Applied Microbiology and Biotechnology (2024) 108:477
https://doi.org/10.1007/s00253-024-13310-3

APPLIED GENETICS AND MOLECULAR BIOTECHNOLOGY

Extracytoplasmic polysaccharides control cellulosomal and non-cellulosomal systems in *Herbivorax saccincola* A7

Sirilak Baramee[1,2] · Pakinee Thianheng[1,3] · Ayaka Uke[1] · Pattsarun Cheawchanlertfa[1,3] · Chakrit Tachaapaikoon[2,3] · Rattiya Waeonukul[2,3] · Patthra Pason[2,3] · Khanok Ratanakhanokchai[3] · Ya-Jun Liu[4,5,6] · Akihiko Kosugi[1]

Received: 29 June 2024 / Revised: 9 September 2024 / Accepted: 10 September 2024 / Published online: 1 October 2024

Abstract

Herbivorax saccincola A7 is an anaerobic alkali-thermophilic lignocellulolytic bacterium that possesses a cellulosome and high xylan degradation ability. To understand the expression profile of extracellular enzymes by carbon sources, quantitative real-time PCR was performed on all cellulosomal and non-cellulosomal enzyme genes of *H. saccincola* A7 using cellulose and xylan as carbon sources. The results confirmed that the scaffolding proteins of *H. saccincola* A7 were expressed. In general, the cellulosomal genes belonging to the glycoside hydrolase families 9, 10, 11, and 48 were repressed when xylan was the sole carbon source, but these genes were significantly induced in the presence of cellulose. These results indicate that cellulose, not xylan, is a key inducer of cellulosomal genes in *H. saccincola* A7. The RsgI-like proteins, which regulate a carbohydrate-sensing mechanism in *Clostridium thermocellum*, were also found to be encoded in the *H. saccincola* A7 genome. To confirm the regulation by RsgI-like proteins, the relative expression of σI1–σI4 factors was analyzed on both carbon sources. The expression of alternative σI1 and σI2 factors was enhanced by the presence of cellulose. By contrast, the expression of σI3 and σI4 factors was activated by both cellulose and xylan. Taken together, the results reveal that the cellulosomal and non-cellulosomal genes of *H. saccincola* A7 are regulated through a carbohydrate-sensing mechanism involving anti-σ regulator RsgI-like proteins.

Key points

- *qRT-PCR performed on cellulosomal and non-cellulosomal genes of H. saccincola A7*
- *Cellulose is a key inducer of the cellulosome of H. saccincola A7*
- *H. saccincola A7 possesses a similar system of anti-σ regulator RsgI-like proteins*

Keywords *Herbivorax saccincola* · *Acetivibrio saccincola* · Cellulosome · Non-cellulosome · Carbohydrate-sensing mechanism · Lignocellulose

Introduction

Lignocellulosic biomasses such as rice straw, corn stover, and empty fruit bunches are some of the most abundant renewable polysaccharide resources that have great potential for biofuel and biochemical conversion. The polysaccharides within lignocellulosic biomass mainly comprise cellulose and hemicellulose (Sun and Cheng 2002).

The cellulosome, a multienzyme complex produced by anaerobic cellulolytic bacteria such as *Clostridium* spp. (including *C. thermocellum*, *C. cellulolyticum*, and *C. cellulovorans*), is efficient at degrading plant biomass (Bayer et al. 2004; Lynd et al. 2002). The cellulosome of *C. thermocellum* is well characterized and consists of a large (197 kDa), non-catalytic, multimodular scaffolding protein CipA, which includes nine cohesins, four hydrophilic modules, a family 3 carbohydrate-binding module (CBM3), and more than 40 cellulosomal enzymes (Bayer et al. 2004). The catalytic units of the cellulosomal enzyme are non-covalently attached to CipA via high-affinity type I interactions between the dockerin domains of the catalytic units and the cohesins on CipA. Biological saccharification and consolidated bioprocessing using *C. thermocellum* with the cellulosome system has been shown to be cost-effective and useful for lignocellulosic biomass degradation (Bayer et al. 2004; Demain et al. 2005; Lynd et al. 2002).

Regulation studies in *C. thermocellum* (Dror et al. 2003; Raman et al. 2011; Riederer et al. 2011; Stevenson and Weimer 2005) demonstrated that the expression level and

Extended author information available on the last page of the article

the composition of the cellulosomal proteins vary with the availability of the carbon source (cellobiose and cellulose) and the presence of extracellular plant cell wall-derived polysaccharides. *C. thermocellum* encodes and regulates many hemicellulase genes; however, it cannot utilize the end-products such as pentose, because it lacks the essential genes for the pentose catabolism pathway (Demain et al. 2005; Lynd et al. 2002). Recently, Kahel-Raifer et al. and Nataf et al. (Kahel-Raifer et al. 2010; Nataf et al. 2010) identified a novel carbohydrate regulatory system in *C. thermocellum* that allows this organism to sense and react to the presence of high molecular weight polysaccharides in the extracellular environment without importing their low molecular weight degradation products. The majority of cellulosomal genes of *C. thermocellum* are regulated by a carbohydrate sensory system (Nataf et al. 2010). In this system, the different components of the lignocellulosic biomass would be extracellularly detected by RsgI-borne binding elements such as cellulose binding modules (CBMs) and glycoside hydrolases (GHs), and appropriate signals are intracellularly transmitted (Kahel-Raifer et al. 2010; Nataf et al. 2010). Thus, it would be useful to study the comprehensive regulation of cellulosomal or non-cellulosomal enzyme expression in hemicellulose-assimilating cellulolytic thermophilic bacteria other than *C. thermocellum*, such as *Herbivorax saccincola* (*Acetivibrio saccincola*), to understand their lignocellulosic biomass degradation mechanisms.

H. saccincola A7, an anaerobic alkali-thermophilic bacterium, was newly identified from bovine manure compost (Aikawa et al. 2018). *H. saccincola* A7 can utilize a wide range of carbon sources, including cellulose, xylan, and natural plant biomass (Aikawa et al. 2020). The high-quality draft genome sequence revealed that this organism produces a cellulosome-like multienzyme complex, containing many enzymatic subunits (Aikawa et al. 2020). However, the mechanism by which *H. saccincola* A7 flexibly responds to different polysaccharides, other than cellulose, and regulates the expression of biomass degradation-related genes is unclear.

In this study, a quantitative real-time PCR (qRT-PCR) technique was carried out to demonstrate the expression profiles of the genes encoding cellulosomal scaffold protein and GHs on cellulose, xylan, or both cellulose and xylan. Cellulose, as an extracellular polysaccharide, is the main inducer for cellulosomal and non-cellulosomal enzyme subunits of *H. saccincola* A7. We also confirmed the presence of an external carbohydrate-sensing system in *H. saccincola* A7, similar to that found in *C. thermocellum*. These results provide valuable insight into the biomass degradation mechanism of *H. saccincola* A7, a hemicellulose-assimilating bacterium with a different sugar metabolism from that of *C. thermocellum*.

Materials and methods

Substrates and carbon source preparation

Carboxymethyl cellulose (CMC, low viscosity) and microcrystalline cellulose powder (Sigmacell Type 20) were purchased from Sigma-Aldrich (St. Louis, MO, USA). Oat-spelt xylan was purchased from Tokyo Chemical Industry Co., Ltd. (TCI, Tokyo, Japan). Cellobiose was purchased from FUJIFILM Wako Pure Chemicals. Phosphoric acid swollen cellulose (PASC) was prepared from Sigmacell Type 20, as previously described by Walseth (1952).

Microorganisms and culture media

H. saccincola A7 was isolated from the cellulose-degrading bacterial community inhabiting bovine manure compost on Ishigaki Island, Japan (Aikawa et al. 2018). *Clostridium thermocellum* ATCC 27405^T was obtained from the American Type Culture Collection (ATCC) (Nhim et al. 2022). Basal medium (BM) was prepared with the following composition per liter: 2.9 g K_2HPO_4, 1.5 g KH_2PO_4, 2.1 g urea, 6 g yeast extract, 0.01 g $CaCl_2{\cdot}2H_2O$, 2.5 g Na_2CO_3, 0.5 g L-cysteine-HCl, 0.5 mg resazurin, and 200 µL mineral solution (25.0 g/L $MgCl_2{\cdot}6H_2O$, 0.312 g/L $FeSO_4{\cdot}7H_2O$, and 37.5 g/L $CaCl_2{\cdot}2H_2O$) and was distributed into Hungate tubes (Bellco Glass, Inc., Vineland, NJ, USA) and/or serum bottles. BM containing each carbon source was degassed by heating in boiling water, followed by bubbling with high-purity N_2 (BMN), and high-purity CO_2 (BM7CO) to create anaerobic conditions (Nhim et al. 2022), before being autoclaved.

