Article

Utilization of monosaccharides by *Hungateiclostridium thermocellum* ATCC 27405 through adaptive evolution

Dung Minh Ha-Tran[1,2,3], Trinh Thi My Nguyen [2], Shou-Chen Lo[2*], Chieh-Chen Huang[2,4*]

[1] Molecular and Biological Agricultural Sciences Program, Taiwan International Graduate Program, Academia Sinica and National Chung Hsing University, Taipei 11529, Taiwan
[2] Department of Life Sciences, National Chung Hsing University, Taichung 40227, Taiwan
[3] Graduate Institute of Biotechnology, National Chung Hsing University, Taichung 40227, Taiwan
[4] Innovation and Development Center of Sustainable Agriculture, National Chung Hsing University, Taichung 40227, Taiwan
***Cocorresponding authors: Chieh-Chen Huang, Ph.D and Shou-Chen Lo, Ph.D**
E-mail addresses: hatranminhdung@gmail.com (DMH-T), mytrinhnguyen0410@gmail.com (TTMN), scl@dragon.nchu.edu.tw (SCL), and cchuang@dragon.nchu.edu.tw (CCH)

Abstract: ***Hungateiclostridium thermocellum*** **ATCC 27405 is a promising bacterium with a robust ability to degrade lignocellulosic biomass complexes, including crystalline cellulose components, through a multienzyme cellulosomal system. In contrast, it exhibits poor growth on simple monosaccharides such as fructose and glucose. This phenomenon raises many important questions concerning its glycolytic pathways and sugar transport systems. Until now, the detailed mechanisms of *H. thermocellum* adaptation to growth on monosaccharides have been poorly explored. In this study, adaptive laboratory evolution was applied to train the bacterium on monosaccharides, and genome resequencing was used to detect the genes that had mutated during adaptation. RNA-seq data of the 1st-generation culture growing on either fructose or glucose revealed that several glycolytic genes in the EMP pathway were expressed at lower levels in these cells than in cellobiose-grown cells. After 8 generations of culture on fructose and glucose, the evolved *H. thermocellum* strains grew faster and yielded greater biomass than the nonadapted strains. Genomic screening also revealed several mutation events in the genomes of the evolved strains, especially in genes responsible for sugar transport and central carbon metabolism. Consequently, these genes could be applied as targets for further metabolic engineering to improve this bacterium for bioindustrial usage.**

Keywords: ***Hungateiclostridium thermocellum*****, adaptive laboratory evolution, RNA-seq, cellulosomal genes, EMP pathway, monosaccharides.**

1. Introduction

Hungateiclostridium thermocellum (previously *Clostridium thermocellum*), a thermophilic, gram-positive, anaerobic bacterium, exhibits a strong capacity to efficiently degrade crystalline cellulose [1–3]. *H. thermocellum* produces a number of industrially important fermentation products, such as ethanol, acetic acid, lactic acid, and hydrogen, and understanding its responses to different carbon sources will help improve the production of the desired end-products. To date, several studies focusing on the regulatory mechanisms of specific genes of interest [4–10] and genomic [11, 12], transcriptomic [8, 9,13–17], metabolomic [18, 19], proteomic [20–23], and integrated omics [24–27] studies have been

 doi:10.20944/preprints202012.0496.v1

performed to expand our knowledge of *H. thermocellum* genetics, gene expression, metabolism and physiology. However, most previous studies utilized the preferred substrates of *H. thermocellum,* such as cellobiose [12, 24, 27], crystalline cellulose [13, 22], cellobiose and crystalline cellulose in combination [20, 28, 29], pretreated switchgrass and *Populus* [15], and pretreated yellow poplar and cellobiose [17]. Since *H. thermocellum* always demonstrates poor growth on comparatively simple monosaccharides such as glucose, fructose, or sorbitol [30–32], some research groups were interested in investigating the production of glycolytic enzymes [33], the production of cellulase [30], sugar transport systems [34, 35], mutation of the bacterium [31], and carbon metabolism [32] in this species when growing on unfavored substrates. However, the mechanisms of adaptation to monosaccharide utilization in *H. thermocellum* are still poorly understood, since those studies emphasized single individual aspects rather than the entire system. *H. thermocellum* exhibits a long lag phase when growing on these substrates, which Nochur *et al.* (1990) assumed might be due to the time needed for genetic changes in the bacterial genome to allow it to grow on fructose or glucose. However, these authors did not determine which genes were mutated during the lag period. In another study, Nochur *et al.* speculated that the long lag phase of fructose-adapted cells (FAs1) and glucose-adapted cells (GAs1) could be due to the lower intracellular pH in these cells than in cellobiose-grown cells (CGs1) [32]. However, this could not explain the newly acquired characteristics of the evolved phenotypes, especially in terms of genetic and mRNA expression changes.

Adaptive laboratory evolution (ALE) has been widely used to better understand the basic mechanisms of molecular evolution and the genomic changes that accumulate in microbial populations during long-term selection under specific growth conditions [36]. With rapid advances in transcriptomic profiling and next-generation sequencing (NGS), phenotype–genotype correlations can be easily obtained [37]. In a previous study, *H. thermocellum* was cultured in *Populus* hydrolysate, which is a medium that contains various compounds that are toxic to the bacterium. After 117 transfers from 5-17.5% (v/v) *Populus* hydrolysate, 73 mutations were identified [11]. These mutations were found to be related to cellular repair and energy metabolism, which helped the bacterium grow better than the wild type in the toxic medium [11]. Recently, Holwerda *et al.* found that a mutation in the *adhE* gene induced by ALE allowed *H. thermocellum* to grow better and achieve higher ethanol yield [12]. In the present study, the ALE approach followed by genome resequencing was carried out to identify different types of mutations that occurred in the bacterial genome after multiple generations of growth on its less-preferred carbon sources, such as fructose and glucose. Moreover, RNA-seq was used to identify the total differentially expressed genes (DEGs) between cellobiose-grown cells (CGs), which were used as the control, and FAs and GAs to clarify the responses of the bacterium to nutritional stress. We particularly focused on the Embden-Meyerhof-Parnas (EMP) pathway, as this is the predominant glycolytic route in *H. thermocellum* [19]. To the best of our knowledge, no previously published studies have used transcriptome profiling combined with genomic analysis to elucidate the mechanism of adaptation of *H. thermocellum* to monosaccharides. Thus, the advanced RNA-seq and genomic sequencing techniques used in this study allow us to confirm and re-reexamine the results of previous studies, shedding more light on the aforementioned issues and broadening our knowledge in this area.

2. Materials and Methods

2.1. Hungateiclostridium thermocellum ATCC 27405 *growth conditions and medium preparation*

To prepare the seed inoculum, *H. thermocellum* was grown in modified GS-2 medium supplemented with cellobiose (5 g/L) until the stationary phase was reached (OD_{660} ~ 0.7) and then inoculated 1% (v/v) into the same medium containing either glucose (10 g/L), fructose (10 g/L), or cellobiose (5 g/L). Three independent sets of CGs1, FAs1, and GAs1

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 December 2020 doi:10.20944/preprints202012.0496.v1

were performed for growth pattern determination, gas and ethanol measurements, residual reducing sugar tests, and cellulosomal enzyme assays. Cells were grown in batch culture at 60 °C in 250-mL serum bottles containing 100 mL of modified GS-2 medium. Cell growth was monitored daily based on the measurement of optical density at A_{660} using a GeneQuant 1300 spectrophotometer (Biochrom, MA, USA). The cell mass was quantified from the pellet protein as described by Zhang and Lynd [38], except that the Bradford assay, with bovine serum albumin (BSA) (Sigma-Aldrich) as the standard, was used instead of Peterson's method. Briefly, 1 mL of culture broth was centrifuged using a Hitachi Tabletop centrifuge CT15RE (Hitachi Koki Co., Ltd, Japan) at 12,800 x g for 10 min. The pellet was washed with distilled water once, and 1 mL of 0.2 M NaOH and 1% SDS were added to the pellet. The sample was incubated at 28 °C for 30 min with occasional stirring. The pellet was resuspended, and 0.25 mL of 0.8 M HCl was added and mixed well to neutralize the sample. After centrifugation at 12,800 x g for 5 min, the protein in the supernatant was determined. One A_{660} unit corresponded to 0.85 g/L of cell mass. Carbon molarity in cell biomass was calculated based on the elemental composition of cells ($C_5H_8O_2N$), which corresponds to a molecular weight of 114.12 g/mol [39]. One liter of the modified GS-2 medium contained 1.5 g KH_2PO_4, 2.9 g K_2HPO_4, 3 g sodium citrate tribasic dihydrate ($Na_3C_6H_5O_7$), 2.1 g urea, 5 g 3-(N-morpholino) propanesulfonic acid (MOPS), 2 mg resazurin, and 1 g L-cysteine. A 10-fold trace salt solution contained 26 mg $MgCl_2$, 11.3 mg $CaCl_2$, and 0.125 mg $FeSO_4.7H_2O$. A 100-fold vitamin solution contained 2 mg pyridoxine hydrochloride, 0.2 mg biotin, 0.4 mg p-aminobenzoic acid and 0.2 mg vitamin B12 [40]. Trace salt and vitamin solutions were sterilized using a 0.22-µm filter (StarTech, Taiwan). All solutions were prepared with distilled water from an EcoQ Combo (LionBio, Taiwan). The pH of the GS-2 medium was adjusted to 7.2 using 5 M NaOH and purged extensively with pure nitrogen gas to create an anaerobic environment. GS-2 medium was autoclaved at 121 °C for 20 min, and 1% (v/v) trace salt solution and 1% (v/v) vitamin solution were added to the modified GS-2 when the medium was cooled to 50 °C.

2.2. Phosphoric acid swollen cellulose and cellulosome isolation

Amorphous cellulose was prepared by phosphoric acid swelling as described by Zhang *et al.* [41]. Cellulosomes were isolated from 100 mL of cellobiose-, glucose- and fructose-fermentation broth collected at late stationary phase using the affinity digestion protocol described by St Brice *et al.* [42]. Briefly, cell-free broth was recovered by centrifuging the fermentation broth at 12,800 x g for 40 min at 4 °C using a Hermle Z326K centrifuge (Benchmark, Wehingen, Germany). The supernatant was decanted, and the pH was adjusted to 7.0 with 2 M NaOH as evaluated with a S20 SevenEasy™ pH meter (Mettler-Toledo, PoMe, Australia). The cell-free broth was incubated with amorphous cellulose (10 mg/100 mL of cell-free broth) overnight at 4 °C to allow the binding of the cellulosomes to the cellulose. Amorphous cellulose with bound enzymes was centrifuged at 12,800 x g for 30 min at 4 °C, and the pellet was resuspended in 20 mL dialysis buffer (50 mM Tris-base, 10 mM $CaCl_2$, 5 mM dithiothreitol (DTT), pH 7.0). The amorphous cellulose suspension was then dialyzed in membrane bags (Spectra/POR1 MWCO 6-8000, USA) at 55 °C against two liters of distilled water to initiate amorphous cellulose degradation by the enzyme. Distilled water was changed every 60 min to avoid inhibition of the cellulosomes by the degradation products. After 5 h of incubation, the transparent suspension was centrifuged at 3,000 x g for 30 min at 4 °C, decanted and stored at -20 °C for protein and activity analysis. Total protein was measured with a Bio-Rad Bradford protein assay with BSA as a standard.

