Appl Microbiol Biotechnol
DOI 10.1007/s00253-017-8091-1

BIOENERGY AND BIOFUELS

Characterization of four endophytic fungi as potential consolidated bioprocessing hosts for conversion of lignocellulose into advanced biofuels

Weihua Wu[1] · Ryan W. Davis[1] · Mary Bao Tran-Gyamfi[1] · Alan Kuo[2] · Kurt LaButti[2] · Sirma Mihaltcheva[2] · Hope Hundley[2] · Mansi Chovatia[2] · Erika Lindquist[2] · Kerrie Barry[2] · Igor V. Grigoriev[2] · Bernard Henrissat[3,4] · John M. Gladden[1,5]

Received: 19 August 2016 / Revised: 13 December 2016 / Accepted: 21 December 2016

Abstract Recently, several endophytic fungi have been demonstrated to produce volatile organic compounds (VOCs) with properties similar to fossil fuels, called "mycodiesel," while growing on lignocellulosic plant and agricultural residues. The fact that endophytes are plant symbionts suggests that some may be able to produce lignocellulolytic enzymes, making them capable of both deconstructing lignocellulose and converting it into mycodiesel, two properties that indicate that these strains may be useful consolidated bioprocessing (CBP) hosts for the biofuel production. In this study, four endophytes *Hypoxylon* sp. CI4A, *Hypoxylon* sp. EC38, *Hypoxylon* sp. CO27, and *Daldinia eschscholzii* EC12 were selected and evaluated for their CBP potential. Analysis of their genomes indicates that these endophytes have a rich reservoir of biomass-deconstructing carbohydrate-active enzymes (CAZys), which includes enzymes active on both polysaccharides and lignin, as well as terpene synthases (TPSs), enzymes that may produce fuel-like molecules, suggesting that they do indeed have CBP potential. GC-MS analyses of their VOCs when grown on four representative lignocellulosic feedstocks revealed that these endophytes produce a wide spectrum of hydrocarbons, the majority of which are monoterpenes and sesquiterpenes, including some known biofuel candidates. Analysis of their cellulase activity when grown under the same conditions revealed that these endophytes actively produce endoglucanases, exoglucanases, and β-glucosidases. The richness of CAZymes as well as terpene synthases identified in these four endophytic fungi suggests that they are great candidates to pursue for development into platform CBP organisms.

Keywords Endophyte · Volatile organic compound · Terpene · Consolidated bioprocessing · CAZy · Ligninolytic enzymes

Electronic supplementary material The online version of this article (doi:10.1007/s00253-017-8091-1) contains supplementary material, which is available to authorized users.

✉ Weihua Wu
whwu@ucdavis.edu; wwu@sandia.gov

✉ John M. Gladden
jmgladden@lbl.gov; jmgladd@sandia.gov

[1] Biomass Science and Conversion Technologies, Sandia National Laboratories, 7011 East Avenue, Livermore, CA 94551, USA

[2] Department of Energy Joint Genome Institute, Walnut Creek, CA, USA

[3] CNRS, Aix-Marseille University, AFMB, Marseille, France

[4] Department of Biological Sciences, King Abdulaziz University, Jeddah, Saudi Arabia

[5] Joint BioEnergy Institute, Emeryville, CA, USA

Introduction

Rising demand for transportation fuels and concerns over fossil fuel-derived environmental pollution and greenhouse gas emissions have created a compelling need for the development of new sustainable alternative energy sources (Lynd et al. 2005). Biofuels have been considered a promising alternative to fossil fuels as they can overcome some of these issues. Many of the current state-of-the-art biofuel production platforms in deployment or under development are based on the bioconversion of lignocellulosic biomass into various liquid fuels, such as ethanol (Schuster and Chinn 2013; Wu 2013; Wu et al. 2013a), isobutanol (Higashide et al. 2012; Lan and Liao 2013; Wu et al. 2016a), biodiesel (Scott et al. 2010), and others. These platforms utilize various configurations of

Published online: 12 January 2017

multiple unit operations, such as biomass collection and storage, thermochemical, physical or biological pretreatment, saccharification, fermentation, and product recovery (Balat 2011; Haghighi Mood et al. 2013). Recent technoeconomic analysis suggests that pretreatment and saccharification dominate operational costs and thereby play a large part in determining the economic feasibility of these technologies (Lei et al. 2013).

An alternative strategy, consolidated bioprocessing (CBP), attempts to address cost associated with these multiunit operations by combining lignocellulolytic enzyme production, substrate hydrolysis, and fermentation into a single unit operation (Parisutham et al. 2014; Schuster and Chinn 2013). Significant efforts have been made to develop CBP organisms and consortia that can convert lignocellulosic biomass into biofuels and other valuable commodity chemicals (Akinosho et al. 2014; Fan et al. 2012; Favaro et al. 2015; Wei et al. 2014; Wu et al. 2013a, b, 2016b; Wu and Davis 2016), and to date, CBP organism development strategies fall into two major categories: (1) engineering a highly cellulolytic organism to produce ethanol and other bioproducts or (2) engineering non-cellulolytic biofuel-producing organisms to produce cellulolytic enzymes to hydrolyze pretreated lignocellulose and consume both hexose and pentose sugars (Lynd et al. 2005). Although significant progress has been made in engineering CBP strains of bacteria (Akinosho et al. 2014; Yamada et al. 2013), yeast (Favaro et al. 2015; Hasunuma and Kondo 2012; Wei et al. 2014), and filamentous fungi (Anasontzis and Christakopoulos 2014; Fan et al. 2012; Wu et al. 2013b), those efforts have not reached the level of maturity needed for commercial deployment.

Endophytes are relatively unstudied microbes that live endosymbiotically within plants without causing any overt harm to their hosts. Many endophytic fungi are known to produce a variety of biologically active secondary metabolites with the potential applications in anticancer (Tanasova and Sturla 2012), antimicrobial (Kovac et al. 2015), anti-inflammation (Guo et al. 2014), and other pharmaceutical applications (Pant et al. 2014). Recently, several endophytic fungi were shown to produce a wide spectrum of volatile organic compounds (VOCs) with potential fuel applications, called "mycodiesel." The major components of these VOCs are monoterpenes and sesquiterpenes, such as pinene, limonene, caryophyllene, chamigrene, and others (Strobel 2014, 2015). Compared to currently reported biofuel molecules, which have very high oxygen content (up to C/O of 2:1 for ethanol) and incompatibilities with the existing petroleum-based fuel infrastructure (http://www1.eere.energy.gov/bioenergy/pdfs/algal_biofuels_roadmap.pdf), many terpenes contain near-zero oxygen content and have high energy densities, making them particularly promising candidates for "drop-in" fossil fuel replacements, especially for aviation fuels. In addition to these endophytes' ability to produce mycodiesel, their commensal relationships with terrestrial plants suggest that they may also exhibit lignocellulolytic activity, another trait needed for CBP organisms.

In this study, four endophytic fungi *Hypoxylon* sp. CI4A, *Hypoxylon* sp. EC38, *Hypoxylon* sp. CO27, and *Daldinia eschscholzii* EC12 (hereafter called CI4A, EC38, CO27, and EC12) were investigated to assess their potential as CBP organisms. The genomes of the four endophytes were sequenced and annotated, and their lignocellulolytic and biofuel potential was evaluated by mining the genomes for CAZy/ligninolytic enzymes and terpene biosynthetic pathway genes. A functional characterization was also conducted by analyzing both VOC production and cellulolytic activity. The VOC profile of the four endophytes was obtained after growth on four different feedstocks: Avicel (crystalline cellulose), corn stover (CS), switchgrass (SG), and eucalyptus mill powder (eucalyptus). The endoglucanase, exoglucanase, and β-glucosidase cellulolytic activities were characterized using model substrates in liquid assays and by zymography. This study is the first assessment of the CBP potential of endophytic fungi and indicates that they may be suitable CBP production hosts for next-generation biofuels.

Material and methods

Strains, culture medium, and conditions

The endophytes *Hypoxylon* sp. CI4A, EC38, CO27, and *D. eschscholzii* EC12 grown on barley seeds were transferred onto potato dextrose agar plates to develop mycelia for 5 days at room temperature. The mycelia were collected and resuspended into 2 mL of potato dextrose medium (Sigma, St. Louis, MO, USA). One milliliter of potato dextrose with mycelia were transferred into 200 mL of fresh potato dextrose medium and cultured at room temperature for 5 days. The mycelia were then collected by filtering through 30-μm nylon membrane (Thermo, Santa Clara, CA, USA) and rinsed with 1× M9 salt solution three times to remove oligosaccharide residues. The collected mycelia were transferred to 200 mL of 1× M9 medium containing 20 g/L Avicel, switchgrass, corn stover, or eucalyptus mill powder as sole carbon source, respectively, for the induction of cellulase. The samples were taken out every 24 h, and cultural supernatants were prepared for following cellulase activity assay.

Neurospora crassa strain (FGSC #2489) was investigated as well for the comparison with the endophytes. Strain cultivation and cellulose induction was conducted as in previous studies (Fan et al. 2012; Wu et al. 2013b, c). Briefly, 10-day spores of *N. crassa* were collected, and the mycelia were removed by filtering through four layers of cheese cloth. The spores were inoculated into 1× Vogel's medium containing 20 g/L glucose as sole carbon at 10^6 spore/mL medium and cultured at 25 °C for 16 h. The mycelia were collected by filtration through 100-μm nylon membrane and rinsed with 1× Vogel's salts three times to remove glucose residue. The

mycelia were then transferred into 200 mL of 1× Vogel's medium containing 20 g/L Avicel, switchgrass, corn stover, or eucalyptus mill powder as the sole carbon source. All the chemicals were bought from Sigma, St. Louis, MO, USA, unless they were mentioned specifically in this study.

VOC profiling

Five-day-old mycelia of the endophytes *Hypoxylon* sp. CI4A, EC38, CO27, and *D. eschscholzii* EC12 cultivated on potato dextrose agar plates were collected and transferred into 20 mL of 1× M9 media containing 20 g/L Avicel, mill powder of switchgrass (SG), corn stover (CS), and eucalyptus mill powder as the sole carbon source, individually. The cultures were shaken at 25 °C, 180 rpm, for 2 days. The caps of the flasks were then sealed for the accumulation of VOCs for another 3 days. VOCs in the headspace of the flasks were extracted using solid-phase microextraction (SPME) and analyzed by GC-MS. The SPME syringe consists of 50/30 divinylbenzene/carboxen on polydimethylsiloxane on a stable flex fiber (Sigma, St. Louis, MO, USA). The SPME fiber was explored into the headspace of each culture flask for 2 h to absorb the VOCs. The detailed GC-MS methods were adapted from a previous study (Gladden et al. 2013). Spectral components were searched against the NIST 2015 mass spectral library (http://chemdata.nist.gov), and only components with mass spectra match factors >85% were reported as tentatively identified compounds. Furthermore, compounds with peak areas >1% of the total peak area in the chromatogram were reported.

