BIOENGINEERING

Unique genetic cassettes in a *Thermoanaerobacterium* contribute to simultaneous conversion of cellulose and monosugars into butanol

Tinggang Li,[1] Chen Zhang,[1] Kun-Lin Yang,[2] Jianzhong He[1]*

The demand for cellulosic biofuels is on the rise because of the anticipation for sustainable energy and less greenhouse gas emissions in the future. However, production of cellulosic biofuels, especially cellulosic butanol, has been hampered by the lack of potent microbes that are capable of converting cellulosic biomass into biofuels. We report a wild-type *Thermoanaerobacterium thermosaccharolyticum* strain TG57, which is capable of using microcrystalline cellulose directly to produce butanol (1.93 g/liter) as the only final product (without any acetone or ethanol produced), comparable to that of engineered microbes thus far. Strain TG57 exhibits significant advances including unique genes responsible for a new butyrate synthesis pathway, no carbon catabolite repression, and the absence of genes responsible for acetone synthesis (which is observed as the main by-product in most *Clostridium* strains known today). Furthermore, the use of glucose analog 2-deoxyglucose posed a selection pressure to facilitate isolation of strain TG57 with deletion/silencing of carbon catabolite repressor genes—the *ccr* and *xylR* genes—and thus is able to simultaneously ferment glucose, xylose, and arabinose to produce butanol (7.33 g/liter) as the sole solvent. Combined analysis of genomic and transcriptomic data revealed unusual aspects of genome organization, numerous determinants for unique bioconversions, regulation of central metabolic pathways, and distinct transcriptomic profiles. This study provides a genome-level understanding of how cellulose is metabolized by *T. thermosaccharolyticum* and sheds light on the potential of competitive and sustainable biofuel production.

INTRODUCTION

Biofuels produced from renewable lignocellulosic biomass are expected to meet growing energy demands without increasing greenhouse gas emissions as fossil fuels (*1*, *2*). Among all biofuels, butanol is one of the most promising biofuels because of its high energy density (29.2 MJ/liter for butanol versus 19.6 MJ/liter for ethanol and 32 MJ/liter for gasoline) and is more similar to gasoline (*3*, *4*). As a natural reservoir for biomass-based carbon and the most abundant biomass on Earth (*5*), cellulose has become a major feedstock for butanol production in biorefinery processes. Currently, bioconversion of cellulose to biofuels in industries requires several steps including pretreatment, enzymatic saccharification, detoxification, and fermentation (*6*–*8*). Therefore, it is desirable to develop a bioconversion technology for the direct conversion of cellulosic biomass into biofuels without entailing any of the pretreatment steps mentioned above (*9*). However, this effort has been hampered by recalcitrance of cellulose and the lack of potent microbes. A key step to address the issue is to discover novel microorganisms having unique genetic cassettes to convert cellulosic materials into biofuels (*10*, *11*). As reported previously, a number of solventogenic strains from the genus *Clostridium* have been exploited to generate butanol from monosaccharides (for example, glucose and xylose) and starch (for example, corn and cassava). However, these strains cannot use polysaccharides, such as cellulose, for butanol generation; they can do so only for ethanol or hydrogen generation (*12*). Recent research attempts have been directed toward engineering microorganisms that can directly convert cellulose into biofuels in consolidated bioprocesses (*9*, *13*, *14*). For example, genes encoding for biosynthetic pathways of biofuels have been engineered into natural cellulolytic microorganisms such as *Caldicellulosiruptor bescii*, *Clostridium cellulolyticum*, and *Clostridium thermocellum*, which were capable of producing ethanol [0.64 g/liter (*15*) and 22.4 g/liter (*16*)] and isobutanol [0.66 g/liter (*17*) and 5.4 g/liter (*18*)] from cellulose. Similar attempts on *Clostridium acetobutylicum* have also been made by introducing cellulosome genes; however, the engineered strain was unable to use cellulose because of the complex assembly and the expressional stability of functional minicellulosomes (*19*).

Until now, few wild-type species can produce biofuels, particularly butanol, directly from cellulosic biomass at high concentrations and yields. Thus, there is still a need to develop a consolidated bioprocessing technology for butanol production. Note that sustainable butanol production is also impeded by (i) product inhibition and carbon catabolite repression (CCR) (*20*) and (ii) complexities of downstream purification of butanol from other by-products such as acetone and ethanol (*21*). Hence, to achieve and simplify the consolidated bioprocessing, it is desirable to discover bacterial strains that can directly ferment cellulose and hemicellulose to butanol as the main product.

Here, we report the discovery of the first cellulolytic wild-type bacterium (*Thermoanaerobacterium thermosaccharolyticum* strain TG57) that can directly convert cellulose and xylan to butanol. Studies on the genomic characteristics of cellulose-assimilating butanologenic lifestyle are limited thus far (*22*). In *Clostridium*, the genes responsible for the formation of butyrate and acetone encode phosphotransbutyrylase and butyrate kinase. However, similar genes are absent in the genome of *T. thermosaccharolyticum* strain TG57. Instead, the genome of strain TG57 contains novel genes encoding butanol dehydrogenase (Bdh), endocellulase, and cellobiohydrolase, which were not found in *T. thermosaccharolyticum* DSM 571 (*22*), a strain closely related to TG57. Thus, to reveal the genetic repertoire of *T. thermosaccharolyticum* strain TG57, we carried out genomic, transcriptomic, functional, and biochemical characterization of this new isolate to explore the relationship among genetics, metabolism, and molecular regulation.

[1]Department of Civil and Environmental Engineering, National University of Singapore, Singapore 117576. [2]Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore 117576.
*Corresponding author. Email: jianzhong.he@nus.edu.sg

RESULTS

Direct conversion of cellulose to biobutanol

An enriched butanologenic consortium capable of assimilating microcrystalline cellulose was obtained by consecutive transfers and then was treated with a glucose analog—2-deoxyglucose (2-DG; 1.0 g/liter)—to select a bacterium resistant to CCR. Among the 98 isolates, one designated *T. thermosaccharolyticum* strain TG57 (fig. S1A) showed the highest Bdh activity (0.11 U/mg protein) and cellulase activity (2.6 U/mg protein) when fed with cellulose (table S1). The purity of culture TG57 was confirmed by obtaining a single genotype sequence of the 16*S* ribosomal RNA (rRNA) gene from clone library colonies and relative constant ratios of quantitative gene copies of the 16*S* rRNA gene and other housekeeping genes (*rpoA* and *atpD*) (see the Supplementary Materials). Strain TG57 grows optimally at a pH of 5.3 ~ 6.0 and at a temperature of 51° ~ 59°C (fig. S2). From the observed coordinate gene expression, strain TG57 appears to be able to use multiple monosaccharides, oligosaccharides, and polysaccharides (fig. S3). Strain TG57 could produce butanol (1.93 g/liter) directly from microcrystalline cellulose (Fig. 1A). Acetone and ethanol were not detected throughout the fermentation as usually produced by *Clostridium* species. The other bioacids (for example, acetate and butyrate) appeared at low titers, indicating that metabolic patterns were redirected to more reduced butanol with a higher butanol/bioacid ratio (116%), compared with those of gene-modified *Clostridium cellulovorans* (40%) or a coculture of *Clostridium celevecrescens* N3-2 and *C. acetobutylicum* ATCC 824 (55.4%) when using cellulose (*23*, *24*). Strain TG57 consumed cellobiose faster than cellulose, suggesting that the breakdown of cellobiose by the potent activity of glucosidase is not a rate-limiting step during cellulose bioconversion (table S1 and Fig. 1, A to C), the phenomenon of which was also reported previously (*8*, *25*, *26*). Strain TG57 produced butanol (5.17 and 3.63 g/liter) from cellobiose and xylan with a yield of 0.29 and 0.23 g/g, respectively (Fig. 1, B and D). The overall carbon recovery ranged from 92.3 to 95.4% in the fermentation process by strain TG57 (table S2). Results from this first wild-type strain that is capable of producing butanol directly from microcrystalline cellulose are slightly higher than those from previously reported wild-type [ethanol (0.46 g/liter)] or gene-modified [butanol (1.42 g/liter)] strains using cellulose as a substrate (Table 1) (*23*, *27*–*35*).

Resistance to CCR and product inhibition

Monosaccharides (for example, glucose and xylose) and oligosaccharides (for example, cellobiose) are known to be inhibitory to polysaccharide utilization in microbial metabolisms (*20*, *36*, *37*). Product inhibition of cellulolytic enzymes is one of the major factors affecting the cellulolytic hydrolysis rate and product yields; however, limited studies have been done to address this issue. To eliminate the effect of inhibition, we performed the selection of strains by growing them on cellulose but in the presence of 2-DG. The resulting strain TG57 produces twice as much extracellular cellulase as the enrichment before 2-DG treatment. As expected, adding glucose or cellobiose (30 g/liter for both) to the medium BTM1 (biobutanol thermoanaerobacterial medium 1) supplemented with cellulose exhibited negligible inhibition on the hydrolysis rate and cellulase and glucosidase activities as compared to the control (Fig. 2A). Notably, strain TG57 is capable of producing butanol (6.21 g/liter; yield, 0.28 g/g) via fermenting xylose and cellobiose simultaneously (Fig. 3A), which could lower the overall enzyme usage and thus the cost of cellulose saccharification (usually with low activity of β-glucosidase when using fungal cellulase mixtures) (*25*, *26*, *38*). In

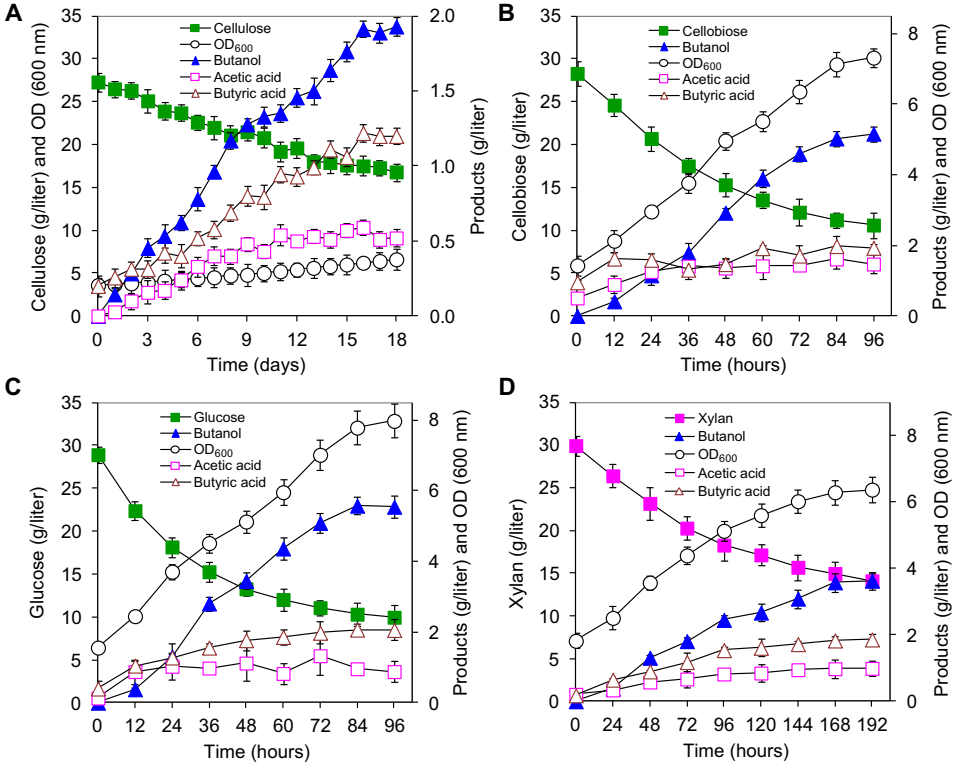

Fig. 1. Fermentation of sugars by *T. thermosaccharolyticum* TG57. (**A**) Cellulose, (**B**) cellobiose, (**C**) glucose, and (**D**) xylan as the sole carbon source. Error bars represent SD of biological triplicates.

