Journal of Applied Microbiology, 2024, 135, lxad319
https://doi.org/10.1093/jambio/lxad319
Advance access publication date: 30 December 2023
Research Article

JOURNAL OF APPLIED MICROBIOLOGY
OXFORD UNIVERSITY PRESS

Transcriptome analysis reveals the underlying mechanism for over-accumulation of alkaline protease in *Bacillus licheniformis*

Anying Ji[1,‡], Xianliang Zheng[2,3,4,5,‡], Wei Yang[6], Ming Chen[2,3,4], Aimin Ma[1], Yongfeng Liu[6,*], Xuetuan Wei[1,*]

[1]State Key Laboratory of Agricultural Microbiology, Huazhong Agricultural University, Wuhan 430070, China
[2]AngelYeast Co., Ltd, Yichang 443003, China
[3]Hubei Provincial Key Laboratory of Yeast Function, Yichang 443003, China
[4]National Key Laboratory of Agricultural Microbiology, Yichang 443003, China
[5]State Key Laboratory of Biocatalysis and Enzyme Engineering, School of Life Sciences, Hubei University, Wuhan 430062, China
[6]GeneMind Biosciences Company Limited, Shenzhen 518001, China

[*]Corresponding authors. No. 1 Shizishan Street, Hongshan District, Wuhan 430070, Hubei, China. E-mails: liuyongfeng@genemind.com; weixuetuan@mail.hzau.edu.cn
[‡]A.J. and X.Z. contributed equally to this work.

Abstract

Aims: *Bacillus licheniformis* AQ is an industrial strain with high production of alkaline protease (AprE), which has great industrial application value. However, how to regulate the production of AprE in the process of industrial fermentation is still not completely clear. Therefore, it is important to understand the metabolic process of AprE production in the industrial fermentation medium.

Methods and results: In this study, transcriptome sequencing of the whole fermentation course was performed to explore the synthesis and regulation mechanism of AprE in *B. licheniformis* AQ. During the fermentation process, the AprE got continuously accumulated, reaching a peak of 42 020 U/mL at the fermentation endpoint (48 h). Meanwhile, the highly expressed genes were observed. Compared with the fermentation endpoint, there were 61 genes in the intersection of differentially expressed genes, functioning as catabolic processes, peptidases and inhibitors, chaperones, and folding catalysts. Furthermore, the protein–protein interactions network of AprE was constructed.

Conclusion: This study provides important transcriptome information for *B. licheniformis* AQ and potential molecular targets for further improving the production of AprE.

Impact Statement

These results contribute to understanding the underlying metabolic mechanism of alkaline protease production in *Bacillus licheniformis* AQ.

Keywords: alkaline protease; *Bacillus licheniformis*; transcriptomics; gene expression; gene regulation

Introduction

Alkaline protease (AprE) is an important industrial enzyme, and it has been widely applied in food processing, detergent industry, leather finishing, biotransformation reactions, waste treatment, bioactive peptides synthesis, etc. (Bougatef et al. 2012, Annamalai et al. 2013, Contesini et al. 2018, Barzkar 2020, Matkawala et al. 2021, Zhou et al. 2021a). At present, AprE accounts for >50% of the total enzyme yield in the world. However, it still cannot meet the industrial demand, leading to the shortage of large-scale industrial enzymes (Banerjee and Ray 2017, Barzkar 2020, Jiang et al. 2022). Therefore, further improving the yield of AprE has broad industrial application prospects.

Previous studies have reported that many microorganisms can produce AprE, such as *Aspergillus Niger* (Abdel Wahab and Ahmed 2018), *Aspergillus flavus* (Damare et al. 2020), *Bacillus amyloliquefaciens* (Jiang et al. 2022), *Bacillus subtilis* (Shafique et al. 2021), *Bacillus licheniformis* (Zhou et al. 2020b), *Bacillus proreolyticus* (Banerjee and Ray 2017), *Serratia marcescens* (Nam et al. 2013), *Aureobasidium pullulans* (Chi et al. 2007), *Pichia pastoris* (Guo and Ma 2008), and so on. Therein, *Bacillus* species are considered to be excellent host strains due to their superior protein secretion ability as well as the advantages of (Generally Recognized as Safe) GRAS status, easy cultivation, convenient gene modification, short fermentation cycle, and strong robustness in industrial fermentation (Liu et al. 2013, van Dijl and Hecker 2013, Mathew and Gunathilaka 2015, Cui et al. 2018, Cai et al. 2019, Zhou et al. 2019). These characteristics make *Bacillus* species the major producer of AprE in industrial fermentation, such as *B. licheniformis* 2709 (Zhou et al. 2019, 2020a, 2021a).

Currently, researchers have developed various genetic modification strategies to improve the yield of AprE, including optimization of expression elements (promoters, signal peptides, and ribosome binding sites), overexpression of molecular chaperone genes, and modification of the physiological state of host strains (Cai et al. 2019, Jiang et al. 2022). For

Received 31 August 2023; **revised** 18 December 2023; **accepted** 29 December 2023
© The Author(s) 2023. Published by Oxford University Press on behalf of Applied Microbiology International. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

Downloaded from https://academic.oup.com/jambio/article/135/2/lxad319/7504755 by Huazhong Agriculture University user on 16 February 2024

example, one AprE gene was cloned from *B. subtilis* and over-expressed in *B. amyloliquefaciens* BAX-9. After further promoter replacement and fermentation medium optimization, the activity of AprE reached 524.12 U/mL (Jiang et al. 2022). Deleting the regulatory gene *sigF* to hinder spore formation, the AprE activity of *B. licheniformis* reaches 29 494 ± 1053 U/mL, 20% higher than that of the wild-type strain (Zhou et al. 2019). Through binding transcription activator Spo0A to the regulatory region of AprE, the production of AprE was increased by 1.46 times in *B. licheniformis* (Zhou et al. 2020a). In conclusion, these modification strategies have been widely used to enhance AprE production. Although some achievements have been obtained in previous reports, the further improvement of fermentation yield is hindered by the unclear regulation mechanism and genetic operation targets. Besides, most strategies are only partial reinforcement, rather than starting from a global perspective (Zhou et al. 2021b). The ability of protease secretion and expression is a very complex process (Zhang et al. 2020). Although a single local reinforcement can increase production, it is more likely to cause intracellular metabolic imbalance, which makes it possible to increase production through continuous investment of time and cost. Therefore, it is necessary to understand the overall state of the fermentation process of *Bacillus* strains.

With the great progress of genomics, more and more efficient *Bacillus* strains with potential practical value have been selected for genome sequencing, while their actual physiological and metabolic states during the fermentation process cannot be obtained directly from the genome information (Wiegand et al. 2013, Han et al. 2017). Transcriptomics, proteomics, and metabolomics have been used as complementary tools to gain insight into the physiological changes of microbial cells (Han et al. 2017), and transcriptomics analysis has been developed as a necessary step to provide important information on the functional elements of gene expression and regulation (Sorek and Cossart 2010). Currently, transcriptomic analysis has been performed to understand the physiological changes of *B. licheniformis* in response to several fermentation processes, while the transcriptional change of *B. licheniformis* during practical industrial fermentation was rarely detected to study the synthesis mechanism of AprE (Zhou et al. 2021b, Wu et al. 2022). In this study, transcriptome analysis of different production stages in the practical industrial fermentation process was performed to reveal the possible synthesis and regulation mechanism of AprE.

Materials and methods

Strains and culture conditions

Bacillus licheniformis AQ was used as the original bacteria. A single clone from a fresh LB agar plate was cultivated in 500 mL of LB liquid medium (10 g/L NaCl, 10 g/L tryptone, 5 g/L yeast extract, with the pH 7.2–7.4) at 37°C with shaking at 220 rpm. After 48 h, the seed culture (8%) was inoculated into a 50-L bioreactor with 25 L fermentation medium containing corn starch (30 g/L), soybean meal (40 g/L), Na_2HPO_4 (2 g/L), Na_2CO_3 (1.25 g/L), and thermostable amylase (0.7 g/L). The pH was maintained at 6.3–7.8 by automatically adding ammonia. The dissolved oxygen in the fermentation process was controlled at 0%–50%. The rotation speed was maintained at 200–600 rpm, and the ventilation ratio was maintained at 5:4–5:6 (V/V). Fermentation time and sampling time vary in each set according to experimental purposes.