Cellulase activity assay

Endoglucanase activity was measured using the dinitrosalicyclic (DNS) colorimetric method as described previously with minor modifications (Wu et al. 2013b, c). Briefly, a mixture of 1.5 mL of culture supernatant and 0.5 mL of carboxymethylcellulose (CM-cellulose) in 50 mM citrate buffer (pH 4.8) was incubated at 50 °C for 1 h. The reactions were quenched by adding 100 μl of reaction mixture into 600 μL of DNS reagent, and the resulted mixtures were incubated at 100 °C for 10 min. One unit of endoglucanase activity is defined as the amount of enzyme that produced 1 μmol of glucose in 1 min under the assay conditions described above.

Exoglucanase activities were measured using *p*-nitrophenyl-β-D-lactoside (*p*NPL) as the substrate described (Wu et al. 2013b, c) as below. The enzyme reaction solution contains 1.5 mL of culture supernatant and 0.5 mL of 1 mg/mL *p*NPL in 50 mM citrate buffer (pH 4.8). The reaction mixtures were incubated at 50 °C for 30 min and quenched by adding 0.9 mL of 50 mM NaOH into 0.6 mL reaction mixture immediately.

β-Glucosidase (BGL) activity was measured using *p*-nitrophenyl-β-D-glucopyranoside (*p*NPG) as the substrate. Similar to the exoglucanase activity assay, the enzyme reaction mixture containing 1 mL of culture supernatant, 0.5 mL of 50 mM citrate butter (pH 4.8), and 0.5 mL of 5 mM *p*NPG in 50 mM citrate buffer (pH 4.8) was incubated at 50 °C for 10 min and quenched by adding 0.9 mL of 50 mM NaOH into 0.6 mL of reaction mixture as well. One unit of exoglucanase and β-glucosidase activities is defined as the amount of enzyme that produced 1 μmol of nitrophenol in 1 min under the assay conditions described above. All the activities of three enzymes were normalized to the initial mycelial cell mass inoculated U/g mycelia (Wu et al. 2013b, c). All the activity assays were performed in triplicate. The values presented are the mean value of triplicate, and errors bars represent one standard deviation.

Zymography assay of secretome of four endophytes on Avicel

Zymography assays were conducted as described previously with modifications (Gladden et al. 2011). The culture broth of each endophyte on Avicel was centrifuged at 10000 rpm at 4 °C for 30 min. The supernatants (200 mL) of the cultures grown for 3 days were concentrated by filtering through Varian spin columns (Varian, Palo Alto, CA, USA) with MWCO of 3 kDa. The concentrated culture broth (20 μL from 2 mL of concentrate) was loaded onto 4–12% Novex polyacrylamide gel for SDS-PAGE. Once PAGE finished, the gel was washed twice with distilled (DI) water for 5 min and washed again four times for 15 min each in 25% (*v*/*v*) isopropanol. Then, the gel was immersed into 50 mM citrate buffer (pH 4.8) four times for 10 min each, followed by incubating into 50 mM citrate buffer (pH 4.8) containing 0.2% CM-cellulose at room temperature for 16 h and another 2 h at 50 °C. Then, the gel was washed with fresh DI water twice. The gel was stained with 0.1% Congo red in 50 mM citrate buffer, pH 7.0, for 10 min and fixed in the 1.0 M NaCl buffer.

DNA purification for sequencing

The mycelia from cultures grown for 5 days was collected as described above and immediately frozen in liquid nitrogen. The biomass was then grinded in liquid nitrogen using mortar and pestle until the mycelia became a fine powder. The powder was transferred into pre-cool 2-mL vials (0.5 g/vial) and suspended in plant DNAzol reagent (ThermoFisher, CA, USA), and genomic DNA was extracted following the manufacturer's protocol. DNA pellets were solubilized in TE buffer. The concentration of the genomic DNA was measured by Nanodrop (ThermoFisher, CA, USA). The integrity of the genomic DNA was determined using an agarose gel.

Genome sequencing and annotation

The Illumina platform was used for sequencing both transcriptomes and genomes of all four species (see Table 1 for statistics). The genomes were sequenced using a combination of Illumina fragment 300 bp insert size and 4 kbp long mate pair (LMP) libraries. The fragment libraries were produced from genomic DNA sheared to 270 bp using the covaris E210 (Covaris, Woburn, MA, USA) and size selected using SPRI beads (Beckman Coulter, Brea, CA, USA). The fragments were treated with end repair, A-tailing, and ligation of Illumina adapters (Eurofins MWG Operon, Louisville, KY, USA) using NEB Genome Sample Prep Kit (NEB, Lpswich, MA, USA). For the LMP libraries, genomic DNA was sheared using HydroShear (Genomic Solutions, Ann Arbor, MI, USA) and gel size selected for 4 kb. The size selected DNA was ligated to adaptors containing Loxp and the Illumina platform specific primer sequence, circularized via recombination by a *Cre* (NEB, Lpswich, MA, USA) excision reaction, and then digested with a cocktail of four-base cutter restriction enzymes, *Nla*III, *Mse*I, and *Hyp*CH4IV (NEB, Lpswich, MA, USA), followed by self-ligation. The paired end library was then amplified via inverted PCR using Illumina specific primer set. The size of the amplified paired end library was selected by running it on a 1.8% agarose gel followed by gel purification of the desired fragment (300–600 bp).

All prepared libraries were quantified using KAPA Biosystem's next-generation sequencing library qPCR kit (KAPA Biosystems, Wilmington, MA, USA) and run on a Roche LightCycler 480 real-time PCR instrument (Roche, Pleasanton, CA, USA). The quantified libraries were then prepared for sequencing on the Illumina HiSeq sequencing platform utilizing a TruSeq paired-end cluster kit, v3 or v4, and Illumina's cBot instrument to generate a clustered flowcell for sequencing. Sequencing of the flow cells was performed on the Illumina GAIIx or HiSeq2000 sequencers using a TruSeq SBS sequencing kit, v3 or v4, following a 2 × 100 or 2 × 150 bp indexed run recipes. For each genome, all genomic reads were quality control (QC) filtered for artifact/process contamination and subsequently assembled with AllPathsLG (Gnerre et al. 2011). In addition, the genome of EC38 strain was further improved by closing gaps with Pacific Biosciences reads using PBJelly (English et al. 2012).

The genomes were annotated using the JGI Annotation Pipeline and made available via the JGI fungal portal MycoCosm (Grigoriev et al. 2014). The genomic data are available at GenBank with the accessions MDGY00000000 (CI4A), MDCK00000000 (EC38), MDCL00000000 (CO27), and MDGZ00000000 (EC12), respectively. The strains *Hypoxylon* sp. CI4A, EC38, CO27, and *D. eschscholzii* EC12 were deposited in the USDA ARS culture collection (NRRL) with the accessions NRRL50501, NRRL50503, NRRL50500, and NRRL66509, respectively.

Identification of CAZy and terpene synthases genes in the endophyte genomes

CAZy gene annotations were assigned in each of the genomes of *Hypoxylon* sp. CI4A, EC38, CO27, and *D. eschscholzii* EC12 as described under "user annotations" in the JGI Mycocosom database with the methods used for the updates of the carbohydrate-active enzyme database (www.cazy.org; Lombard et al. 2014). Mevalonate (MEV) pathway enzymes and putative terpene synthases were identified in the endophyte genomes by searching annotated genes for Pfam functional domains found in these enzymes. As summary of the CAZy, annotations are summarized in the Supplementary Tables S1–S6.

Results

Endophyte genome statistics

The genomes of the four endophytic fungi *Hypoxylon* sp. CI4A, CO27, EC38, and *D. eschscholzii* EC12 were sequenced, assembled, and annotated at the Joint Genome Institute (JGI, Walnut Creek, CA, USA). The genome sizes of the four endophytes range between 37.5 and 47.3 Mb; they contain approximately 11,000–12,000 genes, and their average gene size ranges between 1700 and 1900 bp. The gene density of each genome is between 250 and 320 genes/Mbp (Table 1). This density is fairly high relative to more well-known fungi, such as *N. crassa*, which has a similar-sized genome (41.04 Mb) but approximately 15% fewer genes than these endophytes (Borkovich et al. 2004; Galagan et al. 2003). An examination of functionally related genes revealed that that the two most abundant gene categories in these fungi are associated with carbohydrate and lipid metabolism (Fig. 1). Also, compared to *N. crassa*, all four genomes had a relatively large number of genes involved in glycan biosynthesis and metabolism, polyketide biosynthesis, secondary metabolite biosynthesis, and xenobiotic biodegradation, indicating that that may be a harbor reservoir of biosynthetic pathways with applications relevant to biorefineries.

Genome analysis of endophytic CAZy and terpene biosynthetic genes

A deeper analysis of the endophyte genomes revealed that they are rich in carbohydrate-active enzymes (CAZys), which include degradation enzymes for carbohydrate and lignin (glycoside hydrolases (GH), polysaccharide lyases (PL), carbohydrate esterases (CE), auxiliary activities (AA)), carbohydrate-binding proteins and modules (CBM), as well as glycosyltransferases (GT), which assemble glycans and glycoconjugates. In agreement with earlier observations, the

Table 1 Summary of genome statistics

	Daldinia EC12	*Hypoxylon* CI4A	*Hypoxylon* CO27	*Hypoxylon* EC38
Genome size (Mb)	37.5	37.7	46.5	47.3
Contigs	641	1044	505	1168
Number of genes	11,173	11,712	12,256	12,261
Gene length (avg. nt)	1811	1761	1800	1757
Transcript length (avg. nt)	1638	1600	1590	1576
Protein length (avg. aa)	467	459	459	465
Exon length (avg. nt)	568	564	548	552
Intron length (avg. nt)	94	89	112	100
Exon frequency (no./gene)	2.89	2.84	2.9	2.86
Gene density (no./Mbp)	298	311	263	259

number of GT genes is well conserved (average 92, max 94, min 88) among the five fungi we have examined (Supplementary Table S1). On the other hand, the number of degradation enzymes (GH, PL, CE, and AA categories) reflects lifestyle and preferred nutrient source and varies more extensively, ranging from 271 for *N. crassa* to 407 for EC38, shown in Fig. 2a. Indeed, the four endophytes are closer to each other (average 391; max 407, min 376) and have 38 to 50% more degradation genes than *N. crassa* (Supplementary Table S1). Only for the CBM category is *N. crassa* richer than the four endophytes, the latter having fewer cellulose-binding modules of family CBM1 and α-glucan-binding modules of family CBM24 (Supplementary Table S1).