Enzyme production and enzyme activities

H. saccincola A7 and *C. thermocellum* ATCC 27405^T were cultured in BMN or BM7CO medium supplemented with a 1% carbon source for 4 days, respectively. The culture supernatant was collected by centrifugation and affinity was concentrated by mixing the culture supernatant with 1% PASC in 0.15 M sodium chloride and stirring gently at 4 °C overnight. Washing and elution steps were performed using centrifugation at 10,019 × *g* and 4 °C for 10 min. The PASC-binding proteins were washed twice with cooled phosphate-buffered saline (pH 7.0) and then eluted with 0.25% sucrose. The total volume of eluent was concentrated using 5000 MWCO PES-ultrafiltration Vivaspin columns (Sartorius, Göttingen, Germany). Desalting was performed using Econo-Pac® 10 DG columns (Bio-Rad Laboratories, Hercules, CA, USA). Protein concentration was measured by the BCA method (Wiechelman et al. 1988) using bovine serum albumin as a standard. Enzyme activities were

determined using 100 µg of the PASC-binding protein mixed with the substrate in 0.1 M sodium phosphate buffer, pH 7.0, and incubated at 60 °C for 10 min. One percent (w/v) PASC and oat-spelt xylan were used as substrates for PASCase and xylanase activities, respectively. The release of reducing sugars was determined by the Somogyi–Nelson method (Nelson 1944) using xylose or glucose as the standard. One unit (U) of enzyme activity was defined as the amount of enzyme required to release 1 µmol of reducing sugar per minute under the assay conditions.

Growth quantitation

The growth quantitation method used was adapted from the method described by Tang et al. (2015). This method is unaffected by the mixing of insoluble substrates such as cellulose and xylan, and it provides a sensitive and reproducible measure of the growth of cultured bacteria. *H. saccincola* A7 was cultured in the BMN medium in which a cellobiose, cellulose, xylan, and cellulose-xylan (1:1) mixture at a concentration of 1% (w/v) was used as the carbon source. A 5-mL aliquot was sampled at 6, 12, 24, 48, 72, and 96 h, and then cells were harvested by centrifugation at 10,000 rpm and stored at – 20 °C for subsequent DNA extraction. The protein content in the culture supernatant was determined as described above. DNAs were isolated using the phenol–chloroform extraction method and the purity and concentration were verified using a NanoDrop spectrophotometer (ND-ONE-W, Thermo Fisher Scientific, Waltham, MA, USA). To generate a standard curve of the 16S rRNA gene, PCR was performed to amplify the 16S rRNA domain of *H. saccincola* A7 using the primers listed in Supplementary Table S1. The PCR product (103 bp) was used to calculate the copy number of the 16S rRNA gene using the equation indicated below and was then subjected to tenfold serial dilution from 10^2 to 10^8 copies.

$$\text{Copy number per } \mu\text{L} = \frac{[\text{DNA concentration (ng}/\mu\text{L}) \times 10^{-9}\ \text{(g/ng)}] \times [6.02 \times 10^{23}\ (\text{mol}^{-1})]}{[\text{gene size (bp)} \times 660\ \text{(Dalton/bp)}]}$$

Total RNA extraction and cDNA synthesis

Bacterial cultures with cellobiose, cellulose, xylan, or the cellulose and xylan mixture as a carbon source were harvested at mid-log phase, treated with RNA Protect Bacteria reagent (Qiagen, Hilden, Germany) immediately according to the manufacturer's instructions, and stored at − 80 °C until extraction. Total RNA was extracted using the RNeasy mini kit (Qiagen), including on-column DNase treatment. The quality of RNA was verified from the A260/280 (1.9 to 2.1) readings using a NanoDrop spectrophotometer. The cDNA was synthesized from 40 ng of RNA using the QuantiTect Reverse Transcription Kit (Qiagen). The cDNA was diluted 10 times and stored at − 20 °C prior to use.

Primer design

Primers used in this study were designed using the Primer-BLAST primer design tools on the National Center for Biotechnology Information (NCBI) website (https://www.ncbi.nlm.nih.gov/tools/primer-blast/). The primers were synthesized by Hokkaido System Science Co., Ltd. (Hokkaido, Japan). The synthetic primers were diluted in sterilized MilliQ water to the appropriate concentration before use. The primer sequences are listed in Table S1.

Analysis of qRT-PCR and candidate gene expression

The qRT-PCR reaction, in a total volume of 10 µL, was performed using the CFX96 Touch Real-Time PCR Detection System and SsoAdvanced™ Universal SYBR® Green Supermix (Bio-Rad Laboratories) with the following conditions: denaturation at 95 °C for 30 s, followed by 40 cycles of denaturation at 95 °C for 5 s and annealing at 58 °C for 10 s. A final extension step was conducted at 72 °C for 5 min. To generate the standard curve, the cycle threshold (Ct) of each standard sample was plotted against the log of the copy number of the 16S rRNA gene. The growth curve for each sample was generated using qRT-PCR and the conditions stated above. The cycle threshold (Ct) of each growth curve sample was then compared with a standard curve. Growth curves were plotted between the copy number of the 16S rRNA gene and the cultivation time. The relative quantification of each transcript was calculated by the equation (Schmittgen and Livak 2008) indicated below using the *tpi* gene as an internal calibrator gene.

$$\text{Expression level} = -\Delta\text{Ct}$$

$$\text{Comparative expression} = 2^{-[\Delta\text{Ct of cellulose, xylan, or cellulose and xylan} - \Delta\text{Ct of cellobiose}]}$$

$$\text{When } \Delta\text{Ct} = [\text{Ct of target gene} - \text{Ct of } \textit{tpi} \text{ gene}]$$

Determination of stable calibrator genes

The stability of the 10 candidate calibrator genes, namely *16S rRNA*, *adk* (adenylate kinase), *gluD* (glutamate dehydrogenase), *gukL* (guanylate kinase), *gyrA* (gyrase subunit A), *recA* (recombinase A), *rho* (transcription termination factor), *rpoA* (DNA-directed RNA polymerase subunit), *rpsJ* (30S ribosomal protein S10), and *tpi* (triosephosphate isomerase), were analyzed using RefFinder (https://www.heartcure.com.au/reffinder/, February 16, 2021) (Xie et al. 2023), geNorm (Vandesompele et al. 2002), NormFinder (Andersen et al. 2004), BestKeeper (Pechtl et al. 2018) and the comparative Δ-Ct method (Silver et al. 2006), to recommend a comprehensive ranking of all calibrator genes. In quadruplicate, cDNA from the mid-log phase of growth in cellobiose, cellulose powder, xylan, and the cellulose-xylan mixture was subjected to qRT-PCR for each candidate calibrator gene. The Ct data were subjected to analysis on RefFinder software and geomean ranking values were estimated along with the four different statistical algorithms including geNorm, NormFinder, and BestKeeper, and the comparative Δ-Ct method. The geNorm can express the stability of the calibrator genes by the *M*-value, i.e., the smaller the *M*-value, the more stable the calibrator gene (Vandesompele et al. 2002). NormFinder can express the stability of the calibrator genes by means of the stability value. The smaller the stability value, the more stable the calibrator gene (Andersen et al. 2004). BestKeeper can express the stability of the calibrator genes by the standard deviation, coefficient of variation (CV), and geomean (Pfaffl et al. 2004). When the standard deviation is less than 1, the gene can be used as an internal calibrator gene (Pfaffl et al. 2004). The most and least stable calibrator genes in the different carbon sources were validated by evaluating the expression of the scaffolding gene (accession number: PQQ66558.1) (*scaff*) in *H. saccincola* A7. The relative expression of the target genes was calculated using the $2^{-\Delta\Delta Ct}$ method. Cellobiose is set as a base carbon source for comparison with other carbon sources. Four biological replicates were used in this analysis.

Repositories

Japan Collection of Microorganisms (JCM): *Herbivorax saccincola* A7 (JCM 31827)

Deutsche Sammlung von Mikroorganismen und Zellkulturen (DSM): *Herbivorax saccincola* A7 (DSM 104321)

The National Center for Biotechnology Information (NCBI) accession number for the draft genome sequence of *Herbivorax saccincola* A7 is PRJNA384108.

Gene Expression Omnibus (GEO) accession number for the raw and normalized qRT-PCR data is GSE165680.