Table 1. Performance comparison of strain TG57 with previous microorganisms capable of converting lignocellulose to biofuels. Avicel, crystalline cellulose; PASC, phosphoric acid–swollen cellulose.

Organism	Genotype	Carbon source	Product	Concentration (g/liter)	Yield (g/g)	Productivity [g/(liter·hour)]	Reference
Caldicellulosiruptor bescii	Recombinant	Switchgrass	Ethanol	0.59	0.0294	0.0049	(*15*)
C. thermocellum	Recombinant	Crystalline cellulose	Ethanol	22.4	0.39	0.162	(*16*)
C. cellulolyticum	Recombinant	Avicel	Isobutanol	0.66	0.066	—	(*17*)
C. thermocellum	Recombinant	Avicel	Isobutanol	5.4	0.17	0.072	(*18*)
C. acetobutylicum	Recombinant	Avicel	Butanol	Nil	Nil	Nil	(*19*)
C. cellulovorans	Recombinant	Crystalline cellulose	Butanol	1.42	0.2	0.0059	(*23*)
			Ethanol	1.60	0.19	0.0067	
C. phytofermentans ATCC 700394	Wild type	Birchwood xylan	Ethanol	0.46	0.153	0.0192	(*27*)
		Cellulose	Ethanol	2.76	0.438	0.0041	
Saccharomyces cerevisiae	Recombinant	PASC	Ethanol	2.1	0.525	0.035	(*28*)
C. cellulolyticum	Recombinant	Crystalline cellulose	Ethanol	2.7	0.27	0.008	(*29*)
C. thermocellum	Recombinant	Avicel	Ethanol	5.46	0.14	—	(*30*)
Cellvibrio japonicus	Recombinant	Avicel	Ethanol	2.6	0.26	0.052	(*31*)
S. cerevisiae	Recombinant	PASC	Ethanol	1.0	0.372	0.0052	(*32*)
C. phytofermentans ATCC 700394	Wild type	AFEX-pretreated corn stover	Ethanol	7.0	0.191	0.0265	(*33*)
S. cerevisiae	Recombinant	PASC	Ethanol	2.7	0.27	—	(*34*)
C. cellulolyticum	Recombinant	Avicel	Ethanol	0.8	0.025	—	(*35*)
Clostridium sp. strain MF28	Wild type	Beechwood xylan	Butanol	3.2	0.177	0.0222	(*41*)
T. thermosaccharolyticum TG57	Wild type	Crystalline cellulose	Butanol	1.93	0.201	0.0045	This study
		Beechwood xylan	Butanol	3.63	0.229	0.0189	

addition, strain TG57 used glucose, arabinose, and xylose simultaneously to produce butanol (7.33 g/liter) (Fig. 3B) with a butanol/bioacid ratio of 132 to 139%; acetone or ethanol was not detected in the fermentation broth (Fig. 3, B and C, and fig. S4). An important selection feature of *T. thermosaccharolyticum* strain TG57 was the growth of its enrichment under a selection pressure of a glucose analog, 2-DG. Accumulation of the intermediate of 2-DG–6-phosphate in cells from 2-DG degradation inhibited cell growth rapidly (*39*, *40*), showing CCR; thus, only the non–carbon catabolite–repressing bacterial strain could survive. Therefore, 2-DG was added to increase the isolation efficiency of strain TG57, which contains genes inactivating CCR (for example, the *ccp* gene encoding catabolite control protein A and the *xylR* gene encoding the transcriptional repressor of the xylose operon/ROK family regulator) for the carbon catabolite repressor cascade and upregulates gene (for example, *sec* and *t2ss*) overexpression for the enhancement of cellulolytic cellulase secretion (Figs. 2B and 3C). In addition, the elevated expression of gene *gpi* stimulates to cycle fructose 6-phosphate (F6P) via the oxidative pentose phosphate pathway (PPP) while Pfk takes F6P into glycolysis (Fig. 3D). An additional selective feature of strain TG57 is that it was capable of co-fermenting a combination of glucose and xylose, two typical and major reducing sugars derived from lignocellulosic biomass. During the co-fermentation of glucose and xylose by strain TG57, the highly transcribed xylulokinase (Xylk), xylose isomerases (XylmA and XylmB), and xylose transporter (XylT2) (Fig. 3D) indicate (i) elimination of the CCR and (ii) overcoming the low efficiency of xylose utilization when sugar mixtures were used as substrates. The presence of the glycolysis pathway and PPP, both of which involved the transporter permease, explains no xylose utilization repression via the PPP. Actually, the application of the glucose analog 2-DG played an important role for the isolation of a strain with deletion/silencing of carbon catabolite repressor *ccr* and *xylR* genes and overexpressed secretion system genes (for example, *sec*, *t2ss*, and *pilT*) to secrete high amounts of cellulolytic enzymes (for example, cellulases and glucosidases) and thus exhibited resistance to product inhibition and catabolite repression (Figs. 2 and 3). The above features distinguished *T. thermosaccharolyticum* strain TG57 from *Clostridium* sp. strain MF28 that was regulated by *gidA* and *gidB* genes (encoding glucose-inhibited division protein) when 2-DG was used as a selection pressure (*41*). This

A

Substrate*	Cellulase activity (U/mg)	Maximum cell biomass (OD_{600})	Cellulose hydrolysis rate [g/(g·hour)]	Extracellular protein (g/L)	Degree of hydrolysis†	
Cellulose (control)	1.29 ± 0.09	4.9 ± 0.26	0.025 ± 0.0013	1.16 ± 0.063	1.00‡	1.00§
Cellulose+glucose	3.01 ± 0.16	7.7 ± 0.48	0.038 ± 0.0011	1.83 ± 0.092	1.52‡	0.54§
Cellulose+cellobiose	2.98 ± 0.18	7.2 ± 0.39	0.036 ± 0.0027	1.71 ± 0.075	1.44‡	0.62§

*Substrate concentration: control, 10 g/L cellulose; mixture of substrates, 10 g/L + 30 g/L.
†The degree of hydrolysis was expressed by the ratio between the hydrolysis rate of cellulose in the presence of a possible inhibitor and the hydrolysis rate of a control cellulose hydrolysis without any sugar supplementation. ‡data from this study. §data from the work of Teugjas and Väljamäe (*36*).

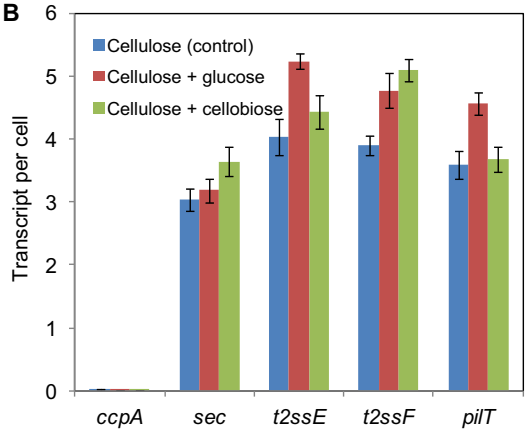

Fig. 2. Strain TG57 exhibits resistance to product inhibition. (**A**) Cellulose conversion in the presence of a typical co-substrate glucose or cellobiose. Substrate concentration: control, cellulose (10 g/liter); co-substrates, cellulose (10 g/liter) + glucose (30 g/liter) or cellobiose (30 g/liter). (**B**) Transcriptional expression of genes responsible for the catabolite control protein and the cellulase secretory system of strain TG57. *ccp* encodes the catabolite control protein, *sec* encodes general secretion pathway protein A, *t2ssE* and *t2ssF* encode type II secretion system proteins E and F, and *pilT* encodes the Tfp pilus assembly protein.

important characteristic of strain TG57 indicates the potential of TG57 for biofuel production by using lignocellulosic hydrolysates in the future.

Genetic basis of conversion of cellulose to butanol—Cellulosome, hemicellulosome, and carbohydratases

The genome of *T. thermosaccharolyticum* strain TG57 harbors a single circular 2,895,726–base pair (bp) chromosome with 34.11% of guanine-cytosine (GC) content, which is much smaller than that of the solventogenic *C. acetobutylicum*, primarily owing to the absence of a large plasmid and genetic materials related to the by-product acetone and lengthy butanoate synthesis pathways (Fig. 4, A and B). A total of 2660 predicted protein coding DNA sequences (CDSs) were annotated to biological functions, covering 89% of the genome, which represented a significant difference from those of the previously sequenced *C. acetobutylicum* (Fig. 4B and fig. S5). In addition, 211 (7%) conserved hypothetical proteins and 115 (4%) proteins of unknown functions are assigned to the remaining 326 CDSs. Strain TG57 shares 99% identity with *T. thermosaccharolyticum* DSM 571 over 2.89 Mbp with an 80% sequence cover, whereas the key gene *bdh* (Thert_02648), encoding Bdh, shares 82.9% similarity and the cellulase CelA belonging to the GH5 family (Thert_00325) and the CelB belonging to the GH51 family (Thert_00869) share 93 and 26.5% similarities, respectively. The putative cellulases in strain TG57 comprise two glycoside hydrolase families (GH5 and GH51). The genome of strain TG57 contains more than 88 putative carbohydratases comprising 62 glycoside hydrolases, 18 carbohydrate esterases, and 6 glycosyltransferases with at least 17 distinct families (table S3). Moreover, large families of closely related genes are responsible for redundant roles especially for extracellular hydrolases. The cellulosome in *T. thermosaccharolyticum* strain TG57 includes at least two putative endoglucanase, three exocellobiohydrolase, and at least eight glucosidases (α-glucosidases, β-glucosidases, 6-phospho-β-glucosidase, and 6-phospho-α-glucosidase). Many β-glucosidase genes were found to encode two enzymes of family GH1 and three members of GH3 (table S3). In contrast to tightly clustered genes encoding hemicellulolytic enzymes, genes for the cellulase of *T. thermosaccharolyticum* strain TG57 are individually dispersed in the genome. In addition to the genes for cellulase, 15 other genes for polysaccharide (for example, hemicellulose, chitin, glycogen, and starch) hydrolysis are found in the genome of TG57. For example, one such operon (Thert_02763-Thert_02782, 2,295,217 to 2,318,716 bp) consists of a cluster of 18 predicted genes associated with hemicellulosic xylan hydrolysis (fig. S3 and Fig. 4C).