The five most abundant GH family enzyme categories of the endophytes are GH43, GH3, GH16, GH18, and GH5, as shown in Fig. 2e. The GH43 family enzymes are primarily responsible for the hydrolysis of xylan/arabinoxylan and include β-xylosidase, α-L-arabinosidase, α-galactosidase, and α-arabinofuranosidase among others (Newis et al. 2016). The GH16 enzymes cleave a large variety of β-linked polysaccharides but not cellulose, including β-galactanase, (1,4)-β-galactanase/κ-carrageenase, "non-specific" (1,3/1,4)-β-D-glucan endohydrolase, and (1,3/1,4)-β-D-glucan endohydrolase, which primarily cleave the galactans or β-D-glucans containing 1,3 or 1,4-linkage (Kawai et al. 2006). The GH5 family enzymes comprises 21 different enzyme activities, including endo-β-1,4-glucanase and exo-β-1,4-glucanase, endo-β-1,4-xylanase, β-mannanase and mannosidase, etc. that act on β-linked oligosaccharides and polysaccharides (Couturier et al. 2013; Zheng and Ding 2013; Aspeborg et al. 2012). The GH3 family encompasses 11 different enzyme activities but are primarily represented by β-glucosidase and β-xylosidase enzymes (Dodd et al. 2010; Lima et al. 2013), which are responsible

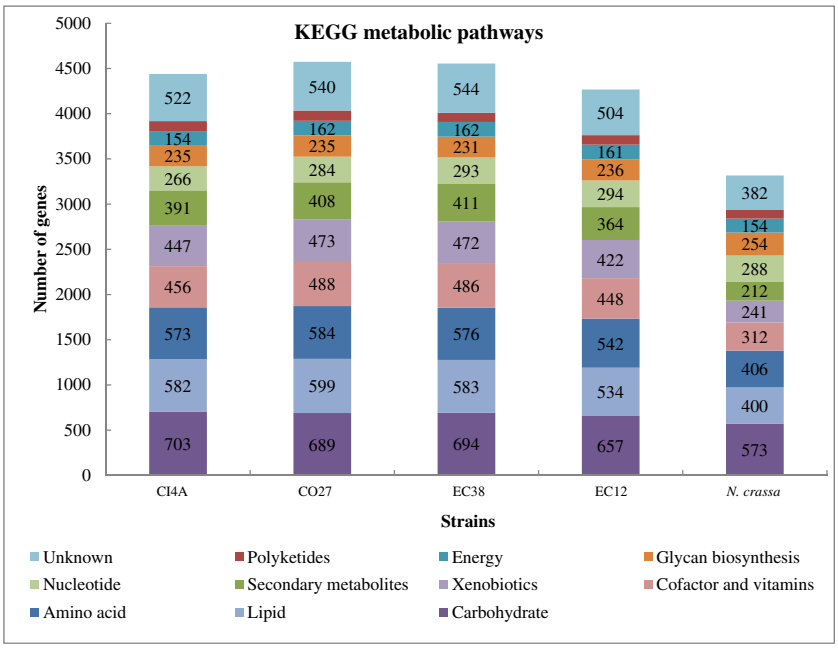
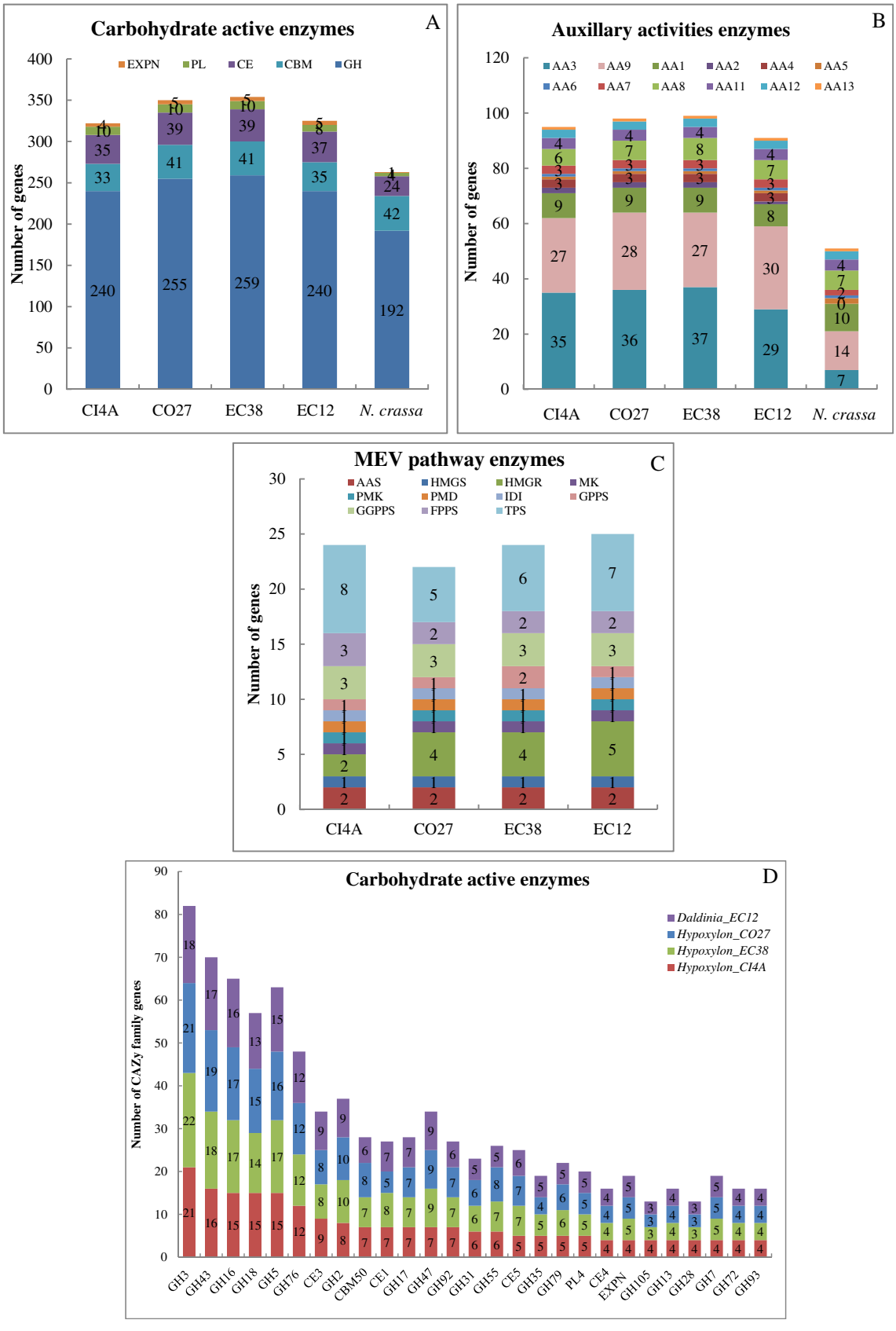


Fig. 1 KEGG metabolic pathways of four endophyte and *N. crassa* genomes


A
Carbohydrate active enzymes
EXPN PL CE CBM GH
Number of genes
CI4A CO27 EC38 EC12 N. crassa
B
Auxillary activities enzymes
AA3 AA9 AA1 AA2 AA4 AA5
AA6 AA7 AA8 AA11 AA12 AA13
Number of genes
CI4A CO27 EC38 EC12 N. crassa
C
MEV pathway enzymes
AAS HMGS HMGR MK
PMK PMD IDI GPPS
GGPPS FPPS TPS
Number of genes
CI4A CO27 EC38 EC12
D
Carbohydrate active enzymes
Daldinia_EC12
Hypoxylon_CO27
Hypoxylon_EC38
Hypoxylon_CI4A
Number of CAZy family genes
GH3 GH43 GH16 GH18 GH5 GH76 CE3 GH2 CBM50 CE1 GH17 GH47 GH92 GH31 GH55 CE5 GH35 GH79 PL4 CE4 EXPN GH105 GH13 GH28 GH7 GH72 GH93

◀ **Fig. 2** Classification of endophyte CAZys and TPSs. *AA* auxiliary activity, *GH* glycosyl hydrolases, *PL* polysaccharide lyases, *CE* carbohydrate esterase, *CBM* carbohydrate-binding module, *EXPN* expansin-like protein, *AAS* acetyl-CoA acetyltransferase, *HMGS* HMG-CoA synthase, *HMGR* HMG-CoA reductase, *MK* mevalonate kinase, *PMK* phosphomevalonate kinase, *PMD* mevalonate diphosphate decarboxylase, *IDI* isoprenyl diphosphate isomerase, *GPPS* geranyl pyrophosphate synthase, *FPPS* farnesyl pyrophosphate synthase, *GGPPS* geranylgeranyl pyrophosphate synthase, *TPSs* terpene synthases, *MEV* mevalonate

for the cleavage of oligosaccharides into sugar monomers. Family GH18 groups together enzymes that cleave β-linked *N*-acetylglucosamine residues in chitin, chitooligosaccharides, peptidoglycan, and *N*-glycans (www.cazy.org/GH18.html). Compared to these five GH families, which perform cleavage of polysaccharides or oligosaccharides to release monomer sugars, the enzymes of the AA9 family are described as copper-dependent lytic polysaccharide monooxygenases that cleave cellulose by an oxidative mechanism and enhance cellulase activity in the degradation of cellulosic biomass (Harris et al. 2010; Frandsen et al. 2016). Each endophytic genome contains more than 26 AA9 genes, indicating that they have significant potential to degrade lignocellulosic biomass. In addition, each endophyte genome contains four to five copies of GH7 cellulases and two copies of GH6 cellulases, which are responsible for the cleavage of cellulose chains either in a random (endoglucanase) or a processive (cellobiohydrolase) mode from either the reducing or the non-reducing ends, releasing primarily cellobiose (Sandgren et al. 2013; Mba Medie et al. 2012).

In addition to GH CAZys, the four endophytes are also rich in ligninolytic enzymes, as shown in Fig. 2c, which indicates the potential to degrade lignin in addition to the plant cell wall polysaccharides. Each endophyte contains 9–10 copies of genes encoding laccase (AA1), cellobiose dehydrogenase (AA3_1), chloro, manganese, and lignin peroxidase (AA2). Besides these major ligninolytic enzymes, the four endophytes also harbor more than 40 copies of genes encoding redox auxiliary enzymes for lignin degradation, which is much higher than *N. crassa* (Fig. 2c).

Another interesting characteristic of these endophytes is that they produce a wide range of VOCs as secondary metabolites (Gladden et al. 2013). One of major components of these VOCs are terpenes, which have been considered potential biofuels due to their high energy density, near-zero oxygen content, and 2:1 H/C ratio (Strobel 2014). As expected, genome analysis revealed that the four endophytes harbor a complete mevalonate (MEV) terpene biosynthetic pathway, shown in Fig. 2d. More than two copies of genes encoding acetyl-CoA acetyltransferase (AAS), HMG-CoA reductase (HMGR), farnesyl pyrophosphate synthase (FPPS), and geranylgeranyl pyrophosphate synthase (GGPPS) were identified in the genomes. The endophytes also have an abundance of terpene synthase (TPS) genes. In the four genomes, a total 26 putative TPS genes were identified, which helps explain their production of a wide spectrum of VOCs (Strobel 2014). Interestingly, although three copies of putative GGPPS genes were identified in each endophytic genome, no diterpenes were detected under the cultivation conditions described in this study.