2.3. Avicelase, CMCase, and Xylanase Assays

The assays were performed as described by Zhang *et al.* [43]. Briefly, 4.1 ml of the well-suspended Avicel solution (2.44 g of dry Avicel/100 mL distilled water) and 0.5 ml of Tris-HCl buffer (0.5 M Tris; pH 7.0; 0.1 M $CaCl_2$, and optional 1.5% (w/v) sodium azide (NaN_3)

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 December 2020 doi:10.20944/preprints202012.0496.v1

were added into 16 x 125 mm anaerobic culture Hungate tubes. The tubes were sealed with 9-mm screw caps with butyl rubber stoppers, held under vacuum, and flushed with pure nitrogen at least 3 times for 5 min each. DTT at 0.5 M was added to the tubes (0.1 mL/tube) before the enzyme activity assay. The tubes were prewarmed in a water bath at 60 °C for 10 min, and 0.3 mL of the enzyme solution dilution series was added to the tubes. After the first 10 min of adsorption and reaction, 0.5 mL of sample (well-mixed Avicel suspension) was withdrawn as the starting point, followed by removal of an additional 0.5 mL every 60 min. The reactions were stopped by transferring the tubes to an ice water bath, and then the samples were placed in precooled 1.5-mL microcentrifuge tubes. The samples were centrifuged at 13,000 x g for 3 min. The net soluble sugar released during the hydrolysis process was calculated by subtracting the amount of sugar measured at the starting point, and enzyme activity was determined based on the linear correlation between sugar released and enzyme concentration. In the CMCase assay, 2% (w/v) carboxymethylcellulose (CMC) was dissolved in citrate buffer (50 mM, pH 4.8). A volume of 0.5 mL of the diluted enzyme was added into the Hungate tubes. One half mL of the CMC solution was then added to the test tubes, and the tube contents were mixed well by pipetting. The mixture was incubated at 50 °C for 30 min, and 3 mL of DNS solution was added into the test tubes to stop the reaction. The Hungate test tubes were boiled for 5 min in boiling water and then placed in an ice water bath. The xylanase activity of purified cellulosomes was measured by the colorimetric assay described by Ribeiro *et al.* [44] using birchwood xylan as the substrate, and the reducing sugars were determined using the DNS method with xylose as a control. Briefly, the reaction mixture (0.05 mL substrate (1% w/v) in 50 mM sodium acetate buffer, pH 5, and 0.01 ml enzyme solution) was incubated at 70 °C in a water bath for 5 min. The reaction was stopped by adding 0.1 mL of DNS and immediately boiling for 5 min. Quantification of the reducing sugars released as a result of enzyme activity was estimated by A_{540} measurements using an xMark™ microplate absorbance spectrophotometer (Hercules, CA, USA), where one unit of enzymatic activity was defined as the amount of enzyme that produced 1 μmol/min of reducing sugars.

2.4. Measurement of residual reducing sugar, hydrogen, carbon dioxide, and ethanol

Reducing sugars were measured using the DNS method [45], and the standard curves were constructed using 5-fold serial dilutions of a pure chemical sample. The standard curves for glucose, fructose and cellobiose assays were built separately, as these sugars have different reactivities with DNS reagent [46]. To prepare samples for ethanol and reducing sugar determination, 1 mL of fermentation broth was centrifuged at 12,800 x g for 10 min at 4 °C. The pellet was discarded, and the supernatant was filtered using a 0.22-μm filter (StarTech, Taiwan). H_2 and CO_2 were measured daily, but ethanol and reducing sugars were analyzed using samples that were taken at the sampling points indicated in Figure 3A–H and Figure 4A–H. These major fermentation end-products were determined by a GC Agilent 7890A (Agilent Technologies, CA, USA) equipped with a J&W 122-3232: 30 m x 250 μm x 0.25 μm DB-FFAP column, with nitrogen as the carrier gas at a flow rate of 30 mL/min. For ethanol measurement, the front inlet was used, the detector was kept at 225 °C, and the oven was heated from 50 °C to 150 °C (50–100 °C at a ramp rate of 30 °C/min and 100–150 °C at a ramp rate of 20 °C/min). For H_2 and CO_2 measurements, the back inlet was used, the detector was kept at 225 °C, and the oven was operated isothermally at 50 °C.

2.5. RNA isolation and library preparation and sequencing

Two independent sets of CGs1, FAs1 and GAs1 were harvested at the late stationary phase for RNA-seq: 60 h for both sets of CGs1, 312 h for both sets of FAs1, and 372 h for the 1st set of GAs1 and 360 h for the 2nd set. All 3 sets of RNA preparations were used for RT-qPCR of selected genes. Total RNA was extracted using TRIzol (R) Reagent (Invitrogen, MA, USA) according to the instruction manual. Purified RNA was quantified at

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 December 2020 doi:10.20944/preprints202012.0496.v1

OD_{260nm} using an ND-1000 spectrophotometer (NanoDrop Technology, USA) and quality checked using a Bioanalyzer 2100 instrument (Agilent Technologies, CA, USA) with an RNA 6000 LabChip kit (Agilent Technologies, CA, USA).

All RNA sample preparation procedures were carried out according to Illumina's official protocol. Agilent's SureSelect Strand-Specific RNA Library Preparation Kit was used for library construction, followed by size selection with AMPure XP beads (Beckman Coulter, CA, USA). The sequences were determined using Illumina sequencing-by-synthesis (SBS) technology (Illumina, CA, USA). Sequencing data (FASTQ reads) were generated based on the Illumina Conversion program bcl2fastq v2.20.

2.6. Differential gene expression analysis

All RNA-seq data analyses and graphics production were performed using R (ver. 3.6.1), except for the Venn diagrams, which were drawn using the Venny tool (ver. 2.1) [47]. The expression levels were calculated as reads per kilobase per million reads (RPKM). For differential expression analysis, three Bioconductor packages were used, namely, edgeR [48], DESeq2 [49], and NOISeq [50]. [51]. Two of these Bioconductor packages, edgeR (ver. 3.28.1) [48] and DESeq2 (ver. 1.26.0) [49], are the most widely used differential expression (DE) methods for RNA-seq. Both edgeR and DESeq2 are parametric methods that perform DE analysis based on the assumption that the inputted raw data follow a negative binomial distribution. The Bioconductor package NOISeq (ver. 2.30.0) is a nonparametric method that uses an empirical Bayes approach to improve the handling of biological variability specific to each gene [50]. Genes with very low expression levels (raw count < 3 or RPKM value < 0.3) in either or both of the treated samples (FAs and GAs) and the control sample (CGs) were excluded. Genes with ≥ 2-fold change or ≤ 0.5-fold change (log_2-fold change ≥ 1 or log2-fold change ≤ -1) and a false discovery rate (FDR) of $p < 0.01$ were considered significantly differentially expressed.

2.7. Sample preparation for genomic analysis

The evolved populations that had been serially passaged on fructose and glucose were used for genome resequencing. Specifically, two independent sets of evolved strains, FAs and GAs (8^{th} transfer), considered the endpoint strains, were harvested and pooled together and stored at -80 °C. Genomic DNA was extracted by a WelPrep DNA kit (Welgene Biotech, Taiwan) according to the manufacturer's instructions. DNA samples with OD 260/280 in the range of 1.8~2.0 were quantity/quality assessed using the Agilent Genomic DNA ScreenTape assay in conjunction with the 4200 TapeStation system (Agilent Technologies, CA, USA).

2.8. Library construction and sequencing

Total DNA (10 μg) was sonicated using a Covaris M220 focused ultrasonicator (Covaris, MA, USA) to sizes ranging from 400 to 500 bp. DNA sizing was checked using the Agilent D1000 ScreenTape assay with the 4200 TapeStation system. Approximately 0.5 μg fragmented DNA was end-repaired, A-tailed and adaptor-ligated following the Illumina TruSeq DNA preparation protocol. After library construction, samples were mixed with MiSeq Reagent Kit v3 (600 cycles) and loaded onto a MiSeq cartridge. Then, a 2 x 300 bp paired-end sequencing run was performed on the MiSeq sequencer platform (Illumina, CA, USA).

2.9. Read mapping, variant calling and annotation

The raw sequences generated from the sequencer first went through a filtering process to obtain good quality reads. FASTP was implemented to trim bases or remove reads

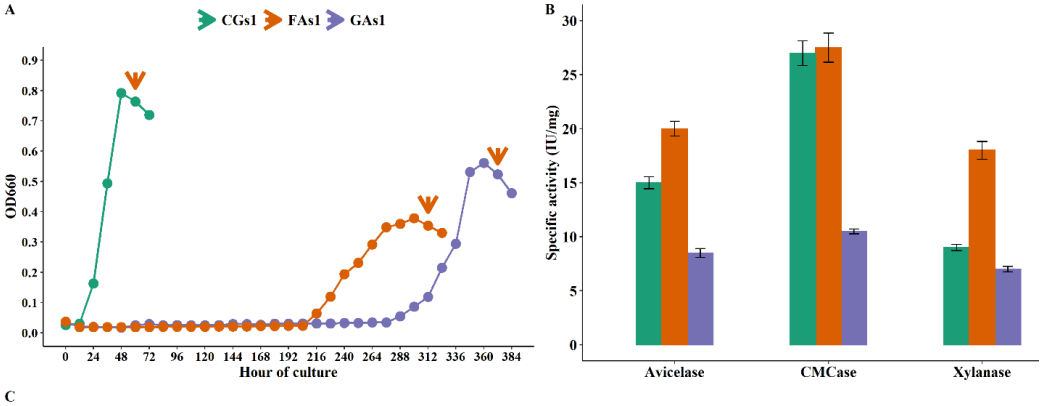
Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 December 2020 doi:10.20944/preprints202012.0496.v1

according to quality scores. The reference genome sequence and annotation GFF file were downloaded from the NCBI RefSeq genome database. The trimmed reads were aligned against the reference genome using the bwa aligner with the BWA-MEM algorithm. Then, the Picard tool MarkDuplicates was used to remove duplicate alignments. The GATK tool HaplotypeCaller was used for variant detection. Finally, snpEff was used to perform variant annotation.

