S promoter, TrCel5A produced dwarfed tobacco even with apoplast targeting. However, using an ethanol-inducible, *alcA*, promoter, TrCel5A was produced upon induction and the enzyme was indistinguishable from that in 35S-construct plants (Klose et al., 2013). Similar to the research described above in maize, tobacco progeny produced more TrCel5A than their parents (Klose et al., 2015). In this study, TrCel5A was produced in apoplasts or the ER. In both cases, T2 progeny had wrinkled leaves with spotted necrosis, which would be considered important off-target effects (Klose et al., 2015).

The recent additions of synthetically creating genes by combing databases and preparing sequences *in silico* has been used to generate EG. A proprietary, synthetically-designed EG (psEG) was derived from a bacteria population and introduced into sugarcane. The psEG was targeted to chloroplast, ER, and vacuole with the chloroplast being the highest accumulator near 0.05% total soluble protein. The psEG was driven by a constitutive maize phosphoenolpyruvate carboxylase (*Zm-PEPc*) promoter which expressed psEG in the leaves. Unlike the observation in E1 in maize embryos, no truncation of the psEG protein was detected (Hood et al., 2007; Harrison et al., 2011).

## CELLOBIOHYDROLASES IN TRANSGENIC PLANTS

Heterologous overproduction of cellobiohydrolases (CBHs) has not been explored as much as EG in plants, but CBHs are necessary for complete hydrolysis (Table 2). The first CBH evaluated in plants was Cel6B (formerly named E3), from the bacterium *T. fusca* in the cytosol of alfalfa, tobacco, and potato (Ziegelhoffer et al., 1999). While Cel6B was detectable by western blotting and was not deleterious to plant growth, its activity was negligible and the authors declined further characterization (Ziegelhoffer et al., 1999). Although the *Mac* promoter was used for Cel6B relatively low yield was achieved. The reason for low yield could be the lack of a signal peptide to target to an organelle

or the need to codon optimize the proteins, which showed improvement in future CBH studies (Hood et al., 2007). In a later study, Cel6B was targeted to tobacco chloroplasts, and no morphological changes were observed in the host plant (Yu et al., 2007). Cel6B had a total soluble protein yield of 0.6–4%. Enzyme activity was measured using crystalline cellulose as the substrate instead of a soluble cellulose derivative substrate (i.e., CMC or MUC) which could explain the low activity reported.

Initial experiments on CBH1 from the fungus *T. reesei* driven by the constitutive 35S promoter, yielded 0.11–0.82% total soluble protein from tobacco leaves and callus and showed no deleterious effect on plant growth (Dai et al., 1999). Further experiments with CBH1 used the maize embryo promoter, globulin-1, and targeted to the cell walls, ER, and vacuoles for increased accumulation (Hood et al., 2007). The first 40 codons were optimized for improved maize protein synthesis. The combination of tissue, organelle, and codon optimization resulted in up to 16% total soluble protein, which is the highest yield reported for CBH1 (Hood et al., 2007). An issue noted while lines for CBH1 targeted to the ER had enzyme activity, they failed to be present on Western blots. However, the cell wall targeted CBH1 was a full length variant. Further evaluation demonstrated the CBH1 was truncated. In future studies, it may be necessary to fully codon optimize GHs to reduce the risk of truncation, especially in cases where the catalytic domain is near the C-terminus of the protein. Field experiments of maize progeny carrying CBH I, showed no yield or growth performance difference compared to wild type counterparts under field conditions (Garda et al., 2015).

Proprietary synthetic CBH1 and CBHII (psCBH1 and psCBHII) genes were generated using multiple bacterial and fungal sources as templates and were transformed into sugar cane (Harrison et al., 2011). Each protein was targeted to either chloroplast, ER, or vacuole with none reporting any morphological defects. The highest activity was measured from senesced leaves of psCBH1 targeted to the vacuole (Harrison et al., 2011). Both proteins were successfully localized to the leaves by the *Zm-PEPc* promoter and neither of them were truncated (Harrison et al., 2011).

In another study, recombinant synthesis of an endogenous rice CBH (EXG1) resulted in increased glucose release in rice shoots (Nigorikawa et al., 2012). The EXG1 construct utilized its native signal peptide. Some transgenic events were not observably morphologically different than the non-transgenic parent, but three events had deformed, split leaves and extra lacunae. The authors hypothesized that the altered leaf phenotype was from weakened cell walls imbued by the EXG1. Of the 28 EXG1 rice events recovered, 12 were completely sterile, and 14 were partially sterile (defined as <40 seeds produced; Nigorikawa et al., 2012).

## β-GLUCOSIDASES IN TRANSGENIC PLANTS

β-glucosidases (BGs) are necessary for degradation of cellobiose into glucose, and various BG genes have been overexpressed in plants (Table 3). Heterologous expression of the *T. maritime* BglB enzyme resulted in no altered plant morphology in

**TABLE 2 | Cellobiohydrolase transgenic plant studies in which various host plant species and subcellular localization was targeted for enzyme overproduction.**

Enzyme name	Source organism	Host plant	Subcellular localization	GHF	Temp	pH	Substrate	Activity	Notes	Reference
CBH I	<i>Trichoderma reesei</i>	<i>Zea mays</i>	Cell wall, ER, vacuole	7	50°C	5.0	MUC	16% TSP	First 40 codons optimized	Hood et al., 2007
CBH I	Fungal-like synthetic	<i>Saccharum officinarum</i>	Chloroplast, ER, vacuole	7	40°C	4.75	MUL	12.25 nmol 4-MU /min/mg protein		Harrison et al., 2011
CBH II	Fungal-like synthetic	<i>Saccharum officinarum</i>	ChloroplastER, vacuole	6	40°C	4.75	Avicel PH-101	8.68 nmol reducing sugars/min/mg protein		Harrison et al., 2011
Cel6A	<i>Thermobifida fusca</i>	<i>Medicago trunculata</i> , <i>Nicotiana tabacum</i> , <i>Solanum tuberosum</i>	Cytosol	6	55°C	5.5	CMC	0.1% TSP		Ziegelhoffer et al., 1999
Cel6B	<i>Thermobifida fusca</i>	<i>Medicago trunculata</i> , <i>Nicotiana tabacum</i> , <i>Solanum tuberosum</i>	Cytosol	6	55°C	5.5	CMC	0.02% TSP		Ziegelhoffer et al., 1999
Cel6B	<i>Thermobifida fusca</i>	<i>Nicotiana tabacum</i>	Chloroplast	6	50°C	5.5	BMCC	0.55 μm glucose released¥		Yu et al., 2007
EXG1	<i>Oryza sativa</i>	<i>Oryza sativa</i>	Native target	6	37°C	7.0	MUC	45 μg/ml 4-MU concentration		Nigorikawa et al., 2012

GHF represents glycosyl hydrolase family as defined by CAZY.org. The tested temperature and pH are listed. Substrate abbreviations are listed below. CMC, carboxymethylcellulose; BMCC, bacterial microcrystalline cellulose; MUL, 4-methylumbelliferyl-β-D-lactopyranoside; MUC, 4-methylumbelliferyl-β-D-cellobioside; TSP, total soluble protein; ¥, estimated values from published figure when exact numbers not provided.

tobacco or Arabidopsis, and the recombinant protein had the characteristic thermophilic properties (Jung et al., 2010, 2013). Crude extract from BglB tobacco was applied to transgenic rice expressing Cel5A endoglucanase, which resulted in improved saccharification in an expected synergistic fashion (Jung et al., 2010). Over-expression of the endogenous BEG1 rice BG in transgenic rice also showed no morphological effects, however no enzyme activity was reported (Nigorikawa et al., 2012). The lack of morphological effects on host plants were hypothesized to be the result of predominant cellobiose degradation and not complete cellulose chains (Jung et al., 2010; Nigorikawa et al., 2012). BGs are the least reported on of the essential CWDs and more investigation into their functional role for a complete autohydrolysis system will be required.

## BACTERIAL AND FUNGAL XYLANASES IN TRANSGENIC PLANTS

Xylans are the major carbohydrates in the hemicellulose portion of the plant cell wall. Xylanases break down the bonds of β-1,4-xylan, resulting in simpler pentose sugars that can be fermented to produce biofuels. Dicots and non-commelinoid monocots have Type I cell walls predominately made up of xyloglucan, while grasses are made up of Type II cell walls that contain high levels of arabinoxylan (Scheller and Ulvskov, 2010). Because of the inherent differences, xylanase efficacy for the different cell wall types in the host organism should be considered.

Several plants have been the hosts to heterologously-produced xylanases in plants, including Arabidopsis, barley, potato, rice, sunflower, and tobacco. In these cases the enzymes have maintained their thermophilic and mild acidic to neutral pH properties (Table 4). Overexpression of xylanases have not yielded any plants with observable morphological differences compared to controls.

A codon-optimized *xynA* gene from the fungus *Neocallimastix patriciarum* was expressed in barley under the control of two separate endosperm-specific promoters (*GluB-1* & *Hor2-4*) (Patel et al., 2000). The *GluB-1* version had twice the expression and activity of the other version in mature seed, with no expression observed in leaf, stem, or root tissues. Dried and stored seed maintained xylanase activity, which would enable biofuel catalyst production or direct feeding to animals. Two *Clostridium thermocellum* xylanases have been produced in transgenic plants. The first was a truncated xylanase (XynZ) produced in tobacco apoplasts, in which thermostability was maintained (Herbers et al., 1995). XynZ was active against multiple xylan substrates at high temperature and slightly acidic pH. The second was XynA produced in rice, which accumulated in the cytoplasm. The native signal peptide was removed, but the catalytic domain was left intact. The recombinant enzyme had native-like activity at neutral pH and thermophilic temperature of 60°C (Kimura et al., 2003). In these experiments, the enzyme accumulated in seed was active after dry storage. In Arabidopsis, two different xylanases from the fungus *T. reesei* were used to overproduce high levels of enzyme. The first, XYLII, accumulated in the cytosol,

**TABLE 3 |  $\beta$ -glucosidase (BG) transgenic plant studies in which various host plant species and subcellular localization was targeted for enzyme overproduction.**

Enzyme name	Source organism	Host plant	Organelle targeting	GHF	Temp.	pH	Substrate	Activity	Notes	Reference
BEG1	<i>Oryza sativa</i>	<i>Oryza sativa</i>	Native	1	n.d.	n.d.	n.d.	n.d.		Nigorikawa et al., 2012
BglB	<i>Thermotoga maritima</i>	<i>Nicotiana tabacum</i> and <i>Arabidopsis thaliana</i>	Chloroplast Cytosol	3	40-90°C (80°C)	3-6 (4.5)	pNPG	180 mmol/ml $\text{h}^{-1}$		Jung et al., 2010
BglB	<i>Thermotoga maritima</i>	<i>Nicotiana tabacum</i> and <i>Arabidopsis thaliana</i>	Chloroplast	3	80°C	4.5	pNPG	2090.9 mg/ml	New chloroplast promoter	Jung et al., 2013

GHF represents glycosyl hydrolase family as defined by CAZY.org. The temperatures and pH ranges tested and (optimum) are listed. Substrate abbreviations are listed below. pNPG, *p*-nitrophenyl  $\beta$ -D-glucopyranoside;  $\text{h}^{-1}$ , estimated values from published figure when exact numbers not provided.

