You *et al.*
Biotechnology for Biofuels and Bioproducts (2025) 18:22
https://doi.org/10.1186/s13068-025-02619-4

Biotechnology for Biofuels and Bioproducts

RESEARCH

Open Access

Combining transcriptomic and metabolomic insights into carbohydrate utilization by *Ruminiclostridium papyrosolvens* DSM2782

Mengcheng You[1,2†], Zhenxing Ren[3†], Letian Ye[1], Qiuyun Zhao[1], Ziyi Liu[1], Houhui Song[1*] and Chenggang Xu[1*]

Abstract

Background Lignocellulose is the most abundant renewable bioresource on earth, and its biodegradation and utilization would contribute to the sustainable development of the global environment. *Ruminiclostridium papyrosolvens*, an anaerobic, mesophilic, and cellulolytic bacterium, produces an enzymatic complex known as the cellulosome. As one of the most highly evolved species among *Ruminiclostridium*-type species, *R. papyrosolvens* is particularly relevant for understanding how cellulolytic clostridia modulate their biomass degradation mechanisms in response to diverse carbon sources.

Results Our study investigates the transcriptional responses of *Ruminiclostridium papyrosolvens* to different carbon sources to understand its lignocellulose utilization. Using RNA-seq, we analyzed gene expression under glucose, cellobiose, xylan, cellulose, and corn stover, identifying distinct metabolic preferences and regulatory responses. We found significant gene expression changes under corn stover compared to other carbon sources, with enrichment in ABC transporters and cell growth pathways. CAZyme gene expression was regulated by TCSs, affecting sugar transporter systems. Metabolic profiling showed *R. papyrosolvens* produced more complex metabolites during corn stover fermentation, revealing its adaptability to various carbon sources and implications for metabolic engineering.

Conclusion This study not only uncovers the intricate response mechanisms of *R. papyrosolvens* to lignocellulose and its hydrolysates, but it also outlines the strategy for using *R. papyrosolvens* as a cellulolytic chassis in genetic engineering.

Keywords Transcriptome, Metabolome, CAZymes, ABC transporters, Two-component systems, *Ruminiclostridium papyrosolvens*

†Mengcheng You and Zhenxing Ren have contributed equally to this work.

*Correspondence:
Houhui Song
songhh@zafu.edu.cn
Chenggang Xu
xucg@zafu.edu.cn
[1] Laboratory of Applied Technology on Green-Eco-Healthy Animal Husbandry of Zhejiang Province, Zhejiang Provincial Engineering Research Center for Animal Health Diagnostics & Advanced Technology, Zhejiang International Science and Technology Cooperation Base for Veterinary Medicine and Health Management, China-Australia Joint Laboratory for Animal Health Big Data Analytics, College of Animal Science and Technology & College of Veterinary Medicine of Zhejiang Agriculture and Forestry University, Hangzhou 311300, Zhejiang Province, China
[2] Key Laboratory of Chemical Biology and Molecular Engineering of Ministry of Education, Institute of Biotechnology, Shanxi University, Taiyuan 030006, Shanxi Province, China
[3] Institute of Applied Chemistry, Shanxi University, Taiyuan 030006, Shanxi Province, China

Background

Lignocellulose is the most abundant biomass in the earth, capable of being converted by a wide range of microorganisms. Specifically, anaerobic and cellulolytic clostridia represent a major paradigm for efficient biological degradation of cellulosic biomass [1–6]. Many of these anaerobes produce an elaborate supramolecular enzymatic complex, termed the cellulosome [7, 8], where primarily catalytic components, such as glycoside hydrolases [9], carbohydrate esterases [10], and polysaccharide lyases [11, 12], are integrated onto a non-catalytic scaffoldin. The cellulosome affords enhanced synergistic activity among the resident enzymes to efficiently hydrolyze recalcitrant crystalline lignocellulosic substrates. Thus, it is crucial to understand how cellulolytic clostridium species regulate their biomass degradation mechanisms in response to different carbon sources for developing natural or engineered cellulases and their host cells for efficient production of cellulose-based biofuels.

Ruminiclostridium papyrosolvens is a mesophilic, low-GC, Gram-positive, spore-forming, and cellulosome-producing anaerobe bacteria that can be grown on a variety of carbohydrates [3]. This bacterium stands out as one of the most highly evolved species within the *Ruminiclostridium* genus. In our previous study, we sequenced and completed the genome of *R. papyrosolvens* DSM2782 to understand the mechanism of lignocellulose degradation [5]. Its genome also harbors two large cellulosomal clusters: *cip-cel* operon [13, 14] encoding cellulases and *xyl-doc* gene cluster encoding hemicellulases. The *xyl-doc* gene cluster is controlled by its upstream two-component system (TCS) XydS/R [15, 16]. Secretomes showed *R. papyrosolvens* secrets different sets of enzymes to breakdown different types of complex biomass [3]. In addition, we have successfully developed the mature genetic manipulation system for *R. papyrosolvens*, including electrotransformation [15] and ClosTron methods [17, 18]. These studies have laid a solid foundation for further exploration of *R. papyrosolvens* in degradation of lignocellulose.

In this study, to understand the mechanisms of biomass degradation and metabolism of *R. papyrosolvens*, we first compared the transcriptomic profiles of *R. papyrosolvens* grown on various carbon sources. Subsequently, we performed the cluster analysis for expression patterns of CAZymes including cellulosomal subunits and identified the ABC transporters responsible for the uptake of various sugars. Following this, we confirmed the regulatory role of two-component systems (TCSs) in the expression of both CAZymes and ABC transporters. Finally, we compared the metabolomes between cellobiose and corn stover. This study provides valuable insights into the transcriptome and metabolome information of *R. papyrosolvens*, which would be useful for exploring its application in biodegradation and utilization of lignocellulose resources.

Materials and methods

Bacterial strain growth

R. papyrosolvens was anaerobically cultured at 35 °C in 100 mL flasks with 50 mL working volume of GS-2 medium supplemented with 3.0 g/L of glucose, cellobiose, xylan, cellulose, and corn stover. Cell growth on glucose, cellobiose, and xylan was monitored by optical density of the culture at 600 nm (OD600), while that on cellulose and corn stover was measured based on increase of cellular proteins in the culture using the bicinchoninic acid assay [19]. All cultivations were performed in triplicate.

Total RNA extraction

The total RNA was extracted from *R. papyrosolvens* samples cultured under different carbon sources and harvested at the mid-exponential-phase using Total RNA Extraction Reagent (Yeasen, Shanghai) and EZ-10 Total RNA Mini-Prep Kit (Sangon, Shanghai). Then the total RNA samples were quantified using a NanoDrop 2000 spectrophotometer (Thermo, USA).

Transcriptomic analysis

To study the gene expression of *R. papyrosolvens* in different carbon sources, the total RNA was extracted from glucose, cellobiose, xylan, cellulose, and corn stover cultures. Three parallel experiments were set in each group. RNA of the samples was sequenced using an Illumina NovaSeq 6000 platform (2 × 150 bp paired ends). The raw reads were filtered with fastp (v0.23.4) [20] to remove the low-quality reads and adapter sequences. Bowtie2 (v2.5.1) [21] was used to remove the rRNA sequences. Next, the generated clean reads were mapped to the genome of *R. papyrosolvens* using Hisat2 (v2.2.1) [22]. The number of reads mapped to each gene was counted using featureCounts (v2.0.6) [23]. We used the FPKM values for normalization to make the expression levels of different genes and samples comparable, FPKM values were used to assess the levels of gene expression in different carbon sources.

Differential gene expression and principal component analysis

The transcriptome samples were further analyzed for differential expression, enrichment, and clustering. The samples were clustered based on gene expression data by performing principle component analysis (PCA). The expression levels of genes were calculated using FPKM values. Differential expression analysis was performed using DESeq2 package [24]. Differentially

expressed genes (DEGs) were identified with a threshold of the padj < 0.05 and |log2FC| > 1. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis of DEGs was performed with a threshold of $p \leq 0.05$.

Construction of the *R. papyrosolvens* mutants

Plasmids for targeted disruption were constructed based on ClosTron [17, 18]. Targeting sites for disruption and the intron retargeting primers were designed with the online tool based on Perutka algorithm (http://clostron.com/). Gene SOEing method was performed to achieve the 353 bp targeting regions using pSY6-Pxyl as template. Reconstruction plasmid was constructed by inserting the corresponding targeted regions into *Xho*I and *Bsr*GI sites of pSY6-Pxyl, and targeted gene disruption in *R. papyrosolvens*. Electroporation was performed based on the previously described method [18]. Transformants carrying target plasmids were selected on GS-2 liquid medium supplemented with erythromycin. The mutants targeted for gene destruction were isolated on GS-2 solid medium supplemented with erythromycin. Because ClosTron would result in gene disruption by inserting an about 900-bp intron into the targeting site, the mutants were verified and distinguished from the wild-type by PCR using corresponding primer sets upstream and downstream targeting sites. Then the mutants were inoculated and cultivated in medium without erythromycin pressure to loss of the plasmid. All manipulations were performed under anaerobic conditions.

Quantitative reverse transcription-PCR (qRT-PCR) experiment

We measured the transcript level of target genes via qRT-PCR. RNA was reverse transcribed using the SPARKscript II One Step RT-PCR Kit (SparkJade, China), and the amplification was performed with 2×SYBR Green qPCR Mix (SparkJade, China) using the CFX96 real-time PCR detection systems (Bio-Rad, USA). 16s rRNA were used as house-keeping genes for validation of target genes in *R. papyrosolvens*. Three biological and three technical replicates were performed for each sample. Gene expression levels were estimated using the $2^{-\Delta\Delta Ct}$ method [25].