The richness of both GH and ligninolytic CAZys (shown in Supplementary Tables S1–S6) is consistent with the species' habit and potential role as a lignocellulose degrader. In addition, the abundance of terpene synthases and other potential VOC biosynthetic pathways indicate that these organisms have the potential to directly convert lignocellulosic biomass into a myriad of products with industrial applications, such as fragrances, food additives, pharmaceuticals, and biofuels. To better understand this potential, we examined both the lignocellulolytic enzymes and VOCs produced by these organisms when grown on a variety of lignocellulosic feedstocks, as described in the following sections.

Endophyte VOC production when grown on lignocellulosic feedstocks

In previous studies, Booth et al. (2011) reported that the endophyte *Hypoxylon* sp. CI4A produced a wide spectrum of volatile organic compounds on potato dextrose medium, mostly monoterpenes and sesquiterpenes. In this study, we further investigated the VOC profile of *Hypoxylon* sp. CI4A and the three other endophytes on four different lignocellulosic feedstocks (Avicel, corn stover, switchgrass, and eucalyptus mill powder). The mycelia of the endophytes were grown in minimal medium containing each feedstock as the sole carbon source. The VOCs were extracted by SPME and analyzed by GC-MS. Results confirm that terpenes, including monoterpenes and sesquiterpenes, were one of the most abundant components of VOCs, but the percentage varied between strains and feedstocks (Fig. 3a–d). All four strains produced a higher abundance of terpenes on Avicel (cellulose) than other three feedstocks, particularly *Hypoxylon* sp. CI4A, which produced a VOC profile with 63% terpenes (Fig. 3a). The percentage of terpenes in the VOC profiles decreased when the endophytes were grown on other three feedstocks, except for strain EC12, which produced a slightly increased abundance of terpenes in its VOC profile when grown on SG, CS, and eucalyptus mill powder compared to Avicel.

Other than terpenes, the endophytes produced cyclic alcohols (terpene derivatives), phenolic alcohols, ketones, alkanes/alkenes, and aromatic compounds when grown on the four feedstocks (Fig. 3). The alkane/alkenes were detected in the cultures of all four strains on all four cellulosic feedstocks, ranging from 1% (EC12 on SG) to 13% (EC12 on Avicel). Relative to alkanes and alkenes, the strains produced more

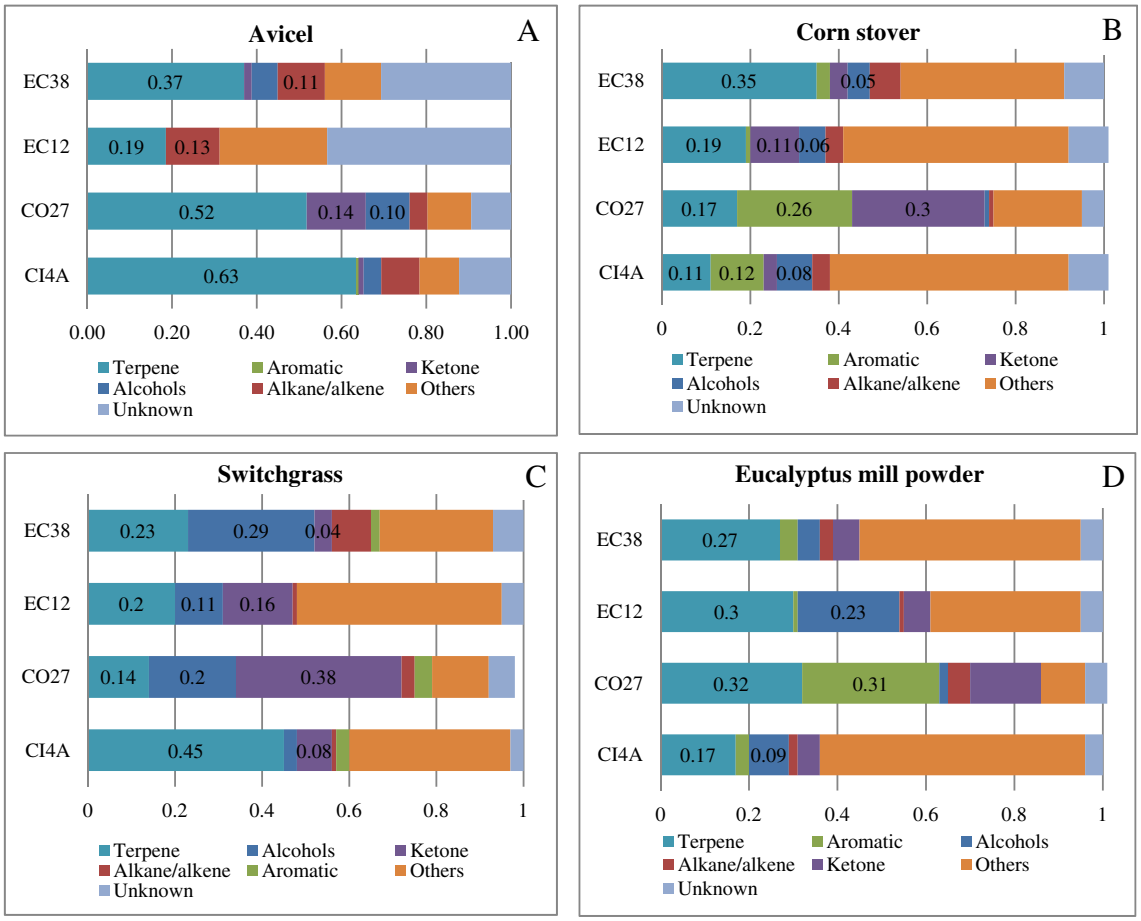


Fig. 3 VOC profiles of four endophytes cultivated on Avicel, corn stover (CS), switchgrass (SG), and eucalyptus mill powder (eucalyptus)

ketones and aromatic compounds on these substrates. Strain CO27 yielded the highest abundance of ketones (38% of the VOCs) on SG as well as the highest abundance of aromatic compounds (31% of VOCs) on eucalyptus mill powder. Various cyclic or phenolic alcohols were detected in the cultures of the endophytes as well. The GC-MS results showed that the SG induced production of higher levels of alcohols in the VOCs than on the other three feedstocks. Specifically, the strain EC38 and CO27 yielded about 30 and 20% of cyclic or phenolic alcohols in the VOCs when grown on SG. Also, strain EC12 produced more than 20% alcohols in the VOC fraction. In addition to these classes of compounds, there was still a large number peaks that could not be identified by MS spectrum, indicating an even greater diversity of VOC that discussed above.

Endophyte endoglucanase, cellobiohydrolase, and β-glucosidase activities when grown on lignocellulosic feedstocks

To induce production of CAZy enzymes, the four endophytes were cultured in minimal salts medium containing Avicel, CS, SG, and eucalyptus mill powder, respectively, as the sole carbon source. As a reference, *N. crassa* was also cultured on the same carbon sources. Endoglucanase, exoglucanase, and β-glucosidase activities were measured individually using the model substrates CM-cellulose, *p*NPL, and *p*NPG (Soares et al. 2013; Xiao et al. 2005). Results reveal that all of the endophytes produced significant endoglucanase activities on four different feedstocks, as shown in Fig. 4a–d. However, each strain had different endoglucanase expression patterns on the four feedstocks. All strains produced higher endoglucanase activities on Avicel than the other three feedstocks except the strain EC38, which produced similar levels of endoglucanase activity on all four feedstocks, ranging from 20 U/g mycelia (SG) to 46 U/g mycelia (CS). Among four strains, the EC12 produced the highest endoglucanase activity on Avicel, up to 450 U/g mycelia, followed by CO27 (129 U/g mycelia), CI4A (47 U/g mycelia), and EC38 (32 U/g mycelia). The model substrate endoglucanase activities for each strain grown on Avicel are also consistent with enzymatic activities detected by zymography, shown in Fig. 5. Eucalyptus mill powder was good at inducing cellulolytic enzyme activities; strains EC38 and CI4A produced similar levels of endoglucanase activity, more than CO27 and CI4A. The maximum endoglucanase activities of CI4A and EC38 grown on

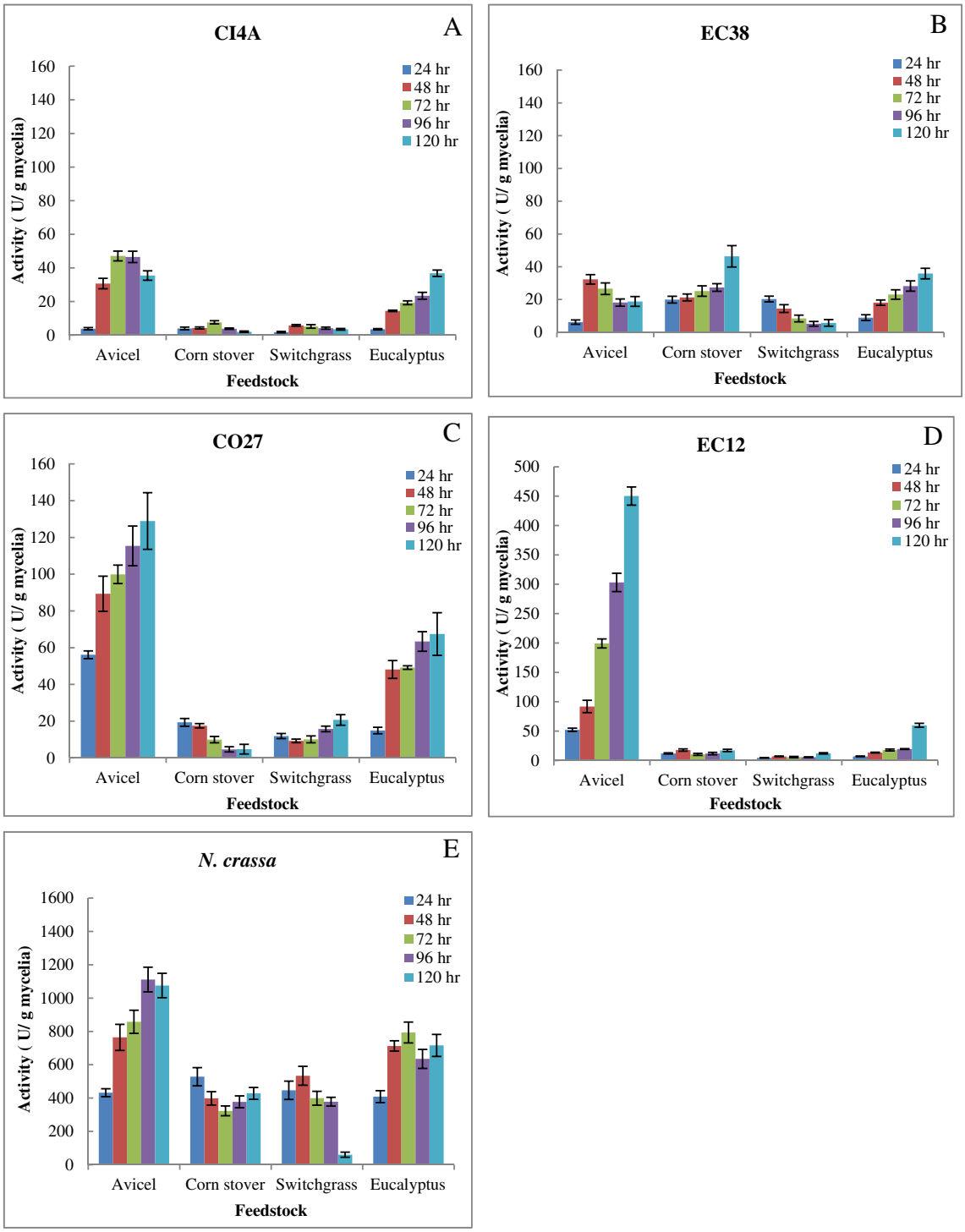


Fig. 4 **a**–**e** Endoglucanase activities of endophytes and *N. crassa* when cultivated on different feedstocks

