Supporting Information

Znameroski et al. 10.1073/pnas.1118440109

SI Materials and Methods

Strains. Strains obtained from the Fungal Genetics Stock Center (FGSC) include the WT (FGSC 2489), and deletion strains for the intracellular β-glucosidase NCU00130 (FGSC 11822 and FGSC 11823), and extracellular β-glucosidases: NCU08755 (FGSC 18387 and FGSC 18388) and NCU04952 (FGSC 13731 and FGSC 13732). The homokaryon *cre-1* deletion strain (NCU08807) is described in ref. 1. Multiple deletion strains were made by performing sequential crosses. The genotype of each multiple deletion strain was confirmed using a gene-specific primer and a common primer for the hygromycin (hph) cassette. The primer for *hph* was 5′-CGA CAG ACG TCG CGG TGA GTT CAG-3′. Reverse primers were NCU00130: 5′- TAG TGT ACA AAC CCC AAG C -3′; NCU004953: 5′-AAC ACA CAC ACA CAC ACT GG-3′; NCU08755: 5′-ACA GTG GAG GTG AGA AAG G-3′; and NCU08807: 5′-GTA CTT ACG CAG TAG CGT GG-3′.

Transcriptional Studies. Conidia from strains were inoculated at a concentration equal to 10^6 conidia per milliliter in 50 mL Vogel's salts (2) with 2% wt/vol sucrose in a 250 mL Erlenmeyer flask and grown under constant light at 200 rpm for 16 h. Biomass was then spun at 1,000 × g for 10 min and washed in Vogel's salts twice to remove any excess sucrose. Biomass was then added to a new flask containing 50 mL Vogel's salts supplemented with 1% wt/vol sucrose, 0.2% wt/vol cellobiose (Sigma), or 1% wt/vol Avicel PH 101 (Sigma). Cultures were induced for 4 h under constant light at 200 rpm. The culture biomass was then harvested by filtration over a Whatman glass microfiber filter (GF/F) on a Buchner funnel and washed with 50 mL Vogel's salts. The biomass was flash frozen in liquid nitrogen and stored at −80 °C. Three independent biological replicates (flasks) were evaluated for each time point.

RNA Isolation. Total RNA from frozen samples was isolated using Zirconia/Silica beads (0.5 mm diameter; Biospec) and a Mini-Beadbeater-96 (Biospec) with 1 mL TRIzol reagent (Invitrogen) according to the manufacturer's instructions. The total RNA was further purified by digestion with TURBO DNA-free (Ambion) and an RNeasy kit (Qiagen). RNA concentration and integrity was checked by Nanodrop and agarose gel electrophoresis.

RT-PCR. Quantitative RT-PCR was performed using the EXPRESS One-Step SYBR GreenER Kit (Invitrogen) and the StepOnePlus Real-Time PCR System (Applied Biosystems). Reactions were performed in triplicate with a total reaction volume of 10 μL including 300 nM each forward and reverse primers and 75 ng template RNA. Data Analysis was performed by the StepOne Software (Applied Biosystems) using the Relative Quantitation/Comparative CT (ΔΔCT) setting. Data was normalized to the endogenous control actin with expression on sucrose as the reference sample.

RT-PCR Primers. The primers for actin (NCU04173) were forward 5′-TGA TCT TAC CGA CTA CCT-3′ and reverse 5′-CAG AGC TTC TCC TTG ATG-3′. The primers for *cbh-1* (NCU07340) were forward 5′-ATC TGG GAA GCG AAC AAA G-3′ and reverse 5′-TAG CGG TCG TCG GAA TAG-3′. The primers for *gh6-2* (NCU09680) were forward 5′-CCC ATC ACC ACT ACT ACC-3′ and reverse 5′-CCA GCC CTG AAC ACC AAG-3′. The primers for *gh5-1* (NCU00762) were forward 5′- GAG TTC ACA TTC CCT GAC A-3′ and reverse 5′-CGA AGC CAA CAC GGA AGA-3′. RT-PCR primers were previously identified and optimized in Tian, et al. (3) and Dementhon, et al. (4).