Detection of protease activity

AprE was determined using the Folin phenol method listed by the national standardization administration commission (Zhou et al. 2020a, Chen et al. 2022), using casein as the substrate (10 g/L). The boric acid buffer (9.54 g/L $Na_2B_4O_7 \cdot 10H_2O$, 1.60 g/L NaOH, pH = 10.5) was applied to dissolve the AprE. One milliliter casein solution was mixed with 1 mL diluted enzyme solution in a water bath (40°C) for 10 min, then 2 mL trichloroacetic acid was added into the mixture and reacted for 10 min. After centrifugation at 9500 × *g* for 10 min, 1 mL supernatant was mixed with 5 mL sodium carbonate solution and 1 mL Folin reagent, and the mixture was incubated in a water bath at 40°C for 20 min. Then, the absorbance of the mixture was measured at 680 nm. One unit of protease activity was defined as the amount of enzyme that hydrolyzed casein to produce 1 μg of tyrosine per minute. The protease activity was computed using the following equation: protaese activity U/mL = $(A \times 4 \times n)/10$, where A and n indicate the absorbance and dilution factor, respectively.

RNA extraction, library construction, and transcriptome sequencing

Cells were harvested by centrifugation at different fermentation phases in the 50-L fermentation tank. Total RNAs were extracted by TRIzol® Reagent (Invitrogen, Waltham, MA, USA) and purified by the Rio-Zero rRNA Removal Kit (Illumina, San Diego, CA, USA) following the manufacturers' instructions. The quantity and quality of the RNA were determined by NanoDrop™ 2000 (Thermo Scientific, Hampton, NH, USA) and 1% agarose gels. Then, the RNA libraries were constructed by TruSeq™ RNA sample preparation Kit (Illumina, San Diego, CA, USA) and sequenced with the GenoLab M (GeneMind Biosciences Co., Ltd) at SE50 bp read length (Liu et al. 2021). The sequencing data were filtered (Adapter removing, remaining reads with N% < 10%, Qphred > 10, length > 30bp) to produce clean reads. The raw GenoLab M sequencing dataset of the *B. licheniformis* AQ strain is available on the CNGB Sequence Archive (https://db.cngb.org/cnsa/) under project accession number CNP0004346.

Analyses of alignments, quantification, and differential expression

Cleansing reads were mapped against the genome of *B. licheniformis* ATCC 14580 (GCF_006094335.1) using HISAT2 v2.2.1 with no spliced and unique mapping reads (Kim et al. 2019). The resulting mapping files were quantified using StringTie v2.2.1 with default parameters. The differential expression analysis was carried out using the R packages edgR v3.40.2 with significantly differentially expressing genes being defined as having a *P*-value ≤ 0.05 and |log (Fold Change) ≥ 0.5|.

Analysis of Short Time-series Expression Miner, GO, and KEGG

Short Time-series Expression Miner (STEM) v1.3.13 was used for clustering and visualizing gene expression from short time series microarray experiments (Ernst and Bar-Joseph 2006). With the default parameters, those clustering gene sets that

Downloaded from https://academic.oup.com/jambio/article/135/2/lxad319/7504755 by Huazhong Agriculture University user on 16 February 2024

the number of genes assigned to a model profile compared to the expected number of genes assigned should be considered significant. Then, these gene sets were transformed to the (Gene Ontology) GO and (Kyoto Encyclopedia of Genes and Genomes) KEGG analysis via the PANTHER website (http://www.pantherdb.org/).

Analysis of protein–protein interactions during the AprE production

The amino acid sequences of RNA polymerase sigma factor and regulatory factors were downloaded, including YlaC (GBC64750.1), SigV (GBC66723.1), SigD (VEH78101.1), AbrB (CAA43955.1), ScoC (NP_388880.1), SinR (NP_390341.1), Spo0A (BAA12581.1), AbbA (P_389295.1), SalA (NP_388035.1), TnrA (NP_389214.1), and SinI (NP_390340.1). Then, these proteins were aligned to the protein set of *B. licheniformis* ATCC 14580 by BLASTP to get corresponding protein information. Finally, a protein list (29 proteins) was imported into the online Search Tool for the Retrieval of Interacting Genes/Proteins (STRING) database v11.5 (http://string-db.org) to obtain known and predicted protein–protein interactions (PPIs) (Szklarczyk et al. 2014).

Quantitative real-time (Polymerase Chain Reaction) PCR analysis

For *B. licheniformis* AQ, the expression levels of the selected 21 genes were validated by real-time RT-PCR analysis, including *aprE*, *sigY*, *sigX*, *sigW*, *sigM*, *ylaC*, *sigK*, *sigH*, *sigG*, *sigF*, *sigE*, *sigB*, *sigL*, *sigA*, *sigD*, *codY*, *ccpA*, *spo0A*, *tnrA*, *sinR*, and *abrB*. The cells were collected to extract the total RNA according to the description of the Total RNA Isolation Kit and Prime Script RT Master Mix Kit. DNA removal and RNA reverse transcription were performed simultaneously in one system. The reaction conditions (EasyScript® gDNA Removal and cDNA Synthesis) were 25°C for 10 min, 42°C for 15 min, and 85°C for 10 s. RT-PCR was performed following the conditions: 1 cycle (95°C for 5 min), 40 cycles (95°C for 30 s, 60°C for 30 s), and dissolution curve (1 min at 95°C, 1 min at 65°C) (Jiang et al. 2022). Relative quantitative analysis was performed with the 16S rRNA gene as the internal reference gene, and the expression level was normalized to the internal control 16S rRNA gene (Livak and Schmittgen 2001).

Statistical analysis

Each experiment was designed with three independent replicates. SPSS 20.0 (IBM, Armok, NY, USA) was used for statistical analysis, including calculating the means and standard deviations, and evaluating the significance. GraphPad Prism 8 (GraphPad Software Inc., San Diego, CA, USA) was used to process the data and prepare the graphs.

Comparative transcriptome analysis with *B. licheniformis* ATCC14580

We downloaded raw sequencing data of the transcriptome of *B. licheniformis* ATCC14580 cultured in LB containing 1% (w/v) corn starch at 37°C for 36 h from the SRA database (Liu et al. 2017). The transcriptome data of *B. licheniformis* AQ at 32 h was chosen to perform the comparative transcriptome analysis. After using Cuffdiff v2.2.1 with default parameters to calculate gene expression FPKM (Fragments Per Kilobase of exon model per million mapped fragments), respectively, the relative FPKM was calculated based on each data's FPKM of 16S rRNA. Then, the top 100 highly expression genes were used for GO and KEGG enrichment via EGGNOG-MAPPER v2. The genes were annotated and allocated to associated GO terms and KEGG pathways with the default parameters. We acquired detailed information on linked GO terms and KEGG pathways for those genes and performed GO and KEGG enrichment analysis using in-house scripts.

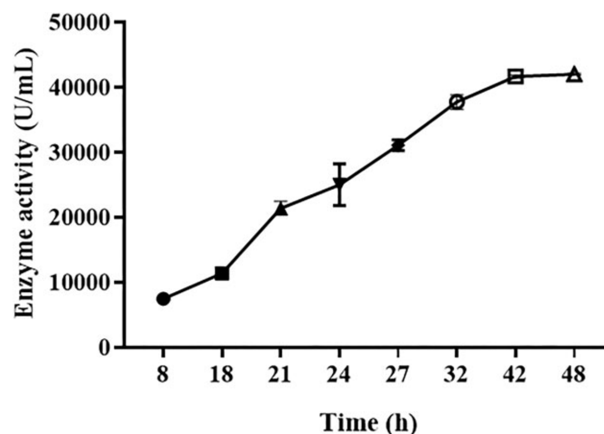

Figure 1. Changes of alkaline protease activity during the fermentation process.