eucalyptus mill powder were around 36 U/g mycelia, while CO27 and EC12 produced 67 (CO27) and 60 U/g mycelia, respectively. All strains produced less than 20 U/g mycelia endoglucanase activity when grown on corn stover and switchgrass, except EC38, which produced the maximum of 46 U/g mycelia endoglucanase activity on corn stover, indicating that they are the two least inductive feedstocks for the endophytes.

For comparison, the endoglucanase activity levels produced by *N. crassa* when grown on the four feedstocks are shown in Fig. 4e. The strain produced the highest endoglucanase activity on Avicel, up to 1075 U/g mycelia,

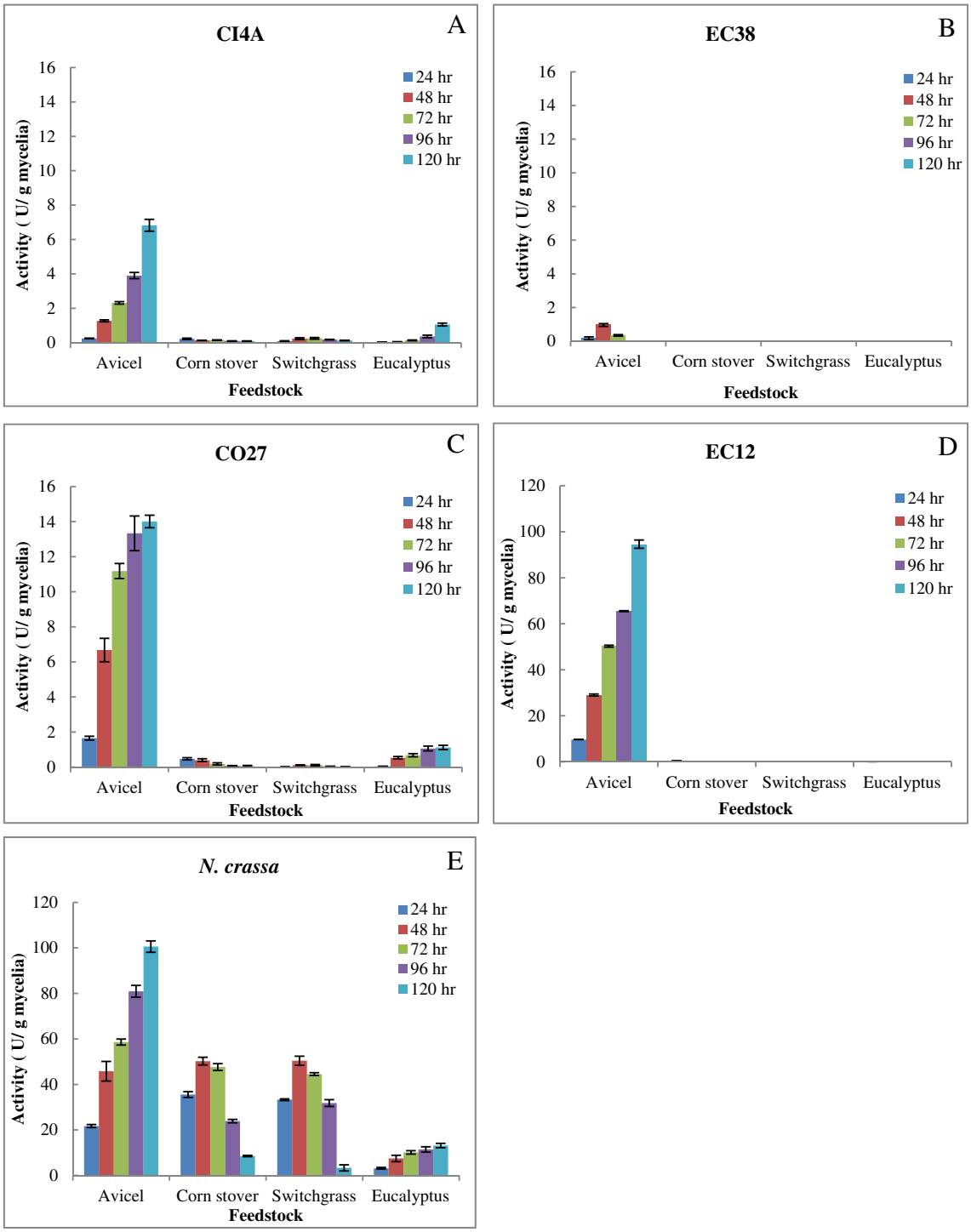


Fig. 5 **a**–**e** Exoglucanase activities of endophytes and *N. crassa* when cultivated on different feedstocks

followed by eucalyptus mill powder (793 U/g mycelia). The strain yielded lower levels of endoglucanase activities on corn stover and switchgrass, which were both around 530 U/g mycelia. This level of activity was much higher on all four feedstocks relative to the four endophytes.

When the four endophytes were grown on CS, SG, and eucalyptus mill powder, only *Hypoxylon* sp. CI4A and CO27 produced detectable CBH activities but the maximal activities were below 1.1 U/g mycelia. However, all four endophytes produced detectable cellobiohydrolase (CBH) activities on Avicel. For each of the endophytes, the CBH activities continually increased during the cultivation, except for strain EC38, which achieved a maximum CBH activity (0.96 U/g mycelia) on day 2. For the other three endophytes, the highest

CBH activity achieved after 5 days was 94.62 U/g mycelia for EC12, followed by CO27 (14.01 U/g mycelia), and CI4A (6.82 U/g mycelia). Compared to the endophytes, *N. crassa* produced detectable CBH activities on all four substrates. On Avicel and eucalyptus mill powder, the maximal CBH activities were achieved after 5 days, 100.59 and 13.14 U/g mycelia, respectively. When grown on CS and SG, the highest values of CBH activity were reached after 2 days.

All four endophytes produced detectable β-glucosidase activities on four of the different feedstocks as shown in Fig. 3a–d. All the endophytes produced higher BGL activity on Avicel than on the other three feedstocks, except EC38, which produced its lowest BGL activity on Avicel compared to the other three feedstocks. EC12 produced the highest BGL activity on Avicel among the four endophytes, up to 753 U/g mycelia. The highest BGL activities on Avicel were achieved after 5 days for endophytes CI4A, CO27, and EC12. For this class of enzymes, the cultivation time to achieve the maximum BGL activity was different for each strain on each different biomass. For instance, the CI4A and EC38 achieved the highest BGL activities after 2 or 3 days on all three biomass feedstocks; however, CO27 yielded the highest BGL activity on corn stover within 1 day but achieved the highest BGL activity on eucalyptus mill powder after 5 days. CI4A and EC12 produced 5–20 times higher BGL activities on Avicel than that on other three biomass feedstocks, while strain EC38 produced the highest BGL activity on corn stover, up to 33 U/g mycelia. CO27 yielded similar amount of BGL activity on four substrates, as shown in Fig. 6c.

All four endophytes produced lower overall BGL activities than that of *N. crassa*. The maximum BGL activities of *N. crassa* achieved was 2486 U/g mycelia, which was 3 (*N. crassa*/EC12) to 1000 times (*N. crassa*/EC38) higher than that of the endophytes. However, *N. crassa* only produced approximately 10 times more BGL activity than endophytes when grown on non-Avicel feedstocks. CI4A produced significantly higher BGL activity on switchgrass and eucalyptus mill powder than on corn stover. CO27, EC12, and *N. crassa* produced similar amounts of BGLs on all three feedstocks. Interestingly, unlike the other fungi, EC38 yielded the highest BGL activities on corn stover rather than Avicel, which indicates that it may have a different inducer for its cellulolytic enzymes.

Zymography of the endophyte endoglucanases

To investigate the complement of endoglucanase enzymes produced by each endophyte, a zymography assay was performed using CMC as the substrate. Only the secretomes of endophytes grown on Avicel were investigated due to the fact that the other three cultures caused a high level of background on the zymogram. The assay showed that the major active endoglucanases secreted by the four endophytes have molecular weights that range between 38 and 49 kDa, as shown in Fig. 7, except for strain EC12, which produces additional enzymes between 62 and 98 kDa. The band intensity on the zymogram, which translates to enzyme activity (more intense bands have more activity), was consistent with the endoglucanase activity assays described above. In order to determine which GH genes in the genome of each endophyte encode the endoglucanases observed by zymography, an attempt was made to compare predicted and measured protein molecular weights, as listed in the Table 2 and Supplementary Table S7.

Discussion

Lignocellulosic biomass is an attractive renewable carbon source for bioenergy applications because of its large-scale availability, low cost, and environmentally benign production. The primary barrier for bioconversion of lignocellulosic biomass to liquid fuels is the absence of the low-cost technologies to overcome the structural recalcitrance of these feedstocks to enzymatic degradation. Consolidated bioprocessing, a process that consolidates lignocellulolytic enzyme production, lignocellulosic biomass deconstruction, and sugar fermentation into a single step, is a promising approach to overcome this barrier as it has the potential to significantly reduce costs in a biorefinery (Olson et al. 2012).