Phylogenetic Analysis. GenBank accession numbers (PID), Joint Genome Institute protein identification (JGI), or Broad Institute Fusarium Comparative Database Genes (FGSG) numbers for B-G's used in phylogenetic analysis are, as follows: NCU08755: *Myceliophthora thermophila*, JGI 80304; *Aspergillus niger*, PID 254674400; *Phanerochaete chrysosporium*, PID 19352194; *Trichoderma reesei*, JGI 121735; *Fusarium graminearum*, FGSG 06605; *Sclerotinia sclerotiorum*, PID 156051478; *Botryotinia fuckeliana*, PID 154301968; *Penicillium chrysogenum*, PID 255942539; *Schizophyllum commune*, JGI 256304; *Postia placenta*, JGI 107557. NCU00130: *Myceliophthora thermophila*, JGI 115968; *Aspergillus niger*, PID 213437; *Phanerochaete chrysosporium*, PID 127920; *Trichoderma reesei*, JGI 120749; *Fusarium graminearum*, FGSG 07274; *Sclerotinia sclerotiorum*, PID 156037816; *Botryotinia fuckeliana*, PID 156037816; *Penicillium chrysogenum*, PID 255941826; *Schizophyllum commune*, JGI 57050; *Postia placenta*, JGI 45922. NCU04952: *Myceliophthora thermophila*, JGI 66804; *Aspergillus terreus*, PID 115401928; *Phanerochaete chrysosporium*, PID 3320413; *Trichoderma reesei*, JGI 76672; *Sclerotinia sclerotiorum*, PID 156050519; *Botryotinia fuckeliana*, PID 154293970; *Penicillium chrysogenum*, PID 255945487; *Schizophyllum commune*, PID 302694815.

All proteins used in the alignments were identified using BLASTp. Homologous proteins sequences were aligned in MEGA5 using ClustalW. Maximum likelihood phylogeny was determined using the Poisson model to estimate distances and the nearest-neighborhood-interchange tree searching strategy with 500 bootstrap replications (5, 6).

mRNA Sequencing. mRNA sequencing was performed using an Illumina kit (RS-100-0801) with RNA isolated as described above. The final cDNA library was quantified by an Agilent bioanalyzer 2000 (Functional Genomics Laboratory, University of California, Berkeley) and sequenced using an Illumina Genome Analyzer-II (Vincent J. Coates Genomic Sequencing Laboratory, University of California, Berkeley) using standard Illumina operating procedures.

Analysis of Differential Expression. To establish biological variation, triplicate cultures were sampled and analyzed for the WT strain on cellulose and sucrose at 4 h after the media shift. For all other strains and conditions, a single RNAseq library was analyzed.

Sequenced libraries were mapped against predicted transcripts from the *Neurospora crassa* OR74A genome (version 10) with Tophat (version 1.1.4) (7). Transcript abundance was estimated with Cufflinks (version 0.9.2) in FPKMs (fragments per kilobase of exon per million fragments mapped) (8) using upper quartile normalization and mapping against reference isoforms from the Broad Institute.

Hierarchical Clustering Analysis. Genes exhibiting statistically significant expression changes between strains or growth conditions were identified with Cuffdiff, using upper quartile normalization and a minimum of mapped reads per locus. These genes were then filtered to select only those exhibiting a twofold change in estimated abundance between all biological replicates of each strain/condition tested and only those genes with an FPKM consistently above 10 in at least one strain/condition.

The hierarchical clustering analysis was performed using Cluster 3.0 (9) according to the FPKMs in the WT strain on Avicel, WT on cellobiose, mutant strains on cellobiose, and mutant strains on Avicel. Prior to clustering, FPKMs were log-transformed, normalized across strains/conditions on a per gene basis, and centered on the mean value across strains/conditions. The Pearson correlation coefficient (uncentered) was used as the similarity metric and average linkage as the clustering method.

Bioreactor Studies. Cellulase production was carried out in a 3.7 L bioreactor (BioEngineering AG) at an operating volume of 1 L. The bioreactor was equipped with one 48 mm Rushton impeller and four equally spaced baffles to provide adequate mixing. Impeller speed was controlled at 200 rpm for 8 h to allow spore germination followed by 500 rpm for the remainder of the experiment. The temperature was maintained at 25 °C, and medium pH was controlled at 5.5 using 40% phosphoric acid and 1:5 diluted ammonium hydroxide. The dissolved oxygen was maintained at a level greater than 20% of the saturation value of the medium by varying the aeration rate between 0.5 and 3 vessels volumes per minute in response to the dissolved oxygen tension. Minimal growth medium with 1% wt/vol sucrose as the sole carbon source (unless otherwise noted) was inoculated with 10^9 conidia. After 24 h initial growth, cellulase production was induced with either cellobiose or Avicel added to a final concentration of 0.2% wt/vol. Supernatant samples were collected when inoculated, 12 h before induction, at induction, as well as 4, 8, 12, 24, and 36 h post induction. Samples were spun at 4,500 × g for 5 min to pellet biomass and the supernatant was filtered through a .2 μm polyethersulfone filter before being stored at −20 °C until all samples were collected.

Enzyme Activity Measurements. Total secreted proteins were measured using the BioRad Protein Assay kit (BioRad) and visualized by running 15 μL of unconcentrated supernatant on a Criterion 4–14% Tris·HCl polyacrylamide gel and stained with Thermo Scientific GelCode Blue Stain Reagent.