3. Results and Discussion

3.1. Growth patterns, sugar utilization, cellulosomal enzyme activities and ethanol productivity of CGs1, FAs1, and GAs1

After a long lag phase of adaptation, *H. thermocellum* ATCC 27405 eventually used glucose or fructose as a single carbon source and grew rapidly (Figure 1A). This finding is in agreement with the results of Hernandez [34] and Johnson *et al.* [30], which showed that the *H. thermocellum* bacterium could grow on fructose and glucose only after an extended lag time. However, the final cell biomass of FAs1 and GAs1 was still lower than that of CGs1, in contrast to the results obtained in the work of Johnson *et al.* [30], who found that the final *H. thermocellum* cell densities of FAs1 and GAs1 were comparable to those of CGs1. As fructose and glucose are unfavored carbon sources for *H. thermocellum* growth, it is not surprising that large amounts of these substrates remained in the culture broths at the end of the fermentation process (Figure 1C). Cellobiose, in contrast, is the preferred carbon source of *H. thermocellum*; thus, this disaccharide was rapidly consumed by the bacterium, enabling robust growth and effective end-product formation. Consequently, the ethanol productivity, sugar consumption, and cell biomass of CGs1 were significantly greater than those of FAs1 and GAs1 (Figure 1C). As the most powerful hydrolytic machinery that has been identified in a microorganism to date, *H. thermocellum* cellulosomal function, composition, and regulation patterns have been a popular research topic for decades [1, 3–7, 10, 15, 17, 20, 21, 28–30, 38, 40, 41, 43, 50–54]. Regarding cellulosomal enzyme activities, FAs1 cellulosomes expressed the highest exoglucanase, endoglucanase, and xylanase activities, followed by CGs1 cellulosomes. Surprisingly, the lowest cellulosomal enzyme activities were found in GCs1, even though *H. thermocellum* growing on glucose is not subject to carbon catabolite repression [30]. The observed GAs1 enzyme activities were in agreement with a study by Zhang *et al.* [59], who used kinetic models and empirical experiments to prove that glucose was an inhibitor of the *H. thermocellum* cellulosome. Other reports also indicated that cellobiose was a strong inhibitor of cellulosome and cellulase [41,42]. These previous studies could partly explain the low activities of cellulosomal enzymes derived from CGs1 and GAs1 compared to those derived from **FAs1.**



H.thermocellum	Cs	Pe	Yp.s	Vep	Cb
CGs1	25.36 ± 0.79*	42.61 ± 1.74	1.68 ± 0.02	0.710 ± 0.003	5.69 ± 0.22
FAs1	30.64 ± 4.19	19.83 ± 2.67	0.65 ± 0.06	0.064 ± 0.007	2.23 ± 0.41
GAs1	27.67 ± 6.75	18.13 ± 4.50	0.66 ± 0.05	0.049 ± 0.006	3.50 ± 0.58

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 December 2020 doi:10.20944/preprints202012.0496.v1

Figure 1. *H. thermocellum* growth patterns, cellulosomal enzyme activity, sugar consumption, and ethanol fermentation. (A) Representative growth profiles of the 1st generation cultures, CGs1, FAS1, and GAs1. The arrowheads indicate the sampling points used for RNA-seq, RT-qPCR, and cellulosome isolation. (B) Specific enzyme activity of the *H. thermocellum* cellulosome. (C) Sugar consumption, ethanol productivity, and cell biomass of CGs1, FAs1, and GAs1. All data in Figure 1B and Figure 1C represent the mean ± sd of biological triplicates.

Abbreviations: Cs: Consumed sugar (mmoles/L), *mmoles glucose equivalents; Pe: Produced ethanol (mmoles/L) at the sampling point; Yp.s: Ethanol yield (mmoles produced ethanol/mmoles consumed sugar); Vep: Volumetric ethanol productivity (mmoles/L/h); Cb: Cell biomass (mmoles/L).

The expression profiles of cellulosomal genes in RNA-seq data may help to confirm the results of enzyme activities. Our RNA-seq data showed that many carbohydrate-active enzyme (CAZyme)-encoding genes in GAs1 remained at low expression levels, while these genes exhibited high expression levels in FAs1 (Table S1). To regulate the expression of cellulosomal genes in response to different substrates (i.e., cellulose, pectin, or xylan) in growth media, *H. thermocellum* has an alternative sigma/antisigma (SigI/RsgI) factor system that works harmoniously via an external carbohydrate-sensing mechanism [6,7,43]. Although the carbon sources used in the present study were not specific carbohydrate targets of the SigI/RsgI system, several pairs of SigI/RsgI factors, including SigI1/RsgI1 (Cthe_0058 [-0.46; -2.08])/(Cthe_0059 [0.16; -2.46]), SigI2/RsgI2 (Cthe_0267 [1.14; -2.78])/(Cthe_0268 [1.32; -2.34]), Sig24C/Rsi24C (Cthe_1470 [2.54; 1])/(Cthe_1471 [2.38; -1.33]), were differentially expressed. This might be explained by the study of Sand *et al.* [58], who found that the regulation of *H. thermocellum* cellulosomal genes was not only mediated by the SigI/RsgI system but also could be governed by the vegetative factor SigA. Given that the cellulosome could be inhibited by glucose and cellobiose, latent regulators of noncellulosomal enzymes might activate the expression of their regulons to hydrolyze any potential substrate in culture medium. As a consequence, of the various noncellulosomal genes known to date, we identified a dozen with differential expression in FAs1 and/or GAs1 (Table S1). In the study of Johnson *et al.* [30], poor growth of *H. thermocellum* on fructose or sorbitol was accompanied by a significant increase (5- or 6-fold higher) in the specific production of cellulase. This result was in accordance with Dror *et al.* [4, 5], who found that the transcription of the exoglucanase CelS, scaffoldin protein CipA, and two anchoring proteins OlpB and Orf2p was affected by the growth rate rather than by the substrate itself. In other words, a slower growth rate of *H. thermocellum* should be associated with higher cellulase activity. In our RNA-seq data, CelS (Cthe_2089), CipA (Cthe_3077), OlpB (Cthe_3078), and Orf2p (Cthe_3079) were highly expressed in FAs1 and GAs1, except that OlpB was downregulated in GAs1 (Table S1). However, the production of specific cellulase decreased when *H. thermocellum* cells were well adapted to fructose, suggesting that sugar catabolism was responsible for the repression of cellulase production. To further clarify this finding from a transcriptomic viewpoint, cell samples from CGs1, FAs1, FAs3, and FAs5 were collected, and sigma factors (SigI1, SigI2, SigI3, SigI4, SigI5, SigI6, SigI7, Sig24C), glycoside hydrolases (CelK, CelR, CelS), scaffoldin CipA, anchoring SdbA, Orf2p, OlpB genes were chosen for RT-qPCR. The list of primers used for RT-qPCR is shown in Table S2. The levels of gene expression determined by RT-qPCR gradually decreased from FAs1 to FAs5 compared to those in CGs1 (Figure S1), in agreement with the enzymatic analysis of Johnson *et al.* [30]. In their study, Johnson and colleagues found that after cells had well adapted to fructose or sorbitol, the repression of cellulase production was stabilized, and after a number of transfers, the cellulase yields of these adapted strains were even lower than those of strains grown on cellobiose or glucose.

3.2 Analysis of DEGs

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 December 2020 doi:10.20944/preprints202012.0496.v1

As a gold standard normalization method has not yet been found [52], and each of the three Bioconductor package algorithms has its own advantage, we considered the overlap of their DE gene lists to be the most reliable DE results for the present study. In the common DE lists of the three packages, 747 genes were upregulated and 575 genes were downregulated in FAs1 vs CGs1, and 638 genes were upregulated and 567 genes were downregulated in GAs1 vs CGs1 (Figure S2, Figure S3). In addition, 25 genes from CGs1, FAs1, and GAs1 representing key cellulosomal genes and sugar transporters were selected for RT-qPCR confirmation (Figure S4). The correlation coefficients for the RT-qPCR and RNA-seq results for FAs1 vs. CGs1 and GAs1 vs. CGs1 were $R = 0.91$ and $R = 0.96$, respectively, confirming the high quality of the RNA-seq data. The list of primers used for RT-qPCR is shown in Table S2.

3.3 EMP pathway

3.3.1 Genes upstream of phosphoenolpyruvate (PEP)

RNA-seq data demonstrated the dynamics of the transcriptome in the EMP pathway. This finding is in line with Xiong *et al.* [19], who found that the EMP pathway was the predominant glycolytic route in *H. thermocellum*. Many key glycolytic genes in the EMP pathway were highly upregulated in CGs1 compared to FAs1 and GAs1 (Figure 2). (Note: The $\log_2$ fold change of FAs1 relative to CGs1 and $\log_2$ fold change of GAs1 relative to CGs1 are presented as Gene name [number1; number2], respectively). As might be expected, the gene Cthe_0275 encoding a CBP (EC 2.4.1.20) that catalyzes the phosphorolysis of cellobiose into Glc1P and Glc was upregulated in CGs1 relative to FAs1 and GAs1. However, another CBP gene (Cthe_1221) was upregulated in FAs1. The roles of this CBP gene in fructose metabolism and the possible existence of a regulator for its high level of expression remain unknown. From Glc to 2PGA, many genes involved in the conversion of intermediate sugar phosphates to their subsequent metabolites were found to be upregulated in CGs1 relative to FAs1 and GAs1. Specifically, Cthe_0390 and Cthe_2938 encode GCKs (EC 2.7.1.2) that convert Glc to Glc6P; Cthe_0217 encodes a PGI (EC 5.3.1.9) responsible for Glc6P to Fru6P conversion; PPi-dependent PFK (EC 2.7.1.11) (Cthe_0347) and ATP-dependent PFK (EC 2.7.1.11) (Cthe_1261) catalyze the conversion of Fru6P to Fru16BP; and Cthe_0349 encodes an FBA (EC 4.1.2.13) that regulates the conversion of Fru16BP to GAP. Gene expression patterns continued their upregulation trend in CGs1, as three genes encoding PGAM (EC 3.1.3.3) (Cthe_0707, Cthe_0946 and Cthe_1292) catalyze the conversion of 3PGA to 2PGA and exhibited high expression levels compared to those in FAs1 and GAs1.