Transcriptome-wide profiles connected to genome CDS, CAZy (carbohydrate-active enzymes), and Kyoto Encyclopedia of Genes and Genomes categories show that expression variation in cellulose culture extends to a variety of metabolisms (Fig. 4C), whereas the expression variation in hemicellulose culture exists in CAZy, PPP, and transporters. In particular, genes for flagellar and type IV secretion and synthesis of fatty acids and numerous housekeeping genes (for example, DNA replication) were down-regulated, whereas genes for glycolysis, carbohydratases, reducing cofactor synthesis (for example, NADH and NADPH), and cellulosome that aim to elevate glycolytic flux were up-regulated in culture TG57 growing on cellulose. For example, the most highly transcribed cellulase gene, Thert_00869 CelB, is predicted to have a cellulose-binding module for covalent attachment to the microcrystalline cellulose wall, which is confirmed by the adhesion of cells to the microcrystalline cellulose as observed by fluorescence microscopy [refer to the "Secretome and transportome (transport systems)" section]. The exocellulase gene Thert_00325 was expressed similarly to Thert_00869, suggesting that these two enzymes synergize the breakdown of cellulose. The gene Thert_00855 in the genome of strain TG57 was the most highly

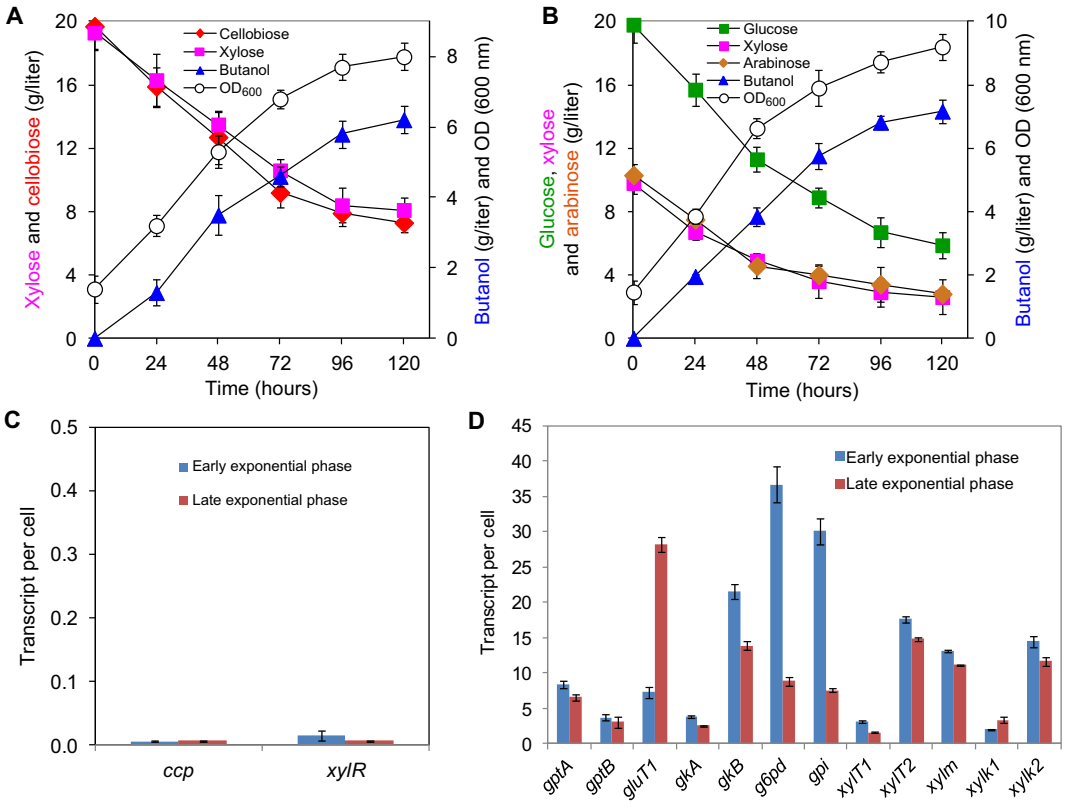

Fig. 3. Co-fermentation and resistance to CCR. (**A**) Co-fermentation of a sugar mixture containing xylose (18.7 g/liter) and cellobiose (18.3 g/liter). (**B**) Co-fermentation of a sugar mixture containing glucose (19.2 g/liter), xylose (8.8 g/liter), and arabinose (9.3 g/liter). Acetate and butyrate data are shown in the Supplementary Materials (see fig. S4). (**C**) Transcriptional levels of the *ccp* gene (encoding the catabolite control protein) and the *xylR* gene (encoding the transcriptional repressor of the xylose operon). (**D**) Comparison of transcriptional levels of genes related to the catabolism of glucose and xylose in strain TG57 at the early exponential phase (12 hours of fermentation) and the late exponential phase (36 hours of fermentation). Functional genes *gptA*, *gptB*, *gluT1*, *gkA*, *gkB*, *g6pd*, *gpi*, *xylT1*, *xylT2*, *xylm*, *xylk1*, and *xylk2* encode the following: glucose-specific phosphotransferase enzyme IIA components, PTS system glucose transporter subunit IIB, a sugar ABC transporter permease, the glucokinase ROK family, glucose-6P dehydrogenase, a glucose-6-phosphate isomerase, a xylose transporter, a xylose isomerase, and xylulokinase. Error bars indicate the SDs calculated from triplicates.

transcribed β-glucosidase responsible for the breakdown of oligosaccharides. In addition to the abovementioned two cellulases and eight glucosidases, cellulosome in strain TG57 contains a number of scaffoldins such as eight S-layer homologies (SLHs), two CBMs (carbohydrate-binding modules), six putative dockerins, and seven putative cohesins according to definitions described previously (*42*). The two highly transcribed genes among SLHs, Thert_00292 and Thert_00726, could secrete a protein coat that covers the cell surface for adhesion and mechanical and osmotic stabilization, and as a molecular sieve and anchor (*27*). The high expression of Thert_00292, which encodes an S-layer in bacteria, suggests that the S-layer may provide anchors for cell surface proteins such as plant degradation enzymes.

The tryptophan biosynthesis genes (Thert_00012, Thert_00013, Thert_03513, and Thert_03514) are up-regulated with cellulose as a substrate. The highly transcribed tryptophan synthesis genes suggest that L-tryptophan may serve as a key precursor of reducing cofactors to enhance the production efficacy of butanol from cellulose. The availability and reducing equivalent of cofactors (for example, NADH and NADPH) may play a crucial role in determining the overall yield in biofuel bioprocesses with cofactor-dependent production systems (*43*–*45*); that is, butanol biosynthesis via Bdh encoded by the *bdh* gene (Thert_02648) requires reducing cofactor NADH (Fig. 4D). L-Tryptophan, a precursor in de novo synthesis of NADH and NADPH, induces redox modulation by the highly transcribed tryptophan biosynthesis genes and triggers the availabilities of NADH and NADPH. Therefore, the NADH-dependent Bdh activities may be enhanced by triggering de novo synthesis of cofactors via supplementation of the precursor L-tryptophan, demonstrating that salvaging cofactors (for example, NADH and NADPH) is a major approach to elevating overexpression of the *bdh* gene and, thus, to redistributing the metabolic flux from the bioacid synthesis pathway to the central NAD/(P)H-dependent butanol pathway (*45*).

Genes involved in distinct central metabolic pathways and their relative transcript abundances

The genes involved in cellulose metabolism pathways in *T. thermosaccharolyticum* TG57 were analyzed on the basis of its genome sequence. The most distinctive feature of strain TG57 is its ability to grow efficiently on microcrystalline cellulose and to produce butanol as the sole solventogenic product, which occurs through cellulose hydrolysis and assimilation, the PPP, glycolysis, and butanologenesis, different from those found in solventogenic *Clostridium* species (Fig. 5). First, the genome of strain TG57 does not contain the genetic determinants for the biosynthesis of acetone (two key genes—*adc*, which encodes acetoacetate decarboxylase, and *ctf*, which encodes coenzyme-A transferase—were absent), which is evidenced by the absence of acetone in the fermentation broth as well. This finding suggests greatly simplified downstream extraction of butanol from the fermentation broth, thus enhancing the

A

	ATCC 824	NCIMB 8052	MF28	DSM 571	TG57
Size (base pairs, bp)	3,940,880	6,000,632	6,148,385	2,785,752	2,895,726
G+C content (%)	30.9	29.9	29.92	34.0	34.11
rRNAs	12	43	54	12	10
tRNAs	73	94	95	57	57
Predicted CDSs	3,602	5,023	5,560	2,601	2,986
Predicted *cel/cbh* genes	2/0	0/0	0/0	0/0	2/3
Predicted *bdh/adh* genes	3/3	2/10	4/9	0/9	1/7
Predicted *adc* genes	1	1	1	0	0