**TABLE 4 | Endo-1,4- $\beta$ -xylanase transgenic plant studies in which various host plant species and subcellular localization was targeted for enzyme overproduction.**

Enzyme name	Source organism	Host plant	Organelle targeting	GHF	Temp.	pH	Substrate	Activity	Notes	Reference
ATX	<i>Thermobifida fusca</i>	<i>Oryza sativa</i>	Cytosol	11	50°C	5.0	Birchwood xylan	3.64 U g <sup>-1</sup> fresh tissue		Weng et al., 2013
XynA	<i>Clostridium thermocellum</i>	<i>Oryza sativa</i>	Cytoplasm	11	60°C	7.0	Birchwood xylan	0.475 Unit/mg protein $\text{h}^{-1}$	Catalytic domain without native signal peptide	Kimura et al., 2003
XynA	<i>Dictyoglomus thermophilum</i>	<i>Arabidopsis thaliana</i> , <i>Nicotiana tabacum</i>	Apoplast	11	85°C	6.5	Wheat arabinoxylan	8.25 $\mu$ mol reducing ends/min/mg $\text{h}^{-1}$	Plant codon optimized	Borkhardt et al., 2010
XynA	<i>Dictyoglomus thermophilum</i>	<i>Arabidopsis thaliana</i> , <i>Nicotiana tabacum</i>	Apoplast	11	85°C	6.5	Wheat arabinoxylan	8.25 $\mu$ mol reducing ends/min/mg $\text{h}^{-1}$	Plant codon optimized	Borkhardt et al., 2010
mXynA	<i>Neocallimastix patriciarum</i>	<i>Hordeum vulgare</i>	Cytoplasm	11	40°C	6.0	AZCL-xylan	27.06 AU/h/grain		Patel et al., 2000
XylII	<i>Trichoderma reesei</i>	<i>Arabidopsis thaliana</i>	Peroxisome, Cytosol, Chloroplast	11	50°C	6.8	Xylan solution	24 U mg <sup>-1</sup>		Hyunjong et al., 2006
XynII	<i>Trichoderma reesei</i>	<i>Arabidopsis thaliana</i>	Cytosol, Chloroplast	11	50°C	6.8	Xylan solution	22.7 U/mg		Bae et al., 2008
XynB	<i>Streptomyces olivaceoviridis</i>	<i>Solanum tuberosum</i>	Cytosol	11	60°C	5.2	Birchwood xylan	87 $\mu$ mol min <sup>-1</sup> g <sup>-1</sup>		Yang et al., 2007
XynZ	<i>Clostridium thermocellum</i>	<i>Nicotiana tabacum</i>	Apoplast	10	60°C	5.4	RBB-Xylan, Birchwood xylan, Oat spelt	46.6 U/mg	Truncated	Herbers et al., 1995

GHF represents glycosyl hydrolase family as defined by CAZY.org. The temperature and pH are values enzyme activity was performed. Substrate and other abbreviations are listed below. AZCL-xylan, azo-crosslinked-xylan; RBB-xylan, remazol brilliant blue R-D-xylan; U, amount of enzyme liberating 1  $\mu$ mol xylose per min; AU, absorbance unit;  $\text{h}^{-1}$ , estimated values from published figure when exact numbers not provided.

peroxisome, and chloroplast, depending on the experiment. Colocalization into chloroplast and peroxisomes resulted in the highest yield of XYLII (Hyunjong et al., 2006). The second xylanase tested in *Arabidopsis* was XYNII, which was targeted to either the chloroplast or the cytosol. Chloroplast-targeted XYNII yielded the higher amounts of protein compared with the non-targeted variant. Plant-derived XYNII was found to be comparable to a commercial xylanase in activity under lab conditions (Bae et al., 2008). Finally, *Streptomyces olivaceoviridis* XynB was overproduced in potato cytosol or apoplast. XynB was stable at 60 and 70°C and subsequent generations of the

transgenic potato produced greater amounts of XynB than the T0 generation (Yang et al., 2007).

## COMBINING GLYCOSYL HYDROLASE GENES IN TRANSGENIC PLANTS FOR IMPROVED CONVERSION PERFORMANCE

Cellulolytic enzymes have been combined to assess potential synergistic effects by gene stacking in transgenic plants (Table 5). Tobacco has been the most often-used host plant, but the

first reported host for stacked GHs was barley (Jensen et al., 1996; Fan and Yuan, 2010; Mahadevan et al., 2011; Lee et al., 2012). The first experiments in barley were accomplished using protein fusions. The EII-hybrid construct is a plant codon-optimized fusion of two bacterial enzymes from *Bacillus spp.* glucanase and a *Bacillus spp.*  $\alpha$ -amylase for expression in barley. EII-hybrid was expressed in endosperm during germination under the control of the endogenous (barley) EII promoter and the the high-pI- $\alpha$ -amylase signal peptide (Jensen et al., 1996). In another study, the same EII-hybrid gene construct was driven by *Hor3-1* endosperm promoter with the *Hor3-1* native signal peptide (Horvath et al., 2000). Neither EII-hybrid construct seemed to cause any morphological changes to plant growth (Jensen et al., 1996; Horvath et al., 2000). EII-hybrid was active at pH 7.4 and 65°C (Horvath et al., 2000; Jensen et al., 1996). Another barley transgenic hydrolase named *cel-hyb1* is a fusion of an EG from the fungus *N. patriciarum* and a 1,3-1,4- $\beta$ -glucanase from *Piromyces sp.* (Xue et al., 2003). The *Cel-hyb1* transgenic lines showed no observable morphological defects to plant health while maintaining equal enzyme activity for fresh or dried biomass. *Cel-hyb1* had a lower temperature activity at 40°C and slightly more acidic pH (Xue et al., 2003). Both cases demonstrated the use of barley as a functional bioreactor and could be used for production of transgenic cellulases or improved feedstocks for animals (Xue et al., 2003).

A hybridized xylanase-arabinase construct (Xyln-ara) was developed by fusing the catalytic domains of a xylanase from *C. thermocellum* and an arabinofuranosidase from *Geobacillus stearothermophilus* and introduced into tobacco. The fusion of the domains revealed novel enzyme character for having an improved activity at pH 6-9 when compared to the original enzyme. The fused enzymes demonstrated activity on carboxymethyl cellulose (CMC) substrate while the native forms did not. The Xyln-ara hybrid was targeted to the cytosol and showed no effect on plant growth (Fan and Yuan, 2010). Use of catalytic domains is advantageous by reducing the size of transformation constructs which improves multi-gene integration experiments and down-stream analysis (Fan and Yuan, 2010).

From *T. maritima*, the *Cel5A* endoglucanase and an engineered CBM6-*Cel5A* fusion hydrolase construct were introduced and compared in transgenic tobacco. CBM6 is a carbohydrate binding module from CBM family 6 of xylanase A in *C. stercorarium*. When the CBM6-*Cel5A* was first engineered and tested in *E. coli* it showed up to an 18-fold increase in activity over non-fused *Cel5A* and was selected for tobacco evaluation (Mahadevan et al., 2008). The CBM6-*Cel5A* was targeted to the chloroplast, apoplast, and cytosol where the chloroplast targeted enzyme had the highest accumulation and no morphological defect was reported (Mahadevan et al., 2011). Autohydrolysis of plant material was performed on biomass and found the enzymes were able to have increased release of glucose in plants with the hybrid and native gene (Mahadevan et al., 2011).

In tobacco, BglB, *Cel5A*, XylIII, or *Cel6B* were connected by the linker from the 2A oligopeptide from the foot-and-mouth disease virus to test a gene stack integration to improve hydrolysis. It

was noted that selected hydrolase genes did have a pH dependent effect on hydrolysis. Combination of these enzymes worked well when on CMC substrate, however they showed low activity on filter paper (Lee et al., 2012). A synergistic effect was observed when BglB:*Cel5A* and *Cel6B*:*Cel5A* were co-expressed, but not observed for XylIII:*Cel5A*.

In sunflower E1 and the xylanase Xyn10A from *A. cellulolyticus* were transiently co-expressed to the leaf apoplast using agroinfiltration for potential applications of biofuel production (Jung et al., 2014). Transient expression of both enzymes was achieved using the *CaMV35S* constitutive promoter, however yields were very low compared to previously reported plant expression systems (Taylor et al., 2008; Jung et al., 2014). Equivalent agroinfiltration experiments using *CMVar* and *TRBO* promoters for both enzymes were failures, but expression was achieved in this system upon methyl jasmonic acid treatment and increasing temperature from 20 to 30°C (Jung et al., 2014).

Future work on gene stacking GHs in plants should combine each enzyme grouping to constitute a complete autohydrolysis system. Selected GHs would need plant codon optimization for improved yield and prevent truncation (Hood et al., 2007). Organelle targeting to either the apoplast or the vacuole would be the best for high protein accumulation (Ziegler et al., 2000; Hood et al., 2007). Use of precision genome editing tools could knock out non-essential growth pathways to increase protein yield (Liu et al., 2014). For example, E1 targeted to maize endosperm cell walls could be mutated to reduce starch storage in favor of storing E1. Adding in CBHs and BGs creates the enzymatic synergy needed for complete autohydrolysis.

## INSECT BIOPROSPECTING AVENUES FOR TRANSGENIC PLANT PRODUCED HYDROLASES

Insects are a relatively recent source for bioprospecting biocatalysts for biofuel production (Carroll and Somerville, 2009; Oppert et al., 2010). Isopteran (termite) digestive tracts are very efficient: 99% of ingested cellulose and 87% of hemicellulose is converted into usable sugars (Ohkuma, 2003). Researchers once believed that insects produced few-to-no endogenous CWD enzymes, but rather relied on those from symbiotic organisms in their guts to break down plant biomass. For some species this is still believed to be the case, however genomic and proteomic analyses have shown that insects produce endogenous enzymes (Watanabe et al., 1998; Tokuda and Watanabe, 2007). For example, termites can survive solely on cellulose, and thus would be a source of cell wall digesting enzymes for bioprospecting (Watanabe and Tokuda, 2010). Heterologous expression of insect cellulases in model organisms have shown them to be potential alternatives to bacterial and fungal-sourced enzymes. Many insect cellulases reported in the literature have temperature optima from 40 to 65°C and perform optimally at alkaline pH (Shi et al., 2011, 2013; Willis et al., 2011). High temperature and variable pH characteristics were

**TABLE 5 | Transgenic plants producing gene stacked cellulolytic enzymes produced in plants as reported in the literature.**