Although much of the work on CBP has been centered on bacteria, several fungi have been reported to possess CBP characteristics, by converting cellulose to ethanol. Unfortunately, these fungi also produced a significant amount of by-products, indicating that further development and exploration needs to be done to identify and develop fungal CBP systems (Anasontzis and Christakopoulos 2014; Abedinifar et al. 2009). The endophytes investigated in this study have been reported to produce various VOCs with potential use as mycodiesel, based on high energy density and other chemical properties important for use as a fuel. In this study, four endophytes were investigated for CBP properties, including possessing biosynthetic pathways to produce fuel molecules and enzymes involved in lignocellulose degradation, with an emphasis on the latter. Genome analysis indicated that these endophytes are rich in terpene biosynthesis pathway genes, a class of compounds that represent many potential fuel molecules. We demonstrated that these endophytes could utilize lignocellulosic feedstocks to produce a wide spectrum of volatile organic compounds (mycodiesel), primarily terpenes, thus supporting their potential as CBP organisms. However, it must be noted that the VOCs examined in this study were extracted using SPME, a sensitive and rapid technique for the detection of compounds from gaseous, liquid, and solid samples. Although SPME has higher extraction efficiency, selectivity, and stability, when compared to other techniques, it is

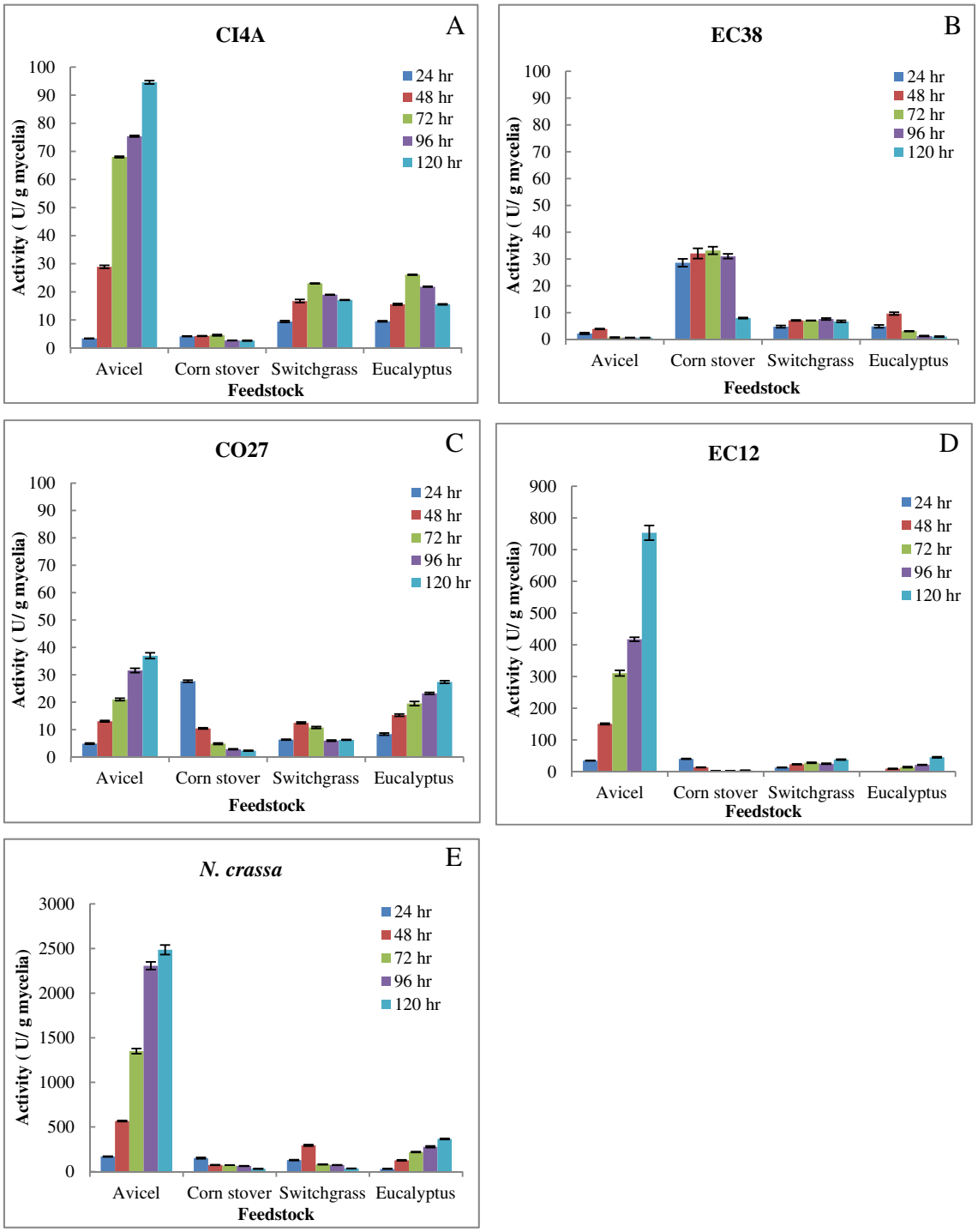


Fig. 6 **a**–**e** β-Glucosidase activities of endophytes on different feedstocks

difficult to quantitate VOC concentrations and correlate those numbers to titers within the culture. Therefore, the titers of terpene in this study were not determined. Most likely, the total terpene concentration would be in the mg/L range, according to other studies (Booth et al. 2011). These titers would need to undergo significant improvement for these organisms to reach their full potential as CBP organisms. Despite that limitation, these endophytes do produce a variety of interesting terpenes with potential biofuel applications. Our analysis showed that the terpene fraction of the VOC consisted primarily of monoterpenes and sesquiterpenes, such as α/β-pine, limonene, cineole, caryophyllene, humulen (Vl), and others. In fact, many of the terpene synthases identified in these endophyte genomes have recently been characterized and were found to produce a similar spectrum of monoterpenes and sesquiterpenes (Gladden et al. 2013; Wu et al.

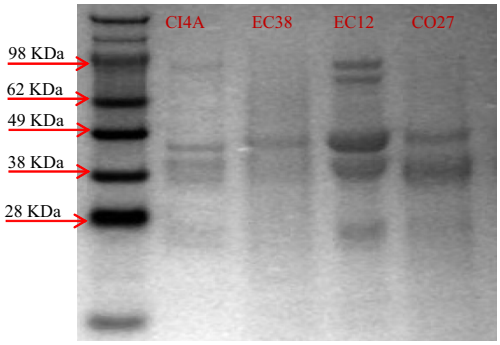


Fig. 7 **a** Zymography of endoglucanase activity in endophyte secretomes. Substrate: CM-cellulose

2016b). Other studies have reported that hydrogenated monoterpenes and sesquiterpenes, such as pinane and carophyllanes, have a moderate cetane number and only moderately high viscosity, suggesting that the hydrocarbons derived from terpenes are functionally similar to the compounds in petroleum distillate fuels and often share similar combustion properties (Harvey et al. 2010; Edwards et al. 2010). Blending of terpanes with synthetic branched paraffins to raise cetane and reduce viscosity could produce biosynthetic fuels that meet applicable jet and diesel specifications (Harvey et al. 2015). The mixture of terpenes the endophytes produced is similar/identical to the compounds in those reports, indicating that they could also be used in these types of applications.

In addition to production of mycodiesel, genome analysis and cellulase activity assays show that these four endophytes can express CAZy lignocellulose-degrading enzymes, supporting direct hydrolysis and conversion of lignocellulosic feedstocks into high energy density terpene compounds. Compared to *N. crassa*, a model fungus that has twice as many cellulase genes (365) as the industrial cellulase producer *Trichoderma reesei*, the four endophytes have 14–20% more CAZy genes (417–437 vs 265 genes) and 27–60% more ligninolytic genes (19–24 vs 15 genes) (Tian et al. 2009; Galagan et al. 2003). The higher number of CAZy and ligninolytic genes in the endophyte genomes suggest that the endophytes have a greater potential for lignocellulosic biomass breakdown than *N. crassa*. However, cellulase activity assays indicated that all of the endophytes have lower endoglucanase, exo-glucanase, and β-glucosidase activity than *N. crassa* when grown on four cellulolytic biomass substrates, indicating that more optimization is required to fully leverage that potential. Zymography assays of the four endophytic secretomes showed that the endophytes *Hypoxylon* sp. CO27 and *D. eschscholzii* EC12 have higher relative cellulase activity compared to the other two endophytes. The genome of each endophyte has more than 61 copies of genes encoding cellulases, as shown in the supplementary tables. GH families 5, 6, 7, 8, 9, 12, and 45 are known to include endoglucanases (Martinez et al. 2008; Miao et al. 2015). In each endophytic genome, we found 15–18 copies of GH5, 2–3 copies of GH6, 4–5 copies of GH7, and 1–2 copies of GH45. Additionally, strain CI4A contains two copies of GH12. The GH families 8 and 9 are unrepresented in the genomes. On cellulose zymograms, all the endophytes produced two intense bands between 38 and 49 kDa, which represent molecular weights and activities from GH 5, 6, and 7 family enzymes. Based on protein molecular weight estimates from the annotated genes, the bands between 62 and 98 kDa may be GH3 family cellulases, while the bands below 28 kDa may be cellulases from GH12, GH30, and GH45 enzyme families. In addition to endoglucanases, these endophytes harbor many β-glucosidases (EC 3.3.1.21). The β-glucosidases are distributed in CAZy GH families 1, 2, 3, 5, 9, 30, and 116. Each endophyte genome contains 1–2 copies of GH1, 8–10 copies of GH2, 18–22 copies of GH3, and 3–4 copies of GH30. GH9 and GH116 are not presented in the genomes. Overall, these organisms have a large complement of enzymes for breaking down lignocellulosic

Table 2 Comparison of measured and predicted endoglucanase protein molecular weights from zymography and annotated genes in the four endophyte genomes

Zymo MW[a]	Predicted CI4A endoglucanases (copy number)[b]	Predicted EC38 endoglucanases (copy number)[b]	Predicted EC12 endoglucanases (copy number)[b]	Predicted CO27 endoglucanases (copy number)[b]
<28 kDa	GH10 (1), GH12 (2), GH45 (1), GH51 (1)	GH5 (1), GH45 (1)	GH45 (1)	GH45 (1)
28 < MW < 38 kDa	GH5 (1), GH7 (1), GH10 (1)	GH5 (1), GH6 (1), GH7 (1), GH10 (1)	GH5 (1), GH10 (1)	GH5 (1), GH7 (1), GH10 (1)
38 < MW < 49 kDa	GH5 (9), GH6 (2), GH7 (3)	GH5 (9), GH6 (1), GH7 (3), GH10 (1)	GH5 (9), GH6 (2), GH7 (4), GH10 (3)	GH5 (9), GH6 (1), GH7 (3), GH10 (1)
49 < MW < 62 kDa	GH5 (3), GH51 (1)	GH5 (5), GH7 (1)	GH5 (3), GH6 (1), GH7 (1), GH51 (1)	GH5 (6), GH6 (1), GH7 (1), GH51 (1)
62 < MW < 98 kDa	GH5 (3), GH51 (1), GH74 (1)	GH5 (2)	GH5 (2), GH51 (1), GH74 (1)	GH5 (1), GH51 (1), GH74 (1)
>98 kDa	None	None	None	None

[a] Molecular weights calculated by comparison to the protein standard

[b] For each annotated GH gene, protein molecular weights were predicted based on amino acid sequence

feedstocks, and there is much potential to leverage this native capacity to develop a robust CBP platform.