Extraction of metabolites of *R. papyrosolvens*

To extract the metabolites produced by *R. papyrosolvens* in cellobiose and corn stover, the bacteria pellets (1×10^7) were mixed with 1000 µL of the culture solution (methanol:acetonitrile:water = 2:2:1; v/v) that contains deuterated internal standards. The mixed solution was homogenized by grinding at 35 Hz for 4 min and ultrasonicated in ice water for 5 min. Subsequently, the samples were sonicated for 10 min in 4 ℃ water bath, and incubated for 1 h at − 40 ℃ to precipitate proteins. Finally, the samples were centrifuged at 12000 rpm for 15 min at 4 ℃. The supernatant was transferred to a fresh glass vial for analysis.

Analysis of non-target metabolomes in *R. papyrosolvens*

Metabolite profiling was performed using Orbitrap Exploris 120 mass spectrometer (Thermo Fisher Scientific, USA). For polar metabolites, a ACQUITY UPLC BEH Amide column (Waters, USA) was used for Vanquish high-performance liquid chromatography (Thermo Fisher Scientific, USA) separation. The mobile phase comprised elution A (25 mmol/L ammonium acetate and 25 mmol/L ammonium hydroxide in water) and elution B (acetonitrile). The auto-sampler temperature was 4 ℃, and the injection volume was 2 µL. For non-polar metabolites, a Kinetex C18 column (Phenomenex, USA) was used for Vanquish high-performance liquid chromatography (Thermo Fisher Scientific, USA) separation. The mobile phase comprised elution A (0.01% acetic acid in water) and elution B (2-propanol:acetonitrile = 1:1; v/v). The auto-sampler temperature was 4 ℃, and the injection volume was 2 µL.

The raw data were converted to mzXML format using ProteoWizard (v3.0.21229), and processed by R package XCMS (v4.1.12) to generate a data matrix consisting of retention time (RT), mass-to-charge ratio (m/z) value, and peak abundance [26]. The metabolites were identified through the R package MetDNA2 and annotated with the BiotreeDB (Biotree Biotech. Co., Ltd., China) [27]. The similarity cutoff was set at 0.7. The measurements were not counted when the detected peaks were less than half of the QC samples or the standard deviation was greater than 30%. In addition, metabolites with a VIP value greater than 1 and a *P* value less than 0.05 were identified as differential metabolites. Volcano plot was used to graphically represent the differences of metabolite expression. Pathway enrichment analysis was analyzed based on the KEGG database, with pathways having a *P* value of 0.05 or lower being considered statistically significant.

Results

Global analysis of transcriptional responses of *R. papyrosolvens* to different carbon source

To identify the components of the lignocelluloses utilization in *R. papyrosolvens*, we characterized the transcriptional profiles of *R. papyrosolvens* under a variety of carbon sources using RNA-seq [28]. The carbohydrate substrates tested included cellulose and its derivatives cellobiose and glucose, xylan as a representative of hemicellulose, and corn stover (a natural plant-derived lignocellulose) (Fig. 1A). In total, we sequenced 15 RNA

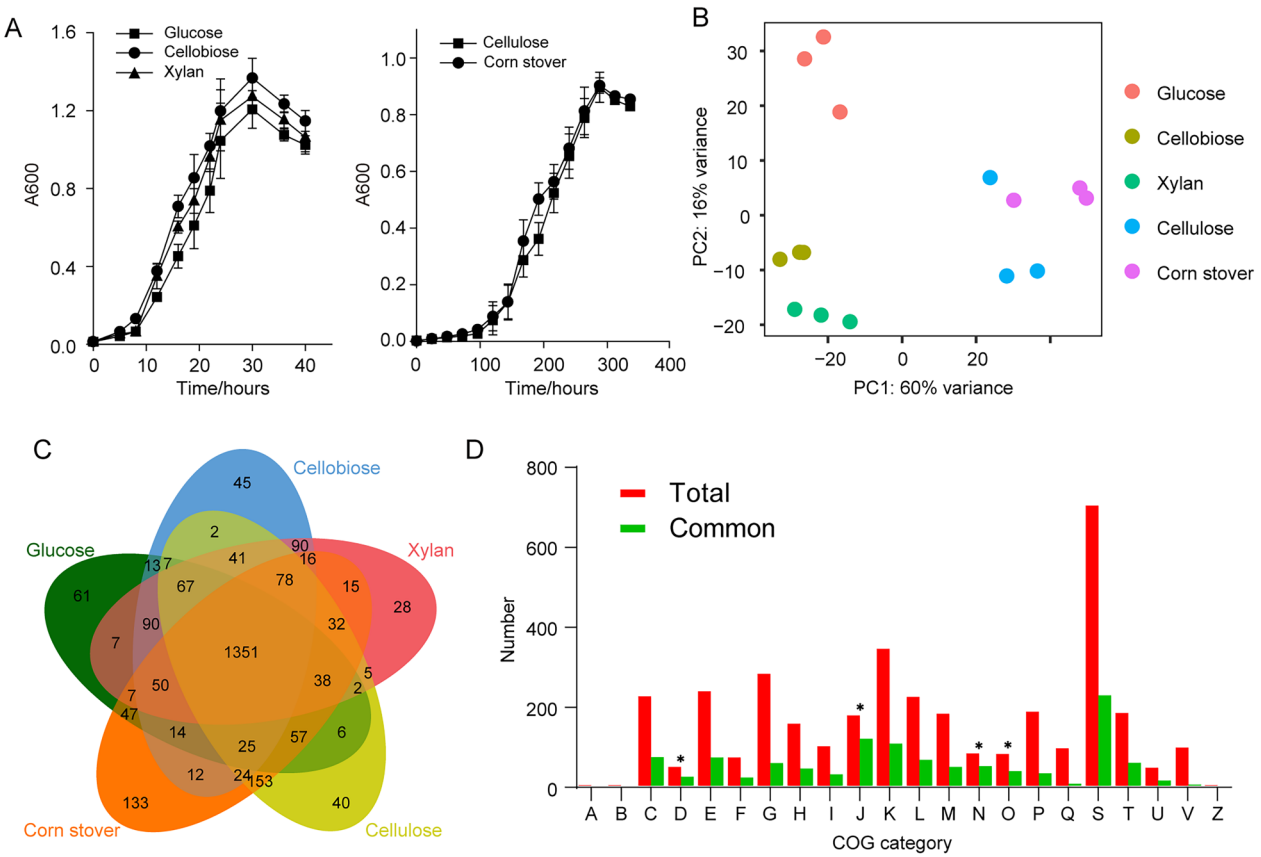

Fig. 1 Growth and gene expression of *R. papyrosolvens* under various carbon sources. **A** Growth curves of *R. papyrosolvens* cultured on 3 g/L glucose, cellobiose, xylan, cellulose, and corn stover. Error bars indicate the standard deviation of the three replicates. **B** Principal component analysis of transcriptomic samples. The color of the scatter plots represents the grouping of samples. **C** Venn diagram of the distribution of the encoding genes expressed as mean FPKM (calculated on 3 biological replicates) values among glucose, cellobiose, xylan, cellulose, and corn stover. **D** Functional annotation of total genes annotated in the genome (Total) and common genes expressed under five carbon sources (Common). Number of the genes in each COG term are shown in columns (D, cell cycle control, cell division, chromosome partitioning; J, translation, ribosomal structure and biogenesis; N, cell motility; O, posttranslational modification, protein turnover, chaperones; A, RNA processing and modification; B, chromatin structure and dynamics; C, energy production and conversion; E, amino acid transport and metabolism; F, nucleotide transport and metabolism; G, carbohydrate transport and metabolism; H, coenzyme transport and metabolism; I, lipid transport and metabolism; K, transcription; L, replication, recombination and repair; M, cell wall/membrane/envelope biogenesis; P, inorganic ion transport and metabolism; Q, secondary metabolites biosynthesis, transport and catabolism; S, function unknown; T, signal transduction mechanisms; U, intracellular trafficking, secretion, and vesicular transport; V, defense mechanisms; Z, cytoskeleton). Genes identified were enriched in D, J, N, and O of COG categories (*$P < 0.01$, hypergeometric test)

libraries of five carbon sources in three biological replicates, yielding a range of 28.3–40.0 million paired-end reads. 27.5–39.7 million of the reads passed the quality control, 0.14–6.65% of the clean reads were aligned to the rRNA sequences. Specifically, 25.9–36.7 million paired-end reads were successfully aligned to the reference genome, with mapping rates that ranged between 93.94 and 97.88%. Among these, 2.2–11.7 million of the clean reads were assigned to the non-coding sequences, while 14.9–33.1 million of the clean reads were assigned to the coding sequences (CDS) (Table S1). We calculated the fragments per kilobase of exon model per million mapped (FPKM) values for each sample to investigate the gene expression differences among different carbon sources in *R. papyrosolvens* (Table S2). Principal component analysis (PCA) revealed that samples from different carbon sources formed distinct groups, clearly separating except for corn stover and cellulose. The expression profiles of *R. papyrosolvens* grown on cellulose and corn stover exhibited strong correlation, as did those grown on cellobiose and xylan. In contrast, the expression profiles on glucose displayed a notably distant clustering (Fig. 1B). This pattern is reflective of the availability of carbon sources: cellulose and corn stover are more

challenging to metabolize, whereas cellobiose and xylan are more readily utilized. Interestingly, glucose, despite being a preferred carbon source for many microorganisms, appears to be less favored for cellulolytic clostridia, indicating a unique metabolic preference or regulatory response in *R. papyrosolvens*.