Enzyme name	Source organisms	Host plant	GHF	Temp. (°C)	pH	Substrate	Organelle targeting	Notes	Reference
EII-hybrid	<i>Bacillus</i> spp.	<i>Hordeum vulgare</i>	16	65	7.4	Azobarley $\beta$ -glucan (40 ng $\pm$ 18 ng) Lichenan $\neq$ ,	Endosperm	Codon optimized	Jensen et al., 1996
EII-hybrid	<i>Bacillus</i> spp.	<i>Hordeum vulgare</i>	16	65	7.4	Azobarley $\beta$ -glucan (Up to 40 $\mu$ g $\cdot$ mg $^{-1}$ ) Lichenan $\neq$ ,	Endosperm	Codon optimized	Horvath et al., 2000
E1 & Xyn10A	<i>Acidothermus cellulolyticus</i>	<i>Helianthus annuus</i>	5 & 10	70	6.0	MUC (up to 2.5 mg/kg) Birchwood xylan (up to 4.1 mg/kg) $\neq$	Apoplast	Plant codon optimized	Jung et al., 2014
BglB: Cel5A	<i>Thermotoga maritime</i>	<i>Nicotiana tabacum</i>	3 & 5	80	5.0	CMC (0.375 mg/ml) $\neq$ pNPG (240 $\mu$ mol/ml) $\neq$	Chloroplast		Lee et al., 2012
Xyll: Cel5A	<i>Trichoderma reesei</i> and <i>Thermotoga maritime</i>	<i>Nicotiana tabacum</i>	11 & 5	50	7.0	CMC (0.275 mg/ml) $\neq$  Oat spelt (0.25 mg/ml) $\neq$	Chloroplast		Lee et al., 2012
Cel6B: Cel5A	<i>Thermobifida fusca</i> and <i>Thermotoga maritime</i>	<i>Nicotiana tabacum</i>	5 & 6	50	7.0	CMC (0.375 mg/ml) $\neq$  Filter paper (0.225 mg/ml) $\neq$	Chloroplast		Lee et al., 2012
Cel-Hyb1 (CelA & Cel6G)	<i>Neocallimastix patriciarum</i> and <i>Piromyces</i> sp.	<i>Hordeum vulgare</i>	& 6	40	6.0	CMC $\neq$ ,  AZCL-glucan (8447 $\pm$ 2059 AU/h/g/ grain) AZCL-HE-Cellulose (4041 $\pm$ 657 AU/h/g/ grain) Lichenan (1585 $\pm$ 195 U/g grain)	Cytoplasm		Xue et al., 2003
TmCel5A & TmCel5A fused CMB6A	<i>Thermotoga maritime</i>	<i>Nicotiana tabacum</i>	5 & CBM6	50	6.0	CMC (up to 2.0 Unit/mg tsp) $\neq$	Apoplast, Chloroplast, Cytosol	Fused with cellulose binding module	Mahadevan et al., 2011
Xyln-ara	<i>Clostridium thermocellum</i> and <i>Geobacillus stearothermophilus</i>	<i>Nicotiana tabacum</i>	10 & 51	60	6.0	RBB-xylan (55.3 $\mu$ mol dye/min/ $\mu$ g) Wheat-AX (22.5 $\mu$ mol xylose/min/ $\mu$ g) $\neq$ Rye-AX (27.5 $\mu$ mol xylose/min/ $\mu$ g) $\neq$	Cytosol	Fusion of catalytic domains	Fan and Yuan, 2010

Enzyme name is the name given by the author, source organisms are where the enzymes originated from, host plant is the plant species which was transformed with the enzyme, GHF is the glycosyl hydrolase family number as defined by CAZY.org, temperature, and pH are values enzyme activity was performed, substrate is the material used to measure enzyme activity followed by parenthesis of reported highest mean activity, organelle targeting is where enzyme was targeted toward in the plant cell, activity is the reported maximum activity provided by authors, notes list other key interests for the reference, reference is the referred paper information was collected.

CBM, carbohydrate binding module; CMC, carboxymethylcellulose; MUC, methylumbelliferyl- $\beta$ -D-cellobioside; pNPG, p-nitrophenol; RBB-xylan, Remazol brilliant blue; AX, arabinoxylan; AZCL= azo-crosslinked; tsp, total soluble protein; AU, absorbance unit;  $\neq$ , substrate was used only for qualitative stain for activity;  $\neq$ , estimated values from published figure when exact numbers not provided.

discovered as researchers compared insect derived hydrolases to microbial hydrolases. These high temperature and variable pH could be from the native structure formed by GHs or from the symbiont ancestors which resided in the insect digestive system.

Insect cellulolytic enzymes fall into the GH families 1, 5, 9, 11, and 45. The majority of insect cellulases discovered are from the Coleoptera (beetle) and Isoptera (termite) taxonomic orders as defined by Misof et al. (2014). In addition, Blattodea (roaches) Phasmatodea (stick insects), Lepidoptera (butterflies/moths), and Orthoptera (grasshopper/crickets) have demonstrated cellulolytic activity (Oppert et al., 2010; Shelomi et al., 2014). Endogenous and insect endosymbiont derived enzymes have similar characteristics to bacterial and fungal derived enzymes in terms of thermostability and variable pH (Oppert et al., 2010). Insect derived enzymes often have a

more alkaline pH range than those from microbial sources, since the insect digestive systems can be up to a pH of 12 in the insect digestive systems (Dow, 1992; Brune, 2014).

The digestive systems of insects are broken into three distinct sections of foregut, midgut, and hindgut. The foregut includes the mouth, salivary glands, and the most anterior chamber of the digestive tract. Cellulolytic enzymes have been discovered in the salivary glands which along with the maceration from the mandibles begins initial degradation. The fore- and midgut contain mostly endogenous cellulases (Slaytor, 1992; Brune, 2014). Malpighian tubules, when present, separate the mid and hindgut. The majority of microbial symbionts have been discovered in the hindgut. Specifically, protist and bacterial cellulases have been isolated from the hindgut of Coleoptera, Isoptera, and Orthoptera insect orders (Watanabe et al., 1998;

Brune, 2014). Both symbiont and endogenous insect GH systems have the potential for use as novel biocatalysts.

A major impediment to identifying and cloning insect CWD enzyme-coding genes is the lack of sequenced and assembled genomes among insect species, especially those that have a diet consisting of plant cell walls. *Apis mellifera*, *Drosophila melanogaster*, *Drosophila pseudoobscura*, *Tribolium castaneum*, *Nasonia* spp., *Acyrtosiphon pisum*, and *Heliconius* spp. are taxa with the most complete genome data (Richards et al., 2005; Weinstock et al., 2006; Tribolium Genome Sequencing et al., 2008; International Aphid Genomics Consortium, 2010; Werren et al., 2010; Heliconius Genome Consortium, 2012). Of these, only *T. castaneum* and *A. pisum* and, *Heliconius* spp., are herbivores. There are efforts, such as the “i5K” to sequence and assemble 5000 insect genomes, of which 28 pilot studies are underway and new phylogenetic trees are being designed (Robinson et al., 2011; Misof et al., 2014; Behura, 2015). To advance insect CWD gene discovery, herbivorous insects and their metagenomes should be sequenced to advance the field beyond its present state (Table 6).

Termites are the archetypical wood-consuming insects. The first endogenous insect cellulase was discovered from the genome of the termite, *Reticulitermes speratus* (Watanabe et al., 1998). Since then many other termite-derived hydrolases have been discovered and even evaluated for heterologous expression. Isoptera families are split into the lower termites (Mastotermitidae, Kalotermitidae, Hodotermitidae, Termopsidae, Rhinotermitidae, and Serritermitidae) and higher termites (Termitidae). The distinction between these families is based on presence of flagellated protist symbionts in the hindgut (lower termites) or their absence (higher termites) (Ni and Tokuda, 2013; Brune, 2014). The “lower vs higher” termite digestion distinction is the basis of evolutionary importance being that the higher termites evolved later (Slaytor, 1992; Watanabe and Tokuda, 2001; Ni and Tokuda, 2013). A prevailing theory on this change is due to ecological changes which caused Isoptera ancestors’ loss of protists symbionts, change of habitat, and/or food source and possible reacquisition of different cellulase-harboring symbionts (Ni and Tokuda, 2013; Slaytor, 1992; Watanabe and Tokuda, 2001). In either case of both higher and lower termites have provided cellulolytic enzymes which could be utilized for autohydrolysis plants.

GHs isolated from two of the lower termite families, Kalotermitidae and Rhinotermitidae, have been evaluated in heterologous systems. From Kalotermitidae, a *Neotermes koshunensis* BG (NkBG) was produced in *E. coli* and in *Aspergillus oryzae*. NkBG was functional in both systems, however it had a higher activity from *E. coli* (156.7 U/mg) compared to *A. oryzae* (12.4 U/mg). NkBG produced in *A. oryzae* had a unique property for maintaining 100% activity in the presence of 0.6 M glucose; most BG enzymes are inhibited by excess glucose (Uchima et al., 2011). Rhinotermitidae enzymes have been heterologously expressed from *Coptotermes formosanus*, *Reticulitermes flavipes*, *Reticulitermes santonensis*, and *R. speratus*. From *C. formosanus* two EGs (CFP-EG1 and CfEG5) were evaluated in *E. coli*. CFP-EG1 is from a protist symbiont and had a thermophilic activity at 70°C with a CMC substrate activity of 105 U/mg (Inoue et al.,

2005). Conversely, CfEG5 had three-fold higher CMC activity at 325 U/mg, but was less thermally active with an optimum of 43°C (Zhang D. et al., 2011). A BG from *C. formosanus* was expressed in *E. coli* and demonstrated an activity of 462.6 U/mg on cellobiose (Zhang D. et al., 2012). From *R. flavipes*, an EG (Cell-1) and a BG (RfBG1) were expressed using the baculovirus expression system in which *Trichoplusia ni* larvae served as bioreactors. Hemolymph from baculovirus-injected larvae were extracted and analyzed for enzyme activity (Scharf et al., 2010). Cell-1 CMC activity of 1.4 U/mg was low while RfBG1 was considerably more effective on cellobiose at 638 U/mg. An opposite trend was seen from an *R. santonensis* BG which had a low activity of 0.441 U/mg against *p*-nitrophenol- $\beta$ -D-glucopyranoside (pNPG) substrate (Mattéotti et al., 2011). Possibly this was due to being produced in *E. coli* instead of *T. ni*. Also from *R. santonensis* is the xylanase, mXylB8, which has the highest reported beechwood xylan activity at 1837 U/mg for an insect derived xylanase with a mildly acidic range of pH 5.0 and temperature of 55°C (Mattéotti et al., 2012). From *R. speratus* two EGs (RsEG and RsSymEG1) were produced in *A. oryzae* both of which were active at 45°C and a CMC activity of 1200 and 605 U/mg, respectively (Hirayama et al., 2010; Todaka et al., 2010). RsEG is an endogenous cellulase and RsSymEG1 is from a protist symbiont. It is interesting to note that the endogenous RsEG is roughly double the activity of the symbiont.

The higher termite family, Termitidae, has also provided useable GHs for plant biocatalysts engineering. The BG, bg1-gs1 from *Globitermes sulphureus*, was produced in *E. coli* and demonstrated a thermophilic activity at 90°C producing 110 U/mg from pNPG substrate (Wang et al., 2012). While not having as high a thermal stability, the MbmGBG1 from *Macrotermes barneyi*, had a higher activity of 206 U/mg on cellobiose with a more acidic pH of 5.0 at temperature of 45°C (Wu et al., 2012). Another thermophilic insect derived BG is the G1mgNTBG from *Nasutitermes takasagoensis*, while optimum temperature is 65°C, the activity was the lowest, 5.83 U/mg, when produced by *Pichia pastoris* (Uchima et al., 2012). Also from *N. takasagoensis*, is the EG, NtEG, which when produced in *A. oryzae* produced the highest CMC activity at 1392 U/mg at a slightly acidic pH of 6.0 at 65°C (Hirayama et al., 2010). A xylanase from *Macrotermes annandalei*, was produced in *E. coli* and evaluated by digestion of beechwood xylan to produce 733 U/mg at a near neutral 7.5 pH and 55°C (Liu et al., 2011). Both higher and lower Isopteran species have proven to be a reliable source of GHs, which should be evaluated for their efficacy in plant based expression systems.