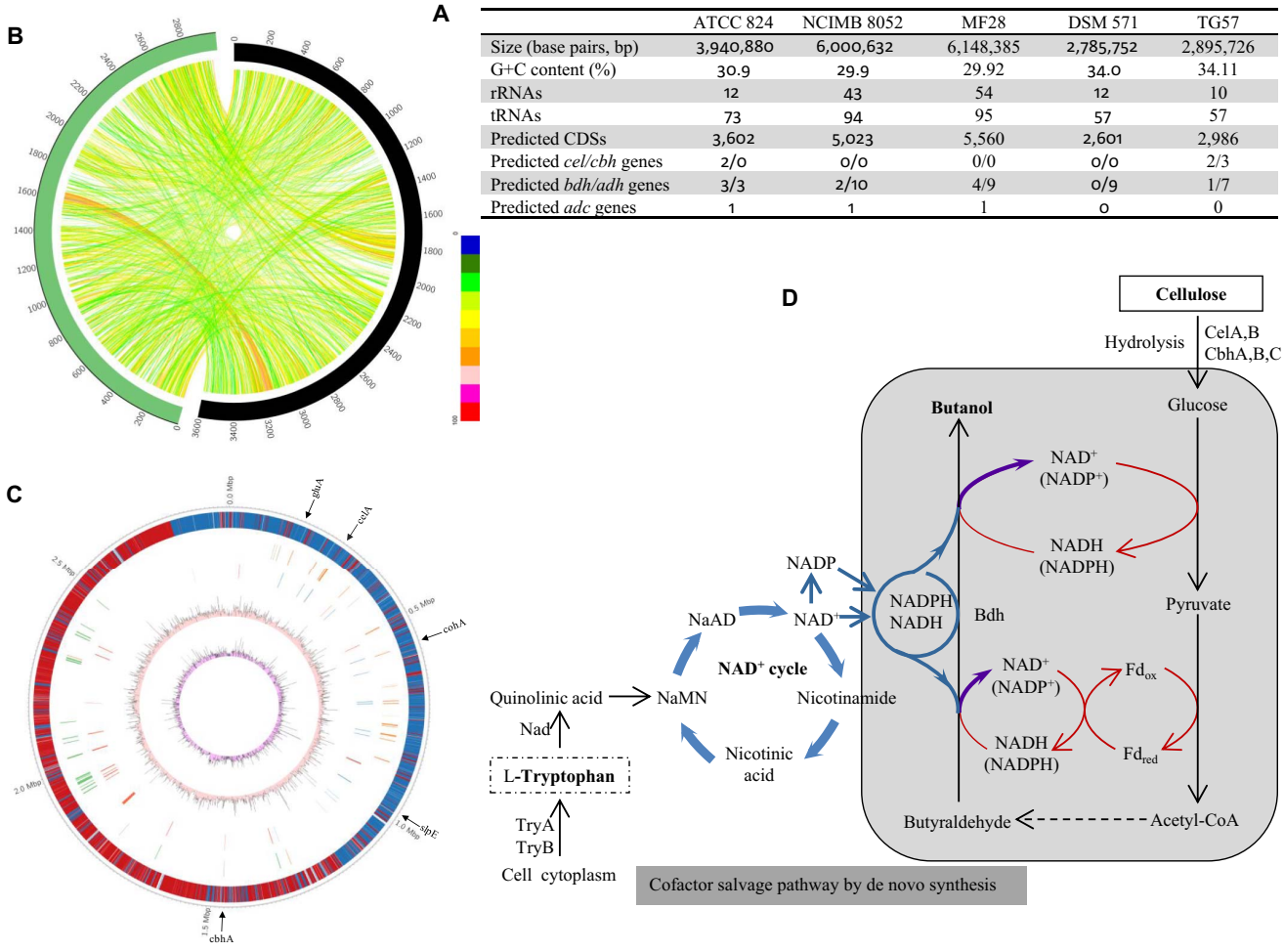

Fig. 4. Genomic and transcriptomic analysis of *T. thermosaccharolyticum* strain TG57. (**A**) Genome overview of strain TG57 with previously genome-sequenced isolates [*C. acetobutylicum* ATCC 824 (accession no. AE001437), *C. beijerinckii* NCIMB 8052 (accession no. CP000721), *Clostridium* sp. MF28 (accession no. CP014331), and *T. thermosaccharolyticum* DSM 571 (accession no. CP002171)] deposited in the GenBank database. (**B**) Circos plot represents similarity between strain TG57 and *C. acetobutylicum* ATCC 824 based on gene sequence and biological annotation. Track (green) on the left, strain TG57; track (black) on the right, strain ATCC 824. Number represents CDS. The links are color-coded by similarity. Red, 100% similarity; blue, 0% similarity. (**C**) Transcriptomic analysis of strain TG57. Tracks from the inside to outside: tracks 1 and 2, the transcription levels of all genes in culture on xylan and cellulose; track 3, cellulosome genes (red, forward strand; blue, reverse strand); track 4, genes for carbohydrates in glycoside hydrolase family (green, forward strand; orange, reverse strand); track 5, CDS (red, forward strand; blue, reverse strand); track 6, genome sequence of *T. thermosaccharolyticum* TG57. (**D**) Butanol synthesis and NADH (reduced form of nicotinamide adenine dinucleotide)/NADPH (reduced form of nicotinamide adenine dinucleotide phosphate) regenerations in *T. thermosaccharolyticum* TG57.

overall economic efficacy for butanol generation as compared to the typical acetone-butanol-ethanol fermentation by prominent solventogenic *Clostridium* species (*21*).

The second feature in the metabolic pathways of strain TG57 is its distinct butanoate metabolism. By overviewing butanol synthesis pathways reported previously, there are four main pathways known for butyrate production, the butyryl-CoA, glutarate, 4-aminobutyrate, and lysine pathways (*46*). The main solventogenic *C. acetobutylicum* and *Clostridium beijerinckii* typically go through the Ptb-Bk-Ctf complex for butyrate to the butyrl-CoA pathway [*ptb* encoding phosphorylation of butyryl-CoA (Ptb) and *bk* encoding butyrate kinase (Bk), followed by *ctfAB* encoding butyryl-CoA/acetate CoA transferase (Ctf)] (*47*), whereas strain TG57 contains atypical transferase where the critical intermediate butyryl-CoA is reversibly converted to butyrate by Ach only (Thert_02978, encoding acetyl-CoA/butyryl-CoA hydrolase) as a terminal enzyme. However, the *ach* gene is not present in the genome of *Clostridium* strains; thus, the butyrate pathway in strain TG57 is different from that in *Clostridium*. It was surprising to find the absence of the *ptb-bk-ctf* genetic complex in the genome of strain TG57. Accordingly, the final step from butyryl-CoA to butanol is catalyzed either by an Ach-Ald1-Bdh pathway or by a mainstream central pathway via Acat-Hbd-Crt-Bcd-Ald1-Bdh (Fig. 5).

In addition, genes in strain TG57 for cellulose hydrolysis and assimilation, the PPP, glycolysis, and subsequent butanologenesis were present in isoforms (that is, the same enzyme is coded by different coding DNA sequences and the transcription variance of these genes' mRNA) (Fig. 5 and data set S1). Transcripts of *celA* and *celB* encoding cellulase for cellulose hydrolysis and those of *xynA* and *xynB* encoding xylanase that is responsible for xylan hydrolysis are highly induced when strain TG57 grows on cellulose and xylan (Fig. 5). Genes for cellobiose and xylobiose hydrolysis, for example, *gluA*, *xysA*, and *xysB*, are moderately induced on cellulose and xylan, whereas *gluB* is not. Transcripts for *gltA* and *gltB* were all induced on cellulose, but *ptsA* and *ptsB* were in low transcripts (Fig. 5), suggesting that strain TG57 appears to prefer the

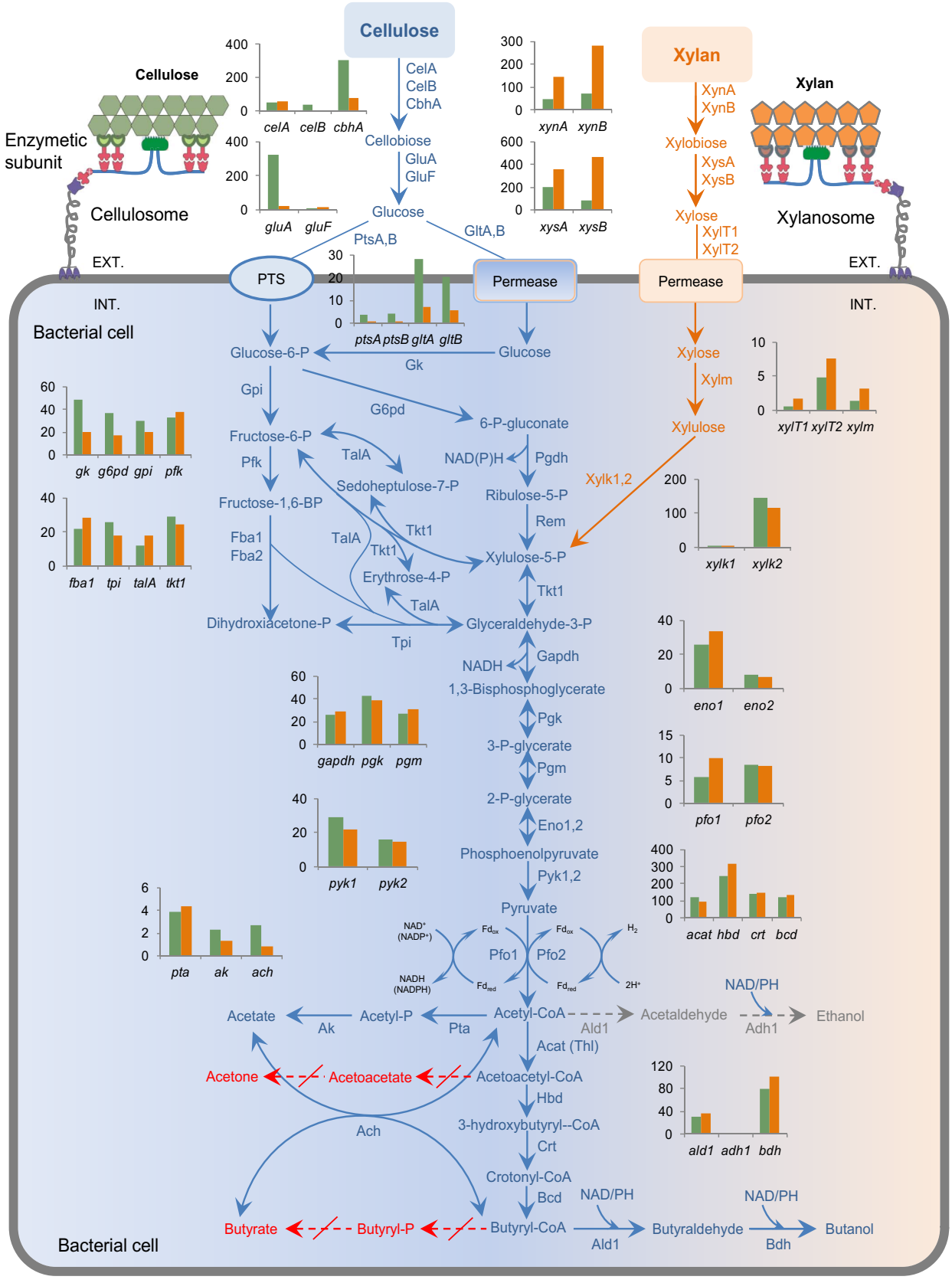

Fig. 5. Distinct central metabolic pathways and relative abundances of transcripts in cellulose and xylan cultures. Solid arrows indicate genes for the reaction. Gray dotted arrows and gray letters indicate genes for the inactive reaction. Red dotted arrows and red letters indicate the absence of genes and reactions in the metabolic pathway.

permease route for sugar uptake to the phosphotransferase system (PTS) regardless of substrate type. Transketolase (Tkt1) in the PPP is induced as one of the top transcripts relative to the genome-wide average in cells grown on cellulose, exhibiting its active metabolic role in linking the glycolysis pathway and PPP. Transcripts for *gk*, *g6pd*, *gpi*, *pfk*, and *tpi* were all induced on cellulose but were low transcripts under xylan conditions (Fig. 5). Genes for xylose uptake (*xylT2*, *xylm*, and *xylk2*) could be transcribed in the presence of glucose, reflecting CCR elimination. Glyceraldehyde-3-phosphate dehydrogenase (Gapdh), the first node for NADH generation, was induced by the presence of cellulose or xylan. Similarly, the transcript for pyruvate/ferredoxin oxidoreductase (Pfo1 and Pfo2), which generates the major reducing cofactors NAD(P)H in the NAD(P)H-linked ferredoxin oxidoreductase reactive cycle, was induced during cultivation on cellulose or xylan.