Enzymatic Hydrolysis. Total Avicelase activity was conducted in 250 mL media bottles incubated at 50 °C on an orbital shaker at 200 rpm. Each bottle contained 1% cellulose (Avicel) and 50 mM (pH 5.0) sodium acetate in a working volume of 50 mL. Tetracycline (10 μg/mL) was added to prevent microbial contamination. Bioreactor culture broth samples were buffer exchanged using a 10 kDa molecular weight cut off centrifugal filter to remove any soluble sugars prior to initiating hydrolysis experiments. After preincubating the hydrolysis mixture to 50 °C, enzyme was added (1 mL filtered culture broth). Samples were taken every 4 h for the first 12 h and then every 12 h thereafter for a total of 48 h. Hydrolysis experiments were performed in triplicate.

Sugar Analysis. Sucrose, fructose, glucose, and cellobiose were measured on a DIONEX ICS-3000 HPLC (Dionex Corporation) using a CarboPac PA20 Analytical Column (3 × 150 mm) and a CarboPac PA20 guard column (3 × 30 mm) at 30 °C. Following injection of 25 μL of diluted samples, elution was performed with 100 mM KOH (isocratic) at 0.4 mL/min. Sugars were detected using pulsed amperometric detection, Four-Potential Carbohydrate Waveform and Peaks were analyzed using the Chromeleon software package.

Mass Spectrometry. Acetonitrile (Fisher Optima grade, 99.9%) and formic acid (Pierce, 1 mL ampules, 99+%) purchased from Fisher Scientific, and water purified to a resistivity of 18.2 MΩ·cm (at 25 °C) using a Milli-Q Gradient ultrapure water purification system (Millipore), were used to prepare mobile phase solvents for liquid chromatography-mass spectrometry.

Trypsin-digested proteins were analyzed using an orthogonal acceleration quadrupole time-of-flight (Q-TOF) mass spectrometer that was connected in-line with an ultraperformance liquid chromatograph (UPLC). Peptides were separated using a nanoAcquity UPLC (Waters) equipped with C_{18} trapping (180 μm × 20 mm) and analytical (100 μm × 100 mm) columns and a 10 μL sample loop. Solvent A was 99.9% water/0.1% formic acid and solvent B was 99.9% acetonitrile/0.1% formic acid (vol/vol). Sample solutions contained in 0.3 mL polypropylene snap-top vials sealed with septa caps (Wheaton Science) were loaded into the nanoAcquity autosampler compartment prior to analysis. Following sample injection (10 μL), trapping was performed for 3 min with 100% A at a flow rate of 15 μL/min. The injection needle was washed with 500 μL each of solvents A and B after injection to avoid cross-contamination between samples. The elution program consisted of a linear gradient from 8% to 35% B over 30 min, a linear gradient to 95% B over 0.33 min, isocratic conditions at 95% B for 3.67 min, a linear gradient to 1% B over 0.33 min, and isocratic conditions at 1% B for 11.67 min, at a flow rate of 500 nL/min. The analytical column and sample compartment were maintained at 35 °C and 8 °C, respectively.

The UPLC column exit was connected to a Universal Nano-Flow Sprayer nanoelectrospray ionization (nanoESI) emitter that was mounted in the nanoflow ion source of the mass spectrometer (Q-TOF Premier, Waters). The nanoESI emitter tip was positioned approximately 3 mm from the sampling cone aperture. The nanoESI source parameters were, as follows: nanoESI voltage 2.4 kV, nebulizing gas (nitrogen) pressure 0.15 mbar, sample cone voltage 35 V, extraction cone and ion guide voltages 4 V, and source block temperature 80 °C. No cone gas was used. The collision cell contained argon gas at a pressure of 8×10^{-3} mbar. The TOF analyzer was operated in V mode. Under these conditions, a mass resolving power (10) of 1.0×10^4 (measured at $m/z = 771$) was routinely achieved, which was sufficient to resolve the isotopic distributions of the singly and multiply charged precursor and fragment ions measured in this study. Thus, an ion's mass and charge were determined independently, i.e., the ion charge was determined from the reciprocal of the spacing between adjacent isotope peaks in the m/z spectrum. External mass calibration was performed immediately prior to analysis using a solution of sodium formate. Survey scans were acquired in the positive ion mode over the range $m/z = 400$–$1{,}500$ using a 0.45-s scan integration and a 0.05-s interscan delay. In the data-dependent mode, up to five precursor ions exceeding an intensity threshold of 20 counts/s (cps) were selected from each survey scan for tandem mass spectrometry (MS/MS) analysis. Real-time deisotoping and charge-state recognition were used to select 2+, 3+, and 4+ charge state precursor ions for MS/MS. Collision energies for collisionally activated dissociation were automatically selected based on the mass and charge state of a given precursor ion. MS/MS spectra were acquired over the range $m/z = 100$–$2{,}000$ using a 0.20-s scan integration and a 0.05-s interscan delay. Ions were fragmented to achieve a minimum total ion current of 30,000 cps in the cumulative MS/MS spectrum for a maximum of 2 s. To avoid the occurrence of redundant MS/MS measurements, real-time dynamic exclusion was used to preclude reselection of previously analyzed precursor ions over an exclusion width of ± 0.2 m/z unit for a period of 300 s.