Coleopteran derived GHs are found predominantly in the larval stages. The wood-boring *Apriona germari* has provided three different EGs all with heterologous activity from Sf9 insect cells. When using CMC as the substrate, AgEG-I, AgEG-II, and AgEG-III produced activities of 992, 812, 1037 U/mg, respectively (Lee et al., 2004, 2005; Wei et al., 2006). A baculovirus transformation system into *Bombyx mori* larvae was used to heterologously express two other wood-boring beetle cellulases. Hemolymph extracted from transformed larvae carried the EGs when tested on CMC produced activity levels of 927 U/mg for *Batocera horsfieldi* EG and 319.22  $\pm$  9.3 U/mg for *Anoplophora malasiaca* EG (Chang et al., 2012; Xia et al., 2013). Larvae

**TABLE 6 | Heterologously produced insect cellulolytic enzymes as reported in the literature.**

Order: family	Genus species	Enzyme (family)	Temp. (°C)	pH	Substrate	Activity	Expression system	Reference
<b>COLEOPTERA</b>								
Cerambycidae	<i>Apriona germari</i>	EG (GH 45)	50	6.0	CMC	992 (U/mg)	Sf9 insect cells	Lee et al., 2004
Cerambycidae	<i>Apriona germari</i>	EG (GH 45)	50	6.0	CMC	812 (U/mg)	Sf9 insect cells	Lee et al., 2005
Cerambycidae	<i>Apriona germari</i>	EG (GH 5)	55	6.0	CMC	1037 (U/mg)	Sf9 insect cells	Wei et al., 2006
Cerambycidae	<i>Batocera horsfieldi</i>	EG (GH 45)	50	6.0	CMC	928 (U/mg)	baculovirus expression in <i>Bombyx mori</i>	Xia et al., 2013
Cerambycidae	<i>Anoplophora malasiaca</i>	EG (GH 45)	50	4.0	CMC	319.22 ± 9.3 (U/mg)	baculovirus expression in <i>Bombyx mori</i>	Chang et al., 2012
Chrysomelidae	<i>Diabrotica virgifera virgifera</i>	EG (GH 45)	45	6.0	CMC	166.67 nmol min <sup>-1</sup> mgprotein <sup>-1</sup>	<i>E. coli</i>	Valencia et al., 2013
Tenebrionidae	<i>Tribolium castaneum</i>	EG (GH 9)	50	9.0	CMC	12.9 (U/mg)	S2 insect cells	Willis et al., 2011
<b>ISOPTERA</b>								
Kalotermitidae	<i>Neotermes koshunensis</i>	BG (GH1)	45	5.0	Cellobiose	156.7 (U/mg)	<i>E. coli</i>	Ni et al., 2007; Uchima et al., 2011
Rhinotermitidae	<i>Coptotermes formosanus</i>	EG (GH5)	70	6.0	pNPG	12.4 (U/mg)	<i>A. oryzae</i>	Inoue et al., 2005
Rhinotermitidae	<i>Coptotermes formosanus</i>	EG (GH9)	43	5.6	CMC, Cellodextrin, Filter paper	325 (U/mg)	<i>E. coli</i>	Zhang D. et al., 2011
Rhinotermitidae	<i>Coptotermes formosanus</i>	BG (GH1)	49	5.6-6.2	Cellobiose	462.6 (U/mg)	<i>E. coli</i>	Zhang D. et al., 2012
Rhinotermitidae	<i>Reticulitermes flavipes</i>	EG (GH9)	50–60	6.5-7.5	CMC	1.3 (U/mg)	baculovirus expression in <i>Trichoplusia ni</i>	Zhou et al., 2010
Rhinotermitidae	<i>Reticulitermes flavipes</i>	BG (GH1)	40	6.5-7.0	Cellobiose	638.0 ± 39.0 (U/mg)	baculovirus expression in <i>Trichoplusia ni</i>	Scharf et al., 2010
Rhinotermitidae	<i>Reticulitermes santonensis</i>	BG (GH1)	40	6.0	pNPG	0.441 (U/mg)	<i>E. coli</i>	Mattéotti et al., 2011
Rhinotermitidae	<i>Reticulitermes santonensis</i>	Xyl (GH11)	55	5.0	AZCL-xylan, Beechwood xylan	1837 (U/mg)	<i>E. coli</i>	Mattéotti et al., 2012
Rhinotermitidae	<i>Reticulitermes speratus</i>	EG (GH9)	45	5.5	CMC	1200 ± 92 (U/mg)	<i>A. oryzae</i>	Hirayama et al., 2010
Rhinotermitidae	<i>Reticulitermes speratus</i>	EG (GH7)	45	6.5	CMC	603 (U/mg)	<i>A. oryzae</i>	Todaka et al., 2010
Termitidae	<i>Globitermes sulphureus</i>	BG (GH1)	90	6.0	pNPG	110 (U/mg)	<i>E. coli</i>	Wang et al., 2012
Termitidae	<i>Macrotermes annandalei</i>	Xyl (GH11)	55	7.5	Beechwood xylan	733 (U/mg)	<i>E. coli</i>	Liu et al., 2011
Termitidae	<i>Macrotermes barneyi</i>	BG (GH1)	45	5.0	Cellobiose, pNPG	206 (U/mg)	<i>E. coli</i>	Wu et al., 2012
Termitidae	<i>Nasutitermes takasagoensis</i>	EG (GH9)	65	6.0	CMC	1392 ± 57 (U/mg)	<i>A. oryzae</i>	Hirayama et al., 2010
Termitidae	<i>Nasutitermes takasagoensis</i>	BG (GH1)	65	5.5	pNPG	5.83 (U/mg)	<i>P. pastris</i>	Uchima et al., 2012

(Continued)

TABLE 6 | Continued

Order: family	Genus species	Enzyme (family)	Temp. (°C)	pH	Substrate	Activity	Expression system	Reference
<b>LEPIDOPTERA</b>								
Noctuidae	<i>Spodoptera frugiperda</i>	BG (GH1)	30	6.0	pNPG	2.4 (mM <sup>-1</sup> s <sup>-1</sup> )	<i>E. coli</i>	Marana et al., 2004
<b>ORTHOPTERA</b>								
Acrididae	Grasshopper (n.r)	BG (GH 1)	60	9.0	pNPG	2.61 ± 0.75 (mM <sup>-1</sup> s <sup>-1</sup> )	<i>E. coli</i>	Shi et al., 2011
Gryllidae	<i>Teleogryllus emma</i>	EG (GH9)	40	5.0	CMC	3118.4 (U/mg)	Sf9 cells	Kim et al., 2008

Insect taxonomic order followed by insect family. Genus species are insect source of enzyme, Enzyme and family represents glycosyl hydrolase and its corresponding family as defined by CAZY.org, temperature and pH are values reported for best performance, substrate is the material used to measure enzyme activity, heterologous expression system defines transformation system evaluated, reference is the referred paper information was collected.

EG, endo-1,4-β-glucanase; BG, β-glucosidase; Xyl, Xylanase; U, units of mMol of substrate released per minute; n.d., not determined; n.r., not reported; n.o., not optimized, used when references reports activity, however was a single temperature or pH testing.

from *Diabrotica virgifera virgifera* were examined as a biocatalyst source due to their diet of maize stems which could be used as a biofuel source. An EG from *D. virgifera* was isolated and produced in *E. coli* with a CMC activity of 0.166 U/mg (Valencia et al., 2013). Unlike the previous Coleoptera species, *T. castaneum*, provided TcEG-1 which has a lower activity compared to other beetle species, but was active at very alkaline pHs (pH 9–11) in both S2 insect cells and yeast (Willis et al., 2011; Shirley et al., 2014). The Coleopteran cellulases were active at 45–50°C similarly to that of mild-thermophilic microbial cellulases.

Two additional insect orders have received attention for their CWD enzymes: Lepidoptera and Orthoptera. *Spodoptera frugiperda*'s larval stage was used as the source of a β-glucosidase BG activity of 2.4 mM<sup>-1</sup>s<sup>-1</sup> pNPG substrate and was effective when produced in *E. coli* (Marana et al., 2004). Orthopteran species are known for their mandibular action for chewing through fibrous plant tissue. The symbiont *Klebsiella* sp. in the gut from an undisclosed grasshopper species was found to have BG with similar activity of 2.61 ± 0.75 mM<sup>-1</sup>s<sup>-1</sup> on pNPG to that of *S. frugiperda* (Shi et al., 2011). From *Teleogryllus emma*, an EG from GH family nine was produced in Sf9 cells showing an activity rate of 3118.4 (U/mg) against CMC (Kim et al., 2008).

Insect GHs have not yet been synthesized in transgenic plants. Under heterologous expression in microbial systems, the majority of insect derived GHs demonstrate similar thermal and acidic pH ranges. A few notable examples from Coleoptera have lower activity, but a uniquely alkaline pH range. Attributes, such as high alkaline affinity, learned from insect derived GHs could be utilized in site-directed mutagenesis studies to enhance non-insect biocatalyst research. Overall plant produced insect GHs are another viable avenue for an autohydrolytic plant systems.

## STRATEGIES FOR ENGINEERING AUTOHYDROLYSIS IN PLANTS

Constitutive gain-of-function hydrolase activity could have negative phenotypic effects in transgenic plants. While these

negative growth effects seldom have been observed in plants harboring GHs with high temperature optima, we envisage host off-effects as a potential issue in heterologous production systems for non-thermophilic enzymes. Use of inducible promoters to attenuate hydrolase activity can reduce off-target effects and decrease chances of lethality (Skena et al., 1991; Taylor et al., 2008). Inducible systems could be utilized at several life cycle/growth stages to specifically induce the hydrolase activity to detect onset of different genes and pathways involved in cell wall biosynthesis. One such inducible approach has been used to control CWD transgene expression in tobacco. An ethanol inducible promoter demonstrated effective cellulase production while preventing dwarfism when compared to constitutively produced cellulase in tobacco plants (Klose et al., 2013).

Protein modification can be used on GHs to control for off-target effects and increase protein yield. Sequestration of proteins by signal peptides to organelles, such as vacuoles, ER, or plastids improves production rate of transgenic proteins, but also isolates them from their intended substrates (Hood et al., 2007; Taylor et al., 2008). The use of hyperthermic proteins, where enzyme activation temperature is above that required for plant growth are useful during biorefinery steps to prevent undesired effects (Mir et al., 2014). A unique protein sequence known as an intein, can be inserted into the portions of enzymes and prevents proper protein folding. Inteins are heat labile and have been proven to successfully prevent undesired enzymatic action until placed in high temperature, which induces intein excision (Shen et al., 2012).

Plants are attractive bioreactors for many recombinant proteins because their post-translational modification systems are analogous to other eukaryotes (Hellwig et al., 2004; Fischer et al., 2012; Zhang, 2015). Plant cell culture systems could also speed the screening process for functional GH production *in planta* prior to generation of time consuming whole transgenic plants (Hellwig et al., 2004; Shen et al., 2013). A multi-well system using a plant cell system for rapid GH stacking of multiple constructs would speed the screening process of GHs.