Although genetic determinants for ethanol synthesis are present in the genome of strain TG57, the *adh1* gene (49.1 and 33.3% similarities to those present in *C. acetobutylicum* ATCC 824 and *C. acetobutylicum* NCIMB 8052 strains) encoding alcohol dehydrogenase (Adh) was not transcribed in the presence of cellulose, xylan, glucose, and xylose. In addition, the independent measurements of enzyme activities in strain TG57 evidenced that the iron-containing Adh was not active on aldehydes (acetaldehyde and butyraldehyde), although the aldehyde dehydrogenase (Ald) was only active on butyryl-CoA and not on acetyl-CoA. It implies that Adh in *T. thermosaccharolyticum* strain TG57 can be different from a bifunctional aldehyde-alcohol dehydrogenase (Aad) gene associated with the formation of alcohols including butanol and ethanol in solventogenic microbes such as *C. acetobutylicum* and *C. beijerinckii* (*48*, *49*). Accordingly, during cellulose and xylan fermentation, butanol

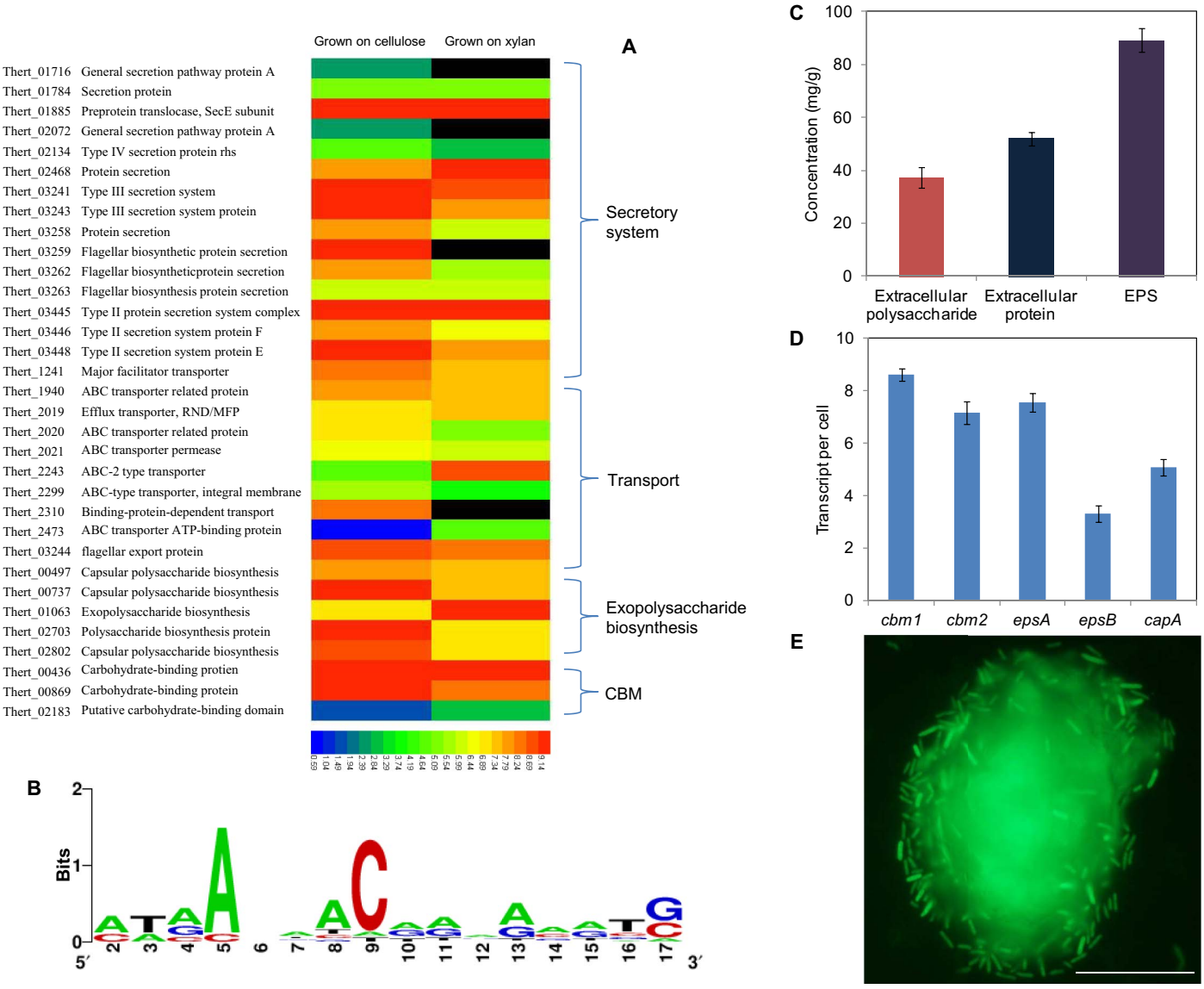

Fig. 6. The secretome and transportome features in *T. thermosaccharolyticum* strain TG57. (**A**) Relative abundances of transcripts of the secretory system, transport and efflux, exopolysaccharide, and carbohydrate-binding protein for biofilm formation. (**B**) Consensus sequences of signal peptide motifs for the secretory system. (**C**) Secreted exopolysaccharide, protein, and extracellular polymeric substance (EPS) contents in culture TG57. (**D**) qPCR analysis of carbohydrate-binding modules and exopolysaccharide biosynthesis proteins in culture TG57. (**E**) Morphology of a fluorescence micrograph showing the adhesion of cells to a crystalline cellulose particle (scale bar, 20 μm).

was detected to be the sole solventogenic product in the absence of ethanol (Fig. 1). In subsequent central butanol synthesis reactions, the related primary genes *acat-hbd-crt-bcd-ald1-bdh* are strongly induced and are the most abundant transcripts in the cells fed with cellulose or xylan, whereas the bioacids synthesis genes *pta*, *ak*, and *ach* showed weak transcripts.

Secretome and transportome (transport systems)

The secretion mechanisms of strain TG57 are of particular interest owing to the effective hydrolysis of insoluble microcrystalline cellulose to fermentative sugars by requiring extracellular hydrolases secretion before producing butanol. Although the genome of strain TG57 contains a number of secretion genes, the Sec system, T2SS, and T3SS act as the primary means of secretion with significant expressions when the cells were fed with cellulose or xylan (Fig. 6A). Determining strain TG57's secretion signals can probe and regulate its secretion system and predict signal motifs. Type I signal motif (EGF_1, EGF-like domain signature 1) and type II signal motif (2Fe2S_FER_1, 2Fe-2S ferredoxin-type iron-sulfur binding region signature) are mainly found in strain TG57's secretion patterns, which contain the conserved cysteine residues for exopolysaccharide. The significant positions and subtle patterns were found in strain TG57's secretion system (Fig. 6B). For instance, the decrease in the height of the first base of the pattern indicates the presence of alternative bases, exhibiting the importance of alternative varying bases in the binding site. It is noted that spaces are significant to aid in the identification of the site being considered. Sec-, T2SS-, and T3SS-dependent secretion correlates with the possibility of many highly expressed carbohydratases (Figs. 4C and 6A). Secretion system determinants act primarily in carbohydrate degradation, cell surface and flagellar assembly operon, and efflux transport systems (Fig. 6A). Because butanol accumulation is known to be highly toxic to the cell (*50*, *51*), feedback regulation of a positively transcribed efflux pump enhanced butanol tolerance to relieve toxicity by way of secretion of the product butanol (*52*).

Adhesion to plant biomass, such as biofilm, is a fundamental means of stress adaptation by cellulolytic microorganisms to elevate cellulolysis (*53*) by secreting enzymes concentrating around cellulosic substrates. The genome of *T. thermosaccharolyticum* strain TG57 contains a large number of transcribed genes for biosynthesis, export, and secretion of extracellular polymeric substances and highly transcribed CBMs putatively involved in the formation of biofilm in addition to type II and type IV secretory mechanisms mediating the formation of biofilms (Fig. 6, C to E).

DISCUSSION

The newly discovered wild-type cellulolytic *T. thermosaccharolyticum* TG57 reveals novel characteristics to convert microcrystalline cellulose to butanol: (i) generating butanol up to 1.93 g/liter with a yield of 0.20 g/g; (ii) simultaneous utilization of hexose and pentose to produce butanol (7.33 g/liter) from glucose, xylose, and arabinose mixtures (Figs. 1 and 3); and (iii) lack of by-product acetone from using cellulose. Among the few available reports on biofuel (for example, butanol and ethanol) production from lignocelluloses, strain TG57 represents the first wild-type strain that can ferment microcrystalline cellulose to butanol (1.93 g/liter), comparable to previously reported wild-type [ethanol (0.46 g/liter)] or gene-modified [butanol (1.42 g/liter)] microbes (Table 1). Collectively, the metabolic characteristic of strain TG57 strengthens the economic feasibility of butanol generation on the basis of lowering substrate costs and simplifying downstream product extraction complexities.

Both genomic and transcriptomic characterizations of *T. thermosaccharolyticum* strain TG57 provide a deep insight into the following: (i) the unusual metabolic characteristics of its genome; (ii) silencing in a CCR-encoding *ccp* gene to eliminate catabolite repression (for example, negligible transcription of the *ccp* gene coding a regulator that mediates high catabolite repression, leading to higher expressions of the *xylm* and *xylk* genes responsible for xylose utilization); (iii) a streamlined genome with a compact butonate pathway, a shortage of transposons, and a lack of genes encoding acetone bypass synthesis, while having plenty of determinants responsible for central metabolic functions with wide carbohydrate assimilation and efficient butanol production capabilities; and (iv) active transcription variations in diverse aspects of metabolism including up-regulated tryptophan synthesis as well as efflux and secretory systems, but down-regulated fatty acid synthesis, suggesting multiple ways of optimizing cellulosic fermentation by dosing growth supplements (for example, L-tryptophan).