Data resulting from liquid chromatography-MS/MS analysis of trypsin-digested proteins were processed using ProteinLynx Global Server software (version 2.3, Waters), which performed background subtraction (threshold 35% and fifth order polynomial), smoothing (Savitzky–Golay, 10 times, over three channels), and centroiding (top 80% of each peak and minimum peak width at half height four channels) of mass spectra and MS/MS spectra. Processed data were searched against the *Neurospora crassa* protein database (Broad Institute). The following criteria were used

for the database search: precursor ion mass tolerance 100 ppm, fragment ion mass tolerance 0.15 Da, digest reagent trypsin, allowing for up to three missed cleavages, and methionine oxidation as a variable modification. The identification of at least three consecutive fragment ions from the same series, i.e., b or y-type fragment ions (11), was required for assignment of a peptide to an MS/MS spectrum. MS/MS spectra were inspected to verify the presence of fragment ions that identify the peptides. A protein was determined to be present if at least one peptide was detected in two out of three biological replicates [whole supernatant, phosphoric acid-swollen cellulose (PASC) bound or PASC unbound].

1. Sun J, Glass NL (2011) Identification of the CRE-1 cellulolytic regulon in *Neurospora crassa*. *PLoS One* 6:e25654.
2. Vogel H (1956) A convenient growth medium for *Neurospora*. *Microb Genet Bull* 13:42–46.
3. Tian C, et al. (2009) Systems analysis of plant cell wall degradation by the model filamentous fungus *Neurospora crassa*. *Proc Natl Acad Sci USA* 106:22157–22162.
4. Dementhon K, Iyer G, Glass NL (2006) VIB-1 is required for expression of genes necessary for programmed cell death in *Neurospora crassa*. *Eukaryot Cell* 5:2161–2173.
5. Hall BG (2008) *Phylogenetic Trees Made Easy: A How-To Manual* (Sinauer Associates, Sunderland, MA) 3rd Ed, pp xiv, 233.
6. Tamura K, et al. (2011) MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods. *Mol Biol Evol* 28:2731–2739.
7. Langmead B, Trapnell C, Pop M, Salzberg SL (2009) Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol* 10:R25.
8. Roberts A, Trapnell C, Donaghey J, Rinn JL, Pachter L (2011) Improving RNA-Seq expression estimates by correcting for fragment bias. *Genome Biol* 12:R22.
9. de Hoon MJL, Imoto S, Nolan J, Miyano S (2004) Open source clustering software. *Bioinformatics* 20:1453–1454.
10. Marshall AG, Hendrickson CL (2008) High-resolution mass spectrometers. *Annu Rev Anal Chem* 1:579–599.
11. Roepstorff P, Fohlman J (1984) Proposal for a common nomenclature for sequence ions in mass spectra of peptides. *Biomed Mass Spectrom* 11:601 (lett).

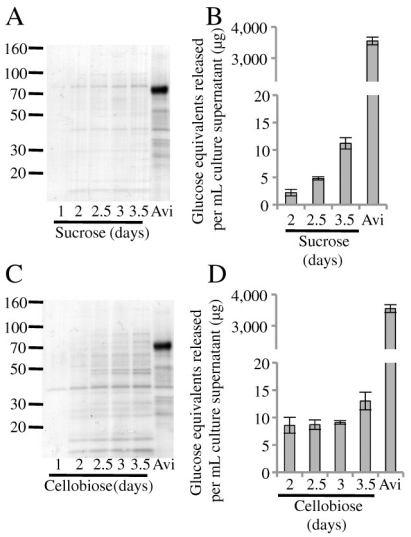

Fig. S1. Protein production and enzyme activity are not induced by starvation. (*A*) SDS-PAGE of secreted proteins in culture filtrates from WT grown on sucrose for 3.5 d and Avicel for 5 d. (*B*) Supernatant activity (from *A*) toward Avicel. (*C*) SDS-PAGE of secreted proteins in culture filtrates from WT grown on cellobiose for 3.5 d and Avicel for 5 d. (*D*) Supernatant activity (from *C*) toward Avicel. Error bars are 1 SD.