A current research problem is the variety of substrates used to measure activity and efficiency. Filter paper is reformed cellulose,

but the different densities and size of the filter paper disc used during assays have effects on enzyme activity (Xiao et al., 2004). Carboxymethylcellulose is the most common substrate used because of its solubility in liquids, which is caused by methylating cellulose (Reese et al., 1950). Avicel (powdered crystalline cellulose) is a close relative to cellulose, but can be purchased at crystalline particle size making comparisons with the soluble form difficult (Yeh et al., 2010). Both of these (and other cellulose derivative) substrates can also be affixed with fluorophores which are activated upon hydrolysis (Megazyme Wicklow, Ireland). The units differ based on the method used to measure, some do a straight absorbance compared to a control, a reduction in viscosity of a solution, use of a standard curve, or percent of a commercial cellulase standard (Taylor et al., 2008; Willis et al., 2010).

An important decision to be made when developing commercial-ready autohydrolytic plants is the host plant. This issue is important because C4 grasses (leading biofuel crops) have different cell wall composition and genetic mechanisms than C3 plants. While current monocot model species, such as rice (*O. sativa*) and Brachypodium (*Brachypodium distachyon*), have extensive genomics toolsets; e.g., completed genomes and activation tagging systems; both are C3 grasses that are distantly related to C4 grasses (Jeon and An, 2001; Ito et al., 2005; Matsumoto et al., 2005; Vogel et al., 2009, 2010). Thus, while rice and Brachypodium have proven to be powerful tools, they are suboptimal to dissect C4 grass metabolism and cell wall biosynthesis (Dal'Molin et al., 2010). This latter point is a crucial consideration for testing autohydrolytic biofuel crop lines. C4 grasses (e.g., maize, sorghum, and switchgrass) possess divergent expression patterns of cell wall-related genes, and indeed likely use different genes, compared to C3 grasses. The specialized vascular anatomy of C4 plants relies on similar genes to C3 grasses, however, their regulation and expression differ, resulting in different structures; their secondary cell walls are simply not well modeled by C3 grasses (Nelson, 2011). The brown midrib (BMR) mutant phenotype in maize, sorghum, and pearl millet is a classic lignin alteration in cell walls, but no C3 plants have been reported with this cell wall phenotype. Putative BMR ortholog mutants in C3 plants seem to have a different basis; BMR rice mutant orthologs are known, but present different phenotypes if altered (Sattler et al., 2010). A digestibility assay comparing two C3 and two C4 grasses native to Australia showed that C3 grass leaves had more facile enzymatic hydrolysis compared to that of the C4 grasses, which was attributed to C3/C4 differences in leaf anatomy (Wilson and Hacker, 1987). *In vitro* evaluations using purified cellulose and lignin polymer models demonstrated that for lignin to affect cellulose saccharification it must be specifically cross linked to cellulose (Jung et al., 2011). Grass arabinoxylans cross link to lignin by ether bonds using ferulate and diferulate molecules (Burr and Fry, 2009). Dicot cell walls lack arabinoxylans for cross linking

to lignin and therefore are inadequate to use as an improved digestibility model for C4 grass crops. Alternatively, dicot cell walls are primarily composed of xyloglucan hemicellulose, which cross link to lignin (Scheller and Ulvskov, 2010; Furtado et al., 2014; Damm et al., 2016). Utilization of C4 grass species is more appropriate for evaluation of CWD enzymes if the target feedstock is a C4 grass. The applied approach for evaluation of novel autohydrolytic lines should be used with recently improved transformation strategies of many C4 crops (e.g., maize, switchgrass, sorghum) to reduce variability to real world marketable biomass (Ishida et al., 2007; Brutnell et al., 2010, 2015; Li and Qu, 2011; Liu and Godwin, 2012; Lambert et al., 2014).

## CONCLUSION

There are many economic pinch points in the production of cellulosic biofuels. Enzyme cost is one of them and is not trivial. Using recombinant systems, including the plant feedstock itself, is likely necessary to produce advanced biofuels in a sustainable manner (Gressel, 2008). Plants, including a few species that are relevant proxies for cellulosic feedstocks, have been used as hosts for overexpression of microbial GH genes. We envisage insect-derived GHs playing a role in reducing cell wall recalcitrance in transgenic plants owing to their high diversity of function and eukaryotic origin. Utilization of multiple enzyme classes should be used in concert in feedstocks, coupled with appropriate subcellular targeting as well as spatial and temporal control of synthesis. While there is much research needed to produce a commercial cellulosic feedstock with autocatalytic properties, field experiments with E1-producing transgenic maize demonstrates that research in this area should be worthwhile. The combination of improved biomass feedstocks and enzyme technologies is a step toward a renewable plant-based fuel system.

## AUTHOR CONTRIBUTIONS

JW drafted the manuscript and MM and CS participated in the drafts and revisions. All authors read and consented to the final version of the manuscript.

## ACKNOWLEDGMENTS

This work was supported by funding from the BioEnergy Science Center (DE-PS02-06ER64304). The BioEnergy Science Center is a U.S. Department of Energy Bioenergy Research Center supported by the Office of Biological and Environmental Research in the DOE Office of Science. We also thank the University of Tennessee, Ivan Racheff Endowment, and USDA HATCH funds for supporting the research program of CNS.

## REFERENCES

- Bae, H. J., Kim, H. J., and Kim, Y. S. (2008). Production of a recombinant xylanase in plants and its potential for pulp bleaching applications. *Bioresour. Technol.* 99, 3513–3519. doi: 10.1016/j.biortech.2007.07.064
- Behura, S. K. (2015). Insect phylogenomics. *Insect Mol. Biol.* 24, 403–411. doi: 10.1111/imb.12174
- Bhalla, A., Bansal, N., Kumar, S., Bischoff, K. M., and Sani, R. K. (2013). Improved lignocellulose conversion to biofuels with thermophilic bacteria and thermostable enzymes. *Bioresour. Technol.* 128, 751–759. doi: 10.1016/j.biortech.2012.10.145
- Bhat, M. K. (2000). Cellulases and related enzymes in biotechnology. *Biotechnol. Adv.* 18, 355–383. doi: 10.1016/S0734-9750(00)00041-0
- Biswas, G. C. G., Ransom, C., and Sticklen, M. (2006). Expression of biologically active *Acidothermus cellulolyticus* endoglucanase in transgenic maize plants. *Plant Sci.* 171, 617–623. doi: 10.1016/j.plantsci.2006.06.004
- Borkhardt, B., Harholt, J., Ulvskov, P., Ahring, B. K., Jørgensen, B., and Brinch-Pedersen, H. (2010). Autohydrolysis of plant xylans by apoplastic expression of thermophilic bacterial endo-xylanases. *Plant Biotechnol. J.* 8, 363–374. doi: 10.1111/j.1467-7652.2010.00506.x
- Boyer, L. J., Vega, J. L., Klasson, K. T., Clausen, E. C., and Gaddy, J. L. (1992). The effects of furfural on ethanol-production by *Saccharomyces cerevisiae* in batch culture. *Biomass Bioenerg.* 3, 41–48. doi: 10.1016/0961-9534(92)90018-L
- Bruce, A. M., and Palfreyman, J. W. (1998). *Forest Products Biotechnology*. London; Bristol, PA: Taylor & Francis.
- Brune, A. (2014). Symbiotic digestion of lignocellulose in termite guts. *Nat. Rev. Microbiol.* 12, 168–180. doi: 10.1038/nrmicro3182
- Brunecky, R., Selig, M. J., Vinzant, T. B., Himmel, M. E., Lee, D., Blaylock, M. J., et al. (2011). In planta expression of *A. cellulolyticus* Cel5A endoglucanase reduces cell wall recalcitrance in tobacco and maize. *Biotechnol. Biofuels* 4, 1–11. doi: 10.1186/1754-6834-4-1
- Brutnell, T. P., Bennetzen, J. L., and Vogel, J. P. (2015). *Brachypodium distachyon* and *Setaria viridis*: model genetic systems for the grasses. *Annu. Rev. Plant Biol.* 66, 465–485. doi: 10.1146/annurev-arplant-042811-105528
- Brutnell, T. P., Wang, L., Swartwood, K., Goldschmidt, A., Jackson, D., Zhu, X. G., et al. (2010). *Setaria viridis*: a model for C4 photosynthesis. *Plant Cell* 22, 2537–2544. doi: 10.1105/tpc.110.075309
- Burr, S. J., and Fry, S. C. (2009). Feruloylated arabinoxylans are oxidatively cross-linked by extracellular maize peroxidase but not by horseradish peroxidase. *Mol. Plant* 2, 883–892. doi: 10.1093/mp/ssp044
- Carroll, A., and Somerville, C. (2009). Cellulosic biofuels. *Annu. Rev. Plant Biol.* 60, 165–182. doi: 10.1146/annurev-arplant.043008.092125
- Chang, C. J., Wu, C. P., Lu, S. C., Chao, A. L., Ho, T. H., Yu, S. M., et al. (2012). A novel exo-cellulase from white spotted longhorn beetle (*Anoplophora malasiaca*). *Insect Biochem. Mol. Biol.* 42, 629–636. doi: 10.1016/j.ibmb.2012.05.002
- Chou, H. L., Dai, Z., Hsieh, C. W., and Ku, M. S. (2011). High level expression of *Acidothermus cellulolyticus* beta-1, 4-endoglucanase in transgenic rice enhances the hydrolysis of its straw by cultured cow gastric fluid. *Biotechnol. Biofuels* 4:58. doi: 10.1186/1754-6834-4-58
- Cosgrove, D. J. (2005). Growth of the plant cell wall. *Nat. Rev. Mol. Cell Biol.* 6, 850–861. doi: 10.1038/nrm1746
- Dai, Z., Hooker, B. S., Anderson, D. B., and Thomas, S. R. (2000). Expression of *Acidothermus cellulolyticus* endoglucanase E1 in transgenic tobacco: biochemical characteristics and physiological effects. *Transgenic Res.* 9, 43–54. doi: 10.1023/A:1008922404834
- Dai, Z., Hooker, B. S., Quesenberry, R. D., and Gao, J. (1999). Expression of *Trichoderma reesei* exo-cellobiohydrolase I in transgenic tobacco leaves and calli. *Appl. Biochem. Biotechnol.* 77–79, 689–699. doi: 10.1385/ABAB:79:1-3:689
- Dai, Z., Hooker, B. S., Quesenberry, R. D., and Thomas, S. R. (2005). Optimization of *Acidothermus cellulolyticus* endoglucanase (E1) production in transgenic tobacco plants by transcriptional, post-transcription and post-translational modification. *Transgenic Res.* 14, 627–643. doi: 10.1007/s11248-005-5695-5
- Dai, Z. Y., Hooker, B. S., Anderson, D. B., and Thomas, S. R. (2000). Improved plant-based production of E1 endoglucanase using potato: expression optimization and tissue targeting. *Mol. Breeding* 6, 277–285. doi: 10.1023/A:1009653011948
- Dal'Molin, C. G., Quek, L. E., Palfreyman, R. W., Brumbley, S. M., and Nielsen, L. K. (2010). C4GEM, a genome-scale metabolic model to study C4 plant metabolism. *Plant Physiol.* 154, 1871–1885. doi: 10.1104/pp.110.166488
- Damm, T., Commandeur, U., Fischer, R., Usadel, B., and Klose, H. (2016). Improving the utilization of lignocellulosic biomass by polysaccharide modification. *Process Biochem.* 51, 288–296. doi: 10.1016/j.procbio.2015.12.003
- Dow, J. A. (1992). pH Gradients in Lepidopteran midgut. *J. Exp. Biol.* 172, 355–375.
- Fan, Z., and Yuan, L. (2010). Production of multifunctional chimaeric enzymes in plants: a promising approach for degrading plant cell wall from within. *Plant Biotechnol. J.* 8, 308–315. doi: 10.1111/j.1467-7652.2009.00484.x
- Fischer, R., Schillberg, S., Hellwig, S., Twyman, R. M., and Drossard, J. (2012). GMP issues for recombinant plant-derived pharmaceutical proteins. *Biotechnol. Adv.* 30, 434–439. doi: 10.1016/j.biotechadv.2011.08.007
- Furtado, A., Lupoi, J. S., Hoang, N. V., Healey, A., Singh, S., Simmons, B. A., et al. (2014). Modifying plants for biofuel and biomaterial production. *Plant Biotechnol. J.* 12, 1246–1258. doi: 10.1111/pbi.12300
- Garda, M., Devaiah, S. P., Vicuna Requesens, D., Chang, Y. K., Dabul, A., Hanson, C., et al. (2015). Assessment of field-grown cellulase-expressing corn. *Transgenic Res.* 24, 185–198. doi: 10.1007/s11248-014-9838-4
- Gressel, J. (2008). Transgenics are imperative for biofuel crops. *Plant Sci.* 174, 246–263. doi: 10.1016/j.plantsci.2007.11.009
- Harrison, M. D., Geijskes, J., Coleman, H. D., Shand, K., Kinkema, M., Palupe, A., et al. (2011). Accumulation of recombinant cellobiohydrolase and endoglucanase in the leaves of mature transgenic sugar cane. *Plant Biotechnol. J.* 9, 884–896. doi: 10.1111/j.1467-7652.2011.00597.x
- Heliconius Genome Consortium (2012). Butterfly genome reveals promiscuous exchange of mimicry adaptations among species. *Nature* 487, 94–98. doi: 10.1038/nature11041
- Hellwig, S., Drossard, J., Twyman, R. M., and Fischer, R. (2004). Plant cell cultures for the production of recombinant proteins. *Nat. Biotechnol.* 22, 1415–1422. doi: 10.1038/nbt1027
- Henrissat, B. (1991). A classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 280(Pt 2), 309–316. doi: 10.1042/bj2800309
- Henrissat, B., and Bairoch, A. (1993). New families in the classification of glycosyl hydrolases based on amino acid sequence similarities. *Biochem. J.* 293(Pt 3), 781–788. doi: 10.1042/bj2930781
- Henrissat, B., and Bairoch, A. (1996). Updating the sequence-based classification of glycosyl hydrolases. *Biochem. J.* 316(Pt 2), 695–696. doi: 10.1042/bj3160695
- Herbers, K., Wilke, I., and Sonnewald, U. (1995). A thermostable xylanase from *Clostridium thermocellum* expressed at high-levels in the apoplast of transgenic tobacco has no detrimental effects and is easily purified. *Bio. Technol.* 13, 63–66. doi: 10.1038/nbt0195-63
- Hirayama, K., Watanabe, H., Tokuda, G., Kitamoto, K., and Arioka, M. (2010). Purification and characterization of termite endogenous beta-1,4-endoglucanases produced in *Aspergillus oryzae*. *Biosci. Biotechnol. Biochem.* 74, 1680–1686. doi: 10.1271/bbb.100296
- Hood, E. E., Devaiah, S. P., Fake, G., Egelkrout, E., Teoh, K. T., Requesens, D. V., et al. (2012). Manipulating corn germplasm to increase recombinant protein accumulation. *Plant Biotechnol. J.* 10, 20–30. doi: 10.1111/j.1467-7652.2011.00627.x
- Hood, E. E., Love, R., Lane, J., Bray, J., Clough, R., Pappu, K., et al. (2007). Subcellular targeting is a key condition for high-level accumulation of cellulase protein in transgenic maize seed. *Plant Biotechnol. J.* 5, 709–719. doi: 10.1111/j.1467-7652.2007.00275.x
- Horvath, H., Huang, J., Wong, O., Kohl, E., Okita, T., Kannangara, C. G., et al. (2000). The production of recombinant proteins in transgenic barley grains. *Proc. Natl. Acad. Sci. U.S.A.* 97, 1914–1919. doi: 10.1073/pnas.030527497
- Hyunjong, B., Lee, D. S., and Hwang, I. (2006). Dual targeting of xylanase to chloroplasts and peroxisomes as a means to increase protein accumulation in plant cells. *J. Exper. Bot.* 57, 161–169. doi: 10.1093/jxb/erj019
- Inoue, T., Moriya, S., Ohkuma, M., and Kudo, T. (2005). Molecular cloning and characterization of a cellulase gene from a symbiotic protist of the lower termite, *Coptotermes formosanus*. *Gene* 349, 67–75. doi: 10.1016/j.gene.2004.11.048