Results

Xylanolytic and cellulolytic enzyme activities of *H. saccincola* A7 in response to carbon sources

H. saccincola can utilize pentose-based polysaccharides such as xylan for cell growth. To evaluate whether the extracellular GHs of *H. saccincola* A7 are influenced by the use of cellulose and xylan as carbon sources, the activities of cellulase and xylanase were measured using PASC-binding proteins prepared from media containing different proportions of cellulose and xylan. The PASCase activity of *C. thermocellum* was found to be slightly increased in the presence of xylan, but decreased rapidly when the percentage of xylan was more than 50% (Fig. 1a). By contrast, the

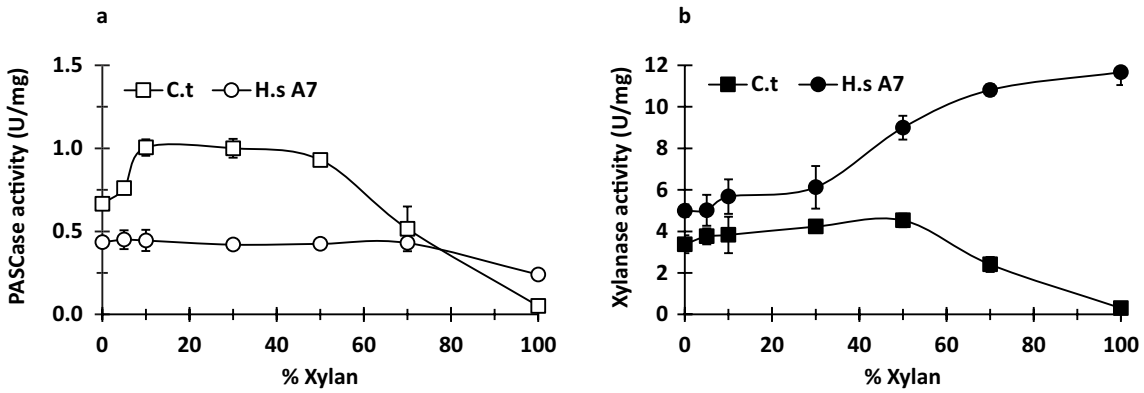

Fig. 1 The effects on cellulase and xylanase activities produced by *C. thermocellum* and *H. saccincola* A7 in mixed carbon source cultures with varying cellulose to xylan ratios. The specific activity of PASCase (cellulase) (**a**) and xylanase (**b**) was measured using PASC-binding protein prepared from *C. thermocellum* and *H. saccincola* A7 grown in medium supplemented with different ratios of cellulose and xylan. The *x*-axis shows the concentration of xylan as a percentage of the 1% carbon source

PASCase activity of *H. saccincola* A7 was not significantly affected by increasing the xylan content (Fig. 1a). The PASCase and xylanase activities of *C. thermocellum* decreased with increasing xylan concentration (Fig. 1a, b). This may be due to poor growth resulting from low xylan assimilation. The xylanase activity of *H. saccincola* A7 was increased along with the xylan concentration in the medium (Fig. 1b), whereas the PASCase activity was unaffected by increasing xylan concentration (Fig. 1a), suggesting the existence of a specific expression mechanism that maintains PASCase activity regardless of cellulose presence. Xylanase activity was enhanced with increasing xylan concentration, suggesting that *C. thermocellum* may have a regulatory xylan-induced expression system (Fig. 1b) and that the response in the presence of extracellular polysaccharides differs from that of *C. thermocellum*.

Determination of sampling points and correction genes of *H. saccincola* A7 for qRT-PCR assays

Gene expression is dependent on the bacterial cell growth phase (Rydzak et al. 2012). To verify the growth characteristics of *H. saccincola* A7, a standard curve was plotted from three replicate reactions based on serial dilutions of the 16S rRNA gene (Bartoš et al. 2024). The results showed that the linear range of the amplified 16S rRNA-based qPCR assay for *H. saccincola* A7 was 10^5 to 10^8 copies (Supplementary Figure S1). When distinct carbon sources, such as cellulose and xylan are utilized, the adjustment of the sampling point for the growth phase of *H. saccincola* A7 on each carbon source is important. The growth curve of *H. saccincola* A7 with cellobiose, cellulose, xylan, and a cellulose-xylan mixture as the carbon source was plotted based on the copy number of the 16S rRNA gene for each culture (Fig. 2). *H. saccincola* A7 grew fastest on cellobiose compared with the other carbon sources, but the cell density did not increase (Fig. 2). When cellulose and the cellulose-xylan mixtures were used as carbon sources, the cell density of *H. saccincola* A7 was increased compared to the other carbon sources. In addition, the cellulose-xylan mixture as a carbon source showed the fastest growth rate with a slight lag phase up to about 12 h (Fig. 2). By contrast, the growth rate and density of *H. saccincola* A7 were slower when xylan was used as the carbon source. Based on these growth curves and cell densities, we next examined gene expression by sampling cells at mid-log phase, i.e., after 12 h for cellobiose, 24 h for cellulose, 48 h for xylan, and 18 h for the cellulose-xylan mixture.

Stability analysis of the candidate calibrator genes in *H. saccincola* A7

A candidate calibrator gene is usually defined by the stability of its expression, i.e., a low coefficient of variation (CV) and a maximum fold change, which is the ratio of the maximum and minimum values observed within the data set less than two, and a mean value less than the maximum value minus two standard deviations (De Jonge et al. 2007). Ten candidate calibrator genes were validated, namely *16S rRNA*, *adk*, *gluD*, *gukL*, *gyrA*, *recA*, *rho*, *rpoA*, *rpsJ*, and *tpi*. The RNA expression levels of the 10 candidate calibrator genes were measured in quadruplicate under the four different carbon sources at the mid-log growth phase described above. The Ct values of the 10 candidate calibrator genes from the qRT-PCR were evaluated using RefFinder software to recommend a comprehensive ranking of all calibrator genes. Table 1 shows the results obtained with the geNorm and NormFinder algorithms, with *tpi* being the most suitable candidate calibrator gene. *16S rRNA* was the top-ranked using BestKeeper, whereas the *adk* gene ranked highest using the comparative Δ-Ct method. Based on the comprehensive rankings of the RefFinder software, the best candidate calibrator gene was *tpi* in *H. saccincola* A7.

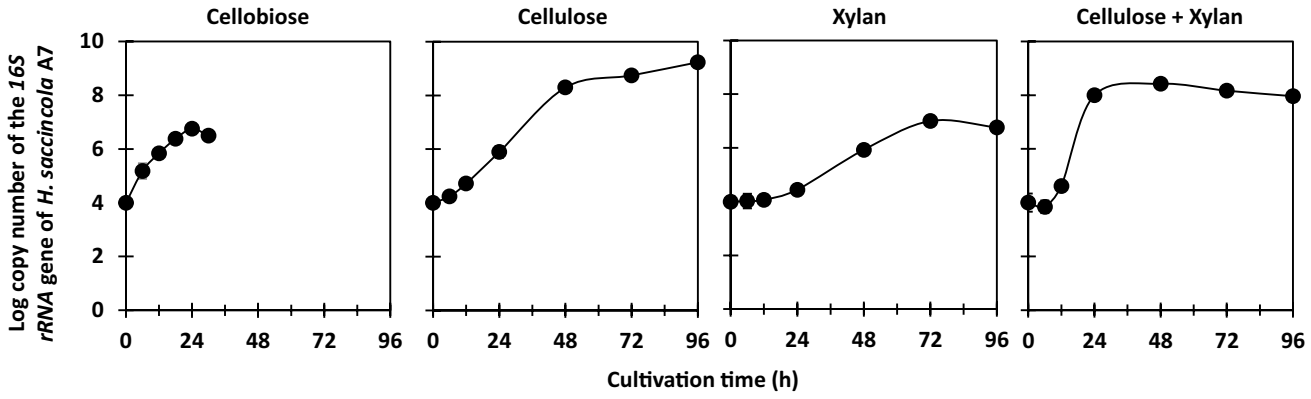

Fig. 2 Growth profile of *H. saccincola* A7 on different carbon sources. The 16S rRNA content (log copy number/μL) of *H. saccincola* A7 grown on each carbon source was measured by quantitative real-time PCR using genomic DNA extracted from the cells at each time point

Table 1 The stability ranking of candidate calibrator genes by geNorm, NormFinder, BestKeeper, the comparative -ΔCt method, and RefFingder