The genome of *T. thermosaccharolyticum* strain TG57 is of potential biotechnological interest because it harbors many of the genes responsible for metabolic processes such as 7 haloacid dehalogenase genes; 11 genes for biotin, thiamine, cobalamin, and riboflavin biosynthesis; 10 various cytochrome genes that function as carriers of electrons; and 37 genes for oxidoreductases such as the NADH/flavin family, the Fe-S family, and related transcriptional regulators and transporters with clustered forms similar to those described previously (*54*). Collectively, *T. thermosaccharolyticum* TG57 extends a biotechnological potential beyond lignocellulosic biofuel. Production of butanol from cellulose exhibits a foundational milestone to realize the ultimate goal of cost-effective production of renewable biofuels and chemicals from lignocellulosic biomass in a consolidated bioprocess.

MATERIALS AND METHODS

Culture medium and growth conditions

Anaerobic medium BTM1 was reduced by L-cysteine (0.2 mM), $Na_2S{\cdot}9H_2O$ (0.2 mM), and DL-dithiothreitol (0.5 mM) and buffered by MES [2-(*N*-morpholino)ethanesulfonic acid; 20 mM], which contained $MgCl_2{\cdot}6H_2O$ (0.2 g/liter), $K_2HPO_4{\cdot}3H_2O$ (2.9 g/liter), $CaCl_2{\cdot}2H_2O$ (0.1 g/liter), NH_4Cl (1.3 g/liter), KH_2PO_4 (1.5 g/liter), $FeSO_4{\cdot}7H_2O$ (1.25 mg/liter), and yeast extract (3 g/liter), plus 0.1 ml of resazurin solution, 1 ml of trace element mixture, and 1 ml of selenite-tungstate solution, as described previously (*55*). The medium pH was adjusted to 6.0 and sterilized by autoclaving. Agar plates (solidified with 2% agar) were prepared with the same medium BTM1 but contained 1% (w/v) microcrystalline cellulose as a sole carbon source. Microcrystalline cellulose (Sigmacell 50) and other chemicals were at least analytical grade and were purchased from Sigma-Aldrich, unless indicated otherwise. All cultures including controls without inocula were incubated at a temperature of 53°C with a shaking speed of 100 rpm.

Enrichment and isolation of butanologenic microorganisms capable of assimilating microcrystalline cellulose

Seeds used for screening butanol-producing strains were 2-year-old spent mushroom substrates provided by a local mushroom farm (MycoFarm). Enrichment of cellulolytic butanologenic microbes exhibiting CCR was performed as previously described (*41*) with slight modifications. First, the cultures were enriched in 50 ml of anaerobic BTM1 media supplemented with microcrystalline cellulose (3%, w/v)

as the sole carbon source. The cultures were incubated for 8 days at 53°C in an incubator shaker (100 rpm). After three generations, the culture was measured for its butanol production capability.

To select a strain without showing any effect of catabolite repression by glucose, we added 2-DG (1.0 g/liter) to BTM1 medium containing microcrystalline cellulose, and the cultures were incubated for 3 days. After three transfers into the fresh BTM1 medium, cultures capable of producing butanol successively were transferred to agar plates and incubated in an anaerobic chamber for 3 days, and then colonies were picked, streaked, and transferred to medium bottles for further tests.

Fermentation kinetics studies

Time courses of *T. thermosaccharolyticum* strain TG57 fermentation using microcrystalline cellulose, xylan, cellobiose, or glucose as a carbon source were carried out in a bioreactor (Sartorius). The bioreactor has a working volume of 1.5 liters, and it was operated under an agitation speed of 100 rpm at 53°C. For the inoculation process, 90 ml of a seed culture [optical density (OD) ~2] was prepared first. The pH dropped from 6.2 to 5.5 when the culture grew. Then, 6 M NaOH or 3 M H_2SO_4 was supplemented to adjust the pH when it was below 5.5. All experiments in this study were carried out in triplicate.

Analytical techniques

Cell density was determined by a Biospec-1601 spectrophotometer (Shimadzu). Cell culture was diluted as necessary when the OD_{600} (optical density at 600 nm) was above 1.0. For the culture medium spiked with crystalline cellulose, total cell mass was determined indirectly by using a DC protein assay kit (BioRad). Cell protein was converted to OD_{600} [OD_{600}:protein (g/liter) = 1:0.126]. Butanol and other fermentation products were quantified with a gas chromatograph following the method previously described (*55*). Residual cellulose and xylan were determined as described previously (*23*, *56*). Cellobiose and other sugars were analyzed by a high-performance liquid chromatography system equipped with a Phenomenex carbohydrate column (*41*). Adhesion of cells to the microcrystalline cellulose was observed using fluorescence microscopy after staining the cells with a SYTO 9 green molecular probe (Life Technologies). The elemental composition of the microbial cell for carbon balance analysis was assumed to be $C_5H_7NO_2$ (*57*).

Enzymatic analyses

Cells were collected from 10 ml of culture by centrifugation (10,000*g*, 5 min, 4°C), and the resultant supernatant was determined for extracellular enzyme activity. Crude cell extracts were obtained from the resultant pellets to assay for the involved Adh, Ald, and Bdh (*45*). All enzymatic assays were carried out in an anaerobic chamber. Cellulase activity was assayed by measuring the reducing sugar released from the enzymatic hydrolysis of microcrystalline. The reaction mixture contained 0.25 ml of supernatant sample, 1.25 ml of 1% cellulose, and 0.5 ml of 0.2 M acetate buffer (pH 5.5). After incubation at 53°C for 120 min, the liberated reducing sugars were measured with the dinitrosalicylic acid method at 540 nm. One unit of cellulase activity was expressed as the amount of enzyme required to release 1 μmol of glucose per minute under the assay conditions. β-Glucosidase activity was assayed using *p*-nitrophenyl-β-D-glucopyranoside as a substrate, as described previously (*58*). Bdh, Ald, and Adh activities were determined as described previously (*55*, *59*, *60*). Specific activities were calculated on the basis of the total protein in the crude extract as described above. All samples were assayed in triplicate.

Molecular analyses and reverse transcription qPCR

Extractions of DNA and RNA for quantitative polymerase chain reaction (qPCR) and genomic analyses from 10 ml of cellulose- and xylan-fed cultures were performed as described previously (*41*, *45*). Clone libraries of the 16*S* rRNA gene were constructed using the TOPO TA Cloning Kit (Invitrogen). To determine the 16*S* rRNA gene copy number of the genome of *T. thermosaccharolyticum* strain TG57, we carried out qPCR (Applied Biosystems) with a universal bacterial primer pair (338qF, 5′-ACTCCTACGGGAGGCAGCAG-3′; 518qR, 5′-ATTACCGCGGCTGCTGG-3′) and two pairs of strain TG57–specific *rpoA* and *atpD* primers (data set S2). Reverse transcription (RT) was conducted using the ABI High Capacity cDNA RT Kit with a random hexamer primer (Promega) and RNasin (RNase inhibitor; Promega) (*41*, *45*). qPCR enumeration of *T. thermosaccharolyticum* strain TG57 cells and genes was carried out using the Bioline SensiFAST SYBR Lo-ROX Kit (*45*). Luciferase cDNA copies were quantified using gene-specific primers LucF (5′-TACAACACCCCAACATCTTCGA-3′) and LucR (5′-GGAAGTTCACCGGCGTCAT-3′). Transcriptional expression of individual genes was quantified using gene-specific primers (data set S2) and normalized by losses of luciferase mRNA and the corresponding cell numbers at each time point.

Genome sequencing, assembly, and annotation

Genomic DNA was extracted from cellulose-fed culture. DNA sequencing libraries were then prepared with the SMRTbell Template Prep Kit following the manufacturer's instructions (Pacific Biosciences). Small fragments (lower than 20 kb of the SMRTbell template) were removed using the BluePippin Size Selection System for large-insert libraries. The SMRTbell library was first sequenced using 1 SMRT cell (Pacific Biosciences) and C4 chemistry (DNA sequencing Reagent Kit 4.0). A 240-min movie was then made for each SMRT cell using the PacBio RS II sequencing platform. De novo assembly was conducted by using the hierarchical genome assembly process (version 2.3) workflow and consensus polishing with Quiver (*61*). Reads were downsampled to around 318× coverage to improve assembly statistics (N50). Finally, we checked the form for each contig using MUMmer (*62*) and trimmed the ones with self-similar ends for the closed circular genome. Putative gene CDSs were identified by Glimmer (*63*). Functional annotations were assigned by blasting predicted CDSs with Blastall alignment against the NCBI (National Center for Biotechnology Information) Non-Redundant Database (nr), Pfam, COG, and Prosite public databases for all species. Noncoding genes for rRNAs and tRNAs were predicted by RNAmmer (*64*) and tRNAscan-SE (*65*), respectively.

RNA sequencing

Total RNA was extracted from cellulose- and xylan-fed cultures, and RNA integrity for transcriptomic analyses was assessed using Bioanalyzer (Agilent). The strand-specific RNA sequencing (RNA-Seq) library construction was conducted by using a commercial TruSeq Stranded mRNA LT Sample Prep Kit (Illumina) according to the manufacturer's protocol. During the library construction, the enriched mRNA was fragmented to obtain inserts with a fragment size of 120 to 200 bp (a median size of 150 bp). The quality of the library (a single peak in the region of ~232 to 299 bp) was checked using Agilent D1000 ScreenTape. Sequencing was carried out using the Illumina Miseq Sequencers with the Illumina Miseq Reagent v2 (50 cycle kit) Kit to generate a single-read, 51-bp reads run. The pass filter reads generated from the sequencing run were ~17.8 million. The resulting RNA-Seq

reads were uniquely mapped to the *T. thermosaccharolyticum* genome using TopHat (*66*), and the differential expression levels of genes and transcript abundances were quantified by the Cuffdiff program (*67*, *68*) as reads per kilobase pair per million. The reads from rRNA and tRNA regions were masked in the cufflinks analysis.

Data availability

The complete genome sequence of *T. thermosaccharolyticum* strain TG57 has been deposited in GenBank under accession number CP016893. RNA-Seq data are available from the corresponding author upon request.