Results

Fermentation characteristics of *B. licheniformis* AQ

Bacillus licheniformis AQ is an industrial producer with high AprE fermentation activities. In order to understand the synthesis process of the target product, the fermentation activities of AprE from *B. licheniformis* AQ were detected in a 50-L fermentation tank. As shown in Fig. 1, the extracellular AprE activity continuously increases as fermentation time is prolonged. At 48 h, the AprE activity peaked at 42 020 U/mL, which is higher than most strains derived from *B. licheniformis* 2709, such as BLΔF (23 910 ± 885 U/mL) (Zhou et al. 2020b), BLΔUEP (14 309.3 ± 486 U/mL) (Zhou et al. 2021a). *Bacillus licheniformis* AQ has an excellent industrial application value, while the regulation mechanism of AprE is still not deeply understood. Transcriptional analysis not only contributed to explaining the regulation mechanism of AprE, but also helped to mine the key factors to improve the production of AprE. Therefore, transcriptional analysis during the whole fermentation process was further carried out in the following section.

Gene expression pattern in the AprE fermentation

Under the industrial fermentation conditions, cells at eight time points were sampled for transcriptional analysis during the fermentation process. A total of 242 M high-quality reads were collected from these eight samples, and 3821 genes (91.74% of all 4165 genes) were expressed during the fermentation process. Obviously, the highly expressed genes were most at the fermentation endpoint (48 h), and the expression genes were in a normal distribution at 24 and 32 h (Fig. 2a).

Downloaded from https://academic.oup.com/jambio/article/135/2/lxad319/7504755 by Huazhong Agriculture University user on 16 February 2024

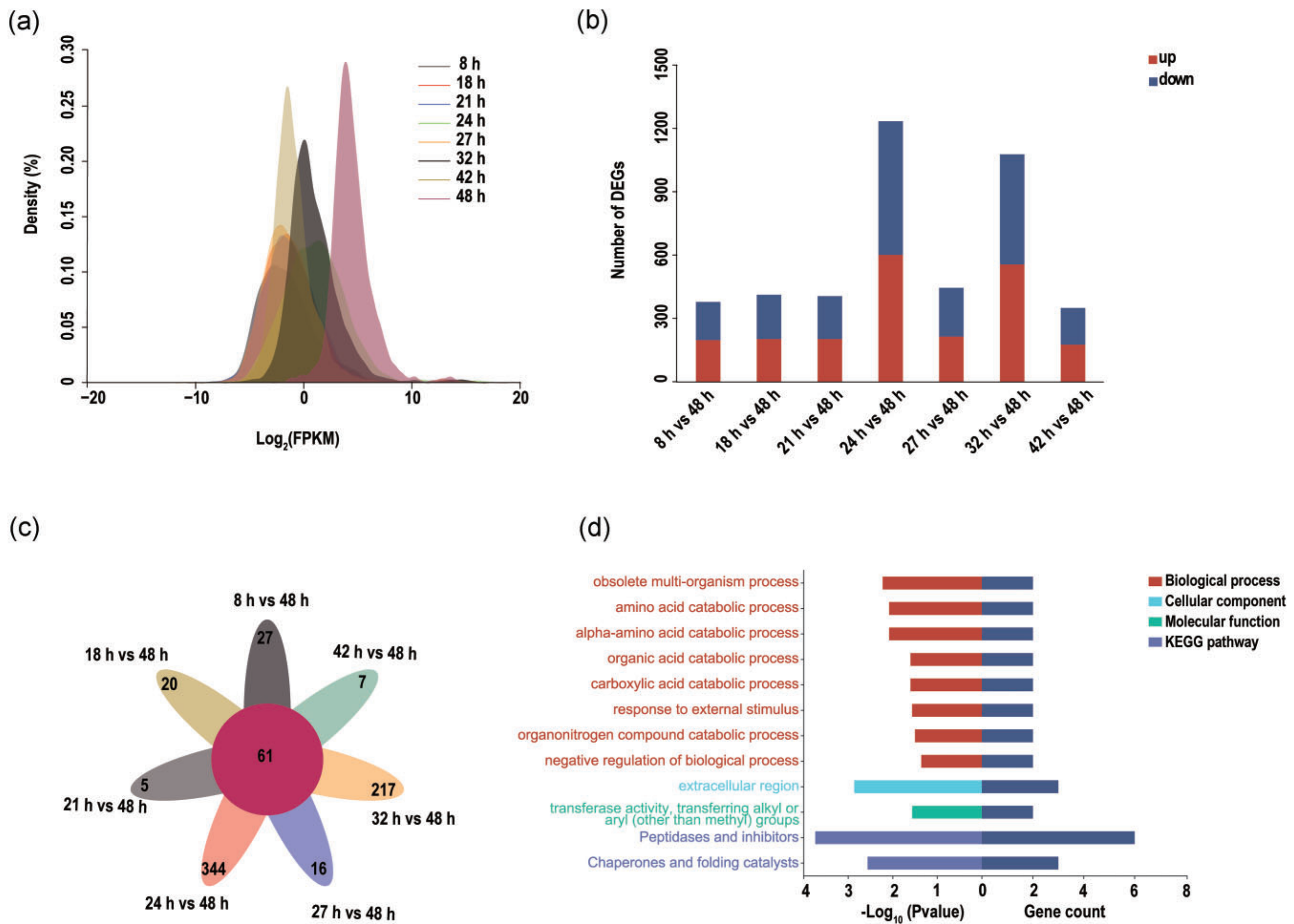

Figure 2. Analysis of the DEGs of *B. licheniformis* AQ over the entire alkaline protease fermentation curve. (a) The density plot of all genes expressing in different fermentation stages. (b) DEGs for all other samples comparing with DY_48 h group. (c) The flow plot shows the unique and common DEGs in each comparison. (d) GO and KEGG enrichment of 61 DEGs.

For other phases, gene expression was modest (Fig. 2a). Because the yield of AprE was highest at the fermentation endpoint (48 h), we compared the other stages to 48 h. At 24 h, a substantial number of differentially expressed genes (DEGs) were detected in *B. licheniformis* AQ, followed by 32 h (Fig. 2b). According to high-expression genes at 48 h, all other stages have more low-expression genes. Based on the common set in two random comparisons, the upset plot of seven comparisons revealed that 32 and 24 h have more close patterns than other stages (Fig. S1). Then, we displayed the 61 DEGs shared by all seven comparisons (Fig. 2c). Following function enrichment analysis, enrichment genes were mainly in biological processes, particularly numerous catabolic processes, such as amino acid, alpha-amino acid, organic acid, and organonitrogen compounds (Fig. 2d). While, peptidases and inhibitors, chaperones and folding catalysts (KEGG), and extracellular region (GO_Cellular component) were the most significantly enriched. The function of most DEGs is peptidases and inhibitors. Efficient protein secretion and folding are crucial during the production of recombinant proteins in *Bacillus*. These processes are assisted by components of the translocation system as well as by both intracellular and extracytoplasmic chaperones (Zhang et al. 2020). Enrichment analysis of this part also confirmed that the secretion and folding activities of AprE were very active during the production process.

The analysis of the short time series gene expression data based on whole-mRNA profile

Secondly, we analyzed the short time series gene expression data using the STEM software based on the whole transcriptome profile. Eventually, five profiles out of a total of 49 model profiles were significant models generated by STEM software (Fig. 3a). The model profiles, clusters 21 and 22, encompassing two essential genes: alkaline serine protease gene, consist of 290 and 2130 genes, respectively (Fig. S2). The expression trend analysis related to those genes in cluster 21 and cluster 22 is depicted in Fig. 3b and d. Because the trend of genes in cluster 41 perfectly matches the increased enzyme activity of the AprE, we simultaneously performed GO and KEGG analyses on clusteres 21, 22, and 41 (Fig. 3c). Although cluster 41 contains the least number of genes (166), it has the most enriched functional entries, only one less than the union of three clusters. The common functions of the two clusters include biosynthetic process, cellular biosynthetic process, small molecule metabolic process, organonitrogen compound metabolic process, organic substance biosynthetic process, transferase activity, and ko02000_Transporters. In KEGG, the term "Transporters" includes ATP-binding cassette (ABC) transporters (prokaryotic type), major facilitator superfamily, phosphotransferase system, and other transporters (https://www.genome.jp/kegg-bin/show_brite?ko02000.keg). Compared to other transporters, ABC transporters are widespread

Downloaded from https://academic.oup.com/jambio/article/135/2/lxad319/7504755 by Huazhong Agriculture University user on 16 February 2024