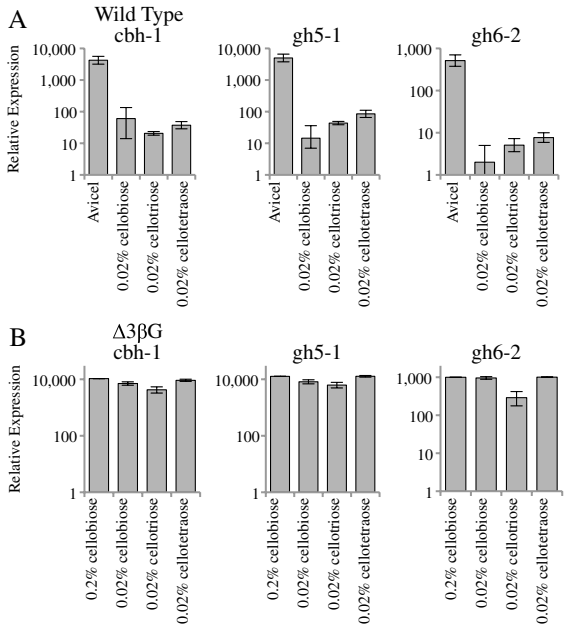

Fig. S2. Cellulase induction in WT and Δ3βG after induction with cellodextrins. (*A*) Expression of *cbh-1*, *gh5-1*, and *gh6-2* in WT after a 4-h induction with Avicel, cellobiose, cellotriose, or cellotetraose. (*B*) Expression of *cbh-1*, *gh5-1*, and *gh6-2* in the Δ3βG mutant after a 4-h induction with Avicel, cellobiose, cellotriose, or cellotetraose. Gene expression levels of *cbh-1*, *gh5-1*, and *gh6-2* were normalized to 1 when induced with 1% sucrose. Actin (NCU04173) gene expression levels were used as an endogenous control in all samples. Error bars indicate 1 SD.

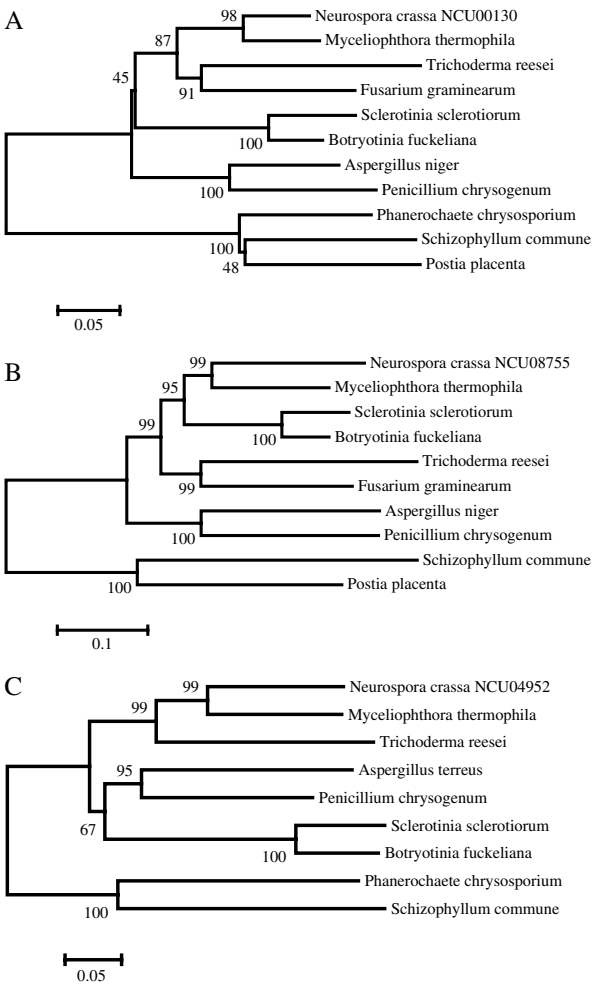

Fig. S3. Maximum likelihood phylogenetic analysis of the deleted β-glucosidases in various filamentous fungi. (*A*) NCU08755, (*B*) NCU04952, and (*C*) NCU00130. Accession numbers for each gene are listed in *SI Materials and Methods*.