Furthermore, we compared the number of highly expressed genes, defined as those with FPKM values greater than 60 [29], among five data sets for different carbon sources using a Venn diagram [30]. First, the number of highly expressed genes was as follows: 1,842 in glucose, 1,925 in cellobiose, 1,917 in xylan, 1,928 in cellulose, and 2,052 in corn stover. Notably, 1,351 genes were found to be shared among all carbon sources (Fig. 1C). This finding suggests that *R. papyrosolvens* exhibits the most extensive gene transcription under corn stover, likely reflecting its capacity to degrade the complex substrate. Moreover, we further analyzed the distribution of COG (Cluster of Orthologous Group) [31] functional category and numbers of these common highly expressed genes under the five carbon sources. A broad spectrum of COGs was found, with these genes being enriched in COG categories such as D (cell cycle control, cell division, chromosome partitioning), N (cell motility), J (translation, ribosomal structure and biogenesis), and O (posttranslational modification, protein turnover, chaperones) (Fig. 1D).

Transcription differences in *R. papyrosolvens* grown on cellulose and corn stover compared to cellobiose

Corn stover, cellulose, and cellobiose were considered as complex, difficult, and easy carbon sources for *R. papyrosolvens*, respectively, that is reflected in the distinct clustering patterns observed in PCA. Thus, differential expression analysis was first performed between cellulose and cellobiose. The differentially expressed genes (DEGs) were identified at the threshold of padj < 0.05 and $|\log_2 FC| > 1$. The analysis revealed 1,065 upregulated and 756 downregulated genes under cellulose compared to cellobiose (Fig. 2A).

To understand the functional implications of these DEGs, a pathway enrichment analysis was performed using the Kyoto Encyclopedia of Genes and Genomes (KEGG) database [32] with the clusterProfiler package [24]. The upregulated genes in cellulose versus cellobiose were found to be significantly enriched in several KEGG

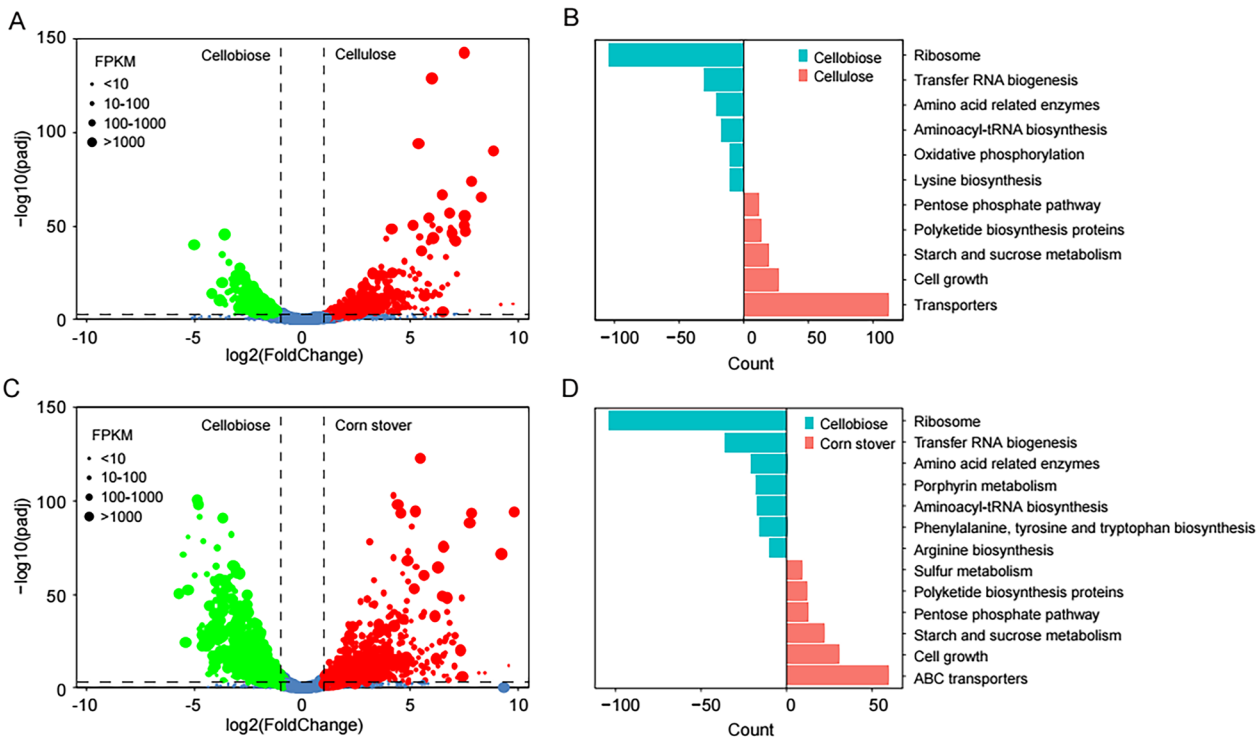

Fig. 2 Difference of gene expression between cellobiose and cellulose or corn stover of *R. papyrosolvens*. **A** and **C** Volcano map of differentially expressed genes (DEGs) in cellulose (**A**) and corn stover (**C**) compared to cellobiose. Red-colored dots indicate significant upregulation of genes with adjusted padj < 0.05 and log2FC > 1, and green-colored dots indicate significant downregulation of genes with padj < 0.05 and log2FC < −1. **B** and **D** KEGG pathway enrichment analysis of DEGs of cellulose (**B**) or corn stover (**D**) compared to cellobiose. The x-axis represents the number of DEGs enriched in each pathway, and the y-axis indicates the pathway classification of KEGG

pathways, including transporters, cell growth, starch and sucrose metabolism, polyketide biosynthesis proteins, and pentose phosphate pathway. Conversely, the downregulated genes were notably enriched in pathways associated with ribosome, transfer RNA biogenesis, amino acid related enzymes, aminoacyl-tRNA biosynthesis, oxidative phosphorylation, and lysine biosynthesis (Fig. 2B).

On the other hand, when differential expression analysis was performed comparing corn stover to cellobiose, we identified 1,312 upregulated and 1,092 downregulated genes (Fig. 2C). The upregulated genes in corn stover versus cellobiose comparison were found to be significantly enriched in several KEGG pathways, including ABC transporters, cell growth, starch and sucrose metabolism, pentose phosphate pathway, polyketide biosynthesis proteins, and sulfur metabolism. Conversely, the downregulated genes were notably enriched in pathways associated with ribosome, transfer RNA biogenesis, amino acid-related enzymes, porphyrin metabolism, aminoacy-tRNA biosynthesis, the biosynthesis of phenylalanine, tyrosine and tryptophan, and arginine biosynthesis (Fig. 2D). This observation is consistent with the inherent complexity of carbon sources. When *R. papyrosolvens* was grown on the complex substrate of corn stover, it necessitates the expression of a broader array of enzymes to degrade the substrates effectively and a variety of transporters to uptake produced sugars. In contrast, when *R. papyrosolvens* was cultivated on the preferred cellobiose, it exhibits rapid growth and a robust metabolism, particularly in the biosynthesis of protein and amino acid.

Expression pattern of CAZymes in *R. papyrosolvens*

Carbohydrate-active enzymes (CAZymes) play a crucial role in breakdown of extracellular carbohydrate sources by cleaving, building, and rearranging oligo- and polysaccharides. They are categorized into various classes within the CAZyme database (http://www.cazy.org/) [33]. To investigate the links between CAZymes and substrates, a total of 191 putative CAZyme genes were identified in *R. papyrosolvens* using the carbohydrate-active enzyme annotation (dbCAN3) database, which contains the HMM profiles for each CAZy category. These CAZymes include 1 auxiliary activity (AA), 56 carbohydrate-binding modules (CBMs), 23 carbohydrate esterases (CEs), 111 glycoside hydrolases (GHs), 70 glycosyltransferases (GTs), and 3 polysaccharide lyases (PLs). In addition, 73 putative cellulosomal subunits contain cohesin subunits (3) and dockerin subunits (70) (Table S3).

To elucidate the influence of diverse carbon sources on the induction of CAZymes, we performed a co-expression analysis of CAZyme genes involved in lignocellulose degradation, including those from GH, CE, PL families, as well as cellulosomal subunits, among different substrates. Based on their substrate-dependent transcription patterns, the 131 CAZyme genes were clustered into seven different groups (Fig. 3A). Group I: This group comprises 21 genes that exhibit higher expression levels under glucose, cellobiose, and xylan, including non-cellulosomal GH43, GH94, and other families. Specifically, members of the GH94 family, involved in intracellular phosphorolytic cleavage of cellodextrin and cellobiose, were highly expressed on cellobiose. Group II: Characterized by ten genes, this group showed high expression specifically in response to cellobiose and xylan. These genes encode non-cellulosomal enzymes from GH42, GH39, CE20, and CE4 families. Group III: This group is primarily composed of cellulosomal components from GH43, 10, 9, 27, and PL11 families, including hemicellulases encoded by the *xyl-doc* gene cluster. They are distinguished by higher expression level under corn stover than other carbon sources. These enzymes are crucial for the degradation of hemicelluloses. Group IV: This group includes 18 genes that predominantly show higher expression levels under corn stover. These genes encode non-cellulosomal GH51, CE1, and other families. The higher expression under corn stover suggests a significant role in the degradation of this complex substrate. Group V: This group is primarily composed of non-cellulosomal components that exhibit high expression under cellulose and corn stover. The genes belong to GH94, CE4, and other families involved in the degradation of cellulose and hemicelluloses. Group VI: This group is primarily composed of cellulosomal components from GH5, 9, 48, 74, and other families, including cellulases encoded by the *cip-cel* gene operon, highlighting the specificity of cellulosomal enzymes in the breakdown of cellulose. Group VII: This group includes 13 genes that show high expression specifically in response to xylan. The genes encode CAZymes from GH10, 11, 43 and other GH and CE families, indicating a specialized role in the degradation of xylan (Fig. 3A). In addition, we observed that genes from GH43 family are induced by variety of different carbon sources. This may indicate a greater emphasis on broad-ranging lignocellulose hydrolysis. CAZymes from families GH5 [34] and GH9 [35], known as cellulases, were highly expressed under cellulose. In contrast, CAZymes from family GH10 [36], recognized as xylanases, showed high expression under xylan and corn stover. The results of gene expression patterns in *R. papyrosolvens* under different carbon sources revealed a substrate-dependent regulation of catalytic components. This insight is crucial for understanding the metabolic strategies employed by this bacterium in response to different carbohydrate resources.