- International Aphid Genomics Consortium (2010). Genome sequence of the pea aphid *Acyrtosiphon pisum*. *PLoS Biol.* 8:e1000313. doi: 10.1371/journal.pbio.1000313
- Ishida, Y., Hiei, Y., and Komari, T. (2007). Agrobacterium-mediated transformation of maize. *Nature Protoc.* 2, 1614–1621. doi: 10.1038/nprot.2007.241
- Ito, Y., Arikawa, K., Antonio, B. A., Ohta, I., Naito, S., Mukai, Y., et al. (2005). Rice Annotation Database (RAD): a contig-oriented database for map-based rice genomics. *Nucleic Acids Res.* 33, D651–D655. doi: 10.1093/nar/gki083
- Jensen, L. G., Olsen, O., Kops, O., Wolf, N., Thomsen, K. K., and von Wettstein, D. (1996). Transgenic barley expressing a protein-engineered, thermostable (1,3-1,4)-beta-glucanase during germination. *Proc. Natl. Acad. Sci. U.S.A.* 93, 3487–3491. doi: 10.1073/pnas.93.8.3487
- Jeon, J. S., and An, G. H. (2001). Gene tagging in rice: a high throughput system for functional genomics. *Plant Sci.* 161, 211–219. doi: 10.1016/S0168-9452(01)00414-9
- Jin, R., Richter, S., Zhong, R., and Lippa, G. K. (2003). Expression and import of an active cellulase from a thermophilic bacterium into the chloroplast both *in vitro* and *in vivo*. *Plant Mol. Biol.* 51, 493–507. doi: 10.1023/A:1022354124741
- Jung, H. G., Mertens, D. R., and Phillips, R. L. (2011). Effect of reduced ferulate-mediated lignin/arabinoxylan cross-linking in corn silage on feed intake, digestibility, and milk production. *J. Dairy Sci.* 94, 5124–5137. doi: 10.3168/jds.2011-4495
- Jung, S., Kim, S., Bae, H., Lim, H. S., and Bae, H. J. (2010). Expression of thermostable bacterial beta-glucosidase (BglB) in transgenic tobacco plants. *Bioresour. Technol.* 101, 7155–7161. doi: 10.1016/j.biortech.2010.03.140
- Jung, S. K., Lindenmuth, B. E., McDonald, K. A., Hwang, H., Bui, M. Q., Falk, B. W., et al. (2014). Agrobacterium tumefaciens mediated transient expression of plant cell wall-degrading enzymes in detached sunflower leaves. *Biotechnol. Prog.* 30, 905–915. doi: 10.1002/btpr.1888
- Jung, S., Lee, D. S., Kim, Y. O., Joshi, C. P., and Bae, H. J. (2013). Improved recombinant cellulase expression in chloroplast of tobacco through promoter engineering and 5' amplification promoting sequence. *Plant Mol. Biol.* 83, 317–328. doi: 10.1007/s11103-013-0088-2
- Kawazu, T., Sun, J. L., Shibata, M., Kimura, T., Sakka, K., and Ohmiya, K. (1999). Expression of a bacterial endoglucanase gene in tobacco increases digestibility of its cell wall fibers. *J. Biosci. Bioeng.* 88, 421–425. doi: 10.1016/S1389-1723(99)80220-5
- Kim, N., Choo, Y. M., Lee, K. S., Hong, S. J., Seol, K. Y., Je, Y. H., et al. (2008). Molecular cloning and characterization of a glycosyl hydrolase family 9 cellulase distributed throughout the digestive tract of the cricket *Teleogryllus emma*. *Comp. Biochem. Physiol. B, Biochem. Mol. Biol.* 150, 368–376. doi: 10.1016/j.cbpb.2008.04.005
- Kimura, T., Mizutani, T., Tanaka, T., Koyama, T., Sakka, K., and Ohmiya, K. (2003). Molecular breeding of transgenic rice expressing a xylanase domain of the *xynA* gene from *Clostridium thermocellum*. *Appl. Microbiol. Biotechnol.* 62, 374–379. doi: 10.1007/s00253-003-1301-z
- Klein-Marcuschamer, D., Oleskowicz-Popiel, P., Simmons, B. A., and Blanch, H. W. (2012). The challenge of enzyme cost in the production of lignocellulosic biofuels. *Biotechnol. Bioeng.* 109, 1083–1087. doi: 10.1002/bit.24370
- Klose, H., Gunl, M., Usadel, B., Fischer, R., and Commandeur, U. (2013). Ethanol inducible expression of a mesophilic cellulase avoids adverse effects on plant development. *Biotechnol. Biofuels* 6:53. doi: 10.1186/1754-6834-6-53
- Klose, H., Gunl, M., Usadel, B., Fischer, R., and Commandeur, U. (2015). Cell wall modification in tobacco by differential targeting of recombinant endoglucanase from *Trichoderma reesei*. *BMC Plant Biol.* 15:54. doi: 10.1186/s12870-015-0443-3
- Klose, H., Röder, J., Girfoglio, M., Fischer, R., and Commandeur, U. (2012). Hyperthermophilic endoglucanase for in planta lignocellulose conversion. *Biotechnol. Biofuels* 5:63. doi: 10.1186/1754-6834-5-63
- Kostylev, M., and Wilson, D. (2012). Synergistic interactions in cellulose hydrolysis. *Biotechnol. Biofuels* 3, 61–70. doi: 10.4155/bfs.11.150
- Lambert, C., Garvey, M., Klinger, J., Heesel, D., Klose, H., Fischer, R., et al. (2014). Challenges and advances in the heterologous expression of cellulolytic enzymes: a review. *Biotechnol. Biofuels* 7, 135–150. doi: 10.1186/s13068-014-0135-5
- Lee, D. S., Lee, K. H., Jung, S., Jo, E. J., Han, K. H., and Bae, H. J. (2012). Synergistic effects of 2A-mediated polyproteins on the production of lignocellulose degradation enzymes in tobacco plants. *J. Exp. Bot.* 63, 4797–4810. doi: 10.1093/jxb/ers159
- Lee, S. J., Kim, S. R., Yoon, H. J., Kim, I., Lee, K. S., Je, Y. H., et al. (2004). cDNA cloning, expression, and enzymatic activity of a cellulase from the mulberry longicorn beetle, *Apriona germari*. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 139, 107–116. doi: 10.1016/j.cbpc.2004.06.015
- Lee, S. J., Lee, K. S., Kim, S. R., Gui, Z. Z., Kim, Y. S., Yoon, H. J., et al. (2005). A novel cellulase gene from the mulberry longicorn beetle, *Apriona germari*: gene structure, expression, and enzymatic activity. *Comp. Biochem. Physiol. B Biochem. Mol. Biol.* 140, 551–560. doi: 10.1016/j.cbpc.2004.12.003
- Li, R. Y., and Qu, R. D. (2011). High throughput Agrobacterium-mediated switchgrass transformation. *Biomass. Bioenerg.* 35, 1046–1054. doi: 10.1016/j.biombioe.2010.11.025
- Liu, G., and Godwin, I. D. (2012). Highly efficient sorghum transformation. *Plant Cell Rep.* 31, 999–1007. doi: 10.1007/s00299-011-1218-4
- Liu, N., Yan, X., Zhang, M., Xie, L., Wang, Q., Huang, Y., et al. (2011). Microbiome of fungus-growing termites: a new reservoir for lignocellulase genes. *Appl. Environ. Microbiol.* 77, 48–56. doi: 10.1128/AEM.01521-10
- Liu, W. S., Yuan, J. S., and Stewart, C. N. Jr. (2014). Advanced genetic tools for plant biotechnology. *Nat. Rev. Genet.* 15, 781–973. doi: 10.1038/nrg3583
- Lombard, V., Ramulu, H. G., Drula, E., Coutinho, P. M., and Henrissat, B. (2014). The carbohydrate-active enzymes database (CAZy) in 2013. *Nucleic Acids Res.* 42, D490–D495. doi: 10.1093/nar/gkt1178
- Lopez-Casado, G., Urbanowicz, B. R., Damasceno, C. M., and Rose, J. K. (2008). Plant glycosyl hydrolases and biofuels: a natural marriage. *Curr. Opin. Plant Biol.* 11, 329–337. doi: 10.1016/j.pbi.2008.02.010
- Mahadevan, S. A., Wi, S. G., Kim, Y. O., Lee, K. H., and Bae, H. J. (2011). In planta differential targeting analysis of *Thermotoga maritima* Cel5A and CBM6-engineered Cel5A for autohydrolysis. *Transgenic Res.* 20, 877–886. doi: 10.1007/s11248-010-9464-8
- Mahadevan, S. A., Wi, S. G., Lee, D. S., and Bae, H. J. (2008). Site-directed mutagenesis and CBM engineering of Cel5A (*Thermotoga maritima*). *FEMS Microbiol. Lett.* 287, 205–211. doi: 10.1111/j.1574-6968.2008.01324.x
- Marana, S. R., Andrade, E. H., Ferreira, C., and Terra, W. R. (2004). Investigation of the substrate specificity of a beta-glycosidase from *Spodoptera frugiperda* using site-directed mutagenesis and bioenergetics analysis. *Eur. J. Biochem.* 271, 4169–4177. doi: 10.1111/j.1432-1033.2004.04354.x
- Matsumoto, T., Wu, J. Z., Kanamori, H., Katayose, Y., Fujisawa, M., Namiki, N., et al. (2005). The map-based sequence of the rice genome. *Nature* 436, 793–800. doi: 10.1038/nature03895
- Mattéotti, C., Bauwens, J., Brasseur, C., Tarayre, C., Thonart, P., Destain, J., et al. (2012). Identification and characterization of a new xylanase from Gram-positive bacteria isolated from termite gut (*Reticulitermes santonensis*). *Protein Expr. Purif.* 83, 117–127. doi: 10.1016/j.pep.2012.03.009
- Mattéotti, C., Haubruge, E., Thonart, P., Francis, F., De Pauw, E., Portetel, D., et al. (2011). Characterization of a new beta-glucosidase/beta-xylosidase from the gut microbiota of the termite (*Reticulitermes santonensis*). *FEMS Microbiol. Lett.* 314, 147–157. doi: 10.1111/j.1574-6968.2010.02161.x
- Mir, B. A., Mewalal, R., Mizrachi, E., Myburg, A. A., and Cowan, D. A. (2014). Recombinant hyperthermophilic enzyme expression in plants: a novel approach for lignocellulose digestion. *Trends Biotechnol.* 32, 281–289. doi: 10.1016/j.tibtech.2014.03.003
- Misof, B., Liu, S. L., Meusemann, K., Peters, R. S., Donath, A., Mayer, C., et al. (2014). Phylogenomics resolves the timing and pattern of insect evolution. *Science* 346, 763–767. doi: 10.1126/science.1257570
- Nelson, T. (2011). The grass leaf developmental gradient as a platform for a systems understanding of the anatomical specialization of C-4 leaves. *J. Exp. Bot.* 62, 3039–3048. doi: 10.1093/jxb/err072
- Ni, J. F., and Tokuda, G. (2013). Lignocellulose-degrading enzymes from termites and their symbiotic microbiota. *Biotechnol. Adv.* 31, 838–850. doi: 10.1016/j.biotechadv.2013.04.005
- Ni, J. F., Tokuda, G., Takehara, M., and Watanabe, H. (2007). Heterologous expression and enzymatic characterization of beta-glucosidase from the drywood-eating termite, *Neotermes kosshunensis*. *Appl. Entomol. Zool.* 42, 457–463. doi: 10.1303/aez.2007.457