In summary, four endophytes were selected to investigate their CBP potential for converting lignocellulosic biomass to biofuels, namely mycodiesel or aviation fuel. Genomic analysis showed that all four endophytes are rich in carbohydrate-active enzymes, fungal oxidative lignin enzymes, and terpene synthases. Cellulase activity assays confirmed that the endophytes secrete many cellulases, providing the ability to directly break down the lignocellulosic feedstocks. GC-MS analysis identified a wide spectrum of mycodiesel VOCs, primarily monoterpenes or sesquiterpenes, which contain high energy density and near-zero oxygen content, and are therefore good potential drop-in fuel compounds. The ability to both deconstruct plant cell walls and to produce mycodiesel show that *Hypoxylon* and *Daldinia* are promising new genera to explore to realize consolidated bioprocessing.

Acknowledgements This work was supported by the Laboratory Directed Research and Development program at Sandia National Laboratories, which is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the US Department of Energy's National Nuclear Security Agency under contract DE-AC04-94AL85000. The authors want to thank Dr. Gary A. Strobel for providing the endophytes. The work by the US Department of Energy Joint Genome Institute, a DOE Office of Science User Facility, is supported by the Office of Science of the US Department of Energy under contract no. DE-AC02-05CH11231.

Author contributions W.W.H. conceived and designed the study, performed the experiments and data analysis, and wrote and revised the manuscript. R.W.D. revised the manuscript. J.M.G. supervised the study and revised the manuscript. M.B.T. maintained the endophyte seed stock. H.H., S.M., and M. C. sequenced genomes and transcriptomes. K.L. and H.S. assembled genomes. E.L. assembled transcriptomes. A.K. annotated genomes. B.H. annotated the CAZymes. K.B. manages the genome project. I.V.G. coordinated the genome projects and revised the manuscript. All authors read and approved the final manuscript.

Compliance with ethical standards

Ethical approval This article does not contain any studies with animals performed by any of the authors.

Competing interests The authors declare that they have no competing interests.

References

Abedinifar S, Karimi K, Khanahmadi M, Taherzadeh MJ (2009) Ethanol production by *Mucor indicus* and *Rhizopus oryzae* from rice straw by separate hydrolysis and fermentation. Biomass Bioenergy 33(5): 828–833

Akinosho H, Yee K, Close D, Ragauskas A (2014) The emergence of *Clostridium thermocellum* as a high utility candidate for consolidated bioprocessing applications. Front Chem 2:66. doi:10.3389/fchem.2014.00066

Anasontzis GE, Christakopoulos P (2014) Challenges in ethanol production with *Fusarium oxysporum* through consolidated bioprocessing. Bioengineered 5(6):393–395. doi:10.4161/bioe.36328

Aspeborg H, Coutinho PM, Wang Y, Brumer H III, Henrissat B (2012) Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). BMC Evol Biol 12:186. doi:10.1186/1471-2148-12-186

Balat M (2011) Production of bioethanol from lignocellulosic materials via the biochemical pathway: a review. Energy Convers Manag 52(2):858–875. doi:10.1016/j.enconman.2010.08.013

Booth E, Strobel G, Knighton B, Sears J, Geary B, Avci R (2011) A rapid column technique for trapping and collecting of volatile fungal hydrocarbons and hydrocarbon derivatives. Biotechnol Lett 33(10): 1963–1972. doi:10.1007/s10529-011-0660-2

Borkovich KA, Alex LA, Yarden O, Freitag M, Turner GE, Read ND, Seiler S, Bell-Pedersen D, Paietta J, Plesofsky N, Plamann M, Goodrich-Tanrikulu M, Schulte U, Mannhaupt G, Nargang FE, Radford A, Selitrennikoff C, Galagan JE, Dunlap JC, Loros JJ, Catcheside D, Inoue H, Aramayo R, Polymenis M, Selker EU, Sachs MS, Marzluf GA, Paulsen I, Davis R, Ebbole DJ, Zelter A, Kalkman ER, O'Rourke R, Bowring F, Yeadon J, Ishii C, Suzuki K, Sakai W, Pratt R (2004) Lessons from the genome sequence of *Neurospora crassa*: tracing the path from genomic blueprint to multicellular organism. Microbiol Mol Biol Rev 68(1):1–108. doi:10.1128/mmbr.68.1.1-108.2004

Couturier M, Roussel A, Rosengren A, Leone P, Stalbrand H, Berrin JG (2013) Structural and biochemical analyses of glycoside hydrolase families 5 and 26 β-(1,4)-mannanases from *Podospora anserina* reveal differences upon manno-oligosaccharide catalysis. J Biol Chem 288(20):14624–14635. doi:10.1074/jbc.M113.459438

Dodd D, Kiyonari S, Mackie RI, Cann IK (2010) Functional diversity of four glycoside hydrolase family 3 enzymes from the rumen bacterium *Prevotella bryantii* B14. J Bacteriol 192(9):2335–2345. doi:10.1128/jb.01654-09

Edwards T, Moses T, Dryer F (2010) Evaluation of combustion performance of alternative aviation fuels. 46th AIAA/ASME/SAE/ASEE Joint Propulsion Conference Exhibit. AIAA:2010–7155

English AC, Richards S, Han Y, Wang M, Vee V, Qu J, Qin X, Muzny DM, Reid JG, Worley KC, Gibbs RA (2012) Mind the gap: upgrading genomes with Pacific biosciences RS long-read sequencing technology. PLoS One 7(11):e47768

Fan Z, Wu W, Hildebrand A, Kasuga T, Zhang R, Xiong X (2012) A novel biochemical route for fuels and chemicals production from cellulosic biomass. PLoS One 7(2):e31693. doi:10.1371/journal.pone.0031693

Favaro L, Viktor MJ, Rose SH, Viljoen-Bloom M, van Zyl WH, Basaglia M, Cagnin L, Casella S (2015) Consolidated bioprocessing of starchy substrates into ethanol by industrial *Saccharomyces cerevisiae* strains secreting fungal amylases. Biotechnol Bioeng 112(9):1751–1760. doi:10.1002/bit.25591

Frandsen KE, Simmons TJ, Dupree P, Poulsen JC, Hemsworth GR, Ciano L, Johnston EM, Tovborg M, Johansen KS, von Freiesleben P, Marmuse L, Fort S, Cottaz S, Driguez H, Henrissat B, Lenfant N, Tuna F, Baldansuren A, Davies GJ, Lo Leggio L, Walton PH (2016) The molecular basis of polysaccharide cleavage by lytic polysaccharide monooxygenases. Nat Chem Biol 12(4):298–303

Galagan JE, Calvo SE, Borkovich KA, Selker EU, Read ND, Jaffe D, FitzHugh W, Ma LJ, Smirnov S, Purcell S, Rehman B, Elkins T, Engels R, Wang SG, Nielsen CB, Butler J, Endrizzi M, Qui DY, Ianakiev P, Pedersen DB, Nelson MA, Werner-Washburne M, Selitrennikoff CP, Kinsey JA, Braun EL, Zelter A, Schulte U, Kothe GO, Jedd G, Mewes W, Staben C, Marcotte E, Greenberg D, Roy A, Foley K, Naylor J, Stabge-Thomann N, Barrett R, Gnerre S, Kamal M, Kamvysselis M, Mauceli E, Bielke C, Rudd S, Frishman D, Krystofova S, Rasmussen C, Metzenberg RL, Perkins DD, Kroken S, Cogoni C, Macino G, Catcheside D, Li WX, Pratt

RJ, Osmani SA, DeSouza CPC, Glass L, Orbach MJ, Berglund JA, Voelker R, Yarden O, Plamann M, Seiler S, Dunlap J, Radford A, Aramayo R, Natvig DO, Alex LA, Mannhaupt G, Ebbole DJ, Freitag M, Paulsen I, Sachs MS, Lander ES, Nusbaum C, Birren B (2003) The genome sequence of the filamentous fungus *Neurospora crassa*. Nature 422(6934):859–868. doi:10.1038/nature01554

Gladden JM, Allgaier M, Miller CS, Hazen TC, VanderGheynst JS, Hugenholtz P, Simmons BA, Singer SW (2011) Glycoside hydrolase activities of thermophilic bacterial consortia adapted to switchgrass. Appl Environ Microbiol 77(16):5804–5812. doi:10.1128/aem.00032-11

Gladden JM, Taatjes AC, Gao C, Bryan GO, Powell AJ, Scheer AM, Turner K, Wu W, Yu ET (2013) Tailoring next-generation biofuels and their combustion in next-generation engines. Sandia Report: 2013–10094

Gnerre S, Maccallum I, Przybylski D, Ribeiro FJ, Burton JN, Walker BJ, Sharpe T, Hall G, Shea TP, Sykes S, Berlin AM, Aird D, Costello M, Daza R, Williams L, Nicol R, Gnirke A, Nusbaum C, Lander ES, Jaffe DB (2011) High-quality draft assemblies of mammalian genomes from massively parallel sequence data. Proc Natl Acad Sci U S A 108(4):1513–1518

Grigoriev IV, Nikitin R, Haridas S, Kuo A, Ohm R, Otillar R, Riley R, Salamov A, Zhao X, Korzeniewski F, Smirnova T, Nordberg H, Dubchak I, Shabalov I (2014) MycoCosm portal: gearing up for 1000 fungal genomes. Nucleic Acids Res 42:D699–D704

Guo K, Mou X, Huang J, Xiong N, Li H (2014) Trans-caryophyllene suppresses hypoxia-induced neuroinflammatory responses by inhibiting NF-κB activation in microglia. J Mol Neurosci 54(1): 41–48. doi:10.1007/s12031-014-0243-5

Haghighi Mood S, Hossein Golfeshan A, Tabatabaei M, Salehi Jouzani G, Najafi GH, Gholami M, Ardjmand M (2013) Lignocellulosic biomass to bioethanol, a comprehensive review with a focus on pretreatment. Ren Sust Energy Rev 27:77–93. doi:10.1016/j.rser.2013.06.033

Harris PV, Welner D, McFarland KC, Re E, Poulsen J-CN, Brown K, Salbo R, Ding H, Vlasenko E, Merino S, Xu F, Cherry J, Larsen S, Lo Leggio L (2010) Stimulation of lignocellulosic biomass hydrolysis by proteins of glycoside hydrolase family 61: structure and function of a large, enigmatic family. Biochemistry 49(15):3305–3316. doi:10.1021/bi100009p