Rank	geNorm		NormFinder		BestKeeper			Comparative -ΔCt		RefFinder	
	Genes	Stability (M)	Genes	Stability	Genes	SD [± Ct]*1	CV [%Ct]*2	Genes	Average of SD	Gene	Recommended comprehensive ranking
1	*tpi*	0.365	*tpi*	0.179	*16S rRNA*	0.605	0.017	*adk*	0.990	*tpi*	1.680
2	*rpoA*	0.365	*adk*	0.246	*adk*	0.955	0.001	*tpi*	1.030	*adk*	1.860
3	*adk*	0.458	*rpsJ*	0.336	*gluD*	0.920	0.001	*gyrA*	1.040	*gyrA*	2.630
4	*gyrA*	0.513	*gyrA*	0.352	*gukL*	0.604	0.017	*rpsJ*	1.090	*rpoA*	3.200
5	*rpsJ*	0.625	*rho*	0.442	*gyrA*	0.915	0.001	*rpoA*	1.110	*rpsJ*	4.360
6	*rho*	0.687	*gluD*	0.506	*recA*	0.402	0.138	*rho*	1.120	*rho*	6.160
7	*gluD*	0.716	*rpoA*	0.538	*rho*	0.953	0.001	*gluD*	1.150	*gluD*	6.740
8	*16S rRNA*	0.834	*16S rRNA*	1.186	*rpoA*	0.829	0.001	*16S rRNA*	1.470	*16S rRNA*	7.110
9	*recA*	1.008	*recA*	1.735	*rpsJ*	0.966	0.001	*recA*	1.870	*recA*	9.000
10	*gukL*	1.370	*gukL*	2.761	*tpi*	0.923	0.001	*gukL*	2.820	*gukL*	10.000

*1; SD [± Ct], standard deviation of the Ct, *2; CV [%Ct], coefficient of variance expressed as a percentage of the Ct level

To validate the calibrator genes selected by the four statistical algorithms, the most stable and unstable calibrator genes were used to confirm the relative expression levels of *scaff*, a scaffold protein known to be more upregulated in the presence of cellulose than cellobiose. The upregulated expression of the scaffold proteins was only detected when the most stable calibrator genes, *tpi* and *adk*, were used for normalization (Fig. 3). Downregulation was detected for the least stable calibrator genes, *16S rRNA* and *recA*. Of *tpi* and *adk*, the most stable calibrator genes, we selected *tpi* as the calibrator gene for qRT-PCR in *H. saccincola* A7.

Expression analysis of the scaffolding proteins, and the cellulosomal and non-cellulosomal-related genes, in *H. saccincola* A7 by carbon source

It has been revealed by genome sequence analysis that *H. saccincola* A7 has cellulosomal and non-cellulosomal systems, similar to *C. thermocellum*. To understand whether cellulose and hemicellulose-degrading enzymes are influenced by extracellular carbon sources, we analyzed the expression levels of all cellulosomal-related genes and all non-cellulosomal GH genes in *H. saccincola* A7 by qRT-PCR following culturing with cellobiose, cellulose, xylan, or a mixture of cellulose and xylan as carbon sources. The cellulosomal scaffold proteins of *H. saccincola* A7 are known to comprise six scaffold proteins and one scaffold-like GH (Fig. 4). The expression levels of each gene were compared with the expression level of the *tpi* gene as a calibrator gene and are presented as a heatmap (Fig. 5). Comparison of the expression levels of these scaffolding protein genes on each carbon source revealed that the major scaffolding protein (PQQ66558.1) and the anchor proteins (PQQ66559.1 and WP_105367906.1) are highly expressed on all carbon sources (cellobiose, cellulose, xylan, and a mixture of cellulose and xylan) (Fig. 5). Although expression of the anchoring proteins (PQQ66559.1 and WP_105367906.1) was slightly higher when cellulose, xylan, and a mixture of cellulose and xylan were used as the carbon source, the expression of other integrated proteins (PQQ67419.1 and PQQ67420.1), anchoring proteins (PQQ67184.1 and PQQ68014.1), and the scaffolding-like GH protein (PQQ66177.1) was poor on all carbon sources (Fig. 5). These results indicate that the cellulosome system of *H. saccincola* A7 is mainly composed of the major scaffold protein (PQQ66558.1) and the two anchor proteins (PQQ66559.1 and WP_105367906.1).

The expression pattern of GH family genes encoding cellulosomal and non-cellulosomal proteins was also compared on different carbon sources (Fig. 5). The expression of 42 out of 52 GH genes was found to be affected by the type of carbon source in *H. saccincola* A7. Among the cellulosomal and non-cellulosomal GH proteins, GH5_4 (PQQ66693.1:

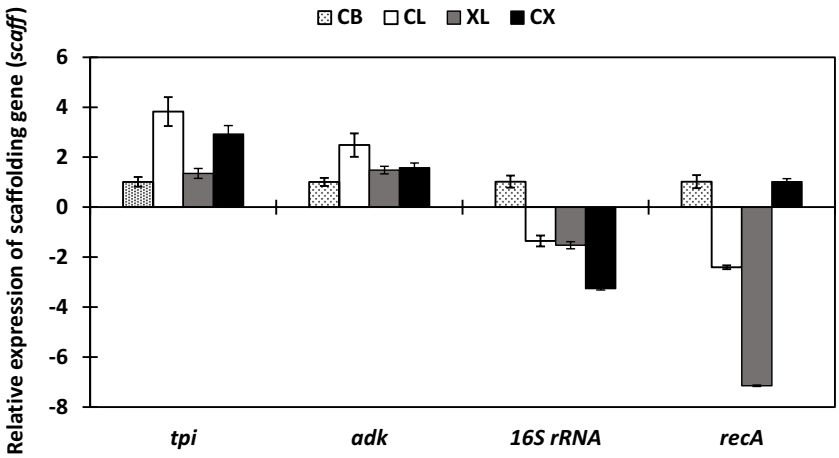

Fig. 3 Validation of *H. saccincola* A7 calibration genes. The relative expression level of the scaffolding gene (*scaff*) of *H. saccincola* A7 grown in different carbon sources using different calibration genes (the most and least stable genes). Samples were taken from the mid-log phase of growth in cellobiose (CB), cellulose (CL), xylan (XL), and a cellulose-xylan mixture (CX)

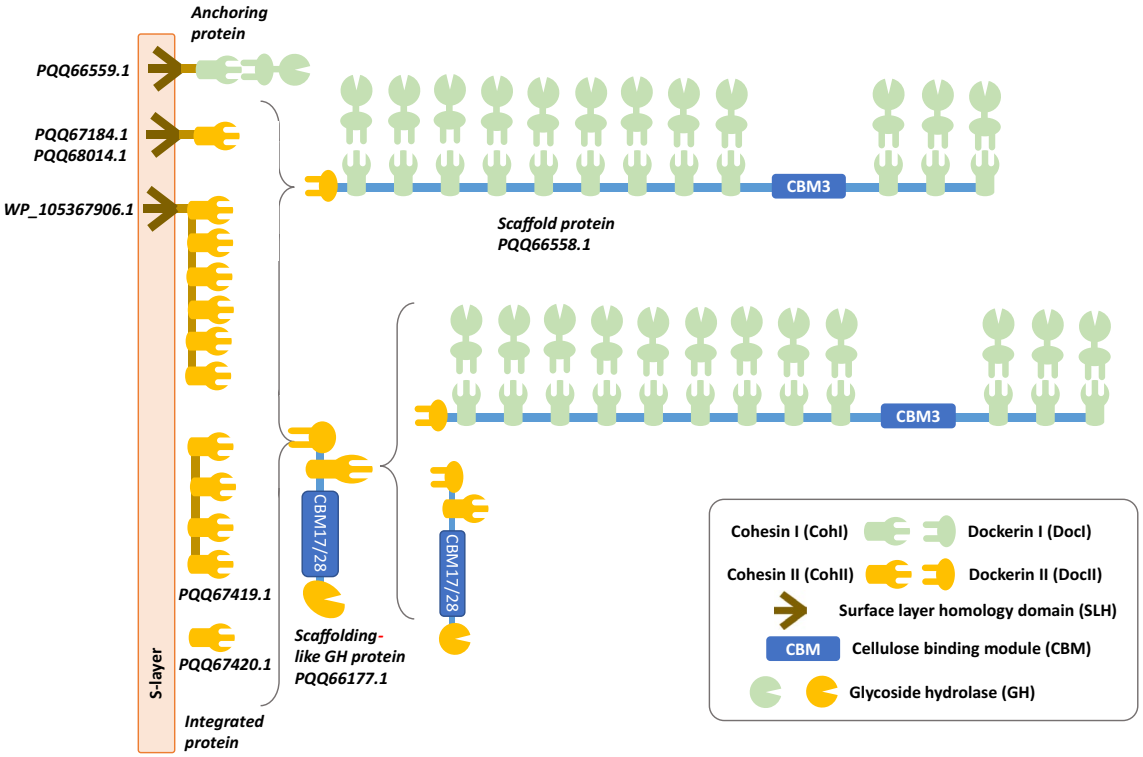

Fig. 4 Schematic diagram of the primary and secondary scaffolding proteins of *H. saccincola* A7