3.3.2 Genes downstream of PEP

The malate shunt is responsible for catalyzing the indirect conversion of PEP to pyruvate via malate. With the absence of an annotated pyruvate kinase sequence in the *H. thermocellum* genome, no PYK (EC 2.7.1.40) enzyme activity was detected in cell extracts, and PPDK (EC 2.7.9.1) is unlikely to be crucial for the conversion of pyruvate from PEP [62]. It could be assumed that pyruvate formation proceeds predominantly via a malate shunt [63]. In our RNA-seq, two of three genes involved in the malate shunt, $NADP^+$-dependent ME (Cthe_0344) and NADH-dependent MDH (Cthe_0345), maintained elevated expression levels in CGs1. Furthermore, as ITP can substitute for GTP in the carboxylase reaction, the upregulation of the inosine 5-monophosphate dehydrogenase-encoding gene (Cthe_0681), which is an enzyme involved in the synthesis of guanine nucleotides, further supported the idea that the malate shunt is the main glycolytic route for producing pyruvate in *H. thermocellum* [22]. Alternatively, OAA can be directly decarboxylated into pyruvate using the membrane-bound enzyme ODC (EC 4.1.1.3). Although previous studies [23, 38] provided evidence of the moderate expression of 2 ODC genes, Clo1313_1523 and Clo1313_1712, our RNA-seq data found that Cthe_0701 [1.47; 2.43]

 doi:10.20944/preprints202012.0496.v1

(~Clo1313_1523) and Cthe_0509 [-0.16; -1.16] (~Clo1313_1712) had differential expression. In contrast, Olson *et al.* [63] could not detect ODC activity in their study. We speculate that the inconsistency may be due to the different strains of *H. thermocellum* used in these works (e.g., Olson *et al.* used *H. thermocellum* DSM 1313, whereas we used *H. thermocellum* ATCC 27405). Patni and Alexander [33] found that the increase in GCK activity in GAs was proportional to the increase in the yield of cell mass. They argued that GCK was likely an inducible enzyme, as its synthesis could be induced by glucose, fructose, or mannose. Intriguingly, the activity of GCK increased in mannitol-grown cells after their suspension in glucose medium, suggesting a fundamental role of GCK in glucose metabolism. Although Ng and Zeikus [64] could not detect GCK activity in *H. thermocellum* during growth on cellobiose, we found the upregulation of GCK genes (Cthe_0390 [-1.13; -1.08] and Cthe_2938 [-1.72; -1]) in CGs1. The elevated expression of GCK genes could be explained by the abundance of intracellular Glc formed from cellobiose, which is easily taken up from cellobiose culture and then undergoes phosphorolytic cleavage by CBP to produce Glc1P and Glc. The high accumulation of sugar phosphate intermediates in FAs1 and GAs1, which was first described in [32], was attributed to the accumulation rate of Fru6P being much greater in fructose medium than in cellobiose medium. This finding suggested that *H. thermocellum* metabolizes cellobiose much faster than fructose. The accumulation of cellobiose, Glc6P, Fru6P, and 3PGA was also found in *H. thermocellum* growing under ethanol stress [24]. This study also indicated that the PFK- and PGAM-encoding genes were downregulated under ethanol stress conditions, similar to what was observed for FAs1 and GAs1 under nutritional stress in the present study. Considering metabolite intermediate data from [32], transcriptomic and metabolomic data from [24] and our RNA-seq together, we concluded that the low expression levels of EMP pathway genes in FAs1 and GAs1 led to less efficient monosaccharide metabolism, eventually causing the poor growth of *H. thermocellum*.

An abundance of ribosomal proteins and other translation-related genes, which were upregulated in CGs1, directly support the potent growth of *H. thermocellum* on its favored carbon source. In FAs1, a large number of genes involved in carbohydrate transport and metabolism (COG category G) experience high levels of regulation; however, these genes mainly encode cellulosomal enzymes and sugar transporters and are not involved in the glycolytic route (Table S3). Under conditions offering only poorly fermentable substrates, the bacterium needs to increase its motility and signal transduction systems to seek available nutrients in the environment. This explains the greater numbers of DEGs in category N and category T in FAs1 and GAs1 relative to CGs1 (Figure S5).

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 December 2020 doi:10.20944/preprints202012.0496.v1

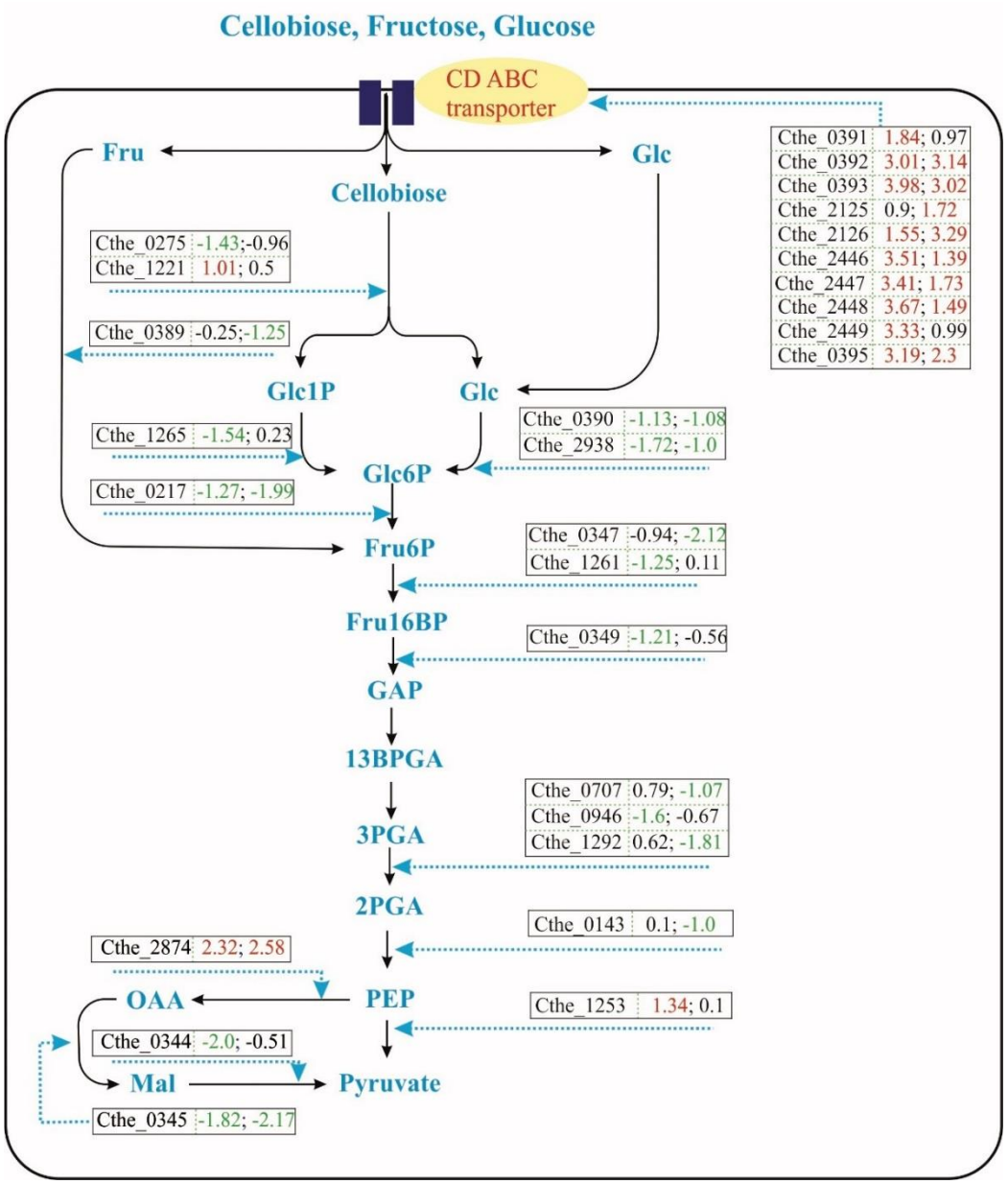


Figure 2. Diagram of the EMP pathway and end-product synthesis using *H. thermocellum* ATCC 27405.

Gene IDs and corresponding $\log_2$ fold change (FC) values (number 1, number 2) are displayed in the box. Text in red with a positive value and green with a negative value represent the upregulated and downregulated genes, respectively, in FAs1 vs. CGs1 and in GAs1 vs. CGs1, whereas text in black represents the genes with no significant transcriptional changes between FAs1 or GAs1 and CGs1 (i.e., $\log_2FC \geq 1$ or $\log_2FC \leq -1$). Nomenclature of metabolites: 2PGA: 2-phosphoglycerate; 3PGA: 3-phosphoglycerate; 13BPGA: 1,3-bisphosphoglycerate; CD ABC transporter: Cellodextrin ATP-binding cassette transporter; Fru: Fructose; Glc: Glucose; CBP: Cellobiose phosphorylase; Fru6P: Fructose 6-phosphate; Fru16BP: Fructose 1,6- biphosphate; FBA: Fructose-bisphosphate aldolase; GA3P: Glyceraldehyde 3-phosphate; Glc: Glucose; GAP: Glyceraldehyde 3-phosphate; Glc1P: Glucose 1-phosphate; Glc6P: Glucose 6-phosphate; GCK: Glucokinase; Mal: Malate; MDH: Malate dehydrogenase; ME: Malic enzyme; OAA: Oxaloacetate; PGI: Glucose-6-phosphate isomerase; PEP: Phosphoenolpyruvate; PFK: Phosphofructokinase; PGAM: Phosphoglycerate mutase; PPDK: Pyruvate phosphate dikinase; PYK: Pyruvate kinase; PEP: Phosphoenolpyruvate; Pi: Inorganic phosphate

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 December 2020 doi:10.20944/preprints202012.0496.v1

3.4 Sugar transporters

Among the five operons encoding cello-oligosaccharide ABC transport proteins identified by Nataf *et al.* [65], four operons, CbpA, CbpB, CbpC, and CbpD, were experimentally proven to be specific to cellodextrins of different lengths (G2–G5). Therefore, it is not expected that these sugar transporters would exhibit high levels of expression during growth on monosaccharides. Due to the absence of cellodextrins in the growth medium, Rydzak *et al*. [23] could not detect proteins expressed from either the CbpC (Cthe_2125–2128) or CbpD (Cthe_2446–2449) operons in *H. thermocellum* cells. However, high expression levels of these sugar transporter-encoding genes were found during growth on fructose and glucose in this work. The mRNA expression levels of ATP-binding protein (nbdA) (Cthe_0391), inner membrane translocator (msdA) (Cthe_0392), and ribose ABC transporter (cbpA) (Cthe_0393) were significantly upregulated in FAs1 and GAs1 (Figure 2). Additionally, two of four genes in the CbpC operon, binding protein-dependent transport system (msdC1) (Cthe_2125) and permease component (msdC2) (Cthe_2126), were highly expressed in FAs1 and GAs1. Four genes of the CbpD operon (Cthe_2446–2449) were upregulated in FAs1 and GAs1. These results suggest that in the absence of a preferred substrate, these sugar transporter-related genes would be induced to facilitate fructose or glucose import into the cells. Furthermore, after adaptation to fructose was well established, the expression levels of genes encoding sugar transporters significantly decreased from FAs1 to FAs5 (Figure S6).