- Nigorikawa, M., Watanabe, A., Furukawa, K., Sonoki, T., and Ito, Y. (2012). Enhanced saccharification of rice straw by overexpression of rice exo-glucanase. *Rice* 5:14. doi: 10.1186/1939-8433-5-14
- Ohkuma, M. (2003). Termite symbiotic systems: efficient bio-recycling of lignocellulose. *Appl. Microbiol. Biotechnol.* 61, 1–9. doi: 10.1007/s00253-002-1189-z
- Oppert, C., Klingeman, W. E., Willis, J. D., Oppert, B., and Jurat-Fuentes, J. L. (2010). Prospecting for cellulolytic activity in insect digestive fluids. *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* 155, 145–154. doi: 10.1016/j.cbpb.2009.10.014
- Oraby, H., Venkatesh, B., Dale, B., Ahmad, R., Ransom, C., Oehmke, J., et al. (2007). Enhanced conversion of plant biomass into glucose using transgenic rice-produced endoglucanase for cellulosic ethanol. *Transgenic Res.* 16, 739–749. doi: 10.1007/s11248-006-9064-9
- Patel, M., Johnson, J. S., Brettell, R. I. S., Jacobsen, J., and Xue, G. P. (2000). Transgenic barley expressing a fungal xylanase gene in the endosperm of the developing grains. *Mol. Breed.* 6, 113–123. doi: 10.1023/A:1009640427515
- Ransom, C., Balan, V., Biswas, G., Dale, B., Crockett, E., and Sticklen, M. (2007). Heterologous *Acidothermus cellulolyticus* 1,4-beta-endoglucanase E1 produced within the corn biomass converts corn stover into glucose. *Appl. Biochem. Biotechnol.* 137–140, 207–219. doi: 10.1007/s12010-007-9053-3
- Reese, E. T., Siu, R. G. H., and Levinson, H. S. (1950). The biological degradation of soluble cellulose derivatives and its relationship to the mechanism of cellulose hydrolysis. *J. Bacteriol.* 59, 485–497.
- Richards, S., Liu, Y., Bettencourt, B. R., Hradecky, P., Letovsky, S., Nielsen, R., et al. (2005). Comparative genome sequencing of *Drosophila pseudoobscura*: chromosomal, gene, and cis-element evolution. *Genome Res.* 15, 1–18. doi: 10.1101/gr.3059305
- Robinson, G. E., Hackett, K. J., Purcell-Miramontes, M., Brown, S. J., Evans, J. D., Goldsmith, M. R., et al. (2011). Creating a buzz about insect genomes. *Science* 331, 1386–1386. doi: 10.1126/science.331.6023.1386
- Saha, B. C. (2003). Hemicellulose bioconversion. *J. Ind. Microbiol. Biotechnol.* 30, 279–291. doi: 10.1007/s10295-003-0049-x
- Sattler, S. E., Funnell-Harris, D. L., and Pedersen, J. F. (2010). Brown midrib mutations and their importance to the utilization of maize, sorghum, and pearl millet lignocellulosic tissues. *Plant Sci.* 178, 229–238. doi: 10.1016/j.plantsci.2010.01.001
- Scharf, M. E., Kovaleva, E. S., Jadhao, S., Campbell, J. H., Buchman, G. W., and Boucias, D. G. (2010). Functional and translational analyses of a beta-glucosidase gene (glycosyl hydrolase family 1) isolated from the gut of the lower termite *Reticulitermes flavipes*. *Insect Biochem. Mol. Biol.* 40, 611–620. doi: 10.1016/j.ibmb.2010.06.002
- Scheller, H. V., and Ulvskov, P. (2010). Hemicelluloses. *Annu. Rev. Plant Biol.* 61, 263–289. doi: 10.1146/annurev-arplant-042809-112315
- Schena, M., Lloyd, A. M., and Davis, R. W. (1991). A steroid-inducible gene expression system for plant cells. *Proc. Natl. Acad. Sci. U.S.A.* 88, 10421–10425. doi: 10.1073/pnas.88.23.10421
- Shelomi, M., Watanabe, H., and Arakawa, G. (2014). Endogenous cellulase enzymes in the stick insect (Phasmatodea) gut. *J. Insect Physiol.* 60, 25–30. doi: 10.1016/j.jinsphys.2013.10.007
- Shen, B., Sun, X., Zuo, X., Shilling, T., Apgar, J., Ross, M., et al. (2012). Engineering a thermoregulated intein-modified xylanase into maize for consolidated lignocellulosic biomass processing. *Nat. Biotechnol.* 30, 1131–1136. doi: 10.1038/nbt.2402
- Shen, H., Mazarei, M., Hisano, H., Escamilla-Trevino, L., Fu, C., Pu, Y., Rudis, M. R., et al. (2013). A genomics approach to deciphering lignin biosynthesis in switchgrass. *Plant Cell* 25, 4342–4361. doi: 10.1105/tpc.113.118828
- Shi, W. B., Ding, S. Y., and Yuan, J. S. (2011). Comparison of insect gut cellulase and xylanase activity across different insect species with distinct food sources. *Bioenerg Res.* 4, 1–10. doi: 10.1007/s12155-010-9096-0
- Shi, W., Xie, S., Chen, X., Sun, S., Zhou, X., Liu, L., et al. (2013). Comparative genomic analysis of the microbiome [corrected] of herbivorous insects reveals eco-environmental adaptations: biotechnology applications. *PLoS Genet.* 9:e1003131. doi: 10.1371/journal.pgen.1003131
- Shirley, D., Oppert, C., Reynolds, T. B., Miracle, B., Oppert, B., Klingeman, W. E., et al. (2014). Expression of an endoglucanase from *Tribolium castaneum* (TcEG1) in *Saccharomyces cerevisiae*. *Insect Sci.* 21, 609–618. doi: 10.1111/1744-7917.12069
- Slaytor, M. (1992). Cellulose digestion in termites and cockroaches - What role do symbionts play? *Comp. Biochem. Phys. B* 103, 775–784. doi: 10.1016/0305-0491(92)90194-V
- Sun, Y., Cheng, J. J., Himmel, M. E., Skory, C. D., Adney, W. S., Thomas, S. R., et al. (2007). Expression and characterization of *Acidothermus cellulolyticus* E1 endoglucanase in transgenic duckweed *Lemna minor* 8627. *Bioresour. Technol.* 98, 2866–2872. doi: 10.1016/j.biortech.2006.09.055
- Taylor, L. E. II, Dai, Z., Decker, S. R., Brunecy, R., Adney, W. S., Ding, S. Y., et al. (2008). Heterologous expression of glycosyl hydrolases in planta: a new departure for biofuels. *Trends Biotechnol.* 26, 413–424. doi: 10.1016/j.tibtech.2008.05.002
- Todaka, N., Lopez, C. M., Inoue, T., Saita, K., Maruyama, J., Arioka, M., et al. (2010). Heterologous expression and characterization of an endoglucanase from a symbiotic protist of the lower termite, *Reticulitermes speratus*. *Appl. Microbiol. Biotechnol.* 160, 1168–1178. doi: 10.1007/s12010-009-8626-8
- Tokuda, G., and Watanabe, H. (2007). Hidden cellulases in termites: revision of an old hypothesis. *Biol. Lett.* 3, 336–339. doi: 10.1098/rsbl.2007.0073
- Tribolium Genome Sequencing, C., Richards, S., Gibbs, R. A., Weinstock, G. M., Brown, S. J., Denell, R., et al. (2008). The genome of the model beetle and pest *Tribolium castaneum*. *Nature* 452, 949–955. doi: 10.1038/nature06784
- Uchima, C. A., Tokuda, G., Watanabe, H., Kitamoto, K., and Arioka, M. (2011). Heterologous expression and characterization of a glucose-stimulated beta-glucosidase from the termite *Neotermes koshunensis* in *Aspergillus oryzae*. *Appl. Microbiol. Biotechnol.* 89, 1761–1771. doi: 10.1007/s00253-010-2963-y
- Uchima, C. A., Tokuda, G., Watanabe, H., Kitamoto, K., and Arioka, M. (2012). Heterologous expression in *Pichia pastoris* and characterization of an endogenous thermostable and high-glucose-tolerant beta-glucosidase from the termite *Nasutitermes takasagoensis*. *Appl. Environ. Microbiol.* 78, 4288–4293. doi: 10.1128/AEM.07718-11
- Valencia, A., Alves, A. P., and Siegfried, B. D. (2013). Molecular cloning and functional characterization of an endogenous endoglucanase belonging to GHF45 from the western corn rootworm, *Diabrotica virgifera virgifera*. *Gene* 513, 260–267. doi: 10.1016/j.gene.2012.10.046
- Vogel, J. P., Garvin, D. F., Mockler, T. C., Schmutz, J., Rokhsar, D., Bevan, M. W., et al. (2010). Genome sequencing and analysis of the model grass *Brachypodium distachyon*. *Nature* 463, 763–768. doi: 10.1038/nature08747
- Vogel, J. P., Tuna, M., Budak, H., Huo, N. X., Gu, Y. Q., and Steinwand, M. A. (2009). Development of SSR markers and analysis of diversity in Turkish populations of *Brachypodium distachyon*. *BMC Plant Biol.* 9:88. doi: 10.1186/1471-2229-9-88
- Wang, Q., Qian, C., Zhang, X. Z., Liu, N., Yan, X., and Zhou, Z. (2012). Characterization of a novel thermostable beta-glucosidase from a metagenomic library of termite gut. *Enzyme Microb. Technol.* 51, 319–324. doi: 10.1016/j.enzymictec.2012.07.015
- Watanabe, H., Noda, H., Tokuda, G., and Lo, N. (1998). A cellulase gene of termite origin. *Nature* 394, 330–331. doi: 10.1038/28527
- Watanabe, H., and Tokuda, G. (2001). Animal cellulases. *Cell. Mol. Life Sci.* 58, 1167–1178. doi: 10.1007/PL00000931
- Watanabe, H., and Tokuda, G. (2010). Cellulolytic systems in insects. *Annu. Rev. Entomol.* 55, 609–632. doi: 10.1146/annurev-ento-112408-085319
- Wei, Y. D., Lee, K. S., Gui, Z. Z., Yoon, H. J., Kim, I., Zhang, G. Z., et al. (2006). Molecular cloning, expression, and enzymatic activity of a novel endogenous cellulase from the mulberry longicorn beetle, *Apriona germari*. *Comp. Biochem Physiol. B Biochem. Mol. Biol.* 145, 220–229. doi: 10.1016/j.cbpb.2006.07.007
- Weinstock, G. M., Robinson, G. E., Gibbs, R. A., Worley, K. C., Evans, J. D., Maleszka, R., et al. (2006). Insights into social insects from the genome of the honeybee *Apis mellifera*. *Nature* 443, 931–949. doi: 10.1038/nature05260
- Weng, X., Huang, Y., Hou, C., and Jiang, D. (2013). Effects of an exogenous xylanase gene expression on the growth of transgenic rice and the expression level of endogenous xylanase inhibitor gene RIXI. *J. Sci. Food. Agric.* 93, 173–179. doi: 10.1002/jsfa.5746
- Werren, J. H., Richards, S., Desjardins, C. A., Niehuis, O., Gadau, J., Colbourne, J. K., et al. (2010). Functional and evolutionary insights from the genomes of three parasitoid *Nasonia* species. *Science* 327, 343–348. doi: 10.1126/science.1178028
- Willis, J. D., Oppert, B., Oppert, C., Klingeman, W. E., and Jurat-Fuentes, J. L. (2011). Identification, cloning, and expression of a GHF9 cellulase