GH5-DocI), GH10_3 (PQQ67442.1: CBM4-GH10-DocI), GH30 (PQQ67632.1: DocI-GH30), GH11_5 (PQQ66320.1: GH11-CBM6-DocI-AcXEs), GH11_7 (PQQ66270.1: GH11-CBM6-DocI-XynB), GH10/11 (PQQ65739.1: GH11-CBM6-DocI-GH10), GH10_1 (PQQ66321.1: 4CBM4s-GH10-CBM9-3SLHs), GH94_2 (PQQ67605.1: GH94-GH36), GH11_2 (PQQ67643.1: GH11-2CBM6s-CBM2), and GH11_6 (PQQ66159.1: GH11), were expressed on all carbon sources, indicating that they are permanently expressed GHs in *H. saccincola* A7. By contrast, three cellulosomal genes GH9_5 (PQQ67287.1: GH9-CBM3-DocI), GH5_2 (PQQ67260.1: GH5-DocI), and GH10_4 (PQQ66064.1: DocI-CBM6-GH10), and two non-cellulosomal genes GH43_3 (PQQ66698.1: GH43) and GH94_3 (PQQ66110.1: GH94-GH36) were highly expressed only in the presence of cellulose as the carbon source, and two non-cellulosomal genes GH9_4 (AUG58665.1; GH9-2CBM3s) and GH67 (PQQ66521.1; GH67) were only induced by xylan. These results suggest that most of the GH genes encode cellulosomal and non-cellulosomal enzymes of *H.*

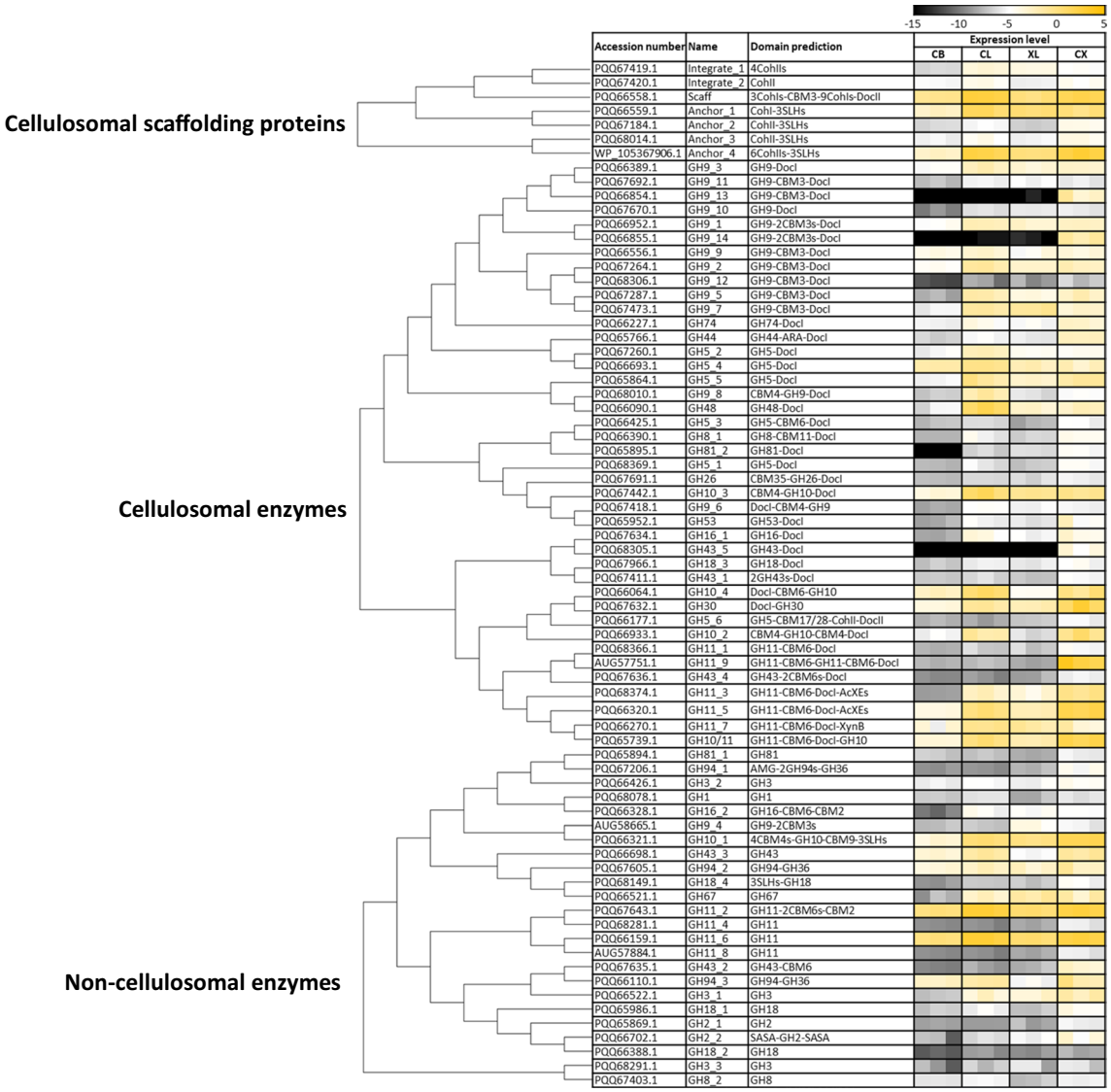

Fig. 5 Heatmap of transcriptional profiling of cellulosomal and non-cellulosomal enzymes in *H. saccincola* A7. *H. saccincola* A7 was grown to mid-log phase on cellobiose (CB), cellulose (CL), xylan (XL), and a cellulose and xylan mixture (CX). Target gene expression was normalized to the expression of the triosephosphate isomerase (*tpi*) gene for each carbon source under the same conditions. Values reported are the mean of at least three independent qRT-PCR experiments from quadruplicate cultures and are presented according to the gradient color bar

saccincola A7, whose gene expression is predominantly regulated by cellulose, and in some cases xylan. Interestingly, the coexistence of cellulose and xylan resulted in the expression of several genes, such as GH9_13 (PQQ66854.1: GH9-CBM3-DocI), GH9_14 (PQQ66855.1: GH9-2CBM3s-DocI), GH43_5 (PQQ68305.1: GH43-DocI), and GH11_9 (AUG57751.1: GH11-CBM6-GH11-CBM6-DocI), although no expression was observed when cellulose or xylan was the sole carbon sources. *H. saccincola* A7 may have a two-component sensor-like entity that senses the coexistence of cellulose and hemicellulose, as a possible explanation for these expression patterns. Based on the expression pattern of each gene by carbon source, *H. saccincola* A7 may have a sensor protein that senses external carbon sources, similar to the extracellular carbohydrate-sensing mechanism identified in *C. thermocellum*, which also functions to regulate gene expression.

Prediction of the carbohydrate-sensing mechanism in *H. saccincola* A7

Recently, it has been demonstrated that the activation of genes encoding polysaccharide-degrading enzymes is regulated by a broad carbohydrate-sensing mechanism. Based

on the expression analysis of cellulosomal and non-cellulosomal genes using cellulose and xylan as carbon sources, a similar carbohydrate-sensing mechanism likely exists in *H. saccincola* A7, especially when cellulose is the carbon source. To confirm the presence of RsgI-like anti-σ factors, we analyzed the *H. saccincola* A7 complete genome sequence, revealing the presence of 10 RsgI-like anti-σ factors in the genome (Fig. 6a). Among the 10 *rsgI*-like genes, *rsgI1* to *rsgI5* appeared to form a bicistronic operon downstream of the gene coding for σI-like factor (SigI), and RsgI-6 was found to contain a putative SigI domain in its N-terminal region (Fig. 6b). In addition, the genes *rsgI7*, *rsgI8*, *rsgI9*, and *rsgI10* are located in the genome and are sigma factor-independent (Fig. 6a, b). The domain structures of these RsgI-like proteins have revealed that the C-terminal region is extracellular and possesses polysaccharide binding and GH enzyme-like domain structures (Fig. 6b). The CBM family 3 module (CBM3) was observed in the C-terminal region of RsgI1 and RsgI2. RsgI3 contained two pectin-binding domains classified in the PA14 superfamily. In addition, RsgI4 contained an α-L-arabinofuranosidase B (AbfB) catalytic domain. No other RsgI-like proteins were recognized to contain characteristic C-terminal functional domains. To understand whether the RsgI-like protein and its complementary *sigI* function as extracellular carbohydrate sensors, the expression levels of four σ-factor genes (*sigI1*, *sigI2*, *sigI3*, and *sigI4*) in *H. saccincola* A7 were compared by qRT-PCR on cellobiose, cellulose, xylan, and a cellulose-xylan mixture as the carbon source (Fig. 7). The expression of *sigI1* and *sigI2* in *H. saccincola* A7 was upregulated 2–3 times in the presence of cellulose compared with xylan (Fig. 7). The presence of CBM3 at the C-terminus of anti-σI factors (RsgI1 and RsgI2) downstream of *sigI1* and *sigI2* suggests that it may serve as a cellulose-specific sensing domain (Fig. 6). The three σI factors, *sigI1*, *sigI2*, and *sigI3*, were further upregulated in the presence of both cellulose and xylan (Fig. 7). Among the anti-σI factors, RsgI3, is located downstream of *sigI3* and contains a tandem protective antigen 14 (PA14) motif that binds pectin or xylan (Grinberg et al. 2019; Izquierdo et al. 2012; Kahel-Raifer et al. 2010) (Fig. 6), and was upregulated in the presence of xylan only. It was also interesting to observe that although