3.5 Energy generation and redox balance

Cellular redox imbalance under certain conditions may cause incomplete substrate utilization and slow growth of bacteria. Redox metabolism of *H. thermocellum* is complex, with multiple reactions that shuttle electrons between reduced ferredoxin, NADH, and NADPH. These reactions use different enzyme complexes, such as reduced ferredoxin:NAD(P) oxidoreductases and several Fe-Fe hydrogenases that are activated by hydrogenase maturase (HydG) [66]. Therefore, *H. thermocellum* produces molecular hydrogen via a hydrogenase-mediated pathway to dispose of the excess reductants generated during carbohydrate catabolism [13]. Hydrogenases have important roles in regenerating oxidized ferredoxin, which has been used as the electron acceptor for pyruvate:ferredoxin oxidoreductase (PFOR) during the conversion of pyruvate to acetyl-CoA [67]. Sander *et al.* [8] reported that four redox-active pathways, sulfate transport and metabolism, ammonia assimilation, porphyrin biosynthesis, and [NiFe] Fd-dependent hydrogenase, in *H. thermocellum* DSM 1313 showed decreased transcription after the addition of methyl viologen, a redox-active chemical, to balance cellular redox. In the present study, FAs1 and GAs1 exhibited upregulated expression in these four pathways, including [NiFe] hydrogenase (Cthe_3013, Cthe_3014, Cthe_3016, Cthe_3018, Cthe_3019, ammonia assimilation-related genes (Cthe_0197–0199), porphyrin biosynthesis (Cthe_2525, Cthe_2527–2529, and sulfate metabolism (Cthe_2531–2534, Cthe_2536–2538) (Table S4). The results suggest that FAs1 and GAs1 suffer from redox imbalance.

3.6 Adaptive laboratory evolution and genomic analysis

3.6.1 Evolved FAs8 and GAs8 strains

The first culture on fructose had a very long lag phase, as shown in Figure 3A. To conduct ALE, 1% (v/v) CGs at exponential phase (OD_{660} ~ 0.6) were inoculated into fresh fructose medium. A dark yellow color appeared in the FAs1 fermentation broth, simultaneously with some motile granules in the medium, suggesting a stress response of the organism to fructose. The color of FAs1 fermentation broth became darker as the pH of the GS-2 medium was increased (Figure S7). Similarly, the dark yellow color appeared in the GAs1 fermentation broth when the pH of the growth medium was increased from 7.2

 doi:10.20944/preprints202012.0496.v1

to 7.8 (Figure S8). However, the color soon became light yellow and then milky white after some serial transfers. Since the modified GS-2 medium supplied with fructose was an unsuitable growth medium for *H. thermocellum* in regard to both sugar transport and glycolytic enzymes, the bacterium needed time to adapt to the harsh growth conditions. From Figure 3B to Figure 3H, the length of the lag phase gradually decreased (from 200 h for the 1st-generation culture (FAs1) to less than 12 h for the 8th-generation culture (FAs8). After some transfers, the bacterial biomass was significantly increased (OD_{660} ~ 0.8–1) compared to that of FAs1 (OD_{660} ~ 0.2–0.3). The ability of the evolved FAs to consume fructose also improved progressively, as displayed in Figure 3I. Consequently, the ethanol yield and H_2 production of FAs8 were higher than those of previous generations (Figure 3J, 3K and Table S5). The preferred carbon source for *E. coli*, as for various other bacteria, is glucose [68]. However, this is not the case for *H. thermocellum*. Although the bacterium used glucose better after being adapted to this carbon source, reflected in both the greater biomass (OD_{660} increased from ~ 0.5 to OD_{660} ~ 1) and the shorter lag phase compared to the first culture GAs1 (decreased from ~ 200 h to ~ 12 h) (Figure 4A–H), the glucose consumed during fermentation (Figure 4I) and the yields of metabolic end-products such as ethanol and hydrogen (Figure 4J, K) were still very low (Table S5). Furthermore, unstable adaptation and the gradual decrease in glucose consumption of the GAs were recorded via growth patterns (Figure 4A–H) and glucose consumption data (Figure 4I). Specifically, we did not observe a gradual decrease in the length of the lag phase, as exhibited in the evolved FA strain; instead, the long lag phase and poor cell growth might reappear unpredictably after many rounds of subculture (Figure 4H). Further studies, therefore, need to be carried out to understand this phenomenon.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 December 2020 doi:10.20944/preprints202012.0496.v1

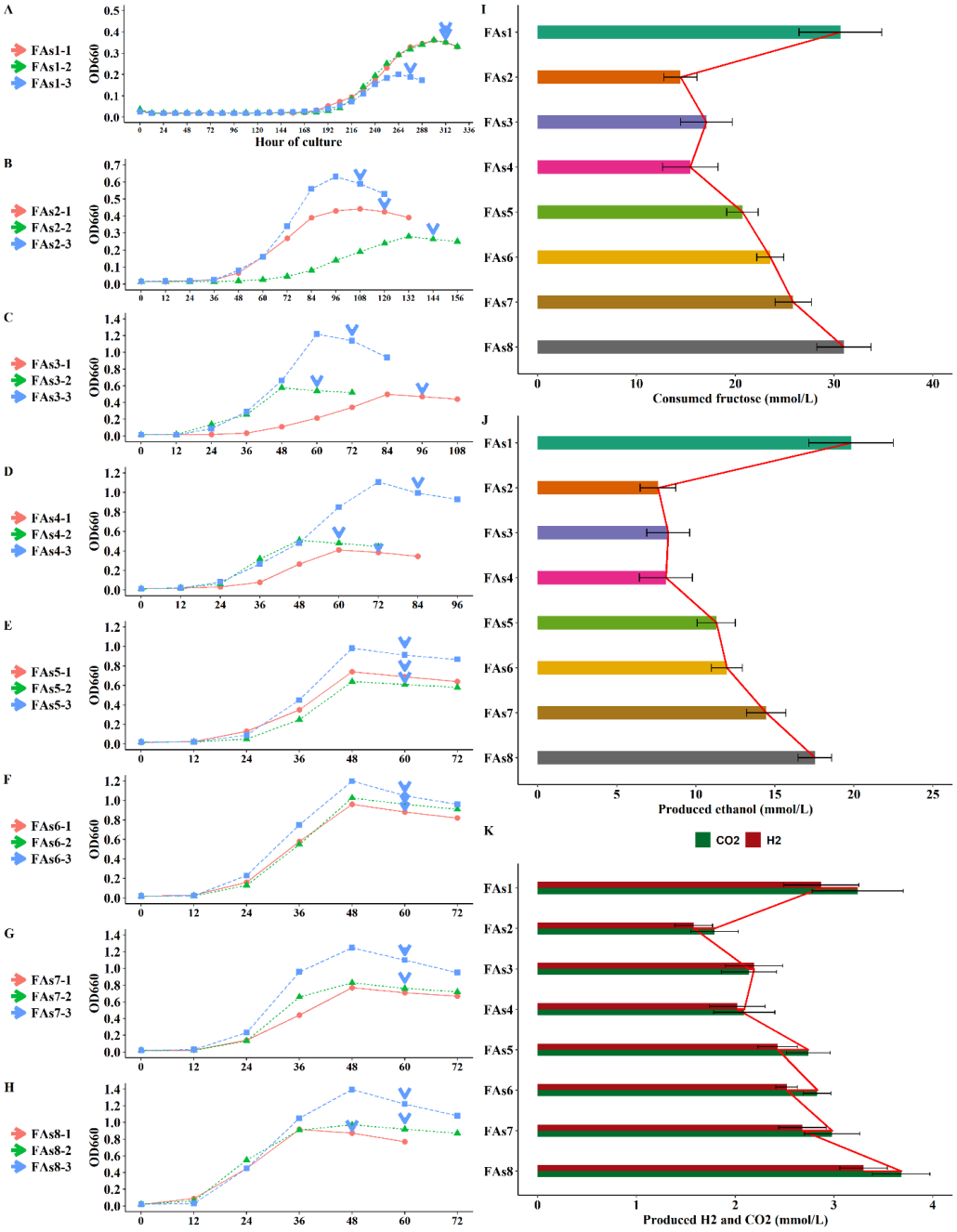


Figure 3. Growth curve of FA (left) and fructose consumption and ethanol and gas production (right) after 8 sequential transfer events. Figures 3A–3H represent the growth curves of FAs1–FAs8, respectively. The arrowheads indicate the sampling points used for ethanol and residual reducing sugar measurements. Figure 3I indicates the fructose consumed during fermentation. Figure 3J shows the ethanol production of FAs1–FAs8, and Figure 3K shows the gas production (CO_2 and H_2) of FAs1–FAs8. The ethanol and gas production results were calculated as averages of three replicates.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 21 December 2020 doi:10.20944/preprints202012.0496.v1

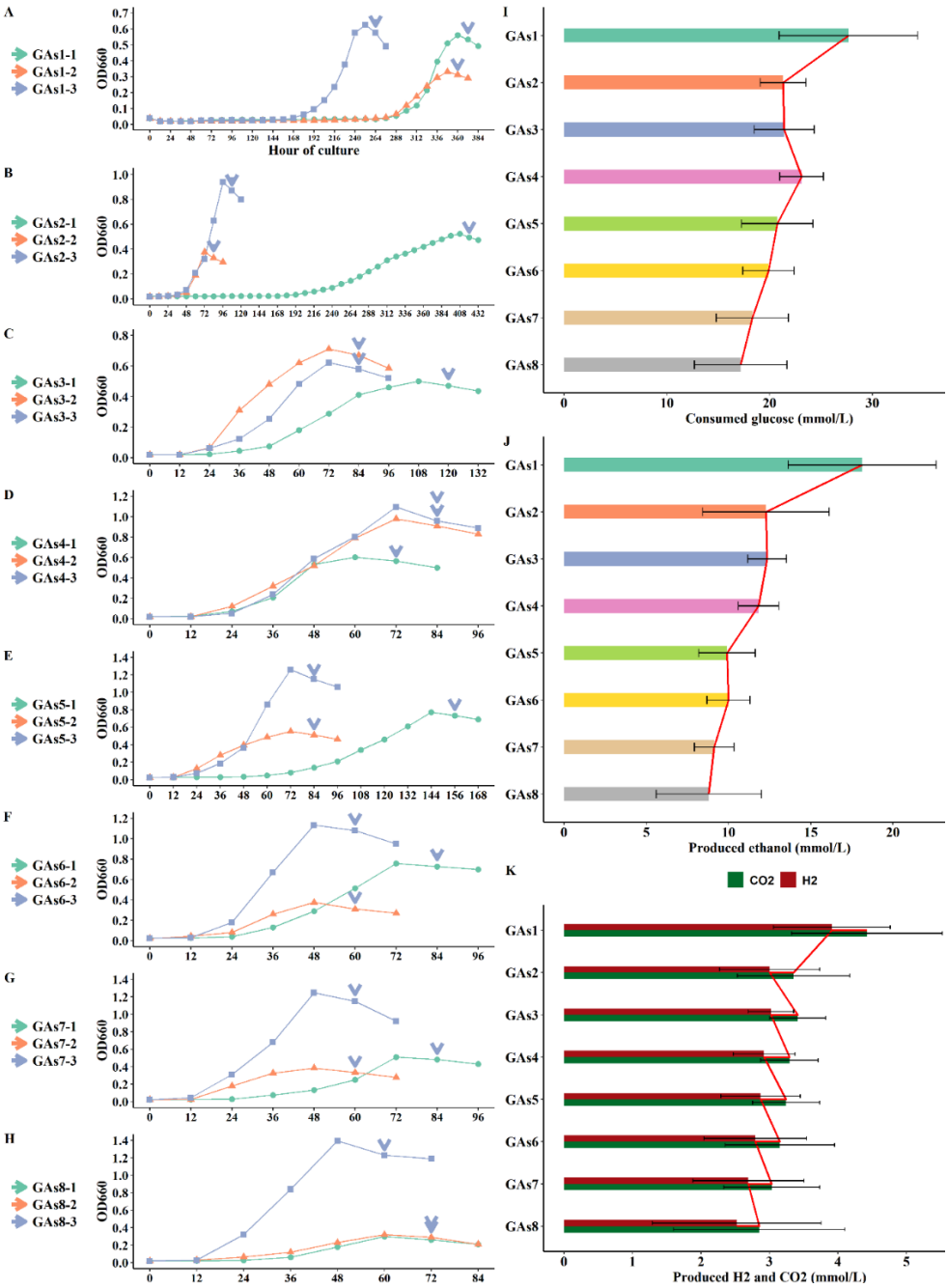


Figure 4. Growth curve of GA (left) and residual fructose, ethanol and gas production (right) after 8 sequential transfer events. Figures 4A to 4H represent the growth curves of GAs1 to GAs8, respectively. The arrowheads indicate the sampling points used for ethanol and residual reducing sugar measurements. Figure 4I shows the glucose consumed during fermentation. Figure 4J shows the ethanol production of GAs1–GAs8, and Figure 4K shows the gas production (CO_2 and H_2) of GAs1–GAs8. The ethanol and the gas production results were calculated as averages of three replicates.