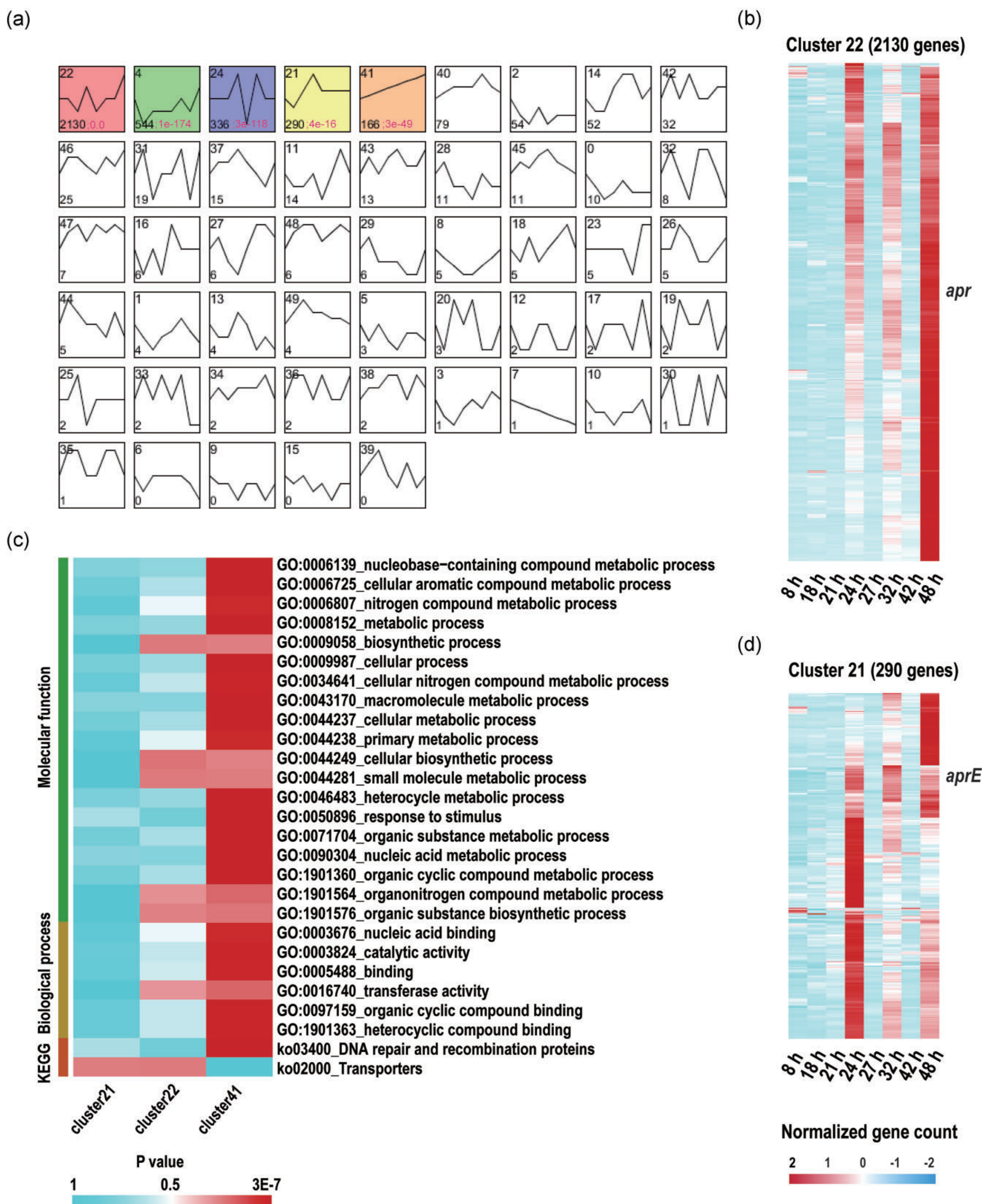

Figure 3. STEM cluster analysis and functional enrichment. (a) The analysis of short time series microarray gene expression of total transcriptoms in *B. licheniformis* AQ. The cluster ID was in the upper left corner, the gene number was in the lower left corner, and the red number was the *P*-value for each profile, (b) the heatmap of cluster 22 genes expression, (c) KEGG and GO enrichment analysis of significant profiles (clusters 21, 22, and 40). Red rectangles represent significant enrichment pathways ($P < 0.05$), (d) the heatmap of cluster 21 genes expression.

Downloaded from https://academic.oup.com/jambio/article/135/2/lxad319/7504755 by Huazhong Agriculture University user on 16 February 2024

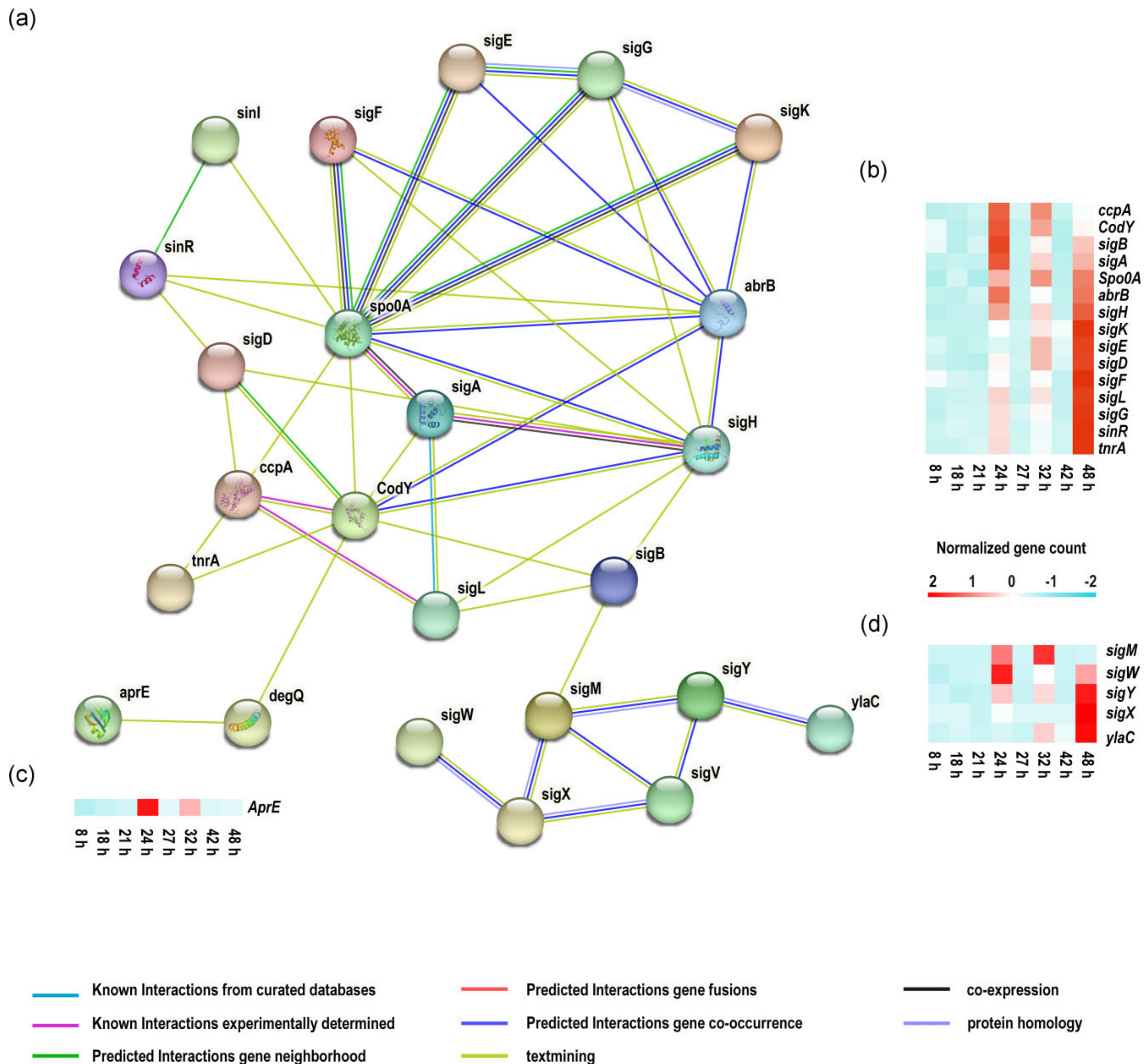

Figure 4. STRING PPI analyses of alkaline serine protease AprE and pleiotropic transcriptional factors and sigma factors. (a) PPI network, expression heatmap of genes encoding proteins in central cluster (b), left bottom cluster (c), and right bottom cluster (d). Clustering coefficient 0.539; enrichment *P*-value 1.0e − 16; average node degree 3.77. PPI legends indicate the eight types of interaction evidence. The heatmap of genes encoding proteins was shown the expression at different times. The color key meaning normalized gene count.