In our analysis of the cellulosomal subunits encoded by the *cip-cel* and *xyl-doc* clusters among different carbon

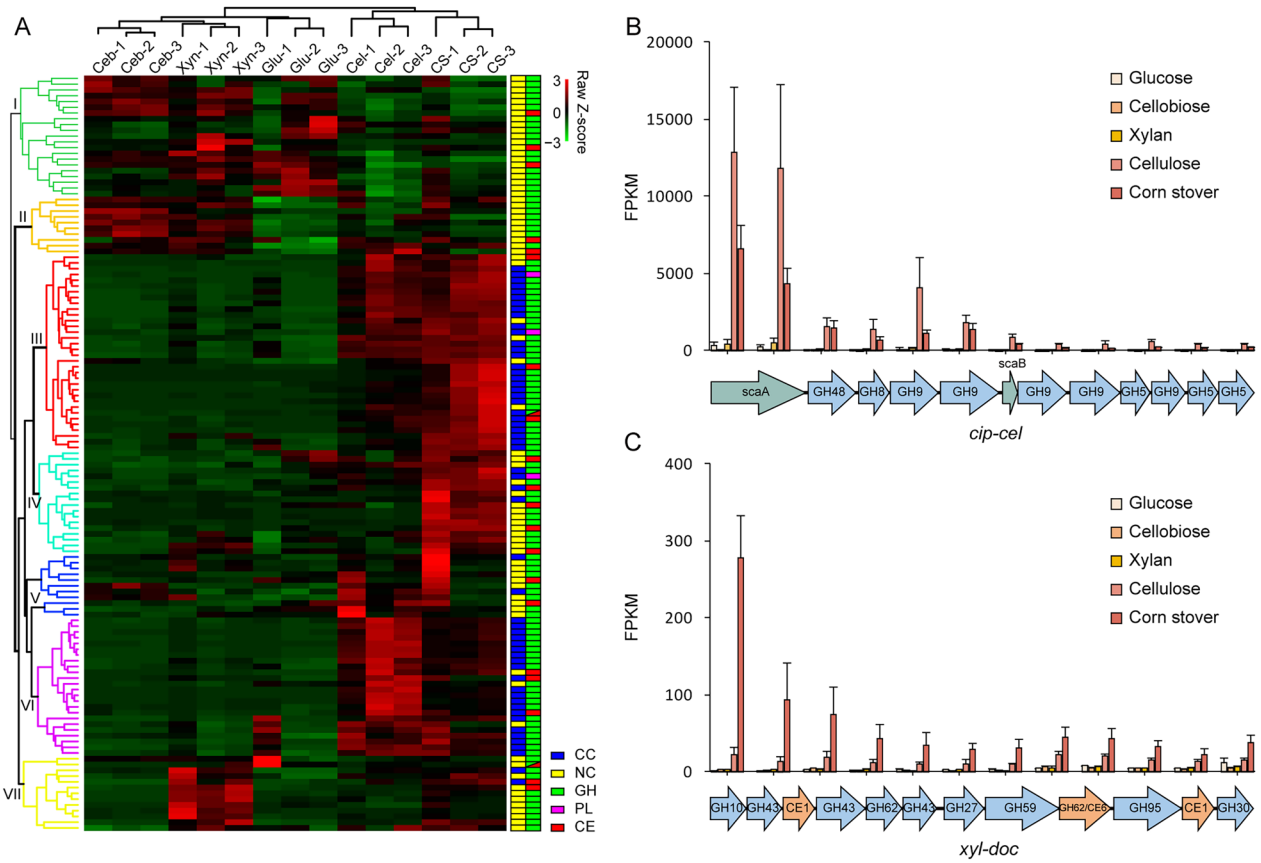

Fig. 3 Expression profile of CAZymes in *R. papyrosolvens* genes in response to five carbon sources. **A** Heatmap of CAZyme gene expression. The expression values of the duplicate sample for five carbon sources (glucose (Glu), cellobiose (Ceb), xylan (Xyn), cellulose (Cel), corn stover (CS)) were used to plot the heatmap by hierarchical cluster analysis. The color scale indicates the relative expression levels, with red representing high expression and green indicating low expression. The structural [cellulosomal component (CC) and non-cellulosomal enzyme (NC)] and functional [glycoside hydrolase (GH), carbohydrate esterase (CE), polysaccharide lyase (PL)] characteristics of CAZymes were distinguished by different color blocks. **B** and **C** The expression comparison of the *cip-cel* (**B**) and *xyl-doc* (**C**) cellulosomal gene clusters under various carbon sources. The values shown are the means of three replicates, and the error bars indicate standard deviations from the mean values

sources, we observed similar patterns of gene expression. The transcriptomic data revealed that both the *cip-cel* and *xyl-doc* gene clusters were significantly upregulated in response to cellulose and corn stover. Notably, the *cip-cel* operon exhibited higher expression under cellulose than corn stover. Conversely, the *xyl-doc* clusters showed higher expression under corn stover than cellulose. In addition, we found that the expression of upstream genes within these clusters was much higher than that of downstream genes (Fig. 3B, C). For example, within the *cip-cel* operon, the expression levels of the first six genes were higher than those of the last six. Specifically, the first two genes (P0092_RS10440, which encodes the scaffoldin ScaA, and P0092_RS10445, encoding an exoglucanase of GH48) showed a marked increase in expression compared to the others (Fig. 3B). The genes in the *xyl-doc* cluster seemed to exhibit a gradual decrease in transcription levels in the direction of transcription (Fig. 3C). In total, the high expression of both the *cip-cel* and *xyl-doc* gene clusters under cellulose and corn stover suggests that they play a crucial role in the degradation of lignocellulosic biomass.

ABC transporters and CAZymes are regulated by two-component systems

In addition to utilizing CAZymes for the degradation of lignocelluloses, bacteria must also deploy sugar transporters to import sugars into cells. Unlike many solventogenic clostridia that transport sugars with numerous phosphotransferase systems (PTS) [5, 37, 38], *R. papyrosolvens* instead adapts to sugars by modulating the expression of different types of ATP-binding cassette (ABC) transporters [5]. The genome analysis of *R. papyrosolvens*, as annotated by KEGG, has revealed a total of 18 gene clusters encoding sugar ABC transporter systems, including 4 Aldouronate, 4 Raffinose/

Stachyose/Melibiose, 1 Galactofuranose, 1 Arabinooligosaccharide, 2 myo-Inositol, 1 cellobiose, 1 Multiple sugar, 1 Oligogalacturonide, 1 Fucose, 1 Ribose/Autoinducer2/D-Xylose, and 1 sn-Glycerol 3-phosphate. Significantly, 14 of these ABC transporter gene clusters are found to be adjacent to two-component systems (TCSs) and/or CAZymes (Fig. 4A). This genetic arrangement suggests a sophisticated regulatory mechanism that allows *R. papyrosolvens* to efficiently adapt to and utilize a diverse range of sugar substrates.

Transcriptomic analysis has revealed that genes encoding ABC transporters are significantly upregulated in corn stover (Fig. 2D). Consequently, a clustering analysis was conducted on all ABC transporters of sugars under various carbon sources. The results showed that various sugar ABC transporter systems responded with increased expression levels to different substrates. For example, the gene clusters P0092_RS13490-13500, P0092_RS14120-14130, and P0092_RS15905-15915 were notably upregulated in the presence of mono- and disaccharides, such as glucose and cellobiose. When