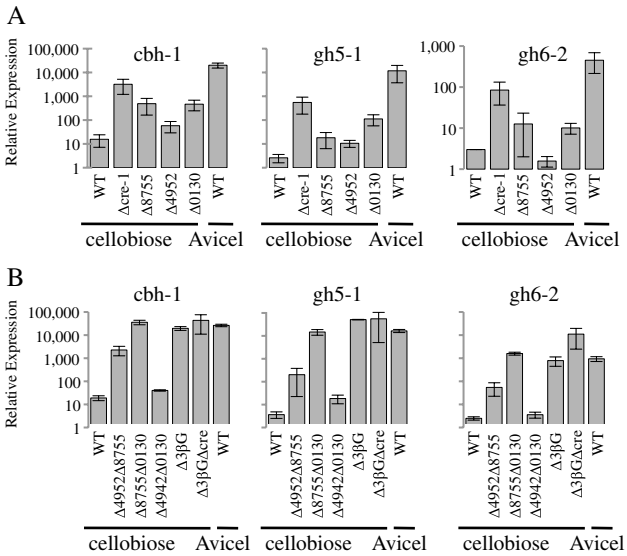

Fig. S4. Cellulase induction in WT and single or double β-glucosidase deletion strains after induction with cellobiose or Avicel. (*A*) Gene expression of select cellulases after 4-h induction with 0.2% cellobiose or 1% Avicel in WT, Δ*cre-1*, ΔNCU08755, ΔNCU04952, and ΔNCU00130, or (*B*) gene expression of select cellulases after 4-h induction with 0.2% cellobiose or 1% Avicel in WT, ΔNCU04952 ΔNCU08755, ΔNCU08755 ΔNCU00130, and ΔNCU04952 ΔNCU00130. Gene expression levels of *cbh-1*, *gh6-2*, and *gh5-1* were normalized to 1 when switched to 1% sucrose. Actin was used as an endogenous control in all samples. Each strain was grown in triplicate and error bars indicate 1 SD.

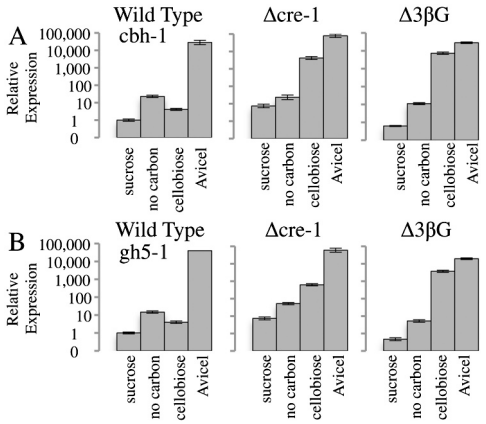

Fig. S5. Cellulase induction in WT, Δ*cre-1*, and Δ3βG after 4-h induction with sucrose, no carbon (starvation), cellobiose, and Avicel. (*A*) Expression of *cbh-1* and (*B*) *gh5-1* in WT, Δ*cre-1*, and Δ3βG after a 4-h induction with 1% sucrose, no carbon (Vogel's salt solution only), 0.2% cellobiose, or 1% Avicel. Expression levels for all genes were normalized to 1 when induced with 1% sucrose. Actin (NCU04173) gene expression levels were used as an endogenous control in all samples. Each reaction was done in triplicate and error bars indicate 1 SD.

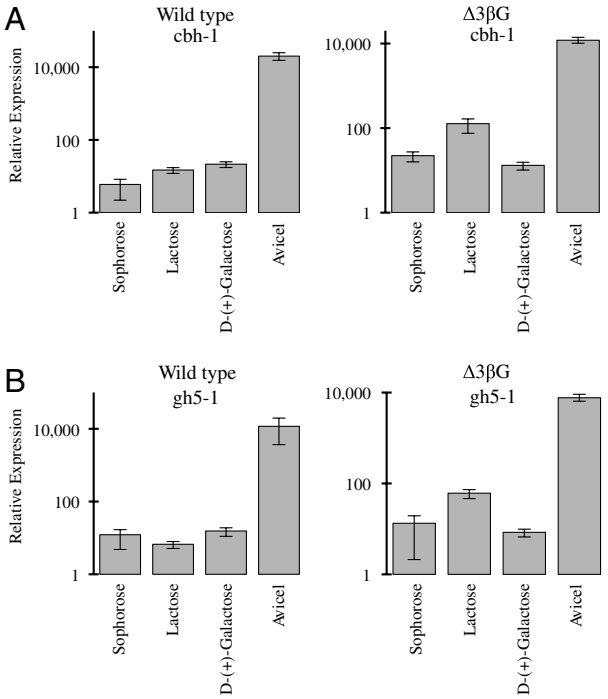

Fig. S6. Cellulase induction in WT and Δ3βG after induction with sophorose, lactose, or D-(+)-galactose. (*A*) Expression of *cbh-1* in WT and Δ3βG after a 4-h induction with 1 mM sophorose, 1 mM lactose, or 1 mM D-(+)-galactose. (*B*) Expression of *gh6-2* in WT and Δ3βG after a 4-h induction with 1 mM sophorose, 1 mM lactose, or 1 mM D-(+)-galactose. Gene expression levels of *cbh-1* and *gh6-2* were normalized to 1 when induced with 1% sucrose. Actin (NCU04173) gene expression levels were used as an endogenous control in all samples. Error bars indicate 1 SD.