3.6.2 Genomic analysis

To overcome environmental stress, the mutation rate in the bacterial genome in the late phase of the adaptation period increases to allow the development of an evolved phenotype [37]. In the present study, which required the bacteria to adapt quickly to harsh nutritional conditions, various mutations were detected in *H. thermocellum* using genomic sequencing. Intriguingly, in the genomes of the evolved strains FAs8 and GAs8, we found

 doi:10.20944/preprints202012.0496.v1

an insertion or deletion (INDEL) mutation in the MDH (Cthe_0345) gene, an INDEL & stop-gain mutation in the GCK (Cthe_0390) gene and a nonsynonymous (Nsy) mutation in the CBP (Cthe_1221) gene (Table S6). We also found various mutations in sugar transporter genes of the evolved strains. Specifically, an INDEL mutation (deletion of GG at position 19) was observed in the msdA gene (Cthe_0392) of FAs8 and GAs8. Both INDEL and stop-gain events occurred in the nbdA gene (Cthe_0391) of FAs8. The cbpB gene (Cthe_1020) in GAs had a Nsy mutation (Ala was replaced by Val^{92}). The gene encoding sugar-binding periplasmic protein (cbpC) (Cthe_2128) of FAs and GAs had a Nsy substitution (Gln was altered by Glu^{355}). Since the sugar-binding protein CbpC operon could bind glucose and cellobiose, we speculate that this mutation might improve the binding affinity of this protein to glucose and fructose. We also found a stop-gain mutation in the nbdD gene (Cthe_2447) of FAs8. Since these genes belong to the CbpA, CbpB, CbpC, and CbpD operons, which are responsible for polysaccharide G2-G5 transport, the mutations found in evolved strains might improve the transport of Fru or Glc into the cells by altering substrate specificity. In regard to cellulosomal genes, we found a synonymous (syn) mutation in OlpB (Cthe_3078) in the evolved FAs8, as A was replaced by G^{5403}. Although this is a syn mutation, its importance should be noted, as the ecological adaptation of bacteria to a new environment can also occur through synonymous changes, as they affect the nature of mRNA/tRNA interactions [68]. In addition, codon selection is common in both prokaryotes and eukaryotes and fundamentally influences the synthesis of a particular polypeptide and/or the accuracy of translation [69]. UvrABC excinuclease plays an important role in bacterial nucleotide excision repair [70]. In the present study, a Nsy mutation was found in the open reading frame of the excinuclease ABC subunit UvrB gene (Cthe_0309), where His was replaced by Arg at amino acid 402 (Table S6). An INDEL mutation (deletion of Ala at position 192 causes a frameshift variant) occurred in the DeoR transcriptional regulator (Cthe_2441) of the evolved FAs8. In *E. coli*, DeoT, a DeoR-type transcriptional regulator, represses the expression of genes involved in a variety of metabolic pathways related to maltose, fatty acid β-oxidation and peptide degradation [71]. In FAs8 and GAs8, AraC (Cthe_3164) acquired a stop mutation at position 227 (this gene is homologous to the AraC gene in *E. coli*, with 22.67% identity). In *E. coli*, AraC functions as a transcription factor regulating the expression of several genes involved in the transport and metabolism of L-arabinose [72], and regulators of the AraC/XylS family are involved in the metabolism of certain carbon sources [73].

5. Conclusions

In summary, the advances in NGS and RNA-seq technologies allowed us to gain insight into the genetic basis and dynamics of adaptation in bacterial populations. RNA-seq showed large changes in core metabolic pathways during growth on different carbon sources, and the most significant transcriptional level differences were related to the EMP pathway genes, ABC sugar transporters and CAZymes. On the other hand, genomic analysis pointed out that mutations in various genes related to carbon transport (ABC transporters) and carbon metabolism occurred in the genomes of the evolved FAs8 and GAs8 strains. These genes could be good candidates for further metabolic engineering approaches to improve biofuel production using this bacterium. This study also has enhanced our understanding of the physiology, metabolism and nutritional adaptation of *H. thermocellum*.

Supplementary Materials: Table S1: List of cellulosomal and noncellulosomal genes with differential expression, Table S2: Primers list for RT-qPCR, Table S3: Differentially expressed genes in the G cluster (FAs1, GAs1 vs CGs1), Table S4: List of four redox active pathways-related genes, Table S5: Consumed sugar, CO2, H2, cell biomass, ethanol yield of fructose- and glucose-adapted cells, Table S6: Genomic analysis of FAs8 and GAs8, Table S7: Raw counts of RNA-seq (CGs1, FAs1, GAs1), Figure S1: Expression of genes from FAs1, FAs3, FAs5 relative to CGs1, Figure S2, Figure S3: Venn

diagrams showing the commonalities between three R packages edgeR, DESeq2, NOISeq, Figure S4: Correlation between RT-qPCR and RNA-seq, Figure S5: Clusters of orthologous groups (COGs), Figure S6: Expression of sugar transporter genes using RT-qPCR, Figure: S7 Fructose-adapted cells, Figure S8: Glucose-adapted cells

Author Contributions: Conceptualization, D.M.H-T. and C.C.H.; Methodology and Investigation, D.M.H-T., S.C.L. and C.C.H.; Software, D.M.H-T. and T.T.M.N.; Data Curation, D.M.H-T., T.T.M.N. and C.C.H.; Writing–Original Draft Preparation, D.M.H-T., S.C.L. and C.C.H.; Writing–Review and Editing, D.M.H-T., T.T.M.N., and C.C.H.; Supervision, C.C. H; Funding Acquisition, S.C.L. and C.C.H. All authors read and approved the final manuscript.

Funding: The study was supported by the Ministry of Science and Technology (104–2621–M–005–003–MY3 and 107–2621–M–005–007–MY3 to C.C.H. and 107–2621–M–005–001 to S.C.L.)

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: The authors would like to deeply acknowledge Dr. John Nghiem, Adjunct Professor, Biosystems Engineering Program, Department of Environmental Engineering and Earth Sciences, Clemson University, South Carolina, USA for his thorough English editing, valuable comments and constructive suggestions.

Conflicts of Interest: The authors declare no conflict of interest.

List of abbreviations

2PGA: 2-phosphoglycerate; 3PGA: 3-phosphoglycerate; 13BPGA: 1,3-bisphosphoglycerate; A: Adenine; ALE: Adaptive laboratory evolution; BSA: Bovine serum albumin; CAZymes: Carbohydrate-active enzymes; CBM: Carbohydrate-binding module; CBP: Cellobiose phosphorylase; CGs: Cellobiose-grown cells; CMC: Carboxymethylcellulose; DE: Differential expression; DEGs: Differentially expressed genes; Fru: Fructose; FAs: Fructose-adapted cells; Fru6P: Fructose 6-phosphate; Fru16BP: Fructose 1,6-biphosphate; FBA: Fructose-bisphosphate aldolase; G: Guanine; Glc: Glucose; GAP: Glyceraldehyde 3-phosphate; GAs: Glucose-adapted cells; Glc1P: Glucose 1-phosphate; Glc6P: Glucose 6-phosphate; GCK: Glucokinase; GH: Glycoside hydrolase; GHnc: Glycoside hydrolase family "Non classified"; GA3P: Glyceraldehyde 3-phosphate; GTP: Guanosine-5'-triphosphate; ITP: Inosine triphosphate; Mal: Malate; MDH: Malate dehydrogenase; ME: Malic enzyme; nbdA: ATP-binding protein; NGS: Next-generation sequencing; OAA: Oxaloacetate; ODC: Oxaloacetate decarboxylase; PGI: Glucose-6-phosphate isomerase; PEP: Phosphoenolpyruvate; PFK: Phosphofructokinase; PGAM: Phosphoglycerate mutase; PPDK: Pyruvate phosphate dikinase; PYK: Pyruvate kinase

References

1. Koeck, D.; Koellmeier, T.; Zverlov, V.; Liebl, W.; Schwarz, W. Differences in biomass degradation between newly isolated environmental strains of Clostridium thermocellum and heterogeneity in the size of the cellulosomal scaffoldin. *Syst. Appl. Microbiol.* 2015, *38*, 424–432, doi:http://dx.doi.org/10.1016/j.syapm.2015.06.005.

2. Paye, J.; Guseva, A.; Hammer, S.; Gjersing, E.; Davis, M.; Davison, B.; Olstad, J.; Donohoe, B.; Nguyen, T.; Wyman, C.; et al. Biological lignocellulose solubilization: comparative evaluation of biocatalysts and enhancement via cotreatment. *Biotechnol. Biofuels* 2016, *9*, doi:DOI 10.1186/s13068-015-0412-y.

3. Xu, Q.; Resch, M.; Podkaminer, K.; Yang, S.; Baker, J.; Donohoe, B.; Wilson, C.; Klingeman, D.; Olson, D.; Decker, S.; et al. Dramatic performance of Clostridium thermocellum explained by its wide range of cellulase modalities. *Sci. Adv* 2016, *2*, doi:e1501254.

4. Dror, T.; Morag, E.; Rolider, A.; Bayer, E.; Lamed, R.; Shoham, Y. Regulation of the cellulosomal celS (cel48A) gene of Clostridium thermocellum is growth rate dependent. *J. Bacteriol.* 2003, *185*, 3042–3048, doi:DOI: 10.1128/JB.185.10.3042–3048.2003.

5. Dror, T.; Rolider, A.; Bayer, E.; Lamed, R.; Shoham, Y. Regulation of expression of scaffoldin-related genes in Clostridium thermocellum. *J. Bacteriol.* 2003, *185*, 5109–5116, doi:DOI: 10.1128/JB.185.17.5109–5116.2003.