among living organisms and comprise one of the largest protein families (Garmory and Titball 2004). It is essential for many cellular processes and can import a variety of allocrites, including sugars and other carbohydrates, amino acids, peptides, polyamines, metal ions, sulfate, iron, and molybdate (Claverys 2001, Detmers et al. 2001, Hosie and Poole 2001, Igarashi et al. 2001, Kertesz 2001, Köster 2001, Schneider 2001, Self et al. 2001, Garmory and Titball 2004). Based on the analysis of transcriptome data of *B. licheniformis* TCCC11148, an important industrial strain to produce AprE, Yuan et al. found that most DEGs belonged to the ABC transporter, and it can indicate that ABC transporter played an important role in the production process of AprE (Yuan et al. 2020). It demonstrated significant changes in primary metabolism at different growth stages during the production of AprE, especially ABC transporters.

Candidate regulation network during the AprE fermentation

Based on the transcriptomic data, there were 17 genes encoding sigma factors involved in sporulation and important global transcription regulation, namely sigA, sigB, sigD, sigE, sigF, sigG, sigH, sigI, sigJ, sigK, sigL, sigM, sigV, sigW, sigX, sigY, and ylaC (Zhou et al. 2021b). We have chosen these 17 proteins from the genome of *B. licheniformis* ATCC 14580. In *B. subtilis*, extracellular protease AprE is directly regulated by several pleiotropic transcriptional factors, including AbrB, DegU, ScoC, and SinR (Barbieri et al. 2016). In addition, AprE is indirectly regulated by other proteins, including phosphorylated Spo0A (a repressor of abrB), AbbA (an inhibitor of AbrB), phosphorylated SalA and TnrA (both of which were reported to be repressors of scoC), SinI (an inhibitor of SinR),

Downloaded from https://academic.oup.com/jambio/article/135/2/lxad319/7504755 by Huazhong Agriculture University user on 16 February 2024

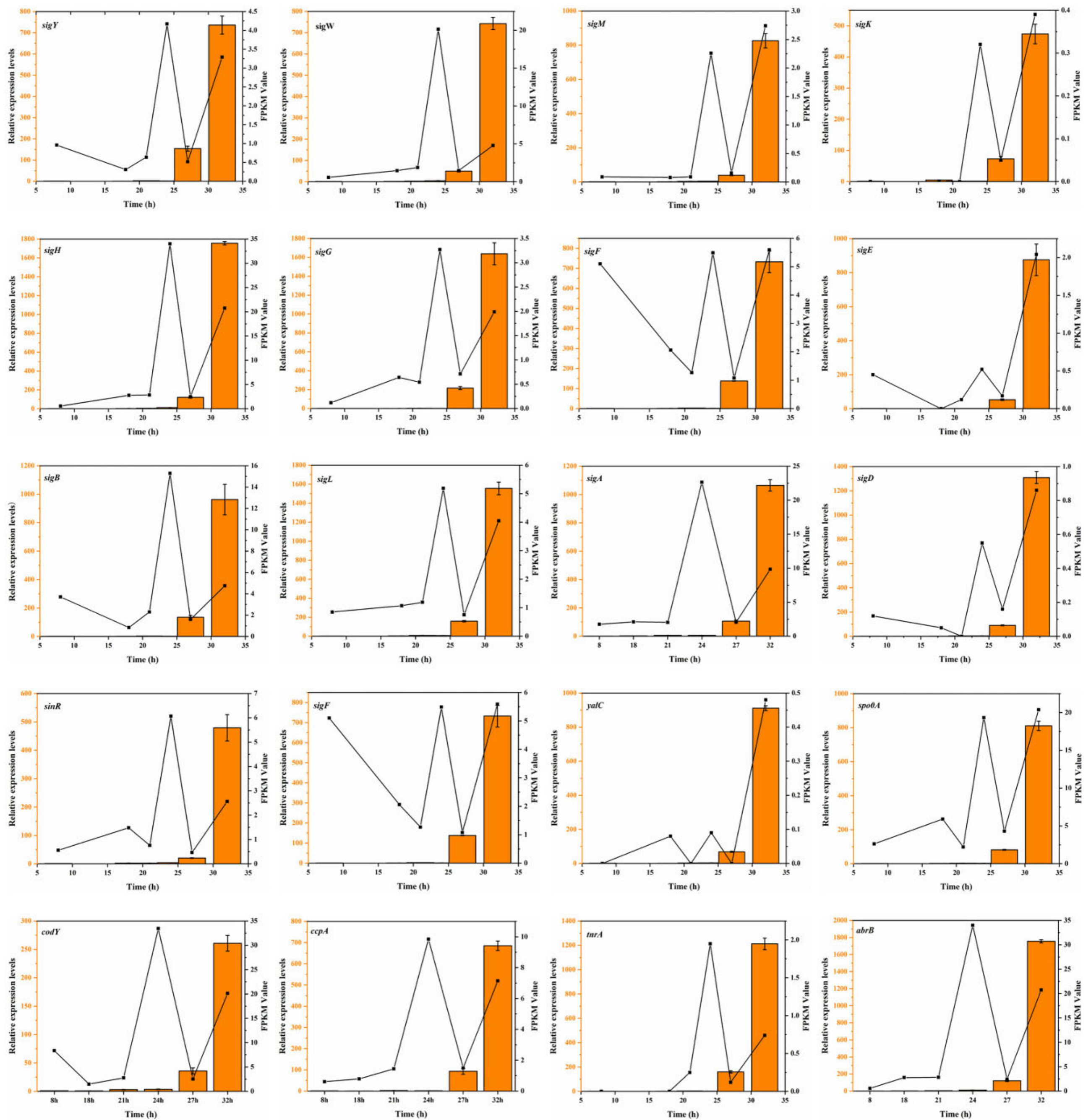

Figure 5. qRT-PCR validation of transcriptome data.

and DegQ (an activator of DegU phosphorylation) (Ogura et al. 2003, Ogura et al. 2004, Kobayashi 2007, Derouiche et al. 2015, Barbieri et al. 2016). CodY, the master transcriptional regulator, controls hundreds of genes in a large regulon in *B. subtilis* (Sonenshein 2007). CcpA, a second global transcriptional regulator that orchestrates fluxes in the central metabolism in *B. subtilis* (Cao et al. 2018). Therefore, we collected these 12 proteins (AbrB, DegU, ScoC, SinR, Spo0A, AbbA, SalA, TnrA, SinI, DegQ, CodY, and CcpA). After the amino acid BLASTP, we obtained the corresponding protein sequences in *B. licheniformis* ATCC 14580. A total of 30 proteins were searched in the STRING to obtain PPIs. Four proteins (ScoC, AbbA, SalA, and sigJ) were not collected in the database, and two proteins DegU and sigI have no relationship with other proteins. So, 22 proteins PPI network was shown in Fig. 4. Spo0A was in the center with the most nodes (11), then, CodY (9), sigH (9), and ArbB (8). Nevertheless, the expression of three genes (degQ, sigV, and sinI) could not be detected in any sample. The expression level of gene AprE peaked at 24 h, then decreased, rose to the second peak at 32 h, then decreased, and slightly increased at the end of fermentation. The expression pattern of six sigma factors in the lower right slightly differs from AprE. Therein, the sigM is the closest, while sigW and sigY have three peaks. Besides, transcriptional regulators CodY and CcpA showed a closer expression pattern to AprE. While, for *abrB* and *Spo0A*, except for 24 and 32 h after the peak expression, they increased again at the end of the fermentation. In this network, it is found

Downloaded from https://academic.oup.com/jambio/article/135/2/lxad319/7504755 by Huazhong Agriculture University user on 16 February 2024