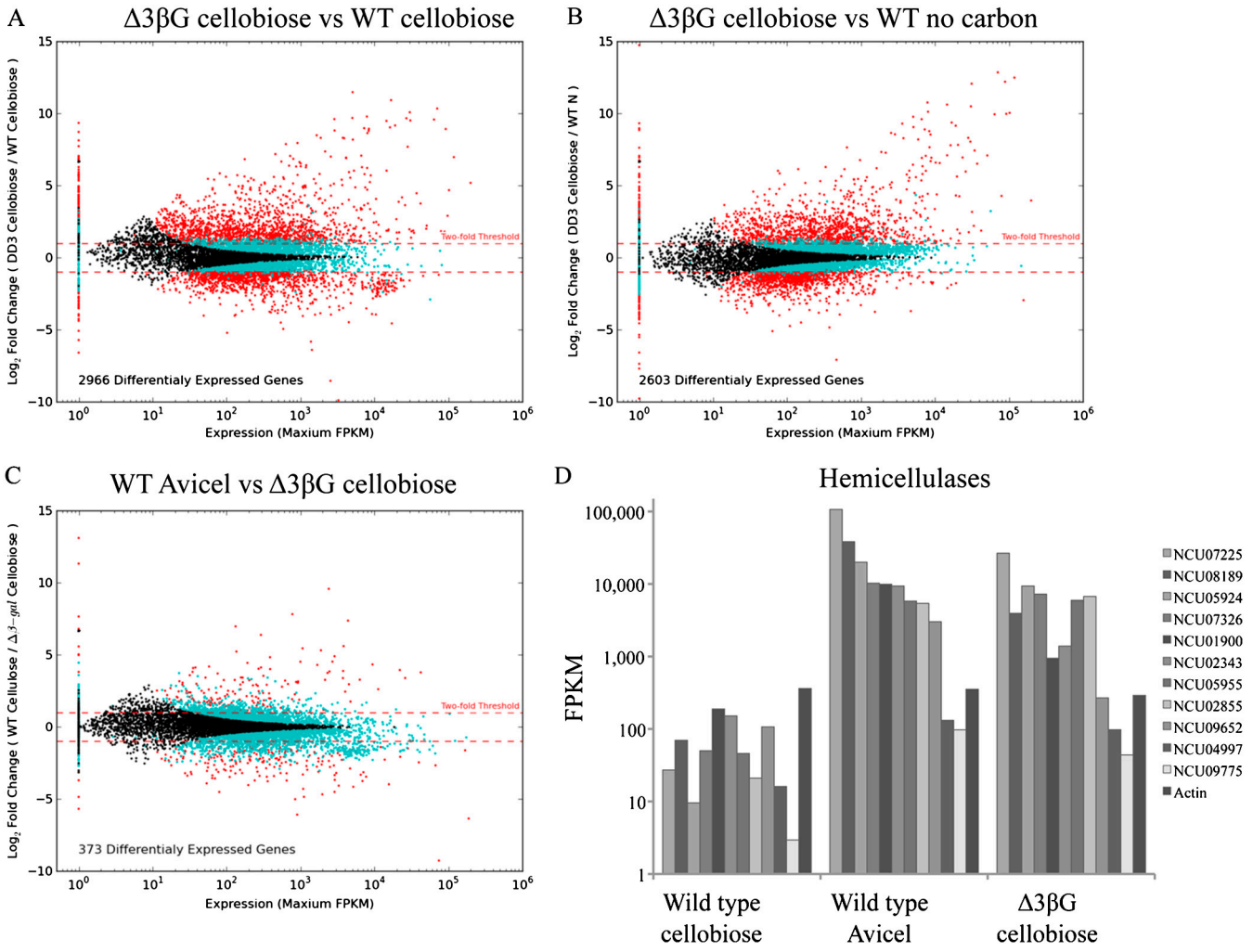

Fig. S7. RNA sequencing of the WT and Δ3βG strains. Scatter plots comparing full genomic patterns of gene expression changes in (*A*) Δ3βG induced with cellobiose versus WT induced with cellobiose, (*B*) Δ3βG induced with cellobiose versus WT starvation, (*C*) WT induced with Avicel versus Δ3βG induced with cellobiose. (*D*) Hemicellulase expression in FPKMs for the WT induced with cellobiose or Avicel as compared to Δ3βG induced with cellobiose. Expression levels for each condition are shown in the order of genes listed in the legend. All strains were grown for 16 h on 2% sucrose, followed by a transfer to no carbon source (Vogel's salt solution only), 0.2% cellobiose, or 1% Avicel for 4 h.