6. Kahel-Raifer, H.; Jindou, S.; Bahari, L.; Nataf, Y.; Shoham, Y.; Bayer, E.; Borovok, I.; Lamed, R. The unique set of putative membrane-associated anti-δ factors in Clostridium thermocellum suggests a novel extracellular carbohydrate-sensing mechanism involved in gene regulation. *FEMS Microbiol. Lett.* 2010, *308*, 84–93, doi:DOI:10.1111/j.1574-6968.2010.01997.x.

7. Bahari L; Gilad, Y.; Borovok, Y.; Kahel-Raifer, H.; Dassa, B.; Nataf, Y.; Shoham, Y.; Lamed, R.; Bayer EA Glycoside hydrolases as components of putative carbohydrate biosensor proteins in Clostridium thermocellum. *Springer* 2011, *38*, 825–832, doi:10.1007/s10295-010-0848-9.

8. Sander, K.; Wilson, C.; Rodriguez Jr, M.; Klingeman, D.; Rydzak, T.; Davison, B.; Brown, S. Clostridium thermocellum DSM 1313 transcriptional responses to redox perturbation. *Biotechnol. Biofuels* 2015, *8*, doi:DOI 10.1186/s13068-015-0394-9.

9. Choi, J.; Klingeman, D.; Brown, S.; Cox, C. The LacI family protein GlyR3 co-regulates the celC operon and manB in Clostridium thermocellum. *Biotechnol. Biofuels* 2017, *10*, 1–11, doi:DOI 10.1186/s13068-017-0849-2.

10. de Ora, L.; Lamed, R.; Liu, Y.; Xu, J.; Cui, Q.; Feng, Y.; Shoham, Y.; Bayer, E.; Munoz-Gutierrez, I. Regulation of biomass degradation by alternative δ factors in cellulolytic clostridia. *Nature* 2018, *8*, 1–11, doi:10.1038/s41598-018-29245-5.

11. Linville, J.; Rodriguez Jr, M.; Land, M.; Syed, M.; Engle, N.; Tschaplinski, T.; Mielenz, J.; Cox, C. Industrial Robustness: Understanding the mechanism of tolerance for the populus hydrolysate-tolerant mutant strain of Clostridium thermocellum. *Plos ONE* 2013, *8*, 1–16, doi:e78829.

12. Holwerda, E.; Olson, D.; Ruppertsberger, N.; Stevenson, DM; Murphy, S.; Maloney, M.; Lanahan, A.; Amador-Noguez, D.; Lynd, L. Metabolic and evolutionary responses of Clostridium thermocellum to genetic interventions aimed at improving ethanol production. *Biotechnol. Biofuels* 2020, *13*, 1–20, doi:https://doi.org/10.1186/s13068-020-01680-5.

13. Raman, B.; McKeown, C.; Rodriguez Jr, M.; Brown, S.; Mielenz, J. Transcriptomic analysis of Clostridium thermocellum ATCC 27405 cellulose fermentation. *BMC Microbiol.* 2011, *11*, doi:http://www.biomedcentral.com/1471-2180/11/134.

14. Wilson, C.; Yang, S.; Rodriguez Jr, M.; Ma, Q.; Johnson, C.; Dice, L.; Xu, Y.; Brown, S. Clostridium thermocellum transcriptomic profiles after exposure to furfural or heat stress. *Biotechnol. Biofuels* 2013, *6*, doi:http://www.biotechnologyforbiofuels.com/content/6/1/131.

15. Wilson, C.; Rodriquez Jr, M.; Johnson, C.; Martin, S.; Chu, T.; Wolfinger, R.; Hauser, L.; Land, M.; Klingeman, D.; Syed, M.; et al. Global transcriptome analysis of Clostridium thermocellum ATCC 27405 during growth on dilute acid pretreated Populus and switchgrass. *Biotechnol. Biofuels* 2013, *6*, 1–18, doi:http://www.biotechnologyforbiofuels.com/content/6/1/179.

16. Linville, J.; Rodriguez Jr, M.; Brown, S.; Mielenz, J.; Cox, C. Transcriptomic analysis of Clostridium thermocellum Populus hydrolysate-tolerant mutant strain shows increased cellular efficiency in response to Populus hydrolysate compared to the wild type strain. *BMC Microbiol.* 2014, *14*, 1–17, doi:http://www.biomedcentral.com/1471-2180/14/215.

17. Wei, H.; Fu, Y.; Magnusson, L.; Baker, J.; Maness, P.; Xu, Q.; Yang, S.; Bowersox, A.; Bogorad, L.; Wang, W.; et al. Comparison of transcriptional profiles of Clostridium thermocellum grown on cellobiose and pretreated yellow poplar using RNA-seq. *Front. Microbiol.* 2014, *5*, doi:doi: 10.3389/fmicb.2014.00142.

18. Tian, L.; Perot, S.; Stevenson, D.; Jacobson, T.; Lanahan, A.; Amador-Noguez, D.; Olson, D.; Lynd, L. Metabolome analysis reveals a role for glyceraldehyde 3-phosphate dehydrogenase in the inhibition of C. thermocellum by ethanol. *Biotechnol. Biofuels* 2017, *10*, doi:DOI 10.1186/s13068-017-0961-3.

19. Xiong, W.; Lo, J.; Chou, K.; Wu, C.; Magnusson, L.; Dong, T.; Maness, P. Isotope-assisted metabolite analysis sheds light on central carbon metabolism of a model cellulolytic bacterium Clostridium thermocellum. *Front. Microbiol.* 2018, *9*, doi:doi: 10.3389/fmicb.2018.01947.

20. Gold, N.; Martin, V.J. Global view of the Clostridium thermocellum cellulosome revealed by quantitative proteomic analysis. *J. Bacteriol.* 2007, *189*, 6787–6795, doi:doi:10.1128/JB.00882-07.

21. Raman, B.; Pan, C.; Hurst, G.; Rodriguez Jr, M.; McKeown, C.; Lankford, P.; Samatova, N.; Mielenz, J. Impact of pretreated switchgrass and biomass carbohydrates on Clostridium thermocellum ATCC 27405 cellulosome composition: A quantitative proteomic analysis. *Plos ONE* 2009, *4*, 1–13, doi:e5271.

22. Burton, E.; Martin, V.J. Proteomic analysis of Clostridium thermocellum ATCC 27405 reveals the upregulation of an alternative transhydrogenase-malate pathway and nitrogen assimilation in cells grown on cellulose. *NRC Res. Press* 2012, *58*, 1378–1388, doi:10.1139/cjm-2012-0412.

23. Rydzak, T.; McQueen, P.; Krokhin, O.; Spicer, V.; Ezzati, P.; Dwivedi, R.; Shamshurin, D.; Levin, D.; Wilkins, J.; Sparling, R. Proteomic analysis of Clostridium thermocellum core metabolism: relative protein expression profiles and growth phase-dependent changes in protein expression. *BMC Microbiol.* 2012, *12*, 1–18, doi:http://www.biomedcentral.com/1471-2180/12/214.

24. Yang, S.; Giannone, R.; Dice, L.; Yang, Z.; Engle, N.; Tschaplinski, T.; Hettich, R.; Brown, S. Clostridium thermocellum ATCC 27405 transcriptomic, metabolomic and proteomic profiles after ethanol stress. *BMC Microbiol.* 2012, *13*, 1–17, doi:http://www.biomedcentral.com/1471-2164/13/336.

25. Dumitrache, A.; Klingeman, D.; Natzke, J.; Rodriguez Jr, M.; Giannone, R.; Hettich, R.; Davison, B.; Brown, S. Specialized activities and expression differences for Clostridium thermocellum biofilm and planktonic cells. *Sci. Rep.* 2017, *7*, doi:DOI: 10.1038/srep43583.

26. Poudel, S.; Giannone, R.; Rodriguez Jr, M.; Raman, B.; Martin, M.; Engle, N.; Mielenz, J.; Nookaew, I.; Brown, S.; Tschaplinski, T.; et al. Integrated omics analyses reveal the details of metabolic adaptation of Clostridium thermocellum to lignocellulose-derived growth inhibitors released during the deconstruction of switchgrass. *Biotechnol. Biofuels* 2017, *10*, doi:DOI 10.1186/s13068-016-0697-5.

27. Whitham, J.; Moon, J.; Rodriguez Jr, M.; Engle, N.; Klingeman, D.; Rydzak, T.; Abel, M.; Tschaplinski, T.; Guss, A.; Brown, S. Clostridium thermocellum LL1210 pH homeostasis mechanisms informed by transcriptomics and metabolomics. *Biotechnol. Biofuels* 2018, *11*, 1–13, doi:https://doi.org/10.1186/s13068-018-1095-y.

28. Stevenson, DM; Weimer, P. Expression of 17 genes in Clostridium thermocellum ATCC 27405 during fermentation of cellulose or cellobiose in continuous culture. *Appl. Environ. Microbiol. 71*, 4672–4678, doi:10.1128/AEM.71.8.4672–4678.2005.

29. Zhang, Y.; Lynd, L. Regulation of cellulase synthesis in batch and continuous cultures of Clostridium thermocellum. *J. Bacteriol.* 2005, *187*, 99–106, doi:10.1128/JB.187.1.99–106.2005.

30. Johnson, E.; Bouchot, F.; Demain, A. Regulation of cellulase formation in Clostridium thermocellum. *J. Gen. Microbiol.* 1985, *131*, 2303–2308.

31. Nochur, S.; Roberts, M.; Demain, A. Mutation of Clostridium thermocellum in the presence of certain carbon sources. *FEMS Microbiol. Lett.* 1990, *71*, 199–204.

32. Nochur, S.; Demain, A.; Roberts, M. Carbohydrate utilization by Clostridium thermocellum: Importance of internal pH in regulating growth. *Enzyme Microb. Technol.* 1992, *14*, 338–349.

33. Patni, N.; Alexander, J. Utilization of glucose by Clostridium thermocellum: Presence of glucokinase and other glycolytic enzymes in cell extracts. *J. Bacteriol.* 1971, *105*, 220–225.

34. Hernandez, P. Transport of D-glucose in Clostridium thermocellum ATCC 27405. *J. Gen. Appl. Microbiol.* 1982, *28*, 469–477.

35. Strobel, H.; Caldwell, F.; Dawson, K. Carbohydrate transport by the anaerobic thermophile Clostridium thermocellum LQRI. *Appl. Environ. Microbiol.* 1995, *61*, 4012–4015.

36. Dragosits, M.; Mattanovich, D. Adaptive laboratory evolution-principles and applications for biotechnology. *Microb. Cell Factories* 2013, *12*, 1–17, doi:http://www.microbialcellfactories.com/content/12/1/64.

37. Conrad, T.; Lewis, N.; Palsson, B. Microbial laboratory evolution in the era of genome-scale science. *Mol. Syst. Biol.* 2011, *7*, doi:doi:10.1038/msb.2011.42.

