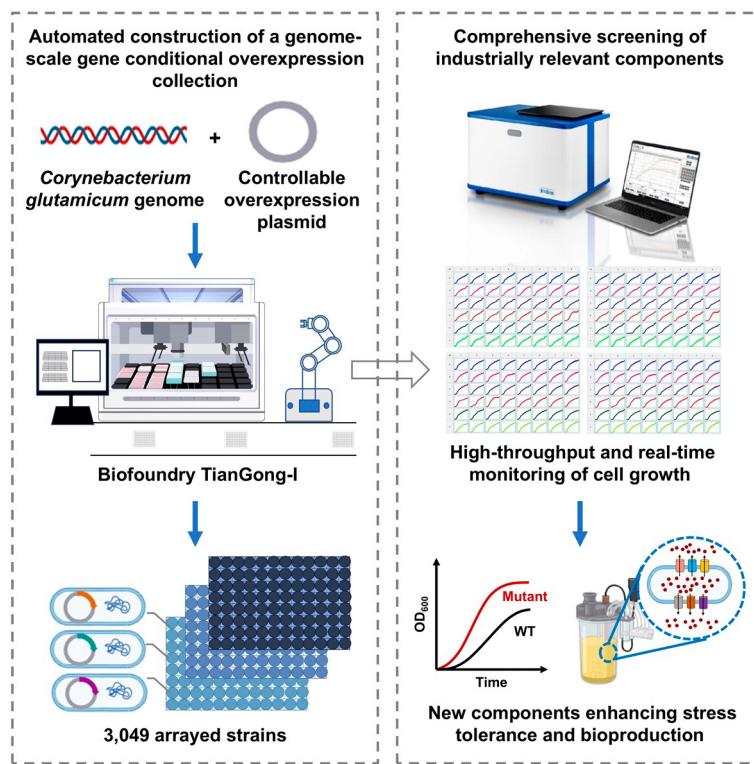


Resource Article

Comprehensive screening of industrially relevant components at genome scale using a high-quality gene overexpression collection of *Corynebacterium glutamicum*



A high-quality gene overexpression collection of the industrial workhorse *Corynebacterium glutamicum* containing 3049 strains was produced. To demonstrate its application, a comprehensive screening of industrially relevant components at a genome scale was conducted. New components for tolerance and transport were identified and used for developing industrial amino acid-producing strains.

Jiao Liu (刘娇), Xiaojia Zhao (赵晓佳), Haijiao Cheng (程海娇), Yanmei Guo (郭艳梅), Xiaomeng Ni (倪晓蒙), Lixian Wang (王丽贤), Guannan Sun (孙冠男), Xiao Wen (温潇), Jiuzhou Chen (陈久洲), Jin Wang (王瑾), Jingjing An (安晶晶), Xuan Guo (郭轩), Zhenkun Shi (史振坤), Haoran Li (李浩然), Ruoyu Wang (王若宇), Muqiang Zhao (赵募强), Xiaoping Liao (廖小平), Yu Wang (王钰), Ping Zheng (郑平), Meng Wang (王猛), Jibin Sun (孙际宾)

wang_y@tib.cas.cn (Y. Wang),
zheng_p@tib.cas.cn (P. Zheng)
wangmeng@tib.cas.cn (M. Wang).

Highlights

A high-quality gene overexpression collection of *Corynebacterium glutamicum* was constructed with a highly automated biofoundry.

Functional components for enhanced stress tolerance and amino acid transport were comprehensively screened at a genome scale.

Transcriptional factors regulating cell division and energy metabolism and DNA repair proteins removing m6G methylations were found to be important for hyperosmotic tolerance.

Application of the newly screened exporters facilitates the highest L-threonine production by *C. glutamicum* achieved thus far.

Trends in Biotechnology, January 2025,
Vol. 43, No. 1
<https://doi.org/10.1016/j.tibtech.2024.09.020>



Resource Article

Comprehensive screening of industrially relevant components at genome scale using a high-quality gene overexpression collection of *Corynebacterium glutamicum*

Jiao Liu (刘娇)^{1,2,5}, Xiaojia Zhao (赵晓佳)^{1,2,3,5}, Haijiao Cheng (程海娇)^{1,2,5}, Yanmei Guo (郭艳梅)^{1,2}, Xiaomeng Ni (倪晓蒙)^{1,2}, Lixian Wang (王丽贤)^{1,2}, Guannan Sun (孙冠男)^{1,2,3}, Xiao Wen (温潇)^{1,2,4}, Jiuzhou Chen (陈久洲)^{1,2}, Jin Wang (王瑾)^{1,2}, Jingjing An (安晶晶)^{1,2}, Xuan Guo (郭轩)^{1,2}, Zhenkun Shi (史振坤)^{1,2}, Haoran Li (李浩然)^{1,2}, Ruoyu Wang (王若宇)^{1,2}, Muqiang Zhao (赵募强)^{1,2}, Xiaoping Liao (廖小平)^{1,2}, Yu Wang (王钰)^{1,2,3,*}, Ping Zheng (郑平)^{1,2,3,*}, Meng Wang (王猛)^{1,2,3,6,*}, and Jibin Sun (孙际宾)^{1,2,3}