SUPPLEMENTARY MATERIALS

Supplementary material for this article is available at http://advances.sciencemag.org/cgi/content/full/4/3/e1701475/DC1

section S1. Taxonomic identification and phylogenetic analysis of strain TG57

section S2. Confirmation of culture purity of *T. thermosaccharolyticum* strain TG57

section S3. Carbohydratases

fig. S1. Neighbor-joining tree of the housekeeping gene sequences after global alignment.

fig. S2. Relative growth of culture TG57 when fed on cellulose at various temperatures and pHs.

fig. S3. Transcription levels of genes for polysaccharide and other carbon utilization in *T. thermosaccharolyticum* TG57.

fig. S4. Bioacid (acetic acid and butyric acid) production in the co-fermentation broth by *T. thermosaccharolyticum* TG57.

fig. S5. Comparative genomic characterization of *T. thermosaccharolyticum* TG57 with previously sequenced isolates.

table S1. Specific activities of functional enzymes involved in *T. thermosaccharolyticum* TG57.

table S2. Carbon balance of fermenting different carbohydrates by *T. thermosaccharolyticum* TG57.

table S3. Comparison of glycoside hydrolases found in *T. thermosaccharolyticum* TG57 with *T. thermosaccharolyticum* DSM 571.

data set S1. Genes in strain TG57 were present in isoforms.

data set S2. Gene-specific primers used for RT-qPCR.

REFERENCES AND NOTES

1. A. Morone, R. A. Pandey, Lignocellulosic biobutanol production: Gridlocks and potential remedies. *Renew. Sust. Energ. Rev.* **37**, 21–35 (2014).
2. C. Xue, X. Q. Zhao, C. G. Liu, L. J. Chen, F. W. Bai, Prospective and development of butanol as an advanced biofuel. *Biotechnol. Adv.* **31**, 1575–1584 (2013).
3. R. Cascone, Biobutanol—A replacement for bioethanol? *Chem. Eng. Prog.* **104**, S4–S9 (2008).
4. S. Y. Lee, J. H. Park, S. H. Jang, L. K. Nielsen, J. Kim, K. S. Jung, Fermentative butanol production by clostridia. *Biotechnol. Bioeng.* **101**, 209–228 (2008).
5. S. B. Leschine, Cellulose degradation in anaerobic environments. *Annu. Rev. Microbiol.* **49**, 399–426 (1995).
6. L. J. Jönsson, B. Alriksson, N.-O. Nilvebrant, Bioconversion of lignocellulose: Inhibitors and detoxification. *Biotechnol. Biofuels* **6**, 16 (2013).
7. Q. Xu, M. G. Resch, K. Podkaminer, S. Yang, J. O. Baker, B. S. Donohoe, C. Wilson, D. M. Klingeman, D. G. Olson, S. R. Decker, R. J. Giannone, R. L. Hettich, S. D. Brown, L. R. Lynd, E. A. Bayer, M. E. Himmel, Y. J. Bomble, Dramatic performance of *Clostridium thermocellum* explained by its wide range of cellulase modalities. *Sci. Adv.* **2**, e1501254 (2016).
8. S. P. S. Chundawat, G. T. Beckham, M. E. Himmel, B. E. Dale, Deconstruction of lignocellulosic biomass to fuels and chemicals. *Annu. Rev. Chem. Biomol. Eng.* **2**, 121–145 (2011).
9. D. G. Olson, J. E. McBride, A. J. Shaw, L. R. Lynd, Recent progress in consolidated bioprocessing. *Curr. Opin. Biotechnol.* **23**, 396–405 (2012).
10. A. L. Demain, Biosolutions to the energy problem. *J. Ind. Microbiol. Biotechnol.* **36**, 319–332 (2009).
11. J. C. Liao, L. Mi, S. Pontrelli, S. Luo, Fuelling the future: Microbial engineering for the production of sustainable biofuels. *Nat. Rev. Microbiol.* **14**, 288–304 (2016).
12. C. Weber, A. Farwick, F. Benisch, D. Brat, H. Dietz, T. Subtil, E. Boles, Trends and challenges in the microbial production of lignocellulosic bioalcohol fuels. *Appl. Microbiol. Biotechnol.* **87**, 1303–1315 (2010).
13. L. R. Lynd, A. M. Guss, M. E. Himmel, D. Beri, C. Herring, E. K. Holwerda, S. J. L. Murphy, D. G. Olson, J. Paye, T. Rydzak, X. Shao, L. Tian, R. Worthen, Advances in consolidated bioprocessing using *Clostridium thermocellum* and *Thermoanaerobacter saccharolyitcum*, in *Industrial Biotechnology: Microorganisms*, C. Wittmand, J. C. Liao, Eds. (Wiley-VCH Verlag GmbH & Co. KGaA, 2016), pp. 365–394.
14. J. Wang, X. Yang, C.-C. Chen, S.-T. Yang, Engineering clostridia for butanol production from biorenewable resources: From cells to process integration. *Curr. Opin. Chem. Eng.* **6**, 43–54 (2014).
15. D. Chung, M. Cha, A. M. Guss, J. Westpheling, Direct conversion of plant biomass to ethanol by engineered *Caldicellulosiruptor bescii*. *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8931–8936 (2014).
16. L. Tian, B. Papanek, D. G. Olson, T. Rydzak, E. K. Holwerda, T. Zheng, J. Zhou, M. Maloney, N. Jiang, R. J. Giannone, R. L. Hettich, A. M. Guss, L. R. Lynd, Simultaneous achievement of high ethanol yield and titer in *Clostridium thermocellum*. *Biotechnol. Biofuels* **9**, 116 (2016).
17. W. Higashide, Y. Li, Y. Yang, J. C. Liao, Metabolic engineering of *Clostridium cellulolyticum* for production of isobutanol from cellulose. *Appl. Environ. Microbiol.* **77**, 2727–2733 (2011).
18. P. P. Lin, L. Mi, A. H. Morioka, K. M. Yoshino, S. Konishi, S. C. Xu, B. A. Papanek, L. A. Riley, A. M. Guss, J. C. Liao, Consolidated bioprocessing of cellulose to isobutanol using *Clostridium thermocellum*. *Metab. Eng.* **31**, 44–52 (2015).
19. K. Kovács, B. J. Willson, K. Schwarz, J. T. Heap, A. Jackson, D. N. Bolam, K. Winzer, N. P. Minton, Secretion and assembly of functional mini-cellulosomes from synthetic chromosomal operons in *Clostridium acetobutylicum* ATCC 824. *Biotechnol. Biofuels* **6**, 117 (2013).
20. L. Yu, M. Xu, I.-C. Tang, S.-T. Yang, Metabolic engineering of *Clostridium tyrobutyricum* for n-butanol production through co-utilization of glucose and xylose. *Biotechnol. Bioeng.* **112**, 2134–2141 (2015).
21. Y. Gu, Y. Jiang, H. Wu, X. Liu, Z. Li, J. Li, H. Xiao, Z. Shen, H. Dong, Y. Yang, Y. Li, W. Jiang, S. Yang, Economical challenges to microbial producers of butanol: Feedstock, butanol ratio and titer. *Biotechnol. J.* **6**, 1348–1357 (2011).
22. C. L. Hemme, H. Mouttaki, Y.-J. Lee, G. Zhang, L. Goodwin, S. Lucas, A. Copeland, A. Lapidus, T. Glavina del Rio, H. Tice, E. Saunders, T. Brettin, J. C. Detter, C. S. Han, S. Pitluck, M. L. Land, L. J. Hauser, N. Kyrpides, N. Mikhailova, Z. He, L. Wu, J. D. Van Nostrand, B. Henrissat, Q. He, P. A. Lawson, R. S. Tanner, L. R. Lynd, J. Wiegel, M. W. Fields, A. P. Arkin, C. W. Schadt, B. S. Stevenson, M. J. Mclnerney, Y. Yang, H. Dong, D. Xing, N. Ren, A. Wang, R. L. Huhnke, J. R. Mielenz, S.-Y. Ding, M. E. Himmel, S. Taghavi, D. van der Lelie, E. M. Rubin, J. Zhou, Sequencing of multiple clostridial genomes related to biomass conversion and biofuel production. *J. Bacteriol.* **192**, 6494–6496 (2010).
23. X. Yang, M. Xu, S.-T. Yang, Metabolic and process engineering of *Clostridium cellulovorans* for biofuel production from cellulose. *Metab. Eng.* **32**, 39–48 (2015).
24. Z. Wang, G. Cao, J. Zheng, D. Fu, J. Song, J. Zhang, L. Zhao, Q. Yang, Developing a mesophilic co-culture for direct conversion of cellulose to butanol in consolidated bioprocess. *Biotechnol. Biofuels* **8**, 84 (2015).
25. S.-J. Ha, J. M. Galazka, S. R. Kim, J.-H. Choi, X. Yang, J.-H. Seo, N. L. Glass, J. H. D. Cate, Y.-S. Jin, Engineered *Saccharomyces cerevisiae* capable of simultaneous cellobiose and xylose fermentation. *Proc. Natl. Acad. Sci. U.S.A.* **108**, 504–509 (2011).
26. C. M. Payne, B. C. Knott, H. B. Mayes, H. Hansson, M. E. Himmel, M. Sandgren, J. Ståhlberg, G. T. Beckham, Fungal cellulases. *Chem. Rev.* **115**, 1308–1448 (2015).
27. A. C. Tolonen, W. Haas, A. C. Chilaka, J. Aach, S. P. Gygi, G. M. Church, Proteome-wide systems analysis of a cellulosic biofuel-producing microbe. *Mol. Syst. Biol.* **7**, 461 (2011).
28. S. Yanase, R. Yamada, S. Kaneko, H. Noda, T. Hasunuma, T. Tanaka, C. Ogino, H. Fukuda, A. Kondo, Ethanol production from cellulosic materials using cellulase-expressing yeast. *Biotechnol. J.* **5**, 449–455 (2010).
29. Y. Li, T. J. Tschaplinski, N. L. Engle, C. Y. Hamilton, M. Rodriguez Jr., J. C. Liao, C. W. Schadt, A. M. Guss, Y. Yang, D. E. Graham, Combined inactivation of the *Clostridium cellulolyticum* lactate and malate dehydrogenase genes substantially increases ethanol yield from cellulose and switchgrass fermentations. *Biotechnol. Biofuels* **4**, 5 (2012).
30. J. Lo, D. G. Olson, S. J. L. Murphy, L. Tian, S. Hon, A. Lanahan, A. M. Guss, L. R. Lynd, Engineering electron metabolism to increase ethanol production in *Clostridium thermocellum*. *Metab. Eng.* **39**, 71–79 (2017).
31. J. G. Gardner, D. H. Keating, Requirement of the type II secretion system for utilization of cellulosic substrates by *Cellvibrio japonicus*. *Appl. Environ. Microbiol.* **76**, 5079–5087 (2010).
32. R. den Haan, S. H. Rose, L. R. Lynd, W. H. van Zyl, Hydrolysis and fermentation of amorphous cellulose by recombinant *Saccharomyces cerevisiae*. *Metab. Eng.* **9**, 87–94 (2007).
33. M. Jin, C. Gunawan, V. Balan, B. E. Dale, Consolidated bioprocessing (CBP) of AFEX™-pretreated corn stover for ethanol production using *Clostridium phytofermentans* at a high solids loading. *Biotechnol. Bioeng.* **109**, 1929–1936 (2012).
34. Y. Liang, T. Si, E. L. Ang, H. Zhao, Engineered pentafunctional minicellulosome for simultaneous saccharification and ethanol fermentation in *Saccharomyces cerevisiae*. *Appl. Environ. Microbiol.* **80**, 6677–6684 (2014).
35. E. Guedon, M. Desvaux, H. Petitdemange, Improvement of cellulolytic properties of *Clostridium cellulolyticum* by metabolic engineering. *Appl. Environ. Microbiol.* **68**, 53–58 (2002).
36. H. Teugjas, P. Väljamäe, Product inhibition of cellulases studied with ^{14}C-labeled cellulose substrates. *Biotechnol. Biofuels* **6**, 104 (2013).