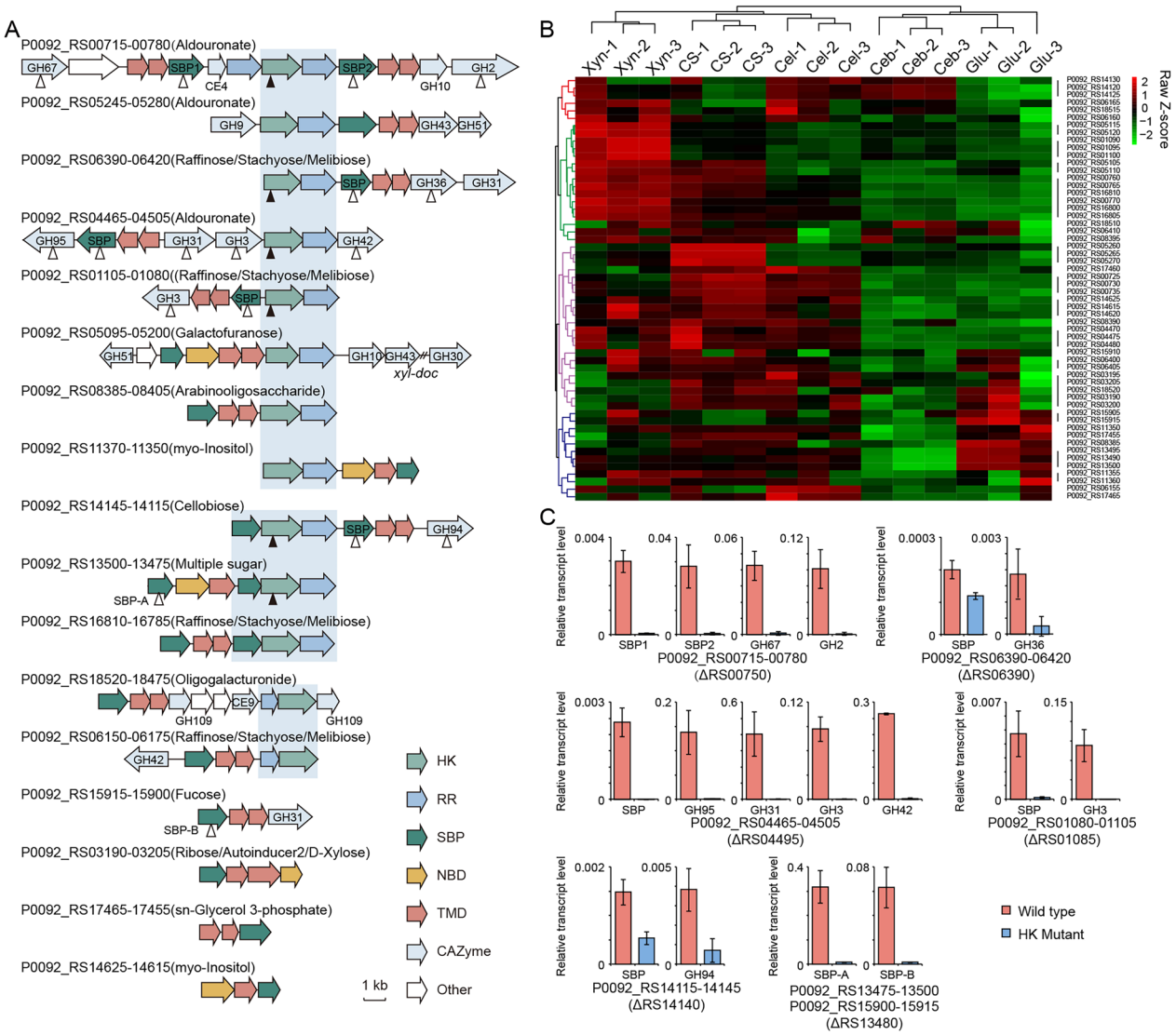

Fig. 4 Regulation of ABC transporters and CAZymes by TCSs in *R. papyrosolvens*. **A** Genetic organization of loci encoding ABC transporters involved in sugar import. Genes encoding ABC transporters are flanked by genes encoding TCSs, and/or CAZymes in 13 loci. Closed triangles indicate the mutated HK genes, while open triangles indicate genes that were analyzed by qPCR. **B** Heatmap of sugar transport-related ABC transport systems. **C** qRT-PCR analysis of relative expression of SBPs and CAZymes in the wild-type and HK mutants. The 16s rRNA gene was used as control. Error bars indicate the standard deviation of triplicate experiments

R. papyrosolvens is grown on polysaccharides, a different set of transporters is activated. Notably, gene clusters P0092_RS00760-00770, P0092_RS01090-01100, P0092_RS05105-05120, and P0092_RS16800-16810 are particularly highly expressed in xylan. For cellulose, the clusters P0092_RS06155-06165, P0092_RS14120-14130, and P0092_RS04470-04480 showed increased expression. In the context of corn stover utilization, a range of transporters are highly expressed, including P0092_RS00725-00735, P0092_RS03190-03205, P0092_RS04470-04480, P0092_RS05260-05270, P0092_RS06400-06410, P0092_RS11350-11360, and P0092_RS14615-14620 (Fig. 4B). These findings underscored the substrate-specific regulation of sugar transporter systems in *R. papyrosolvens*, highlighting its adaptability to diverse carbon sources.

To further confirm the role of TCSs in regulation of ABC transporters and CAZymes [39–41], we engineered disruptions to six genes encoding histidine kinases of TCSs using the ClosTron system (Fig. 4A and S1). Subsequently, we examined the transcriptional level of genes encoding substrate-binding proteins (SBPs) of ABC transporters and CAZymes that are proximal to the mutated TCSs. The results showed that the disruption of the histidine kinase genes would lead to a suppression of transcription in their neighboring ABC transporters and CAZymes. Intriguingly, the gene cluster P0092_RS00715-00780 contains a single histidine kinase gene (P0092_RS00750), which is flanked by two response regulator genes and two ABC transporter genes. Upon disruption of P0092_RS00750, transcription of the proximal SBPs and CAZymes was repressed, indicating that the histidine kinase encoded by P0092_RS00750 sent sugar signals to two response regulators, thereby regulated the transcription of two ABC transporters (Fig. 4C).

On the other hand, some gene clusters encoding ABC transporters that are not in close proximity to TCS still exhibit regulated transcription under different carbon sources. For example, the gene cluster of P0092_RS15915-15900, which encodes an ABC transporter and lacks adjacent TCS, displayed high transcription under glucose and cellulose. This pattern is consistent with the expression profile of the gene cluster P0092_RS13500-13490 that is adjacent to a TCS (Fig. 4B). Remarkably, the disruption of the histidine kinase gene (P0092_RS13480) resulted in the suppression of expression for both the P0092_RS15915-15900 and P0092_RS13500-13490 ABC transporters, suggesting that these two ABC transporters were under the control of the same TCS (Fig. 4C). Thus, the findings underscore the pivotal role of a single TCS in regulating transcription of multiple ABC transporters, either in close proximity or at a distance.

Central carbon metabolism in *R. papyrosolvens*

A variety of lignocellulose-derived sugars are selectively transported into *R. papyrosolvens* cells through their specific ABC transporters. Subsequently, hexoses are converted to fructose-6-phosphate and enter the glycolytic pathway. For example, glucose, mannose, galactose, and cellobiose can directly enter the glycolytic pathway with their respective enzymes. Pentoses such as arabinose and xylose use an isomerase pathway to enter the non-oxidative pentose phosphate pathway (PPP) to form glyceraldehyde 3-phosphate and fructose 6-phosphate, which later enter into the Embden–Meyerhof–Parnas (EMP) pathway (Fig. 5A).

Furthermore, the transcript level of the genes encoding metabolic pathway of hexoses was compared among five carbon sources. Firstly, extracellular cellobiose and cellodextrins are imported into cells by employing cellobiose ABC transport system (P0092_RS14120, P0092_RS14125, P0092_RS14130). These substrates are then further metabolized to generate glucose and glucose-1-phosphate by cellobiose/cellodextrin phosphorylase, which was highly expressed on cellobiose and cellulose. In addition, cellobiose is also taken into cells through PTS system (P0092_RS18685, P0092_RS18690, P0092_RS18695) and is converted into cellobiose-6-phosphate. The compound is subsequently cleaved into glucose and glucose-6-phosphate by 6-phosphate-glucosidase (P0092_RS18670), which was highly expressed on glucose and cellulose. And then glucose and glucose-1-phosphate are transformed into glucose-6-phosphate by glucokinase (P0092_RS10265, P0092_RS18395) and phosphoglucomutase (P0092_RS06665), respectively. Notably, P0092_RS10265, which encodes glucokinase, was found to be highly expressed on cellobiose, whereas its isoenzyme P0092_RS18395 was predominantly expressed on corn stover, indicating distinct functional roles. Secondly, galactose is phosphorylated to galactose-1-phosphate by galactokinase (P0092_RS06435), and is further converted into glucose-1-phosphate by galactose-1-phosphate uridylyltransferases (P0092_RS03730, P0092_RS06430). P0092_RS03730 showed high expression on cellulose and corn stover, while P0092_RS06430 was highly expressed on glucose, cellobiose, and xylan. Lastly, mannose is phosphorylated to mannose-6-phosphate by hexokinase (P0092_RS20815) and is subsequently isomerized to fructose-6-phosphate by mannose-6-phosphate isomerase (P0092_RS16565) (Fig. 5B). These results underscore the complexity and adaptability of the hexose metabolic pathways in response to different carbon sources, highlighting the potential for metabolic engineering to enhance the efficiency of hexose utilization in industrial applications.