Harvey BG, Wright ME, Quintana RL (2010) High-density renewable fuels based on the selective dimerization of pinenes. Energy Fuel 24(1):267–273

Harvey BG, Wright ME, Koontz TA (2015) High-density renewable diesel and jet fuels prepared from multicyclic sesquiterpanes and a 1-hexene-derived synthetic paraffinic kerosene. Energy Fuel 29(4): 2430–2436

Hasunuma T, Kondo A (2012) Consolidated bioprocessing and simultaneous saccharification and fermentation of lignocellulose to ethanol with thermotolerant yeast strains. Proc Biochem 47(9):1287–1294. doi:10.1016/j.procbio.2012.05.004

Higashide W, Li Y, Yang Y, Liao JC (2012) Metabolic engineering of *Clostridium cellulolyticum* for production of isobutanol from cellulose. Appl Environ Microbiol 77(8):2727–2733. doi:10.1128/aem.02335-12

Kawai R, Igarashi K, Yoshida M, Kitaoka M, Samejima M (2006) Hydrolysis of β-1,3/1,6-glucan by glycoside hydrolase family 16 endo-1,3(4)-β-glucanase from the basidiomycete *Phanerochaete chrysosporium*. Appl Microbiol Biotechnol 71(6):898–906. doi:10.1007/s00253-005-0214-4

Kovac J, Simunovic K, Wu Z, Klancnik A, Bucar F, Zhang Q, Mozina SS (2015) Antibiotic resistance modulation and modes of action of (−)-α-pinene in campylobacter jejuni. PLoS One 10(4):e0122871. doi:10.1371/journal.pone.0122871

Lan EI, Liao JC (2013) Microbial synthesis of n-butanol, isobutanol, and other higher alcohols from diverse resources. Bioresour Technol 135:339–349. doi:10.1016/j.biortech.2012.09.104

Lei W, Sharifzadeh M, Templer R, Murphy RJ (2013) Bioethanol production from various waste papers: economic feasibility and sensitivity analysis. Appl Energy 111:1172–1182. doi:10.1016/j.apenergy.2012.08.048

Lima MA, Oliveira-Neto M, Kadowaki MAS, Rosseto FR, Prates ET, Squina FM, Leme AFP, Skaf MS, Polikarpov I (2013) *Aspergillus niger* β-glucosidase has a cellulase-like tadpole molecular shape insights into glycoside hydrolase family 3(GH3) β-glucosidase structure and fuction. J Biol Chem 288(46):32991–33005. doi:10.1074/jbc.M113.479279

Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B (2014) The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res 42(DI):D490–D495

Lynd LR, van Zyl WH, McBride JE, Laser M (2005) Consolidated bioprocessing of cellulosic biomass: an update. Curr Opin Biotechnol 16(5):577–583. doi:10.1016/j.copbio.2005.08.009

Martin J, Bruno VM, Fang Z, Meng X, Blow M, Zhang T, Sherlock G, Snyder M, Wang Z (2010) Rnnotator: an automated de novo transcriptome assembly pipeline from stranded RNA-Seq reads. BMC Genomics 11:663

Martinez D, Berka RM, Henrissat B, Saloheimo M, Arvas M, Baker SE, Chapman J, Chertkov O, Coutinho PM, Cullen D, Danchin EGJ, Grigoriev IV, Harris P, Jackson M, Kubicek CP, Han CS, Ho I, Larrondo LF, de Leon AL, Magnuson JK, Merino S, Misra M, Nelson B, Putnam N, Robbertse B, Salamov AA, Schmoll M, Terry A, Thayer N, Westerholm-Parvinen A, Schoch CL, Yao J, Barabote R, Nelson MA, Detter C, Bruce D, Kuske CR, Xie G, Richardson P, Rokhsar DS, Lucas SM, Rubin EM, Dunn-Coleman N, Ward M, Brettin TS (2008) Genome sequencing and analysis of the biomass-degrading fungus *Trichoderma reesei* (syn. *Hypocrea jecorina*). Nat Biotechnol 26(5):553–560. doi:10.1038/nbt1008-1193a

Mba Medie F, Davies GJ, Drancourt M, Henrissat B (2012) Genome analyses highlight the different biological roles of cellulases. Nat Rev Microbiol 10(3):227–234

Miao Y, Liu D, Li G, Li P, Xu Y, Shen Q, Zhang R (2015) Genome-wide transcriptomic analysis of a superior biomass-degrading strain of *A. fumigatus* revealed active lignocellulose-degrading genes. BMC Genomics 16:459. doi:10.1186/s12864-015-1658-2

Newis K, Lenfant N, Lombard V, Henrissat B (2016) Dividing the large glycoside hydrolase family 43 into subfamilies: a motivation for detailed enzyme characterization. Appl Environ Microbiol 82(6): 1686–1692

Olson DG, McBride JE, Shaw AJ, Lynd LR (2012) Recent progress in consolidated bioprocessing. Curr Opin Biotechnol 23(3):396–405

Pant A, Saikia SK, Shukla V, Asthana J, Akhoon BA, Pandey R (2014) β-Caryophyllene modulates expression of stress response genes and mediates longevity in *Caenorhabditis elegans*. Exp Gerontol 57:81–95. doi:10.1016/j.exger.2014.05.007

Parisutham V, Kim TH, Lee SK (2014) Feasibilities of consolidated bioprocessing microbes: from pretreatment to biofuel production. Bioresour Technol 161:431–440. doi:10.1016/j.biortech.2014.03.114

Parkhomchuk D, Borodina T, Amstislavskiy V, Banaru M, Hallen L, Krobitsch S, Lehrach H, Soldatov A (2009) Transcriptome analysis by strand-specific sequencing of complementary DNA. Nucleic Acids Res 37(18):e123

Sandgren M, Wu M, Karkehabadi S, Mitchinson C, Kelemen BR, Larenas EA, Stahlberg J, Hansson H (2013) The structure of a bacterial cellobiohydrolase: the catalytic core of the *Thermobifida fusca* family GH6 cellobiohydrolase Cel6B. J Mol Biol 425(3):622–635. doi:10.1016/j.jmb.2012.11.039

Schuster BG, Chinn MS (2013) Consolidated bioprocessing of lignocellulosic feedstocks for ethanol fuel production. Bioenerg Res 6(2): 416–435. doi:10.1007/s12155-012-9278-z

Scott SA, Davey MP, Dennis JS, Horst I, Howe CJ, Lea-Smith DJ, Smith AG (2010) Biodiesel from algae: challenges and prospects. Curr Opin Biotechnol 21(3):277–286

Soares Junior FL, Franco Dias AC, Fasanella CC, Taketani RG, de Souza Lima AO, Melo IS, Andreote FD (2013) Endo- and exoglucanase activities in bacteria from mangrove sediment. Braz J Microbiol 44(3):969–976

Strobel G (2014) The story of mycodiesel. Curr Opin Microbiol 19:52–58. doi:10.1016/j.mib.2014.06.003

Strobel GA (2015) Bioprospecting-fuels from fungi. Biotechnol Lett 37(5):973–982. doi:10.1007/s10529-015-1773-9

Tanasova M, Sturla SJ (2012) Chemistry and biology of acylfulvenes: sesquiterpene-derived antitumor agents. Chem Rev 112(6):3578–3610. doi:10.1021/cr2001367

Tian C, Beeson WT, Lavarone AT, Sun J, Marletta MA, Cate JH, Glass NL (2009) Systems analysis of plant cell wall degradation by the model filamentous fungus *Neurospora crassa*. Proc Natl Acad Sci U S A 106(52):22157–22162

Wang L, Sharifzadeh M, Templer R, Murphy RJ (2013) Bioethanol production from various waste papers: economic feasibility and sensitivity analysis. Appl Energy 111:1172–1182. doi:10.1016/j.apenergy.2012.08.048

Wei H, Wang W, Alahuhta M, Wall TV, Baker JO, Taylor LE II, Decker SR, Himmel ME, Zhang M (2014) Engineering towards a complete heterologous cellulase secretome in *Yarrowia lipolytica* reveals its potential for consolidated bioprocessing. Biotechnol Biofuels 7(1):148. doi:10.1186/s13068-014-0148-0

Wu W (2013) Fuel ethanol production using novel carbon sources and fermentation medium optimization with response surface methodology. Int J Agric Biol Eng 6(2):42–53. doi:10.3965/j.ijabe.20130602.006

Wu W, Davis RW (2016) One-pot bioconversion of algae biomass into terpenes for advanced biofuels and bioproducts. Algal Res 17:316–320

Wu W, Fan Z (2013) A general inhibition kinetics model for ethanol production using a novel carbon source: sodium gluconate. Bioprocess Biosyst Eng 36(11):1631–1640. doi:10.1007/s00449-013-0938-y

Wu W, Hildebrand A, Kasuga T, Xiong X, Fan Z (2013a) Direct cellobiose production from cellulose using sextuple β-glucosidase gene deletion *Neurospora crassa* mutants. Enzym Microb Technol 52(3):184–189. doi:10.1016/j.enzmictec.2012.12.010

Wu W, Kasuga T, Xiong X, Ma D, Fan Z (2013b) Location and contribution of individual β-glucosidase from *Neurospora crassa* to total β-glucosidase activity. Arch Microbiol 195(12):823–829

Wu W, Tran-Gyamfi MB, Jaryenneh JD, Davis RW (2016a) Cofactor engineering of ketol-acid reductoisomerase (IlvC) and alcohol dehydrogenase (YqhD) improves the fusel alcohol yield in algal protein anaerobic fermentation. Algal Res 19:162–167

Wu W, Tran W, Taatjes CA, Alonso-Gutierrez J, Lee TS, Gladden JM (2016b) Rapid discovery and functional characterization of terpene synthases from four endophytic xylariaceae. PLoS One 11(2):e0146983

Xiao Z, Storms R, Tsang A (2005) Microplate-based carboxymethylcellulose assay for endoglucanase activity. Anal Biochem 342(1):176–178. doi:10.1016/j.ab.2005.01.052

Yamada R, Hasunuma T, Kondo A (2013) Endowing non-cellulolytic microorganisms with cellulolytic activity aiming for consolidated bioprocessing. Biotechnol Adv 31(6):754–763. doi:10.1016/j.biotechadv.2013.02.007

Zheng F, Ding S (2013) Processivity and enzymatic mode of a glycoside hydrolase family 5 endoglucanase from *Volvariella volvacea*. Appl Environ Microbiol 79(3):989–996. doi:10.1128/aem.02725-12

Zhou P, Liu Y, Yan Q, Chen Z, Qin Z, Jiang Z (2014) Structural insights into the substrate specificity and transglycosylation activity of a fungal glycoside hydrolase family 5 β-mannosidase. Acta Crystallogr D Biol Crystallogr 70(Pt 11):2970–2982. doi:10.1107/s1399004714019762