- from *Tribolium castaneum* (Coleoptera: Tenebrionidae). *J. Insect Physiol.* 57, 300–306. doi: 10.1016/j.jinsphys.2010.11.019
- Willis, J. D., Oppert, C., and Jurat-Fuentes, J. L. (2010). Methods for discovery and characterization of cellulolytic enzymes from insects. *Insect Sci.* 17, 184–198. doi: 10.1111/j.1744-7917.2010.01322.x
- Wilson, J. R., and Hacker, J. B. (1987). Comparative digestibility and anatomy of some sympatric C-3 and C-4 arid zone grasses. *Aust. J. Agr. Res.* 38, 287–295. doi: 10.1071/AR9870287
- Wu, Y., Chi, S., Yun, C., Shen, Y., Tokuda, G., and Ni, J. (2012). Molecular cloning and characterization of an endogenous digestive beta-glucosidase from the midgut of the fungus-growing termite *Macrotermes barneyi*. *Insect Mol. Biol.* 21, 604–614. doi: 10.1111/j.1365-2583.2012.01164.x
- Wyman, C. E. (2007). What is (and is not) vital to advancing cellulosic ethanol. *Trends Biotechnol.* 25, 153–157. doi: 10.1016/j.tibtech.2007.02.009
- Xia, D., Wei, Y., Zhang, G., Zhao, Q., Zhang, Y., Xiang, Z., et al. (2013). cDNA cloning, expression, and enzymatic activity of a novel endogenous cellulase from the beetle *Batocera horsfieldi*. *Gene* 514, 62–68. doi: 10.1016/j.gene.2012.08.044
- Xiao, Z. Z., Storms, R., and Tsang, A. (2004). Microplate-based filter paper assay to measure total cellulase activity. *Biotechnol. Bioeng.* 88, 832–837. doi: 10.1002/bit.20286
- Xue, G. P., Patel, M., Johnson, J. S., Smyth, D. J., and Vickers, C. E. (2003). Selectable marker-free transgenic barley producing a high level of cellulase (1,4-beta-glucanase) in developing grains. *Plant Cell Rep.* 21, 1088–1094. doi: 10.1007/s00299-003-0627-4
- Yang, P. L., Wang, Y. R., Bai, Y. G., Meng, K., Luo, H. Y., Yuan, T. Z., et al. (2007). Expression of xylanase with high specific activity from *Streptomyces olivaceoviridis* A1 in transgenic potato plants (*Solanum tuberosum* L.). *Biotechnol. Lett.* 29, 659–667. doi: 10.1007/s10529-006-9280-7
- Yeh, A. I., Huang, Y. C., and Chen, S. H. (2010). Effect of particle size on the rate of enzymatic hydrolysis of cellulose. *Carbohydr. Polym.* 79, 192–199. doi: 10.1016/j.carbpol.2009.07.049
- Yu, L. X., Gray, B. N., Rutzke, C. J., Walker, L. P., Wilson, D. B., and Hanson, M. R. (2007). Expression of thermostable microbial cellulases in the chloroplasts of nicotine-free tobacco. *J. Biotechnol.* 131, 362–369. doi: 10.1016/j.jbiotec.2007.07.942
- Zhang, Allen, A. B., and Lax, A. R. (2012). Functional analyses of the digestive beta-glucosidase of Formosan subterranean termites (*Coptotermes formosanus*). *J. Insect Physiol.* 58, 205–210. doi: 10.1016/j.jinsphys.2011.11.014
- Zhang, D., Lax, A. R., Bland, J. M., and Allen, A. B. (2011). Characterization of a new endogenous endo-beta-1,4-glucanase of Formosan subterranean termite (*Coptotermes formosanus*). *Insect Biochem. Mol. Biol.* 41, 211–218. doi: 10.1016/j.ibmb.2010.12.006
- Zhang, K., Agrawal, M., Harper, J., Chen, R., and Koros, W. J. (2011). Removal of the fermentation inhibitor, Furfural, using activated carbon in cellulosic-ethanol production. *Ind. Eng. Chem. Res.* 50, 14055–14060. doi: 10.1021/ie2013983
- Zhang, Q., Zhang, W., Lin, C. Y., Xu, X. L., and Shen, Z. C. (2012). Expression of an *Acidotherrmus cellulolyticus* endoglucanase in transgenic rice seeds. *Protein Expr. Purif.* 82, 279–283. doi: 10.1016/j.pep.2012.01.011
- Zhang, Y. H. P. (2015). Production of biofuels and biochemicals by *in vitro* synthetic biosystems: opportunities and challenges. *Biotechnol. Adv.* 33, 1467–1483. doi: 10.1016/j.biotechadv.2014.10.009
- Zhou, X., Kovaleva, E. S., Wu-Scharf, D., Campbell, J. H., Buchman, G. W., Boucias, D. G., et al. (2010). Production and characterization of a recombinant beta-1,4-endoglucanase (glycohydrolase family 9) from the termite *Reticulitermes flavipes*. *Arch. Insect Biochem. Physiol.* 74, 147–162. doi: 10.1002/arch.20368
- Ziegelhoffer, T., Raasch, J. A., and Austin-Phillips, S. (2001). Dramatic effects of truncation and sub-cellular targeting on the accumulation of recombinant microbial cellulase in tobacco. *Mol. Breeding* 8, 147–158. doi: 10.1023/A:1013338312948
- Ziegelhoffer, T., Will, J., and Austin-Phillips, S. (1999). Expression of bacterial cellulase genes in transgenic alfalfa (*Medicago sativa* L.), potato (*Solanum tuberosum* L.) and tobacco (*Nicotiana tabacum* L.). *Mol. Breeding* 5, 309–318. doi: 10.1023/A:1009646830403
- Ziegler, M. T., Thomas, S. R., and Danna, K. J. (2000). Accumulation of a thermostable endo-1,4-beta-D-glucanase in the apoplast of *Arabidopsis thaliana* leaves. *Mol. Breeding* 6, 37–46. doi: 10.1023/A:1009667524690

**Conflict of Interest Statement:** The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Copyright © 2016 Willis, Mazarei and Stewart. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.