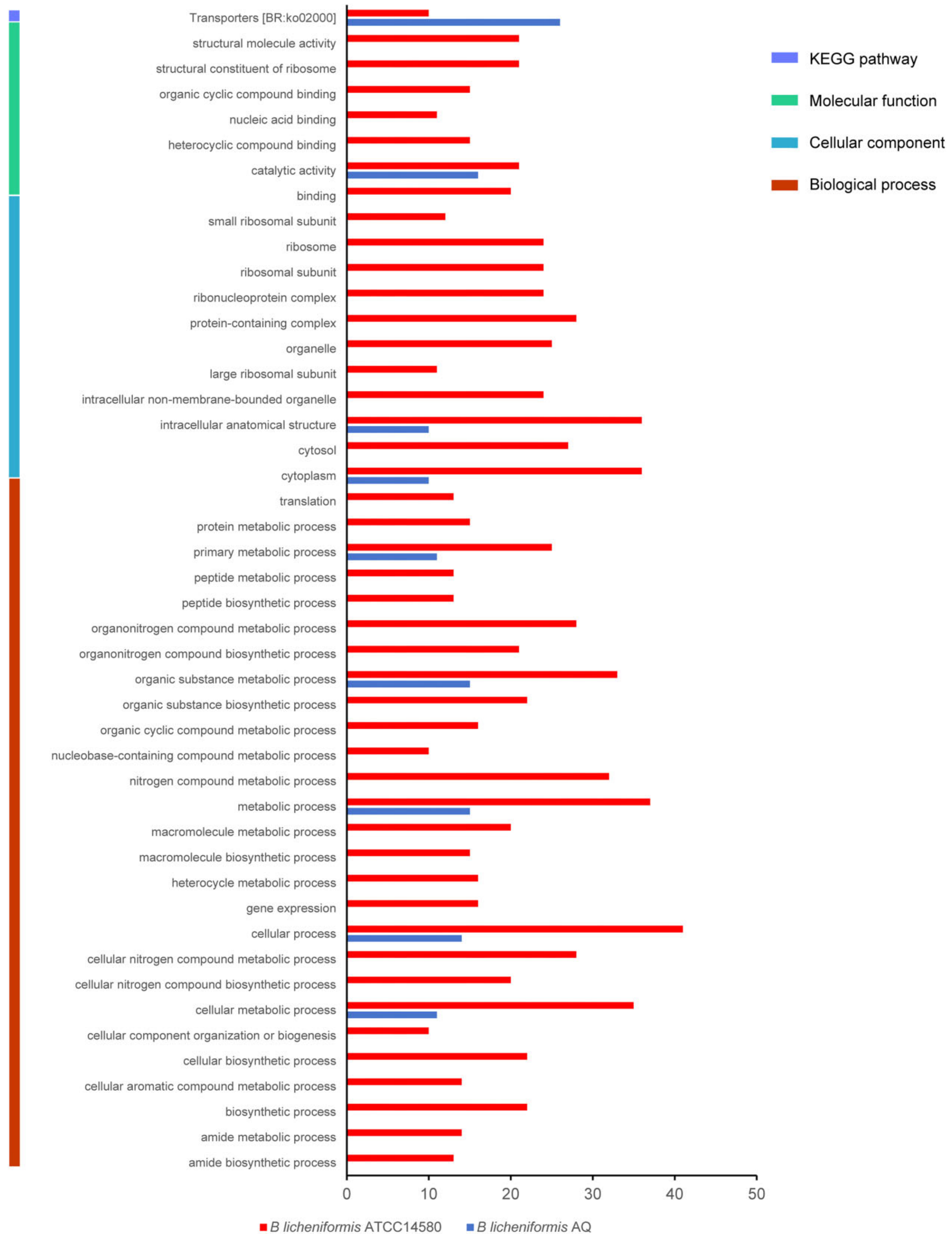

Figure 6. KEGG and GO enrichment analysis of top 100 highly expressed genes in *B. licheniformis* AQ at 32 h and *B. licheniformis* ATCC14580 at 36 h.

that aprE interacts with *degQ* by textmining (Jacobs 1995). DegQ, as an essential transcriptional activator protein, was involved in the regulation of *aprE* (Jacobs 1995). A study in *B. subtilis* has shown that the *aprE* gene is activated by phosphorylated DegU, while DegQ acts as an activator of DegU phosphorylation (Barbieri et al. 2016). Unfortunately, the expression of *degQ* could not be detected by transcriptome sequencing. Therefore, we quantified the expression of *degQ* via RT-PCR and found that the expression level was consistently low before 24 h. It increased after 27 h, peaked at 32 h, and

Downloaded from https://academic.oup.com/jambio/article/135/2/lxad319/7504755 by Huazhong Agriculture University user on 16 February 2024

steadily decreased. In this case, specific genes were missed in the transcriptome sequencing. However, this is a known limitation of transcriptome sequencing. Therefore, it is necessary to perform RT-PCR to validate gene expression levels, as it remains the gold standard (Costa et al. 2013).

qRT-PCR verification of the transcriptome results

As two complementary techniques, qRT-CR and transcriptome data are often combined in many studies. A more comprehensive understanding of the expression of specific molecular genes within cells can be achieved by matching data. Therefore, 20 genes that may have important effects on the production of AprE were selected for qRT-PCR in order to verify the transcriptome results. These genes mainly include AprE responsible for AprE synthesis, spore formation related genes (*sigY*, *sigW*, *sigM*, *sigK*, *sigH*, *sigG*, *sigF*, *sigE*, *sigB*, *sigL*, *sigA*, *sigD*, *sinR*, *sigF*, *yalC*, and *spo0A*), and global regulatory factors (*codY*, *ccpA*, *tnrA*, and *abrB*) (Zhou et al. 2021b). As shown in Fig. 5, qRT-PCR data show that most genes have higher expression levels at 32 h, but fewer genes have lower expression levels. Meanwhile, transcriptome data show that most of genes have higher expression levels at 24 h, but qRT-PCR data were not detected. Although there were some differences between qRT-PCR and RNA-seq data, the expression trends of the majority genes were consistent, which proved the reliability of transcriptome data.

Comparative transcriptome with *B. licheniformis* ATCC14580

After searching the literature, the promoters were investigated based on the transcriptome of *B. licheniformis* ATCC14580, which was a model strain without high production of AprE. By calculating the relative FPKM, it was found that the gene *aprE* showed the highest expression level among 2907 genes in *B. licheniformis* AQ (Table S1). In comparison, the gene *ssrA* displayed the highest level among 977 genes expressed in *B. licheniformis* ATCC14580 (Table S2). GO and KEGG enrichment analysis of 100 highly expressed genes in these two stains found that the KEGG of both strains were only enriched in Transporters [BR: ko02000], and *B. licheniformis* AQ has 16 more expressed genes (Fig. 6). Regarding GO, eight functions, mainly the biological process, including cellular metabolic process, cellular process, metabolic process, organic substance metabolic process, and primary metabolic process in *B. licheniformis* AQ, were enriched and all included in *B. licheniformis* ATCC14580. Besides, *B. licheniformis* ATCC14580 enriched more GO functions (22 biological processes, 9 cellular components, and 6 molecular functions). In summary, the functions of highly expressed genes in *B. licheniformis* AQ are mainly involved in synthesizing and transporting alkaline protease and are more concentrated. For model *B. licheniformis* ATCC14580, the function of highly expressed genes is more extensive and participates in cell growth and metabolism.

Discussion

This study first examined the changes in alkaline protease of *B. licheniformis* AQ at different time points during industrial fermentation. The data showed that alkaline protease was continuously accumulated during the fermentation process, and it could reach its highest value at 48 h. Subsequently, a transcriptome analysis was applied to *B. licheniformis* AQ to monitor its transcriptome profile at different stages during industrial fermentation in this study. It revealed that the most highly expressed genes were found at the endpoint of fermentation (48 h), but were normally distributed at 24 and 32 h. Then, we used STEM to analyze full transcriptome spectroscopy and generated five important models. Furthermore, the hypothesized PPI network was constructed, and the expression levels of key genes during the fermentation of AprE were revalidated using fluorescent quantitative PCR. Next, we will use CRISPR/Cas9, expression *in vitro* to explore the functions of these key genes. The next plan will also include identifying genes positively associated with alkaline protease production and determining the promoters for high-level expression of *AprE*. Finally, we performed the simple comparative transcriptome analysis with *B. licheniformis* ATCC 14580, which does not express *AprE*. We found that highly expressed genes in *B. licheniformis* AQ and ATCC 14580 are mainly involved in the synthesis and transport of alkaline protease, and global, respectively. *Bacillus licheniformis* AQ, one strain that highly expresses alkaline protease, differs from model strain *B. licheniformis* ATCC14580 in transcriptome profiles. Due to the limitation of a 50-L bioreactor, the collection of some experimental parameters is hindered. Biological replicates would be performed to repeat the alkaline protease fermentation process under laboratory conditions. Then, we would measure gene expression and detect metabolites to obtain more information about transcriptional and metabolic processes.