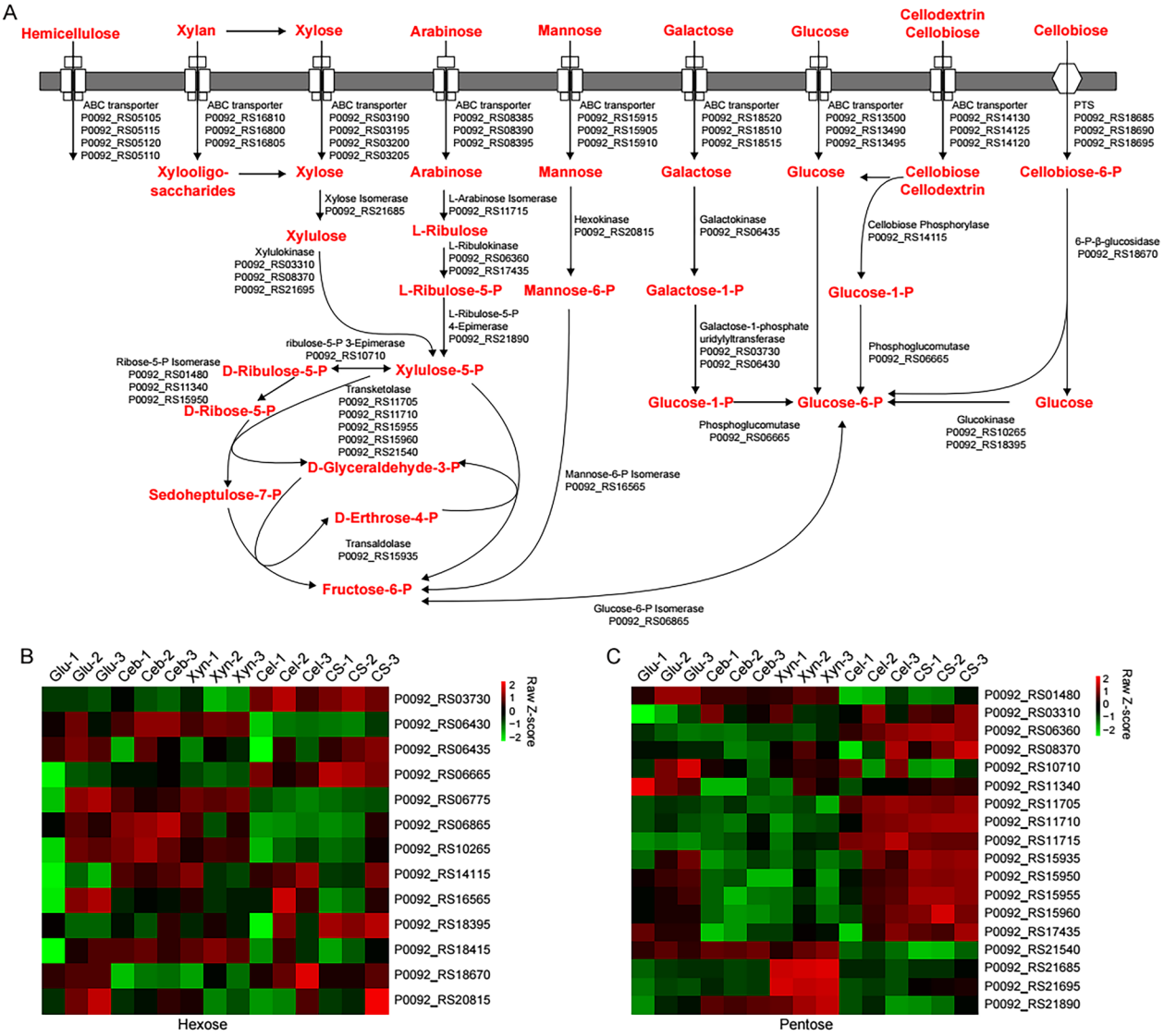

Fig. 5 Differential expression of genes involved in utilization of various sugars derived from lignocelluloses in *R. papyrosolvens*. **A** Schematic representation of the pathways for utilization of different carbon sources. (B and C) Heatmaps of gene expression patterns associated with hexose (**B**) and pentose (**C**) utilization among five different carbon sources

Moreover, the pentoses such as xylose and arabinose derived from hemicellulose are efficiently metabolized by *R. papyrosolvens* via the non-oxidative branch of the pentose phosphate pathway. Specifically, xylose is converted into xylulose by xylose isomerase (P0092_RS21685), and is subsequently phosphorylated to xylulose-5-phosphate by xylulokinases (P0092_RS03310, P0092_RS08370, and P0092_RS21695). Both types of these enzymes exhibited high levels of expression on xylan and corn stover. Similarly, arabinose is isomerized to L-ribulose by arabinose isomerase (P0092_RS11715) and then phosphorylated to L-ribulose-phosphate by ribulokinases (P0092_RS06360 and P0092_RS17435), with both types of enzymes showing heightened activity on corn stover. L-Ribulose-phosphate is further transformed into xylulose-5-phosphate by ribulose-5-phosphate-4-epimerase, an enzyme that is notably expressed on xylan. Xylulose-5-phosphate ultimately enters the Embden–Meyerhof–Parnas (EMP) pathway, in which transketolases (P0092_RS11705, P0092_RS11710, P0092_RS15955, and P0092_RS15960) and transaldolase (P0092_RS15935) were highly expressed on cellulose and corn stover (Fig. 5C). This metabolic pathway highlights the versatility and efficiency of *R. papyrosolvens* in converting hemicelluloses-derived pentoses into intermediates of central metabolism.

Comparative metabolic profiling of *R. papyrosolvens*

To further understand the mechanisms of lignocellulose degradation by *R. papyrosolvens*, we determined the metabolic profiling of *R. papyrosolvens* grown on corn stover compared with cellobiose using non-target metabolomic methods [42]. The metabolites with significant differences were identified in corn stover and cellobiose (Table S4). The results of KEGG pathway enrichment showed that the metabolites were mainly enriched in metabolic pathway, nucleotide metabolism, ABC transporters, phosphotransferase system (PTS), alanine, aspartate and glutamate metabolism, purine metabolism, vitamin B6 metabolism, biosynthesis of cofactors and C5-branched dibasic acid metabolism (Fig. 6A).

Furthermore, we compared the difference in metabolite abundance between corn stover and cellobiose. The significant differential metabolites were identified with a VIP (variable importance in projection) value greater than 1 and a *P* value less than 0.05. The findings revealed a significant upregulation of 349 metabolites and a downregulation of 45 metabolites in corn stover compared with cellobiose, suggesting that *R. papyrosolvens* produced more metabolites on corn stover (Fig. 6B). Our data indicated that during the fermentation process, *R. papyrosolvens* produced more metabolic intermediates of glycolytic pathways such as fructose 1-phosphate, fructose 6-phosphate, glucose 6-phosphate, mannose 6-phosphate, mannose 1-phosphate, lactate, and pyruvate, when utilizing cellobiose compared to corn stover (Fig. 6C). This suggests that *R. papyrosolvens* is capable of releasing more energy and thus grows more rapidly when metabolizing cellobiose. Conversely, *R. papyrosolvens* was observed to produce a greater amount of metabolic intermediates from EMP pathway, such as erythrose 4-phosphate, along with various secondary metabolites like vanillic acid, ferulate, arabitol, xylitol, and glyceric acid, during the fermentation of the corn stover than cellobiose (Fig. 6D). This underscores the differential metabolic strategies employed by *R. papyrosolvens* in response to distinct carbon sources.

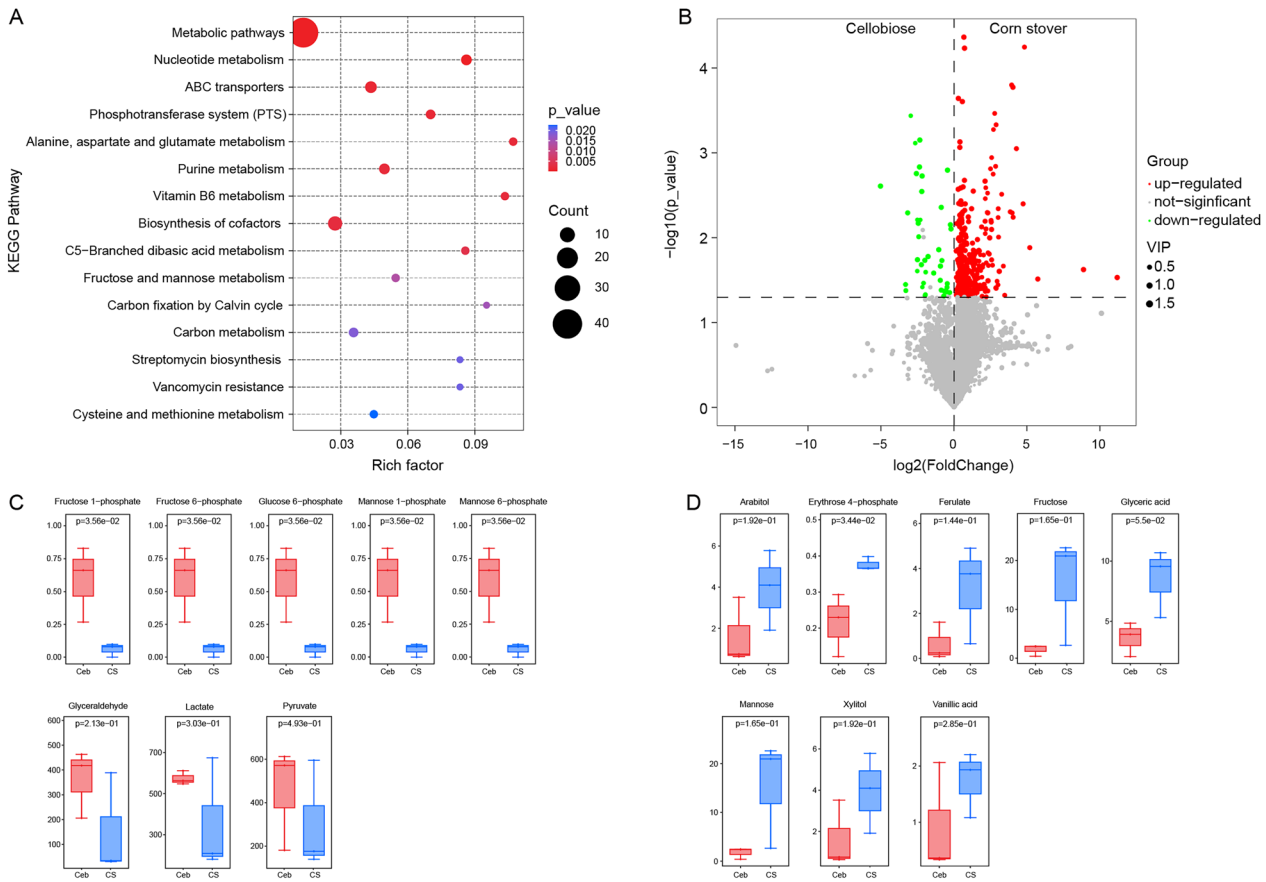

Fig. 6 Comparison of metabolite profile of *R. papyrosolvens* between corn stover and cellobiose. **A** KEGG pathway enrichment analysis of all metabolites identified in both corn stover and cellobiose. **B** Volcano plots showing differentially expressed metabolites in corn stover compared to cellobiose. Red-colored dots indicate significant upregulation of metabolites, and green-colored dots indicate significant downregulation of metabolites. **C** and **D** The box plots showing the compounds that were highly abundant in cellobiose (**C**) and corn stover (**D**)

Discussion

The type species of *Ruminiclostridium*, such as *Ruminiclostridium cellulolyticum*, *Ruminiclostridium josui*, and *R. papyrosolvens* [3, 4, 39], are renowned for their secretion of a repertoire of carbohydrate-active enzymes, including cellulases, hemicellulases such as xylanases, pectinases, and chitinases. Notably, some of these enzymes form cellulosomes for the efficient degradation of lignocellulose biomass. *R. papyrosolvens* stands out as one of the most highly evolved species within the *Ruminiclostridium* genus [15]; thus, it is worth to analyze its mechanisms for the degradation of lignocellulose.