38. Zhang, Y.; Lynd, L. Quantification of cell and cellulase mass concentration during anaerobic cellulose fermentation: Development of an Enzyme-Linked Immunosorbent Assay-Based Method with application to Clostridium thermocellum batch cultures. *Anal. Chem.* 2003, *75*, 219–227.

39. Xiong, W.; Lin, P.; Magnusson, L.; Warner, L.; Liao, J.; Maness, P.; Chou, K. CO2-fixing one-carbon metabolism in a cellulose-degrading bacterium Clostridium thermocellum. *Proc. Natl. Acad. Sci. U. S. A.* 2016, *113*, 13180–13185, doi:www.pnas.org/cgi/doi/10.1073/pnas.1605482113.

40. Johnson, E.; Madia, A.; Demain, A. Chemically defined minimal medium for growth of the anaerobic cellulolytic thermophile Clostridium thermocellum. *Appl. Environ. Microbiol.* 1981, *41*.

41. Zhang, Y.; Cui, J; Lynd, L.; Kuang, L. A transition from cellulose swelling to cellulose dissolution by o-Phosphoric acid: Evidence from enzymatic hydrolysis and supramolecular structure. *Biomacromolecules 7*, 644–648, doi:10.1021/bm050799c.

42. St Brice, L.; Shao, X.; Izquierdo, J.; Lynd, L. Optimization of affinity digestion for the isolation of cellulosomes from Clostridium thermocellum. *Prep. Biochem. Biotechnol.* 2014, *44*, 206–216, doi:10.1080/10826068.2013.829494.

43. Zhang, Y.; Hong, J.; Ye, X. Cellulase Assays. In *Biofuels-Methods and Protocols*; Mielenz, J., Ed.; Methods in Molecular Biology; Humana Press, 2009; pp. 213–231 ISBN 978-1-60761-213-1.

44. Ribeiro, L.; De Lucas, R.; Vitcosque, G.; Ribeiro, L.; Ward, R.; Rubio, M.; Damasio, A.; Squina, F.; Gregory, R.; Walton, P.; et al. A novel thermostable xylanase GH10 from Malbranchea pulchella expressed in Aspergillus nidulans with potential applications in biotechnology. *Biotechnol. Biofuels 7*, doi:http://www.biotechnologyforbiofuels.com/content/7/1/115.

45. Miller, G. Use of dinitrosalicylic acid reagent for determination of reducing sugar. *Anal. Chem.* 1959, *31*.

46. Saqib, A.; Whitney, P. Differential behaviour of the dinitrosalicylic acid (DNS) reagent towards mono- and di-saccharide sugars. *Biomass Bioenergy*, doi:doi:10.1016/j.biombioe.2011.09.013.

47. Oliveros, J. Venny. An interactive tool for comparing lists with Venn's diagrams. *https://bioinfogp.cnb.csic.es/tools/venny/index.html* 2015, doi:https://bioinfogp.cnb.csic.es/tools/venny/index.html.

48. Robinson, M.; McCarthy, D.; Smyth, G. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 2010, *26*, 139–140, doi:DOI: 10.1093/bioinformatics/btp616.

49. Love, M.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 2014, *15*, 1–21, doi:DOI 10.1186/s13059-014-0550-8.

50. Tarazona, S.; Furio-Tari, P.; Turra, D.; Di Pietro, A.; Nueda, M.; Ferrer, A.; Conesa, A. Data quality aware analysis of differential expression in RNA-seq with NOISeq R/Bioc package. *Nucleic Acids Res.* 2015, *43*, doi:10.1093/nar/gkv711.

51. Kvam, V.; Liu, P.; Si, Y. A comparison of statistical methods for detecting differentially expressed genes from RNA-seq data. *Am. J. Bot.* 2012, *99*, 248–256, doi:doi:10.3732/ajb.1100340.

52. Maza, E. In Papyro comparison of TMM (edgeR), RLE (DESeq2), and MRN normalization methods for a simple two-conditions-without-replicates RNA-seq experimental design. *Front. Genet.* 2016, *7*, doi:doi: 10.3389/fgene.2016.00164.

53. Johnson, E.; Sakajoh, M.; Halliwell, G.; Madia, A.; Demain, A. Saccharification of complex cellulosic substrates by the cellulase system from Clostridium thermocellum. *Appl. Environ. Microbiol.* 1982, *43*, 1125–1132, doi:0099-2240/82/051125-08$02.00/0.

54. Bayer, E.; Shoham, Y.; Lamed, R. The cellulosome: an exocellular organelle for degrading plant cell wall polysaccharides. In *Glycomicrobiology*; Doyle, R., Ed.; Kluwer Academic/Plenum Publishers, 2000; pp. 387–439.

55. Nataf, Y.; Bahari, L.; Kahel-Raifer, H.; Borovok, I.; Lamed, R.; Bayer, E.; Sonenshein, A.; Shoham, Y. Clostridium thermocellum cellulosomal genes are regulated by extracytoplasmic polysaccharides via alternative sigma factor. *Proc. Natl. Acad. Sci. U. S. A.* 2010, *107*, 18646–18651, doi:www.pnas.org/cgi/doi/10.1073/pnas.1012175107.

56. Riederer, A.; Takasuka, T.; Makino, S.; Stevenson, D.; Bukhman, Y.; Elsen, N.; Fox, B. Global gene expression patterns in Clostridium thermocellum as determined by microarray analysis of chemostat cultures on cellulose or cellobiose. *Appl. Environ. Microbiol.* 2011, *77*, 1243–1253, doi:doi:10.1128/AEM.02008-10.

57. Waeonukul, R.; Kosugi, A.; Tachaapaikoon, C.; Pason, P.; Ratanakhanokchai, K.; Prawitwong, P.; Deng, L.; Saito, M.; Mori, Y. Efficient saccharification of ammonia soaked rice straw by combination of Clostridium thermocellum cellulosome and Thermoanaerobacter brockii β-glucosidase. *Bioresour. Technol.* 2012, *107*, 352–357, doi:10.1016/j.biortech.2011.12.126.

58. Sand, A.; Holwerda, E.; Ruppertsberger, N.; Maloney, M.; Olson, D.; Nataf, Y.; Borovok, Y.; Sonenshein, A.; Bayer, E.; Lamed, R.; et al. Three cellulosomal xylanase genes in Clostridium thermocellum are regulated by both vegetative SigA (δA) and alternative SigI6 (δI6) factors. *FEBS Lett.* 2015, *589*, 3133–3140, doi:http://dx.doi.org/10.1016/j.febslet.2015.08.026.

59. Zhang, P.; Wang, B.; Xiao, Q.; Wu, S. A kinetics modeling study on the inhibition of glucose on cellulosome of Clostridium thermocellum. *Bioresour. Technol.* 2015, *190*, 36–43, doi:http://dx.doi.org/10.1016/j.biortech.2015.04.037.

60. Verbeke, T.; Garcia, G.; Elkins, J. The effect of switchgrass loadings on feedstock solubilization and biofuel production by Clostridium thermocellum. *Biotechnol. Biofuels* 2017, *10*, 1–9, doi:DOI 10.1186/s13068-017-0917-7.

61. Kim, S.; Himmel, M.; Bomble, Y.; Westpheling, J. Expression of a cellobiose phosphorylase from Thermotoga maritima in Caldicellulosiruptor bescii improves the phosphorolytic pathway and results in a dramatic increase in cellulolytic activity. *Appl. Environ. Microbiol.* 2018, *84*, e02348-17, doi:https://doi .org/10.1128/AEM.02348-17.

62. Zhou, J.; Olson, D.; Argyros, D.; Deng, Y.; van Gulik, W.; van Dijken, J.; Lynd, L. Atypical glycolysis in Clostridium thermocellum. *Appl. Environ. Microbiol.* 2013, *79*, 3000–3008, doi:10.1128/AEM.04037-12.

63. Olson, D.; Horl, M.; Fuhrer, T.; Cui, J.; Zhou, J.; Maloney, M.; Amador-Noguez, D.; Tian, L.; Sauer, U.; Lynd, L. Glycolysis without pyruvate kinase in Clostridium thermocellum. *Metab. Eng.* 2017, *39*, 169–180, doi:http://dx.doi.org/10.1016/j.ymben.2016.11.011.

64. Ng, T.; Zeikus, J. Differential metabolism of cellobiose and glucose by Clostridium thermocellum and Clostridium thermohydrosulfuricum. *J. Bacteriol.* 1982, *150*, 1391–1399.

65. Nataf, Y.; Yaron, S.; Stahl, F.; Lamed, R.; Bayer, E.; Helmut, S.; Sonenshein, A.; Shoham, Y. Cellodextrin and Laminaribiose ABC Transporters in Clostridium thermocellum. *J. Bacteriol.* 2009, *191*, 203–209, doi:doi:10.1128/JB.01190-08.

66. Thompson, R.; Layton, D.; Guss, A.; Olson, D.; Lynd, L.; Trinh, C. Elucidating central metabolic redox obstacles hindering ethanol production in Clostridium thermocellum. *Metab. Eng.* 2015, *32*, 207–219, doi:DOI: 10.1016/j.ymben.2015.10.004.

67. Biswas R; Wilson, C.; Giannone, R.; Klingman, D.; Rydzak, T.; Shah, M.; Hettich, R.; Brown, S.; Guss, A. Improved growth rate in Clostridium thermocellum hydrogenase mutant via perturbed sulfur metabolism. *Biotechnol. Biofuels* 2017, *10*, 1–10, doi:10.1186/s13068-016-0684-x.

68. Retchless, A.; Lawrence, J. Ecological adaptation in bacteria: Speciation driven by codon selection. *Oxf. Univ. Press* 2012, *29*, 3669–3683, doi:10.1093/molbev/mss171.

69. Plotkin, J.; Kudla, G. Synonymous but not the same: the causes and consequences of codon bias. *Macmillan Publ. Ltd.* 2011, *12*, 32–42, doi:10.1038/nrg2899.

70. Thakur, M.; Badugu, S.; Muniyappa, K. UvrA and UvrC subunits of the Mycobacterium tuberculosis UvrABC excinuclease interact independently of UvrB and DNA. *FEBS Lett.* 2020, *594*, 851–863.

71. Elgrably-Weiss, M.; Schlosser-Silverman, E.; Rosenshine, I.; Altuvia, S. DeoT, a DeoR-type transcriptional regulator of multiple target genes. *FEMS Microbiol. Lett.* 2006, *254*, 141–148, doi:doi:10.1111/j.1574-6968.2005.00020.x.

72. Soisson, S.; MacDougall-Shackleton, B.; Schleif, R.; Wolberger, C. The 1.6 Ă crystal structure of the AraC sugar-binding and dimerization domain complexed with D-fucose. *J. Mol. Biol.* 1997, *273*, 226–237.

73. Gallegos, M.; Michan, C.; Ramos, L. The XylS/AraC family of regulators. *Nucleic Acids Res.* 1993, *21*, 807–810.