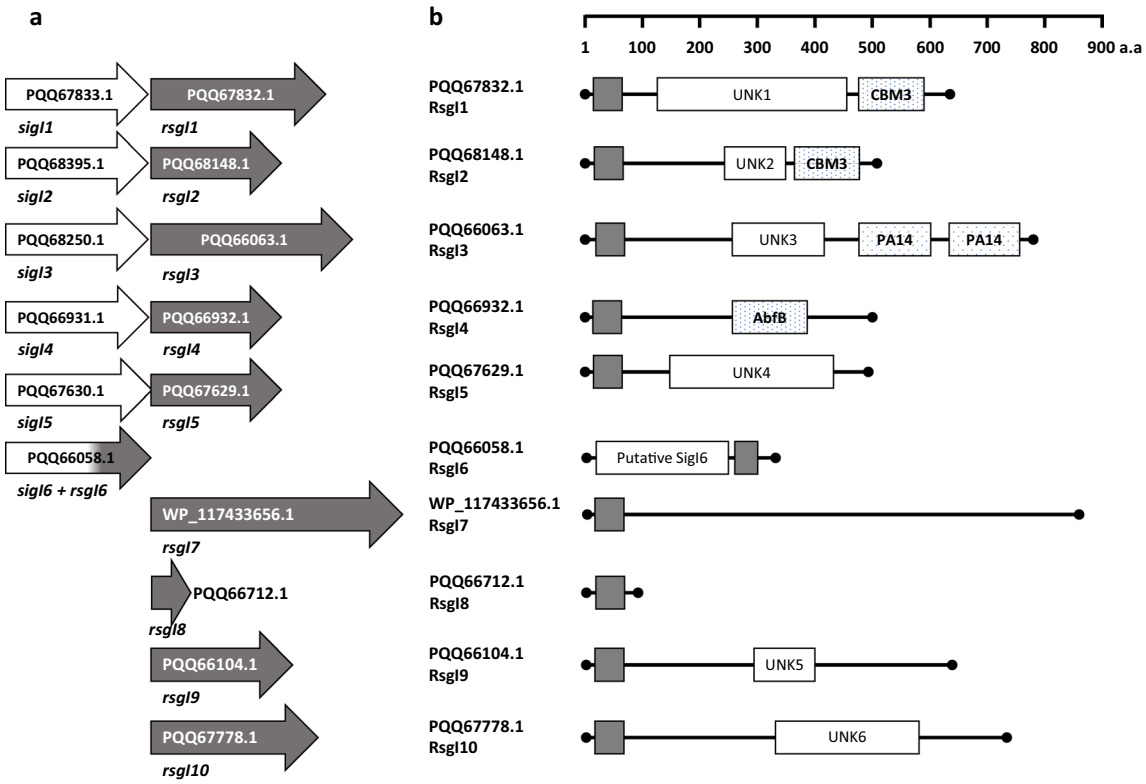

Fig. 6 Schematic diagram of the putative σI and anti-σI regulatory factors of *H. saccincola* A7. **a** White arrows indicate *sigI* genes and gray arrows indicate rsgI-like genes. The *rsgI1–rsgI5* genes form bicistronic operons downstream of *sigI*, the *rsgI-6* gene contains the putative *sigI* domain at the N-terminal region, and the *rsgI7–rsgI10* genes are monocistronic without an upstream *sigI* gene. **b** Schematic representation of the domain organization of the *H. saccincola* A7 RsgI-like proteins. Linkers are shown as lines. Each protein contains an RsgI-like domain (dark gray) with/without various additional domains as follows: CBM3, carbohydrate-binding module family 3; PA14, pectin-binding domain; AbfB, alpha-L-arabinofuranosidase B catalytic module; UNK1-6, domains of unknown function. The length of the proteins (number of amino acids) is indicated by the ruler at the top

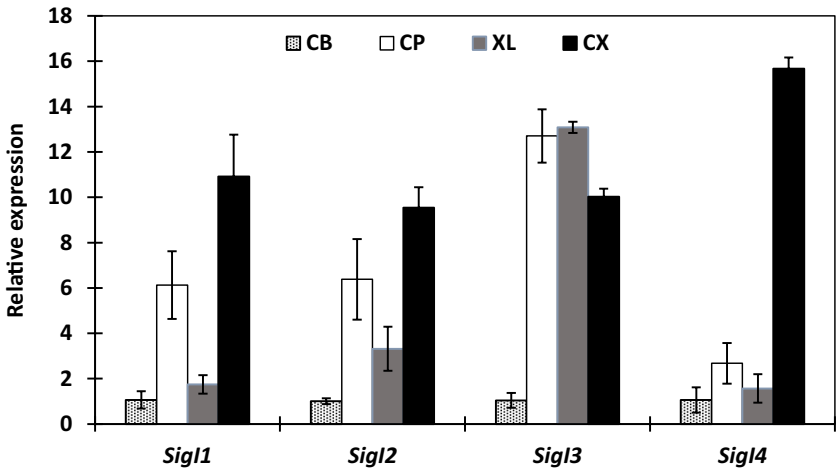

Fig. 7 Relative expression of σ-factors of *H. saccincola* A7 grown with cellobiose, cellulose, or a mixture of cellulose and xylan as carbon sources. Samples were taken from the mid-log phase of growth. Normalization was performed using the *tpi* gene

sigI4 expression levels were low on cellulose or xylan alone, *sigI4* was upregulated sixfold in the presence of both cellulose and xylan compared with cellulose or xylan alone. The C-terminal side of RsgI4, which is downstream of *sigI4*, contains an alpha-L-arabinofuranosidase B (AbfB) catalytic domain. The expression of two σI factors, *sigI1* and *sigI2*, may be mainly upregulated in the presence of cellulose and may control the expression of many cellulosomal and non-cellulosomal enzyme subunits.

By contrast, *sigI3* and *sigI4* were upregulated in the presence of not only cellulose but also xylan. In particular, *sigI4* is uniquely upregulated in the presence of both cellulose and xylan, suggesting that several cellulosomal subunits such as GH9_13, GH9_14, GH43_5, and GH11_9, that are only expressed in the presence of a mixture of cellulose and xylan, may be regulated by *sigI4*.

Discussion

The newly classified *H. saccincola*, similar to *C. thermocellum*, possesses a cellulosome enzyme complex and can degrade xylan as well as cellulose (Aikawa et al. 2018; Koeck et al. 2016). According to proteome (Raman et al. 2011; Stevenson and Weimer 2005; Yoav et al. 2017) and transcriptome (Stevenson and Weimer 2005; Wei et al. 2014) analyzes of *C. thermocellum*, changes in carbon source alter the composition of cellulase and hemicellulase. That is, carbon sources are known to affect cellulosome activity by modulating the cellulosomal enzyme population and its structural composition; the cellulosomal enzymes of mesophilic hemicellulolytic *Clostridium cellulovorans* (Doi and Kosugi 2004) appear to be more affected by the type of carbon source than the non-cellulosomal enzymes (Han et al. 2004; Matsui et al. 2013). However, while *C. thermocellum* cannot grow on xylan as a sole carbon source, *H. saccincola* can grow on xylan alone (Aikawa et al. 2020; Koeck et al. 2016). When *H. saccincola* A7 was grown in media with different concentration ratios of cellulose and xylan as carbon sources, cellulase activity was constant and did not change significantly, while xylanase activity increased with increasing xylan concentration. These results show that in *H. saccincola* A7, the presence of at least xylan enhances xylanase activity but has no effect on cellulase activity.