In our study, we focused on 131 CAZyme genes that are specifically involved in the degradation of lignocellulose, which includes 57 cellulosomal genes and 74 free CAZyme genes. However, there were significant differences in their expression patterns. The cellulosomal genes predominantly exhibited high expression levels in cellulose and corn stover, whereas the genes encoding free CAZymes were found to be highly expressed in carbon sources that are more easily utilized, such as cellobiose and xylan. This pattern aligns with our previous findings in *R. cellulolyticum* [39], suggesting that cellulosomes may play a more crucial role in the degradation of lignocellulosic materials compared to free CAZymes. On the other hand, the enzymes with distinct functions were selectively induced to express in response to their specific substrates. For example, members of the GH94 family involved in intracellular phosphorolytic cleavage of cellodextrin and cellobiose were upregulated on cellobiose. Similarly, GH10 xylanases were found to be upregulated when xylan was present. High levels of GH5 and GH9 glucanases were observed on cellulose biomass. GH43 hemicellulases, which are responsible for the hydrolysis of a variety of different hemicelluloses, were highly expressed on corn stover. In addition, the expression patterns observed for the *cip-cel* and *xyl-doc* gene clusters in *R. papyrosolvens* are remarkably similar to those in *R. cellulolyticum*. This suggests that the regulatory mechanisms found in *R. cellulolyticum*, including carbon catabolite repression (CCR) [39], TCS [39, 43], and selective RNA processing and stabilization (SRPS) [44], are likely to function in *R. papyrosolvens* as well.

Anaerobic, mesophilic, and cellulolytic bacteria, such as *Ruminiclostridium termitidis*, and *R. cellulolyticum*, mainly employ ABC transporter systems to import sugars derived from lignocelluloses [39, 45, 46]. The expression of these ABC transporters is typically controlled by TCSs, which are encoded by genes located in proximity to the transporter genes themselves (Fig. 4). Canonical TCSs are composed of a histidine kinase (HKs) equipped with an extracellular sensor domain and a response regulator (RRs). However, three genomic loci of TCSs harbor an additional gene that encodes sugar-binding proteins (SBPs). These SBPs serve as signal collectors, amplifying the TCS's response to extracellular sugars. As a result, these augmented systems can be classified as three-component systems, which provide an additional layer of regulation and control over the bacterial sugar uptake mechanisms. The three-component system is also found in *R. cellulolyticum* and *Clostridium beijerinckii* [43, 47], indicating that it may play a broad role in these bacteria. Furthermore, our results revealed that TCSs not only control the expression of ABC transporters located adjacent to their genes but also exert control over ABC transporters that are situated at a considerable genomic distance. For example, the TCS encoded by P0092_RS13475-13485 simultaneously control the expression of the ABC transporters encoded by P0092_RS15915-15900 and P0092_RS13500-13490. It underscores the intricate and far-reaching nature of TCS-mediated gene regulation, which could have implications for understanding how bacteria adapt to varying environmental conditions and nutrient availability. It also highlights the potential for TCSs to serve as master regulatory switches for the expression of multiple transport systems, influencing the bacteria's metabolic flexibility and ecological fitness.

Bacteria, like *R. papyrosolvens*, exhibit adaptable metabolic machineries that are able to handle fluctuating environmental carbohydrate availability. The regulation of these processes is complex and appears to be controlled by a combination of CCR and operon-specific regulators [39]. First, *R. papyrosolvens*, like other cellulolytic clostridia such as *R. cellulolyticum* [48] and *Clostridium thermocellum* [49], shows a preference for cellobiose over glucose as a carbon source. Notably, *R. papyrosolvens* harbors an additional PTS for the transport of cellobiose, complementing the ABC transporter system that is a common ortholog among cellulolytic bacteria [5]. This preference suggests that cellobiose may act as a CCR trigger in these bacteria, contrasting with the role of glucose in *Escherichia coli* and *Bacillus subtilis*. Second, genes that encode metabolic pathways of hexoses and pentoses tend to cluster together, such as those involved in the pentose phosphate pathways. Meanwhile, a gene encoding a transcriptional regulator is typically found upstream or downstream of these gene clusters. These gene clusters were co-transcribed and exhibited substrate specificity, suggesting that the regulator acts as an operon-specific regulator to control the expression of these gene clusters. Third, *R. papyrosolvens*, known for their production of major metabolic pathway products of *Ruminiclostridium* species such as acetic acid, formate, ethanol, and lactic acid [6, 45, 50], also exhibits the ability to generate short-chain fatty acids like butyric acid by the butyrate kinase (P0092_RS03275). Thus, it can be

considered as a probiotic for animal intestines, akin to *Clostridium butyricum* [51], due to its butyric acid production, which promotes intestinal health. Furthermore, *R. papyrosolvens* is capable of releasing valuable end-products from corn stover, including xylitol, vanillic acid, and ferulate [52–54]. This study has not only enhanced our understanding of the physiology and metabolism of *R. papyrosolvens*, but also laid a foundation for future metabolic engineering research.

Conclusion

In conclusion, based on transcriptomic and metabolomics techniques, the differential expression of genes and metabolites in *R. papyrosolvens* under different substrate conditions was comprehensively analyzed. The important related functional genes of the lignocellulose degradation process were also analyzed. The key genes of *R. papyrosolvens* degrading lignocellulose model compounds are CAZymes, ABC transporters, and two-component systems. The pathways and metabolic profiling of lignocellulose degradation by *R. papyrosolvens* were determined. This study deepens our understanding of the composition of the lignocellulose-degrading enzymes and metabolic pathways, and laid the foundation for the future development of strategies to systematically regulate the lignocellulose biodegradation.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s13068-025-02619-4.

Supplementary material 1: Figure S1. PCR identification of mutations in six histidine kinase genes that were disrupted by ClosTron in *R. papyrosolvens* DSM2782.

Supplementary material 2: Table S1. Summary of RNA-Seq data of *R. papyrosolvens* DSM2782.

Supplementary material 3: Table S2. The normalized expression values of all genes under different carbon sources in *R. papyrosolvens* DSM2782. In addition, including the differentially expressed genesin cellulose compared to cellobiose and corn stover compared to cellobiose, COG annotation, and KEGG pathway annotation in *R. papyrosolvens* DSM2782.

Supplementary material 4: Table S3. The genes prediction of the CAZymes and cellulosomal components in *R. papyrosolvens* DSM2782.

Supplementary material 5: Table S4. Comparative analysis of differentially expressed metabolites under corn stover and cellobiose in *R. papyrosolvens* DSM2782.

Supplementary material 6: Table S5. Strains, Plasmids and Primers used in this study.

Acknowledgements
Not applicable.

Author contributions
X, S, Y, and R designed experiments; Y, R, Y, L, and Z performed experiments; Y, R, Y, and X analyzed data; Y and X wrote the paper. All authors reviewed the manuscript.

Funding
This work was supported by Grants 32070045 and 32170053 from National Natural Science Foundation of China, Grant Z24C010001 from the Natural Science Foundation of Zhejiang Province and Grant 2022LFR065 from the Science Development Foundation of Zhejiang A&F University.

Availability of data and materials
All RNA-Seq read data has been submitted as sequence read archive (SRA) in NCBI with the BioProject ID PRJNA1112972.

Declarations

Ethics approval and consent to participate
Not applicable.

Consent for publication
Not applicable.

Competing interests
The authors declare no competing interests.

Received: 20 November 2024 Accepted: 11 February 2025
Published online: 22 February 2025

References

1. Munir RI, Schellenberg J, Henrissat B, Verbeke TJ, Sparling R, Levin DB. Comparative analysis of carbohydrate active enzymes in *Clostridium termitidis* CT1112 reveals complex carbohydrate degradation ability. PLoS ONE. 2014;9(8):e104260.
2. Desvaux M. *Clostridium cellulolyticum*: model organism of mesophilic cellulolytic clostridia. FEMS Microbiol Rev. 2005;29(4):741–64.
3. Ren Z, You W, Wu S, Poetsch A, Xu C. Secretomic analyses of *Ruminiclostridium papyrosolvens* reveal its enzymatic basis for lignocellulose degradation. Biotechnol Biofuels. 2019;12:183.
4. Sakka M, Kunitake E, Kimura T, Sakka K. Function of a laminin_G_3 module as a carbohydrate-binding module in an arabinofuranosidase from *Ruminiclostridium josui*. FEBS Lett. 2019;593(1):42–51.
5. You M, Zhao Q, Liu Y, Zhang W, Shen Z, Ren Z, Xu C. Insights into lignocellulose degradation: comparative genomics of anaerobic and cellulolytic *Ruminiclostridium*-type species. Front Microbiol. 2023;14:1288286.
6. Rettenmaier R, Kowollik ML, Klingl A, Liebl W, Zverlov V. *Ruminiclostridium herbifermentans* sp. Nov., a mesophilic and moderately thermophilic cellulolytic and xylanolytic bacterium isolated from a lab-scale biogas fermenter fed with maize silage. Int J Syst Evol Microbiol. 2019;71(3). https://doi.org/10.1099/ijsem.0.004692.
7. Alves VD, Fontes C, Bule P. Cellulosomes: highly efficient cellulolytic complexes. Subcell Biochem. 2021;96:323–54.
8. Bayer EA, Lamed R, White BA, Flint HJ. From cellulosomes to cellulosomics. Chem Rec. 2008;8(6):364–77.
9. Henrissat B, Davies G. Structural and sequence-based classification of glycoside hydrolases. Curr Opin Struct Biol. 1997;7(5):637–44.
10. Correia MA, Prates JA, Bras J, Fontes CM, Newman JA, Lewis RJ, et al. Crystal structure of a cellulosomal family 3 carbohydrate esterase from *Clostridium thermocellum* provides insights into the mechanism of substrate recognition. J Mol Biol. 2008;379(1):64–72.
11. Garron ML, Cygler M. Structural and mechanistic classification of uronic acid-containing polysaccharide lyases. Glycobiology. 2010;20(12):1547–73.
12. Lombard V, Bernard T, Rancurel C, Brumer H, Coutinho PM, Henrissat B. A hierarchical classification of polysaccharide lyases for glycogenomics. Biochem J. 2010;432(3):437–44.
13. Maamar H, Abdou L, Boileau C, Valette O, Tardif C. Transcriptional analysis of the *cip-cel* gene cluster from *Clostridium cellulolyticum*. J Bacteriol. 2006;188(7):2614–24.