To sum up, this study provides important transcriptome information for *B. licheniformis* AQ and potential molecular targets for further improving the production of AprE. It lays a solid foundation for accurately constructing engineering strains with high yield of AprE as well as has important theoretical significance.

Supplementary data

Supplementary data is available at *JAMBIO Journal* online.

Conflict of interest: The authors declare that they have no competing interests.

Funding

This study was supported by the Key Research and Development Program of Hubei Province (No. 2022BBA0040), the National Natural Science Foundation of China (No. 32171423), the National Key Research and Development Program of China (No. 2022YFF1000700), and the Open Project Funding of the State Key Laboratory of Biocatalysis and Enzyme Engineering (No. SKLBEE20220028).

Author contributions

Anying Ji (Data curation, Writing – original draft), Xianliang Zheng (Validation), Wei Yang (Data curation, Software), Ming Chen (Data curation), Aimin Ma (Writing – review & editing), Yongfeng Liu (Conceptualization, Project administration, Resources, Supervision, Writing – review & editing), and Xuetuan Wei (Conceptualization, Project administration, Supervision, Writing – review & editing)

Downloaded from https://academic.oup.com/jambio/article/135/2/lxad319/7504755 by Huazhong Agriculture University user on 16 February 2024

Data availability

The GenoLab M sequencing dataset of the *B. licheniformis* AQ strain is available on the CNGB Sequence Archive (https://db.cngb.org/cnsa/) under project accession number CNP0004346.

References

Abdel Wahab WA, Ahmed SA. Response surface methodology for production, characterization and application of solvent, salt and alkali-tolerant alkaline protease from isolated fungal strain *Aspergillus niger* WA 2017. *Int J Biol Macromol* 2018;**115**:447–58. https://doi.org/10.1016/j.ijbiomac.2018.04.041.

Annamalai N, Rajeswari MV, Thavasi R *et al*. Optimization, purification and characterization of novel thermostable, haloalkaline, solvent stable protease from *Bacillus halodurans* CAS6 using marine shellfish wastes: a potential additive for detergent and antioxidant synthesis. *Bioprocess Biosyst Eng* 2013;**36**:873–83. https://doi.org/10.1007/s00449-012-0820-3.

Banerjee G, Ray AK. Impact of microbial proteases on biotechnological industries. *Biotechnol Genet Eng Rev* 2017;**33**:119–43. https://doi.org/10.1080/02648725.2017.1408256.

Barbieri G, Albertini AM, Ferrari E *et al*. Interplay of CodY and ScoC in the regulation of major extracellular protease genes of *Bacillus subtilis*. *J Bacteriol* 2016;**198**:907–20. https://doi.org/10.1128/JB.00894-15.

Barzkar N. Marine microbial alkaline protease: an efficient and essential tool for various industrial applications. *Int J Biol Macromol* 2020;**161**:1216–29. https://doi.org/10.1016/j.ijbiomac.2020.06.072.

Bougatef A, Balti R, Haddar A *et al*. Protein hydrolysates from Bluefin Tuna (*Thunnus thynnus*) heads as influenced by the extent of enzymatic hydrolysis. *Biotechnol Bioproc E* 2012;**17**:841–52. https://doi.org/10.1007/s12257-012-0053-y.

Cai D, Rao Y, Zhan Y *et al*. Engineering *Bacillus* for efficient production of heterologous protein: current progress, challenge and prospect. *J Appl Microbiol* 2019;**126**:1632–42. https://doi.org/10.1111/jam.14192.

Cao H, Villatoro-Hernandez J, Weme RDO *et al*. Boosting heterologous protein production yield by adjusting global nitrogen and carbon metabolic regulatory networks in *Bacillus subtilis*. *Metab Eng* 2018;**49**:143–52. https://doi.org/10.1016/j.ymben.2018.08.001.

Chen W, Li L, Ye C *et al*. Efficient production of extracellular alkaline protease in *Bacillus amyloliquefaciens* by host strain construction. *LWT* 2022;**163**:113620. https://doi.org/10.1016/j.lwt.2022.113620.

Chi Z, Ma C, Wang P *et al*. Optimization of medium and cultivation conditions for alkaline protease production by the marine yeast *Aureobasidium pullulans*. *Bioresour Technol* 2007;**98**:534–8. https://doi.org/10.1016/j.biortech.2006.02.006.

Claverys J-P. A new family of high-affinity ABC manganese and zinc permeases. *Res Microbiol* 2001;**152**:231–43. https://doi.org/10.1016/S0923-2508(01)01195-0.

Contesini FJ, Melo RR, Sato HH. An overview of *Bacillus* proteases: from production to application. *Crit Rev Biotechnol* 2018;**38**:321–34. https://doi.org/10.1080/07388551.2017.1354354.

Costa C, Giménez-Capitán A, Karachaliou N *et al*. Comprehensive molecular screening: from the RT-PCR to the RNA-seq. *Transl Lung Cancer Res* 2013;**2**:87–91.

Cui W, Han L, Suo F *et al*. Exploitation of *Bacillus subtilis* as a robust workhorse for production of heterologous proteins and beyond. *World J Microbiol Biotechnol* 2018;**34**:145. https://doi.org/10.1007/s11274-018-2531-7.

Damare S, Mishra A, D'Souza-Ticlo-Diniz D *et al*. A deep-sea hydrogen peroxide-stable alkaline serine protease from *Aspergillus flavus*. *3 Biotech* 2020;**10**:528. https://doi.org/10.1007/s13205-020-02520-x.

Derouiche A, Shi L, Bidnenko V *et al*. *Bacillus subtilis* SalA is a phosphorylation-dependent transcription regulator that represses scoC and activates the production of the exoprotease AprE. *Mol Microbiol* 2015;**97**:1195–208. https://doi.org/10.1111/mmi.13098.

Detmers FJM, Lanfermeijer FC, Poolman B. Peptides and ATP binding cassette peptide transporters. *Res Microbiol* 2001;**152**:245–58. https://doi.org/10.1016/S0923-2508(01)01196-2.

Ernst J, Bar-Joseph Z. STEM: a tool for the analysis of short time series gene expression data. *BMC Bioinf* 2006;**7**:1–11. https://doi.org/10.1186/1471-2105-7-191.

Garmory HS, Titball RW. ATP-binding cassette transporters are targets for the development of antibacterial vaccines and therapies. *Infect Immun* 2004;**72**:6757–63. https://doi.org/10.1128/IAI.72.12.6757-6763.2004.

Guo J-P, Ma Y. High-level expression, purification and characterization of recombinant *Aspergillus oryzae* alkaline protease in *Pichia pastoris*. *Protein Expression Purif* 2008;**58**:301–8. https://doi.org/10.1016/j.pep.2007.12.005.

Han L-L, Shao H-H, Liu Y-C *et al*. Transcriptome profiling analysis reveals metabolic changes across various growth phases in *Bacillus pumilus* BA06. *BMC Microbiol* 2017;**17**:1–13. https://doi.org/10.1186/s12866-017-1066-7.

Hosie AHF, Poole PS. Bacterial ABC transporters of amino acids. *Res Microbiol* 2001;**152**:259–70. https://doi.org/10.1016/S0923-2508(01)01197-4.

Igarashi K, Ito K, Kashiwagi K. Polyamine uptake systems in *Escherichia coli*. *Res Microbiol* 2001;**152**:271–8. https://doi.org/10.1016/S0923-2508(01)01198-6.