Development of efficient microbial strains for biomanufacturing requires deep understanding of the biology and functional components responsible for the synthesis, transport, and tolerance of the target compounds. A high-quality controllable gene overexpression strain collection was constructed for the industrial workhorse *Corynebacterium glutamicum* covering 99.7% of its genes. The collection was then used for comprehensive screening of components relevant to biomanufacturing features. In total, 15 components endowing cells with improved hyperosmotic tolerance and L-lysine productivity were identified, including novel transcriptional factors and DNA repair proteins. Systematic interrogation of a subset of the collection revealed efficient and specific exporters functioning in both *C. glutamicum* and *Escherichia coli*. Application of the new exporters was showcased to construct a strain with the highest L-threonine production level reported for *C. glutamicum* (75.1 g/l and 1.5 g/l·h) thus far. The genome-scale gene overexpression collection will serve as a valuable resource for fundamental biological studies and for developing industrial microorganisms for producing amino acids and other biochemicals.

Introduction

Biomanufacturing, which converts renewable feedstocks to useful chemicals under environmentally friendly conditions, holds great potential to address the global challenges of an expanding human population, reducing resources, and environmental change [1]. Microbial strains are key catalysts in industrial biomanufacturing. However, the development of efficient microbial strains that meet the standard of industrialization still requires significant time and investment [2]. The booming **clustered regularly interspaced short palindromic repeats (CRISPR)** (see *Glossary*) technology has considerably accelerated the genetic manipulation of microorganisms [3–6]. Nevertheless, the remaining major challenge in developing industrial microbial strains is to fully understand the biology of the organisms and to discover effective components with specific functions, including catalysis, regulation, transport, and tolerance relevant to the production of target products [7–9].

Technology readiness

This study reports the automated construction of a high-quality arrayed gene overexpression strain collection and high-throughput screening of biological components that can be used for developing industrial strains for biomanufacturing. This strain collection and screening strategy can be readily used to discover new components with various functions. A *Corynebacterium glutamicum* strain with the highest L-threonine production level (75.1 g/l and 1.5 g/l·h) was also engineered using the newly identified biological components and tested in 5-l bioreactors. To meet the requirement of industrialization, further genetic modification and process engineering are necessary to improve the L-threonine production level in *C. glutamicum*. Therefore, the current Technology Readiness Level (TRL) of this production technology lies between 3 and 4.

¹Key Laboratory of Engineering Biology for Low-carbon Manufacturing, Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences, Tianjin 300308, China

²National Center of Technology Innovation for Synthetic Biology, Tianjin 300308, China

³University of Chinese Academy of Sciences, Beijing 100049, China



Amino acids find various applications as flavor enhancers, nutraceuticals, food or feed additives, and building blocks for the chemical and pharmaceutical industries. *C. glutamicum* is an important microbial chassis for the biomanufacturing of amino acids and many other chemicals [10–14]. Over 6 million tons of amino acids per year are produced using *C. glutamicum* strains, including L-lysine [15], L-glutamate [16], L-methionine [17], L-proline [18], and L-ornithine [19], among others. In addition, many functional components found in *C. glutamicum* have also proven effective in other microorganisms [20–25]. Therefore, mining functional components in the genome of *C. glutamicum* is important for the efficient production of amino acids and other chemicals.

However, the physiological function of a considerable proportion of *C. glutamicum* genes remains cryptic. According to the latest annotation, the products of 49.9% genes (1545/3099) of the *C. glutamicum* model strain ATCC 13032 have no well-defined functions (1141 hypothetical, 232 predicted, 142 uncharacterized, and 30 putative proteins) [26]. Therefore, a comprehensive strategy for discovering new functional components is still required. In this study, we constructed a genome-scale gene overexpression collection containing 3057 *C. glutamicum* strains, with each strain conditionally overexpressing a native gene without any tags that may disturb protein function. Systematic examination of the **gain-of-function (GOF)** phenotype of each gene in *C. glutamicum* facilitated the discovery of new functional components capable of enhancing hyperosmotic tolerance, which can be readily used for amino acid-producing strains. Functional mechanisms of these components were investigated and unusual correlations of DNA methylation, energy metabolism, and morphology control with environmental stress were characterized. In addition, four new L-threonine exporters with different characteristics were identified and used for constructing industry-level L-threonine producers. A new record for L-threonine production by *C. glutamicum* (75.1 g/l and 1.5 g/l·h) was achieved. This study provides not only new functional components for developing amino acid-overproducing strains, but also valuable strain resources for the *C. glutamicum* research community.