14. Tao X, Xu T, Kempher ML, Liu J, Zhou J. Precise promoter integration improves cellulose bioconversion and thermotolerance in *Clostridium cellulolyticum*. Metab Eng. 2020;60:110–8.
15. Zou X, Ren Z, Wang N, Cheng Y, Jiang Y, Wang Y, Xu C. Function analysis of 5′-UTR of the cellulosomal *xyl-doc* cluster in *Clostridium papyrosolvens*. Biotechnol Biofuels. 2018;11:43.
16. Celik H, Blouzard JC, Voigt B, Becher D, Trotter V, Fierobe HP, et al. A two-component system (XydS/R) controls the expression of genes encoding CBM6-containing proteins in response to straw in *Clostridium cellulolyticum*. PLoS ONE. 2013;8(2):e56063.
17. Heap JT, Pennington OJ, Cartman ST, Carter GP, Minton NP. The ClosTron: a universal gene knock-out system for the genus *Clostridium*. J Microbiol Methods. 2007;70(3):452–64.
18. Wang D, You M, Qiu Z, Li P, Qiao M, Xu C. Development of an efficient ClosTron system for gene disruption in *Ruminiclostridium papyrosolvens*. Appl Microbiol Biotechnol. 2023;107(5–6):1801–12.
19. Xu C, Qin Y, Li Y, Ji Y, Huang J, Song H, Xu J. Factors influencing cellulosome activity in consolidated bioprocessing of cellulosic ethanol. Bioresour Technol. 2010;101(24):9560–9.
20. Chen S. Ultrafast one-pass FASTQ data preprocessing, quality control, and deduplication using fastp. Imeta. 2023;2(2):e107.
21. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie2. Nat Methods. 2012;9(4):357–9.
22. Kim D, Paggi JM, Park C, Bennett C, Salzberg SL. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. Nat Biotechnol. 2019;37(8):907–15.
23. Liao Y, Smyth GK, Shi W. featureCounts: an efficient general purpose program for assigning sequence reads to genomic features. Bioinformatics. 2014;30(7):923–30.
24. Wu T, Hu E, Xu S, Chen M, Guo P, Dai Z, et al. clusterProfiler 4.0: a universal enrichment tool for interpreting omics data. Innovation. 2021;2(3):100141.
25. Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. Methods. 2001;25(4):402–8.
26. Smith CA, Want EJ, O'Maille G, Abagyan R, Siuzdak G. XCMS: processing mass spectrometry data for metabolite profiling using nonlinear peak alignment, matching, and identification. Anal Chem. 2006;78(3):779–87.
27. Zhou Z, Luo M, Zhang H, Yin Y, Cai Y, Zhu ZJ. Metabolite annotation from knowns to unknowns through knowledge-guided multi-layer metabolic networking. Nat Commun. 2022;13(1):6656.
28. Stark R, Grzelak M, Hadfield J. RNA sequencing: the teenage years. Nat Rev Genet. 2019;20(11):631–56.
29. Luli Y, Zhou S, Li X, Chen Z, Yang Z, Luo H. Differential expression of amanitin biosynthetic genes and novel cyclic peptides in *Amanita molliuscula*. J Fungi. 2021;7(5):384.
30 Jia A, Xu L, Wang Y. Venn diagrams in bioinformatics. Brief Bioinform. 2021;22(5). https://doi.org/10.1093/bib/bbab108.
31. Galperin MY, Kristensen DM, Makarova KS, Wolf YI, Koonin EV. Microbial genome analysis: the COG approach. Brief Bioinform. 2019;20(4):1063–70.
32. Kanehisa M, Furumichi M, Sato Y, Kawashima M, Ishiguro-Watanabe M. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res. 2023;51(D1):D587–92.
33 Lombard V, Golaconda Ramulu H, Drula E, Coutinho PM, Henrissat B. The carbohydrate-active enzymes database (CAZy) in 2013. Nucleic Acids Res. 2014;42:D490-5.
34. Aspeborg H, Coutinho PM, Wang Y, Brumer H 3rd, Henrissat B. Evolution, substrate specificity and subfamily classification of glycoside hydrolase family 5 (GH5). BMC Evol Biol. 2012;12:186.
35. Ravachol J, Borne R, Tardif C, de Philip P, Fierobe HP. Characterization of all family-9 glycoside hydrolases synthesized by the cellulosome-producing bacterium *Clostridium cellulolyticum*. J Biol Chem. 2014;289(11):7335–48.
36. Abedi E, Fatemi F, Sefidbakht Y, Siadat SER. Development and characterization of a thermostable GH11/GH10 xylan degrading chimeric enzyme. Enzyme Microb Technol. 2021;149: 109854.
37. Xu T, Tao X, He H, Kempher ML, Zhang S, Liu X, et al. Functional and structural diversification of incomplete phosphotransferase system in cellulose-degrading clostridia. ISME J. 2023;17(6):823–35.
38. Voigt C, Bahl H, Fischer RJ. Identification of PTS(Fru) as the major fructose uptake system of *Clostridium acetobutylicum*. Appl Microbiol Biotechnol. 2014;98(16):7161–72.
39. Xu C, Huang R, Teng L, Wang D, Hemme CL, Borovok I, et al. Structure and regulation of the cellulose degradome in *Clostridium cellulolyticum*. Biotechnol Biofuels. 2013;6(1):73.
40. Ahmad A, Majaz S, Nouroz F. Two-component systems regulate ABC transporters in antimicrobial peptide production, immunity and resistance. Microbiology (Reading). 2020;166(1):4–20.
41. Alvarez AF, Georgellis D. Environmental adaptation and diversification of bacterial two-component systems. Curr Opin Microbiol. 2023;76:102399.
42. Ribbenstedt A, Ziarrusta H, Benskin JP. Development, characterization and comparisons of targeted and non-targeted metabolomics methods. PLoS ONE. 2018;13(11):e0207082.
43. Fosses A, Franche N, Parsiegla G, Denis Y, Mate M, de Philip P, et al. Role of the solute-binding protein CuaD in the signaling and regulating pathway of cellobiose and cellulose utilization in *Ruminiclostridium cellulolyticum*. Microorganisms. 2023;11(7):1732.
44. Xu C, Huang R, Teng L, Jing X, Hu J, Cui G, et al. Cellulosome stoichiometry in *Clostridium cellulolyticum* is regulated by selective RNA processing and stabilization. Nat Commun. 2015;6:6900.
45. Munir RI, Spicer V, Krokhin OV, Shamshurin D, Zhang X, Taillefer M, et al. Transcriptomic and proteomic analyses of core metabolism in *Clostridium termitidis* CT1112 during growth on alpha-cellulose, xylan, cellobiose and xylose. BMC Microbiol. 2016;16:91.
46. Wu S, You M, Wang N, Ren Z, Xu C. Internal transcription terminators control stoichiometry of ABC transporters in cellulolytic clostridia. Microbiol Spectr. 2022;10(2):e0165621.
47. Sun Z, Chen Y, Yang C, Yang S, Gu Y, Jiang W. A novel three-component system-based regulatory model for D-xylose sensing and transport in *Clostridium beijerinckii*. Mol Microbiol. 2015;95(4):576–89.
48. Fosses A, Mate M, Franche N, Liu N, Denis Y, Borne R, et al. A seven-gene cluster in *Ruminiclostridium cellulolyticum* is essential for signalization, uptake and catabolism of the degradation products of cellulose hydrolysis. Biotechnol Biofuels. 2017;10:250.
49. Yan F, Dong S, Liu YJ, Yao X, Chen C, Xiao Y, et al. Deciphering cellodextrin and glucose uptake in *Clostridium thermocellum*. mBio. 2022;13(5):e0147622.
50. Tao X, Morgan JS, Liu J, Kempher ML, Xu T, Zhou J. Target integration of an exogenous betaglucosidase enhances cellulose degradation and ethanol production in *Clostridium cellulolyticum*. Bioresour Technol. 2023;376:128849.
51. Stoeva MK, Garcia-So J, Justice N, Myers J, Tyagi S, Nemchek M, et al. Butyrate-producing human gut symbiont, *Clostridium butyricum*, and its role in health and disease. Gut Microbes. 2021;13(1):1–28.
52. Lee S, Monnappa AK, Mitchell RJ. Biological activities of lignin hydrolysate-related compounds. BMB Rep. 2012;45(5):265–74.
53. Manishimwe C, Feng Y, Sun J, Pan R, Jiang Y, Jiang W, et al. Biological production of xylitol by using nonconventional microbial strains. World J Microbiol Biotechnol. 2022;38(12):249.
54. Yalameha B, Nejabati HR, Nouri M. Cardioprotective potential of vanillic acid. Clin Exp Pharmacol Physiol. 2023;50(3):193–204.

Publisher's Note