37. Z. Xiao, X. Zhang, D. J. Gregg, J. N. Saddler, Effects of sugar inhibition on cellulases and beta-glucosidase during enzymatic hydrolysis of softwood substrates. *Appl. Biochem. Biotechnol.* **113–116**, 1115–1126 (2004).
38. M. W. Lau, B. Bals, S. P. S. Chundawat, M. Jin, C. Gunawan, V. Balan, A. D. Jones, B. E. Dale, An integrated paradigm for cellulosic biorefineries: Utilization of lignocellulosic biomass as self-sufficient feedstocks for fuel, food precursors and saccharolytic enzyme production. *Energy Environ. Sci.* **5**, 7100–7110 (2012).
39. H. T. Kang, E. S. Hwang, 2-Deoxyglucose: An anticancer and antiviral therapeutic, but not any more a low glucose mimetic. *Life Sci.* **78**, 1392–1399 (2006).
40. A. N. Wick, D. R. Drury, H. I. Nakada, J. B. Wolfe, Localization of the primary metabolic block produced by 2-deoxyglucose. *J. Biol. Chem.* **224**, 963–969 (1957).
41. T. Li, J. He, Simultaneous saccharification and fermentation of hemicellulose to butanol by a non-sporulating *Clostridium* species. *Bioresour. Technol.* **219**, 430–438 (2016).
42. E. A. Bayer, R. Lamed, B. A. White, H. J. Flint, From cellulosomes to cellulosomics. *Chem. Rec.* **8**, 364–377 (2008).
43. A. Knepper, M. Schleicher, M. Klauke, D. Weuster-Botz, Enhancement of the NAD(P)(H) pool in *Saccharomyces cerevisiae*. *Eng. Life Sci.* **8**, 381–289 (2008).
44. T. Zheng, D. G. Olson, L. Tian, Y. J. Bomble, M. E. Himmel, J. Lo, S. Hon, A. J. Shaw, J. P. van Dijken, L. R. Lynd, Cofactor specificity of the bifunctional alcohol and aldehyde dehydrogenase (AdhE) in wild-type and mutant *Clostridium thermocellum* and *Thermoanaerobacterium saccharolyticum*. *J. Bacteriol.* **197**, 2610–2619 (2015).
45. T. Li, Y. Yan, J. He, Enhanced direct fermentation of cassava to butanol by *Clostridium* species strain BOH3 in cofactors-mediated medium. *Biotechnol. Biofuels* **8**, 166 (2015).
46. M. Vital, A. C. Howe, J. M. Tiedje, Revealing the bacterial butyrate synthesis pathways by analyzing (meta)genomic data. *mBio* **5**, e00889-14 (2014).
47. R. Gheshlaghi, J. M. Scharer, M. Moo-Young, C. P. Chou, Metabolic pathways of clostridia for producing butanol. *Biotechnol. Adv.* **27**, 764–781 (2009).
48. S. J. Choi, J. Lee, Y.-S. Jang, J. H. Park, S. Y. Lee, I. H. Kim, Effects of nutritional enrichment on the production of acetone-butanol-ethanol (ABE) by *Clostridium acetobutylicum*. *J. Microbiol.* **50**, 1063–1066 (2012).
49. T. Ezeji, N. Qureshi, H. P. Blaschek, Production of acetone–butanol–ethanol (ABE) in a continuous flow bioreactor using degermed corn and *Clostridium beijerinckii*. *Process Biochem.* **42**, 34–39 (2007).
50. S. Huffer, M. E. Clark, J. C. Ning, H. W. Blanch, D. S. Clark, Role of alcohols in growth, lipid composition, and membrane fluidity of yeasts, bacteria, and archaea. *Appl. Environ. Microbiol.* **77**, 6400–6408 (2011).
51. B. J. Rutherford, R. H. Dahl, R. E. Price, H. L. Szmidt, P. I. Benke, A. Mukhopadhyay, J. D. Keasling, Functional genomic study of exogenous *n*-butanol stress in *Escherichia coli*. *Appl. Environ. Microbiol.* **76**, 1935–1945 (2010).
52. S. Boyarskiy, S. D. López, N. Kong, D. Tullman-Ercek, Transcriptional feedback regulation of efflux protein expression for increased tolerance to and production of *n*-butanol. *Metab. Eng.* **33**, 130–137 (2016).
53. Y. Lu, Y.-H. Zhang, L. R. Lynd, Enzyme–microbe synergy during cellulose hydrolysis by *Clostridium thermocellum*. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 16165–16169 (2006).
54. S. Schneiker, V. A. Martins dos Santos, D. Bartels, T. Bekel, M. Brecht, J. Buhrmester, T. N. Chernikova, R. Denaro, M. Ferrer, C. Gertler, A. Goesmann, O. V. Golyshina, F. Kaminski, A. N. Khachane, S. Lang, B. Linke, A. C. McHardy, F. Meyer, T. Nechitaylo, A. Pühler, D. Regenhardt, O. Rupp, J. S. Sabirova, W. Selbitschka, M. M. Yakimov, K. N. Timmis, F. J. Vorhölter, S. Weidner, O. Kaiser, P. N. Golyshin, Genome sequence of the ubiquitous hydrocarbon-degrading marine bacterium *Alcanivorax borkumensis*. *Nat. Biotechnol.* **24**, 997–1004 (2006).
55. T. Li, Y. Yan, J. He, Reducing cofactors contribute to the increase of butanol production by a wild-type *Clostridium* sp. strain BOH3. *Bioresour. Technol.* **155**, 220–228 (2014).
56. E. J. Wolfrum, R. M. Ness, N. J. Nagle, D. J. Peterson, C. J. Scarlata, "Determination of Structural Carbohydrates and Lignin in Biomass" (Technical Report NREL/TP-510-42618, National Renewable Energy Laboratory, Golden, 2011), pp. 1–18.
57. W. W. Eckenfelder, D. J. O'Connor, *Biological Waste Treatment* (Elsevier, 2013), pp. 36–37.
58. N. J. Parry, D. E. Beever, E. Owen, I. Vandenberghe, J. Van Beeumen, M. K. Bhat, Biochemical characterization and mechanism of action of a thermostable β-glucosidase purified from *Thermoascus aurantiacus*. *Biochem. J.* **353**, 117–127 (2001).
59. P. Dürre, A. Kuhn, M. Gottward, G. Gottschalk, Enzymatic investigations on butanol dehydrogenase and butyraldehyde dehydrogenase in extracts of *Clostridium acetobutylicum*. *Appl. Microbiol. Biotechnol.* **26**, 268–272 (1987).
60. K. P. Stim-Herndon, R. Nair, E. T. Papoutsakis, G. N. Bennett, Analysis of degenerate variants of *Clostridium acetobutylicum* ATCC 824: Biotechnology/Food Microbiology. *Anaerobe* **2**, 11–18 (1996).
61. C.-S. Chin, D. H. Alexander, P. Marks, A. A. Klammer, J. Drake, C. Heiner, A. Clum, A. Copeland, J. Huddleston, E. E. Eichler, S. W. Turner, J. Korlach, Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat. Methods* **10**, 563–569 (2013).
62. S. Kurtz, A. Phillippy, A. L. Delcher, M. Smoot, M. Shumway, C. Antonescu, S. L. Salzberg, Versatile and open software for comparing large genomes. *Genome Biol.* **5**, R12 (2004).
63. A. L. Delcher, K. A. Bratke, E. C. Powers, S. L. Salzberg, Identifying bacterial genes and endosymbiont DNA with Glimmer. *Bioinformatics* **23**, 673–679 (2007).
64. K. Lagesen, P. Hallin, E. A. Rødland, H.-H. Staerfeldt, T. Rognes, D. W. Ussery, RNAmmer: Consistent and rapid annotation of ribosomal RNA genes. *Nucleic Acids Res.* **35**, 3100–3108 (2007).
65. T. M. Lowe, S. R. Eddy, tRNAscan-SE: A program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* **25**, 955–964 (1997).
66. C. Trapnell, L. Pachter, S. L. Salzberg, TopHat: Discovering splice junctions with RNA-Seq. *Bioinformatics* **25**, 1105–1111 (2009).
67. B. Langmead, C. Trapnell, M. Pop, S. L. Salzberg, Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* **10**, R25 (2009).
68. C. Trapnell, B. A. Williams, G. Pertea, A. Mortazavi, G. Kwan, M. J. van Baren, S. L. Salzberg, B. J. Wold, L. Pachter, Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* **28**, 511–515 (2010).

Acknowledgments

Funding: This work was supported by the National Research Foundation, Prime Minister's Office, Singapore, under the Competitive Research Programme (project no. NRF-CRP5-2009-05) and its Campus for Research Excellence and Technological Enterprise programme. This work was also supported by the Singapore Ministry of Education Academic Research Fund Tier 1 under project no. R302000148112. **Author contributions:** T.L. designed and performed experiments, carried out data analysis, and wrote the manuscript. C.Z. assisted in fermentation work. K.-L.Y. assisted in data analysis and revised the manuscript. J.H. designed the work and revised the manuscript. **Competing interests:** The authors declare that they have no competing interests. **Data and materials availability:** All data needed to evaluate the conclusions in the paper are present in the paper and/or the Supplementary Materials. Additional data related to this paper may be requested from the authors and are available from GenBank under accession number CP016893.

Submitted 5 May 2017
Accepted 7 February 2018
Published 23 March 2018
10.1126/sciadv.1701475

Citation: T. Li, C. Zhang, K.-L. Yang, J. He, Unique genetic cassettes in a *Thermoanaerobacterium* contribute to simultaneous conversion of cellulose and monosugars into butanol. *Sci. Adv.* **4**, e1701475 (2018).

Science Advances

Unique genetic cassettes in a *Thermoanaerobacterium* contribute to simultaneous conversion of cellulose and monosugars into butanol

Tinggang Li, Chen Zhang, Kun-Lin Yang and Jianzhong He

Sci Adv **4** (3), e1701475.
DOI: 10.1126/sciadv.1701475