A comprehensive expression analysis of cellulosomal and non-cellulosomal protein-related genes in *H. saccincola* A7 using qRT-PCR showed that cellulose induced the expression of many carbohydrate-degrading genes compared with xylan as the sole carbon source. In lignocellulose, cellulose is surrounded by hemicellulose, including xylan and lignin, forming a complex, solid structure. In the lignocellulosic conversion process, the hemicellulose and lignin components are usually removed in the pretreatment step (Bhattacharya et al. 2015; Gírio et al. 2010). The efficient removal of hemicellulose and lignin is believed to reduce the complexity of the cellulose structure, increase cellulose degradation efficiency, and accelerate the biomass conversion process. In *H. saccincola* A7, it appeared that several xylanases were induced by the presence of xylan; however, expression analysis clearly indicated that cellulose was an important substrate to enhance the expression of cellulosomal and non-cellulosomal enzymes. Therefore, pretreatment techniques that involve exposure to cellulose to increase the expression of carbohydrate-hydrolyzing enzymes, including cellulosomal scaffolding proteins, are important for the efficient degradation of lignocellulose by *H. saccincola* A7.

Recently, it has been demonstrated that the activation of genes encoding polysaccharide-degrading enzymes is regulated by a broad carbohydrate-sensing mechanism. Kahel-Raifer et al. (Kahel-Raifer et al. 2010) identified a group of six putative operons in *C. thermocellum* that encode alternative σ factors and their cognate membrane-associated anti-σ factors that play a role in cellulosome gene regulation. These operons encode a multimodular protein

containing a strongly predicted transmembrane helix, and these transmembrane proteins are homologous to the N-terminal segment of the *Bacillus subtilis* anti-σ factor RsgI, which was found to transmit extracellular polysaccharide-sensing signals into cells for regulation of the cellulosome system, a polysaccharide-degrading multienzyme complex. In addition, Chen et al. recently found that this anti-σ factor RsgI involves an autoproteolytic process in its signaling (Chen et al. 2023). The mechanism revealed that a short peptide together with the transmembrane helix of RsgI, activates the downstream RIP protease RseP, which finally releases bound SigI. This transmembrane signaling process involving autoproteolysis is well-recognized in eukaryotes. However, it has rarely been reported in bacteria, with the exception of *C. thermocellum*. Here, we showed that a similar transmembrane signal transduction process is present in *H. saccincola* A7. The structures of the RsgI-like proteins involved in this signaling are similar between *H. saccincola* and *C. thermocellum*, as is the possession of the extracellular polysaccharide recognition and binding domains of CBM3 and PA14. Conversely, RsgI4 in *H. saccincola* A7 has a unique RsgI-like structure. There is an AbfB domain which hydrolyzes 1,5-alpha, 1,3-alpha, and 1,2-alpha linkages in both oligosaccharides and polysaccharides containing terminal non-reducing L-arabinofuranoses in side chains, on the C-terminal side, and high expression of SigI4 was specifically observed under growth conditions in which the carbon source comprised a mixture of cellulose and xylan. It is likely that SigI4-RsgI4 is specifically upregulated under more native lignocellulose-based culture conditions, such as those containing cellulose and xylan, such as GH9_13, GH9_14, and GH43_5 that are not expressed in cellulose or xylan, but only when cellulose and xylan are mixed, suggesting that *sigI4* affects their expression regulation and may also have potential strategies for lignocellulosic biomass degradation.

In this study, we show for the first time that an alternative σ-factor, previously reported in detail only in *C. thermocellum*, plays an important role as a regulatory mechanism for similar lignocellulose degradation in *H. saccincola* A7, which is specific for hemicellulose degradation. Additional study of these regulatory mechanisms is needed to help design and build robust and environmentally friendly cellular systems for the conversion of lignocellulose into useful products.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00253-024-13310-3.

Acknowledgements S Baramee, P Thianheng, and P Cheawchanlertfa acknowledge the Counterpart Researcher Invitation Program from Japan International Research Center for Agricultural Sciences. We thank Kate Fox, Ph.D., from Edanz (https://jp.edanz.com/ac) for editing a draft of this manuscript.

Authors' contributions Conceived the study: SB, AK. Performed the experiments: SB, PT, AU, PC. Analyzed the data: SB, PT, AU, PC, CT, RW, PP, KR, AK. Curated information: SB, PT, AU, PC, AK. Wrote the initial draft of the manuscript: SB, AU, PC, AK. Review and editing: CT, RW, PP, KR, YL, AK. Supervision: PC, CT, RW, YL, AK. Project administration: SB, CT, KR, AK. Funding acquisition: AU, YL, AK. Commented on working versions of the manuscript and agreed on the final version of the manuscript: all. All authors read and approved the final manuscript.

Funding This work was supported financially by the Science and Technology Research Partnership for Sustainable Development (SATREPS) [grant number JPMJSA1801], the Japan Science and Technology Agency (JST)/Japan International Cooperation Agency (JICA), the Joint Research Projects and Seminars under the Bilateral Programs [grant number JPJSBP120237203 and 323GJHZ2022004MI], the Japan Society for the Promotion of Sciences (JSPS).

Data availability The microbial reads and genome assemblies generated in this study are available at the National Center for Biotechnology Information (NCBI) under the BioProject ID: PRJNA384108. All data generated or analyzed during this study are included in this published article. The raw and normalized qRT-PCR data have been deposited in the Gene Expression Omnibus (GEO) data repository under accession number GSE165680.

Declarations

Ethics approval This article does not contain any studies with human participants or animals performed by any of the authors.

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

References

Aikawa S, Baramee S, Sermsathanaswadi J, Thianheng P, Tachaapaikoon C, Shikata A, Waeonukul R, Pason P, Ratanakhanokchai K, Kosugi A (2018) Characterization and high-quality draft genome sequence of *Herbivorax saccincola* A7, an anaerobic, alkaliphilic, thermophilic, cellulolytic, and xylanolytic bacterium. Syst Appl Microbiol 41(4):261–269

Aikawa S, Thianheng P, Baramee S, Ungkulpasvich U, Tachaapaikoon C, Waeonukul R, Pason P, Ratanakhanokchai K, Kosugi A (2020) Phenotypic characterization and comparative genome

analysis of two strains of thermophilic, anaerobic, cellulolytic-xylanolytic bacterium *Herbivorax saccincola*. Enzyme Microb Technol 136:109517

Andersen CL, Jensen JL, Ørntoft TF (2004) Normalization of real-time quantitative reverse transcription-PCR data: a model-based variance estimation approach to identify genes suited for normalization, applied to bladder and colon cancer data sets. Cancer Res 64(15):5245–5250. https://doi.org/10.1158/0008-5472.Can-04-0496

Bartoš O, Chmel M, Swierczková I (2024) The overlooked evolutionary dynamics of 16S rRNA revises its role as the "gold standard" for bacterial species identification. Sci Rep 14(1):9067. https://doi.org/10.1038/s41598-024-59667-3

Bayer EA, Belaich J-P, Shoham Y, Lamed R (2004) The cellulosomes: multienzyme machines for degradation of plant cell wall polysaccharides. Annu Rev Microbiol 58:521–554

Bhattacharya AS, Bhattacharya A, Pletschke BI (2015) Synergism of fungal and bacterial cellulases and hemicellulases: a novel perspective for enhanced bio-ethanol production. Biotechnol Lett 37(6):1117–1129. https://doi.org/10.1007/s10529-015-1779-3

Chen C, Dong S, Yu Z, Qiao Y, Li J, Ding X, Li R, Lin J, Bayer EA, Liu Y-J, Cui Q, Feng Y (2023) Essential autoproteolysis of bacterial anti-σ factor RsgI for transmembrane signal transduction. Sci Adv 9(27):eadg4846. https://doi.org/10.1126/sciadv.adg4846

De Jonge HJ, Fehrmann RS, de Bont ES, Hofstra RM, Gerbens F, Kamps WA, de Vries EG, van der Zee AG, te Meerman GJ, ter Elst A (2007) Evidence based selection of housekeeping genes. PLoS ONE 2(9):e898. https://doi.org/10.1371/journal.pone.0000898

Demain A, Newcomb M, Wu J (2005) Cellulase, clostridia, and ethanol. Microbiol Mol Biol Rev 69:124–154

Doi RH, Kosugi A (2004) Cellulosomes: plant-cell-wall-degrading enzyme complexes. Nat Rev Microbiol 2(7):541–551

Dror TW, Rolider A, Bayer EA, Lamed R, Shoham Y (2003) Regulation of expression of scaffoldin-related genes in *Clostridium thermocellum*. J Bacteriol 185(17):5109–5116. https://doi.org/10.1128/jb.185.17.5109-5116.2003

Gírio FM, Fonseca C, Carvalheiro F, Duarte LC, Marques S, Bogel-Łukasik R (2010) Hemicelluloses for fuel ethanol: a review. Bioresour Technol 101(13):4775–4800. https://doi.org/10.1016/j.biortech.2010.01.088