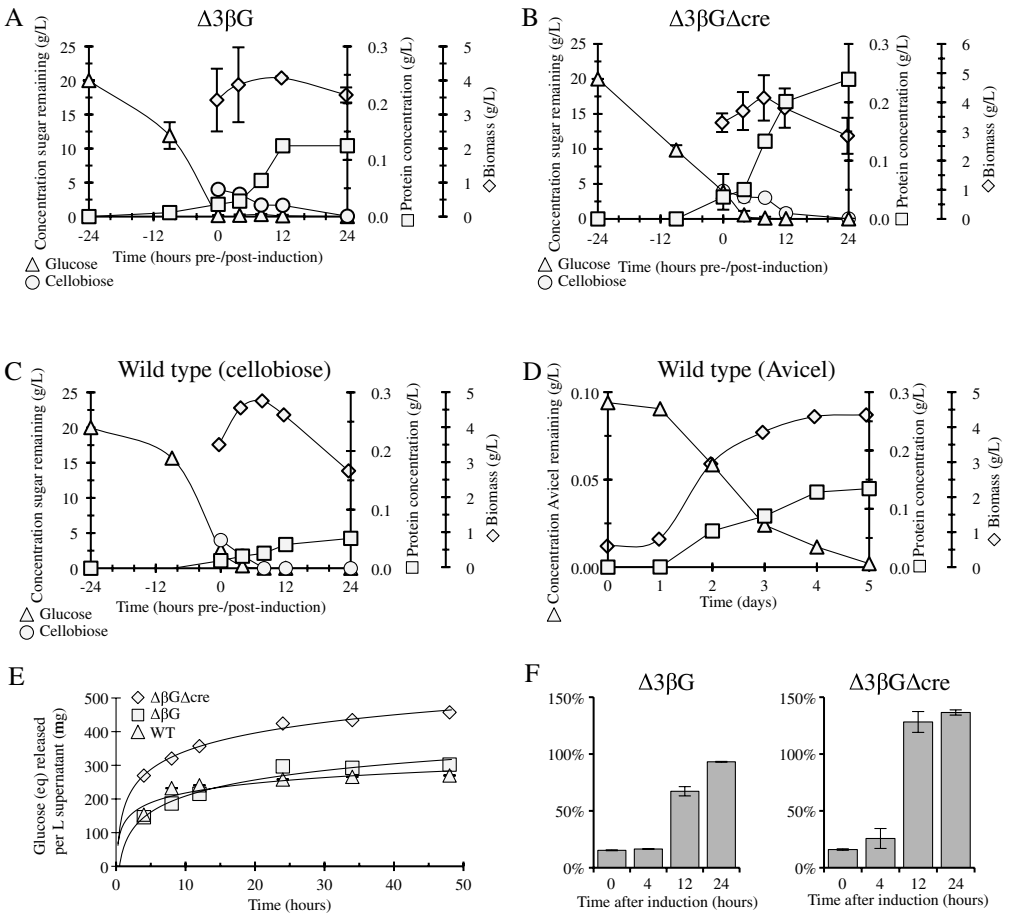

Fig. S8. Protein production and enzyme activity in WT, Δ3βG, and Δ3βGΔcre strains after induction with cellobiose or Avicel. Production of cellulases in a bioreactor using (*A*) Δ3βG induced with cellobiose, (*B*) Δ3βGΔcre induced with cellobiose, (*C*) WT induced with cellobiose, and (*D*) WT grown 5 d on Avicel. Cellobiose-induced strains were pregrown in minimal media with 1% sucrose for 24 h before induction with 0.2% cellobiose for 36 h. The concentration of sucrose, glucose, fructose (in glucose equivalents; △) cellobiose (○), protein production (□), and biomass accumulation (◇) were measured. (*E*) Twenty-four-hour induced supernatant activity from *A*, *B*, and *D* toward Avicel. Cellulase activity of culture supernatant from Δ3βG (□) and Δ3βGΔcre (△) induced with cellobiose for 24 h compared to culture supernatants from WT grown on Avicel for 5 d (◇). Error bars are 1 SD. (*F*) Azo-CMC (endoglucanase) activity time course from bioreactor culture supernatants in *A* and *B*. Azo-CMC activity is expressed as a percentage of activity from WT culture supernatant grown on 2% Avicel for 5 d.

Other Supporting Information Files

Dataset S1 (XLSX)