Jacobs MF. Expression of the subtilisin Carlsberg-encoding gene in *Bacillus licheniformis* and *Bacillus subtilis*. *Gene* 1995;**152**:69–74. https://doi.org/10.1016/0378-1119(94)00655-C.

Jiang C, Ye C, Liu Y *et al*. Genetic engineering for enhanced production of a novel alkaline protease BSP-1 in *Bacillus amyloliquefaciens*. *Front Bioeng Biotechnol* 2022;**10**:977215. https://doi.org/10.3389/fbioe.2022.977215.

Kertesz MA. Bacterial transporters for sulfate and organosulfur compounds. *Res Microbiol* 2001;**152**:279–90. https://doi.org/10.1016/S0923-2508(01)01199-8.

Kim D, Paggi JM, Park C *et al*. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat Biotechnol* 2019;**37**:907–15. https://doi.org/10.1038/s41587-019-0201-4.

Kobayashi K. Gradual activation of the response regulator DegU controls serial expression of genes for flagellum formation and biofilm formation in *Bacillus subtilis*. *Mol Microbiol* 2007;**66**:395–409. https://doi.org/10.1111/j.1365-2958.2007.05923.x.

Köster W. ABC transporter-mediated uptake of iron, siderophores, heme and vitamin B12. *Res Microbiol* 2001;**152**:291–301. https://doi.org/10.1016/S0923-2508(01)01200-1.

Liu L, Liu Y, Shin H *et al*. Developing *Bacillus* spp. as a cell factory for production of microbial enzymes and industrially important biochemicals in the context of systems and synthetic biology. *Appl Microbiol Biotechnol* 2013;**97**:6113–27. https://doi.org/10.1007/s00253-013-4960-4.

Liu X, Yang H, zheng J *et al*. Identification of strong promoters based on the transcriptome of *Bacillus licheniformis*. *Biotechnol Lett* 2017;**39**:873–81. https://doi.org/10.1007/s10529-017-2304-7.

Liu Y, Han R, Zhou L *et al*. Comparative performance of the GenoLab M and NovaSeq 6000 sequencing platforms for transcriptome and LncRNA analysis. *BMC Genom Data* 2021;**22**:1–12. https://doi.org/10.1186/s12863-021-00963-6.

Livak KJ, Schmittgen TD. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods* 2001;**25**:402–8. https://doi.org/10.1006/meth.2001.1262.

Mathew CD, Gunathilaka R. Production, purification and characterization of a thermostable alkaline serine protease from *Bacillus lichniformis* NMS-1. *Int J Biotechnol Mol Biol Res* 2015;**8**:975.

Matkawala F, Nighojkar S, Kumar A *et al*. Microbial alkaline serine proteases: production, properties and applications. *World J Micro-*

Downloaded from https://academic.oup.com/jambio/article/135/2/lxad319/7504755 by Huazhong Agriculture University user on 16 February 2024

biol Biotechnol 2021;**37**:63. https://doi.org/10.1007/s11274-021-03036-z.

Nam MS, Whang KS, Choi SH *et al*. Purification, characterization, and properties of an alkaline protease produced by *Serratia marcescens* S3-R1 inhabiting Korean ginseng rhizosphere. *J Sci Food Agric* 2013;**93**:3876–82. https://doi.org/10.1002/jsfa.6363.

Ogura M, Matsuzawa A, Yoshikawa H *et al*. *Bacillus subtilis* SalA (YbaL) negatively regulates expression of *scoC*, which encodes the repressor for the alkaline exoprotease gene, *aprE*. *J Bacteriol* 2004;**186**:3056–64. https://doi.org/10.1128/JB.186.10.3056-3064.2004.

Ogura M, Shimane K, Asai K *et al*. Binding of response regulator DegU to the aprE promoter is inhibited by RapG, which is counteracted by extracellular PhrG in *Bacillus subtilis*. *Mol Microbiol* 2003;**49**:1685–97. https://doi.org/10.1046/j.1365-2958.2003.03665.x.

Schneider E. ABC transporters catalyzing carbohydrate uptake. *Res Microbiol* 2001;**152**:303–10. https://doi.org/10.1016/S0923-2508(01)01201-3.

Self WT, Grunden AM, Hasona A *et al*. Molybdate transport. *Res Microbiol* 2001;**152**:311–21. https://doi.org/10.1016/S0923-2508(01)01202-5.

Shafique T, Shafique J, Zahid S *et al*. Screening, selection and development of *Bacillus subtilis* apr-IBL04 for hyper production of macromolecule alkaline protease. *Saudi J Biol Sci* 2021;**28**:1494–501. https://doi.org/10.1016/j.sjbs.2020.11.079.

Sonenshein AL. Control of key metabolic intersections in *Bacillus subtilis*. *Nat Rev Micro* 2007;**5**:917–27. https://doi.org/10.1038/nrmicro1772.

Sorek R, Cossart P. Prokaryotic transcriptomics: a new view on regulation, physiology and pathogenicity. *Nat Rev Genet* 2010;**11**:9–16. https://doi.org/10.1038/nrg2695.

Szklarczyk D, Franceschini A, Wyder S *et al*. STRING v10: protein–protein interaction networks, integrated over the tree of life. *Nucleic Acids Res* 2014;**43**:D447–52. https://doi.org/10.1093/nar/gku1003.

van Dijl J, Hecker M. *Bacillus subtilis*: from soil bacterium to super-secreting cell factory. *Microb Cell Fact* 2013;**12**:3. https://doi.org/10.1186/1475-2859-12-3.

Wiegand S, Dietrich S, Hertel R *et al*. RNA-Seq of *Bacillus licheniformis*: active regulatory RNA features expressed within a productive fermentation. *BMC Genomics [Electronic Resource]* 2013;**14**:667. https://doi.org/10.1186/1471-2164-14-667.

Wu Z, Li Y, Xu Y *et al*. Transcriptome analysis of *Bacillus licheniformis* for improving bacitracin production. *ACS Synth Biol* 2022;**11**:1325–35. https://doi.org/10.1021/acssynbio.1c00593.

Yuan F, Li K, Zhou C *et al*. Identification of two novel highly inducible promoters from *Bacillus licheniformis* by screening transcriptomic data. *Genomics* 2020;**112**:1866–71. https://doi.org/10.1016/j.ygeno.2019.10.021.

Zhang K, Su L, Wu J. Recent advances in recombinant protein production by *Bacillus subtilis*. *Annu Rev Food Sci Technol* 2020; **11**:295–318. https://doi.org/10.1146/annurev-food-032519-051750.

Zhou C, Yang G, Zhang L *et al*. Construction of an alkaline protease overproducer strain based on *Bacillus licheniformis* 2709 using an integrative approach. *Int J Biol Macromol* 2021a;**193**:1449–56. https://doi.org/10.1016/j.ijbiomac.2021.10.208.

Zhou C, Zhang H, Fang H *et al*. Transcriptome based functional identification and application of regulator AbrB on alkaline protease synthesis in *Bacillus licheniformis* 2709. *Int J Biol Macromol* 2021b;**166**:1491–8. https://doi.org/10.1016/j.ijbiomac.2020.11.028.

Zhou C, Zhou H, Fang H *et al*. Spo0A can efficiently enhance the expression of the alkaline protease gene *aprE* in *Bacillus licheniformis* by specifically binding to its regulatory region. *Int J Biol Macromol* 2020a;**159**:444–54. https://doi.org/10.1016/j.ijbiomac.2020.05.035.

Zhou C, Zhou H, Li D *et al*. Optimized expression and enhanced production of alkaline protease by genetically modified *Bacillus licheniformis* 2709. *Microb Cell Fact* 2020b;**19**:45. https://doi.org/10.1186/s12934-020-01307-2.

Zhou C, Zhou H, Zhang H *et al*. Optimization of alkaline protease production by rational deletion of sporulation related genes in *Bacillus licheniformis*. *Microb Cell Fact* 2019;**18**:127. https://doi.org/10.1186/s12934-019-1174-1.

Downloaded from https://academic.oup.com/jambio/article/135/2/lxad319/7504755 by Huazhong Agriculture University user on 16 February 2024

Received 31 August 2023; **revised** 18 December 2023; **accepted** 29 December 2023
© The Author(s) 2023. Published by Oxford University Press on behalf of Applied Microbiology International. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com