Grinberg IR, Yaniv O, de Ora LO, Muñoz-Gutiérrez I, Hershko A, Livnah O, Bayer EA, Borovok I, Frolow F, Lamed R (2019) Distinctive ligand-binding specificities of tandem PA14 biomass-sensory elements from *Clostridium thermocellum* and *Clostridium clariflavum*. Proteins 87(11):917–930

Han SO, Cho H-Y, Yukawa H, Inui M, Doi RH (2004) Regulation of expression of cellulosomes and noncellulosomal (hemi) cellulolytic enzymes in *Clostridium cellulovorans* during growth on different carbon sources. J Bacteriol 186(13):4218–4227

Izquierdo JA, Goodwin L, Davenport KW, Teshima H, Bruce D, Detter C, Tapia R, Han S, Land M, Hauser L (2012) Complete genome sequence of *Clostridium clariflavum* DSM 19732. Stand Genomic Sci 6(1):104

Kahel-Raifer H, Jindou S, Bahari L, Nataf Y, Shoham Y, Bayer EA, Borovok I, Lamed R (2010) The unique set of putative membrane-associated anti-σ factors in *Clostridium thermocellum* suggests a novel extracellular carbohydrate-sensing mechanism involved in gene regulation. FEMS Microbiol Lett 308(1):84–93

Koeck DE, Mechelke M, Zverlov VV, Liebl W, Schwarz WH (2016) *Herbivorax saccincola* gen. nov., sp. nov., a cellulolytic, anaerobic, thermophilic bacterium isolated via in sacco enrichments from a lab-scale biogas reactor. Int J Syst Evol Microbiol 66(11):4458–4463. https://doi.org/10.1099/ijsem.0.001374

Lynd LR, Weimer PJ, Van Zyl WH, Pretorius IS (2002) Microbial cellulose utilization: fundamentals and biotechnology. Microbiol Mol Biol Rev 66(3):506–577

Matsui K, Bae J, Esaka K, Morisaka H, Kuroda K, Ueda M (2013) Exoproteome profiles of *Clostridium cellulovorans* grown on various carbon sources. Appl Environ Microbiol 79(21):6576–6584

Nataf Y, Bahari L, Kahel-Raifer H, Borovok I, Lamed R, Bayer EA, Sonenshein AL, Shoham Y (2010) *Clostridium thermocellum* cellulosomal genes are regulated by extracytoplasmic polysaccharides via alternative sigma factors. Proc Natl Acad Sci U S A 107(43):18646–18651

Nelson N (1944) A photometric adaptation of the Somogyi method for the determination of glucose. J Biol Chem 153(2):375–380

Nhim S, Waeonukul R, Uke A, Baramee S, Ratanakhanokchai K, Tachaapaikoon C, Pason P, Liu YJ, Kosugi A (2022) Biological cellulose saccharification using a coculture of *Clostridium thermocellum* and *Thermobrachium celere* strain A9. Appl Microbiol Biotechnol 106(5–6):2133–2145. https://doi.org/10.1007/s00253-022-11818-0

Pechtl A, Rückert C, Maus I, Koeck DE, Trushina N, Kornberger P, Schwarz WH, Schlüter A, Liebl W, Zverlov VV (2018) Complete genome sequence of the novel cellulolytic, anaerobic, thermophilic bacterium *Herbivorax saccincola* type strain GGR1, isolated from a lab scale biogas reactor as established by Illumina and Nanopore MinION sequencing. Genome Announc 6(6):e01493-e1517. https://doi.org/10.1128/genomeA.01493-17

Pfaffl MW, Tichopad A, Prgomet C, Neuvians TP (2004) Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. Biotechnol Lett 26(6):509–515. https://doi.org/10.1023/B:BILE.0000019559.84305.47

Raman B, McKeown CK, Rodriguez M, Brown SD, Mielenz JR (2011) Transcriptomic analysis of *Clostridium thermocellum* ATCC 27405 cellulose fermentation. BMC Microbiol 11(1):134. https://doi.org/10.1186/1471-2180-11-134

Riederer A, Takasuka TE, Makino S, Stevenson DM, Bukhman YV, Elsen NL, Fox BG (2011) Global gene expression patterns in *Clostridium thermocellum* as determined by microarray analysis of chemostat cultures on cellulose or cellobiose. Appl Environ Microbiol 77(4):1243–1253. https://doi.org/10.1128/aem.02008-10

Rydzak T, McQueen PD, Krokhin OV, Spicer V, Ezzati P, Dwivedi RC, Shamshurin D, Levin DB, Wilkins JA, Sparling R (2012) Proteomic analysis of *Clostridium thermocellum* core metabolism: relative protein expression profiles and growth phase-dependent changes in protein expression. BMC Microbiol 12(1):214–231

Schmittgen TD, Livak KJ (2008) Analyzing real-time PCR data by the comparative C_T method. Nat Protoc 3(6):1101–1108

Silver N, Best S, Jiang J, Thein SL (2006) Selection of housekeeping genes for gene expression studies in human reticulocytes using real-time PCR. BMC Mol Biol 7(1):33. https://doi.org/10.1186/1471-2199-7-33

Stevenson DM, Weimer PJ (2005) Expression of 17 genes in *Clostridium thermocellum* ATCC 27405 during fermentation of cellulose or cellobiose in continuous culture. Appl Environ Microbiol 71(8):4672–4678. https://doi.org/10.1128/aem.71.8.4672-4678.2005

Sun Y, Cheng J (2002) Hydrolysis of lignocellulosic materials for ethanol production: a review. Bioresour Technol 83(1):1–11

Tang H, Ou J, Zhu M (2015) Development of a quantitative real-time PCR assay for direct detection of growth of cellulose-degrading bacterium *Clostridium thermocellum* in lignocellulosic degradation. J Appl Microbiol 118(6):1333–1344

Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F (2002) Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal

control genes. Genome Biol 3(7):research0034.1 https://doi.org/10.1186/gb-2002-3-7-research0034
Walseth CS (1952) Occurrence of cellulases in enzyme preparations from microorganisms. Tappi 35(5):228–233
Wei H, Fu Y, Magnusson L, Baker JO, Maness P-C, Xu Q, Yang S, Bowersox A, Bogorad I, Wang W (2014) Comparison of transcriptional profiles of *Clostridium thermocellum* grown on cellobiose and pretreated yellow poplar using RNA-Seq. Front Microbiol 5:142
Wiechelman KJ, Braun RD, Fitzpatrick JD (1988) Investigation of the bicinchoninic acid protein assay: identification of the groups responsible for color formation. Anal Biochem 175(1):231–237
Xie F, Wang J, Zhang B (2023) RefFinder: a web-based tool for comprehensively analyzing and identifying reference genes. Funct Integr Genomics 23(2):125. https://doi.org/10.1007/s10142-023-01055-7
Yoav S, Barak Y, Shamshoum M, Borovok I, Lamed R, Dassa B, Hadar Y, Morag E, Bayer EA (2017) How does cellulosome composition influence deconstruction of lignocellulosic substrates in *Clostridium* (*Ruminiclostridium*) *thermocellum* DSM 1313? Biotechnol Biofuels 10(1):222–237

Authors and Affiliations

Sirilak Baramee[1,2] · Pakinee Thianheng[1,3] · Ayaka Uke[1] · Pattsarun Cheawchanlertfa[1,3] · Chakrit Tachaapaikoon[2,3] · Rattiya Waeonukul[2,3] · Patthra Pason[2,3] · Khanok Ratanakhanokchai[3] · Ya-Jun Liu[4,5,6] · Akihiko Kosugi[1]

✉ Akihiko Kosugi
kosugia0164@jircas.go.jp; akosugi@affrc.go.jp

1 Biological Resources and Post-Harvest Division, Japan International Research Center for Agricultural Sciences (JIRCAS), 1-1 Ohwashi, Tsukuba, Ibaraki 305-8686, Japan

2 Excellent Center of Enzyme Technology and Microbial Utilization, Pilot Plant Development and Training Institute (PDTI), King Mongkut's University of Technology Thonburi (KMUTT), Bangkok 10150, Thailand

3 Enzyme Technology Laboratory, School of Bioresources and Technology, King Mongkut's University of Technology Thonburi (KMUTT), Bangkok 10150, Thailand

4 CAS Key Laboratory of Biofuels, Shandong Provincial Key Laboratory of Synthetic Biology, Shandong Engineering Laboratory of Single Cell Oil, Qingdao Institute of Bioenergy and Bioprocess Technology, Chinese Academy of Sciences, Qingdao 266101, PR China

5 Shandong Energy Institute, Qingdao 266101, PR China

6 Qingdao New Energy Shandong Laboratory, Qingdao 266101, PR China