Results

An arrayed strain collection of *C. glutamicum* with controllable overexpression of native genes

Transcriptomic analyses suggest that the expression levels of many genes in *C. glutamicum* are low and vary dramatically with cultivation conditions [27,28]. For systematic screening of new functional components from *C. glutamicum*, it is better to overexpress each native gene to activate or strengthen its function. However, the construction of a genome-scale gene overexpression collection is labor intensive and costly. Therefore, such a collection is available only for few model microorganisms, including *E. coli* [29], *Saccharomyces cerevisiae* [30,31], and *Schizosaccharomyces pombe* [32]. To facilitate easy localization or purification of encoded proteins, a tag is usually fused to the N or C terminus of a target gene, which may disrupt the structure and function of the encoded protein [33–35]. In this study, to construct a high-quality gene overexpression collection and minimize the labor burden, we used the BioFoundry at Tianjin Institute of Industrial Biotechnology (Tiangong-1) to automate the gene PCR, DNA electrophoresis, plasmid ligation, and competent cell transformation steps, among others (Figure 1A and Table S1 in the supplemental information online). Genes were manipulated with the native sequences without adding any tags.

The strain ATCC 13032 is the model strain of *C. glutamicum* and has been widely engineered for the production of various amino acids and chemicals [36]. Therefore, it was selected as the target in this study. First, the copy of strain ATCC 13032 used in our laboratory was sequenced (Data S1 in the supplemental information online), which revealed 3057 genes. Compared with the genome sequence of strain ATCC 13032 deposited in the National Center for Biotechnology Information (NCBI) database (Accession: BA000036.3, GI: 42602314), 29 **SNPs** in 28 genes (ten synonymous,

⁴School of Life Sciences, Division of Life Sciences and Medicine, University of Science and Technology of China, Hefei 230026, China

⁵These authors contributed equally.

⁶Lead contact.

*Correspondence:
wang_y@tib.cas.cn (Y. Wang),
zheng_p@tib.cas.cn (P. Zheng), and
wangmeng@tib.cas.cn (M. Wang).

15 missense, two frameshift, and two stop-gained mutations) were detected (Table S2 in the supplemental information online). The synonymous and missense SNPs were retained, while the frame-shift and stop-gained SNPs were corrected during construction of the gene overexpression plasmid. In addition, a cluster containing 42 prophage genes (*cgl1848–cgl1889*) was missing in our strain copy and, thus, was not included in the strain collection (Figure S1 in the supplemental information online).

The widely used *E. coli–C. glutamicum* shuttle vector pEC-XK99E was modified for gene cloning and overexpression (Figure S2 and Data S2 in the supplemental information online). A *ccdB* cassette was inserted into the **multiple cloning site (MCS)** to reduce the false positive rate during plasmid construction. An ampicillin selection marker was added to facilitate the transformation of *E. coli* DH5 α competent cells without the heat shock and recovery steps. An isopropyl- β -D-thiogalactoside (IPTG)-inducible P_{trc} promoter was used to control gene expression levels. By testing the expression of a heterologous red fluorescent protein (RFP)-encoding gene and two native genes, the constructed plasmids were demonstrated to show controllable gene overexpression at both the transcript and protein levels. By adjusting the IPTG usage, up to 20-fold induction of gene overexpression was obtained (Figure 1A,C).

As a trial run, the overexpression plasmids for ~500 genes were constructed using the BioFoundry. Primers used for gene amplification are listed in Data S3 in the supplemental information online. The PCR products were analyzed using microchip-based capillary electrophoresis (Figure 1D). After the trial run, we observed a positive correlation between the purity of gene PCR product and the success rate of plasmid construction (Figure 1E). Therefore, in the subsequent experiment, only the PCR products with >50% purity were used for plasmid construction. The constructed plasmid was subjected to two Sanger sequencing reactions to verify the accurate insertion of the target gene into MCS and to detect any mutations. For genes that needed more than two Sanger sequencing reactions to cover the whole gene, the plasmids were mixed and analyzed by **next-generation sequencing (NGS)**, which reduced the sequencing costs. A total of 3039 gene overexpression plasmids were successfully constructed.

Although the remaining 18 target genes were successfully amplified, either no transformants were obtained or mutations were frequently detected in the recombinant plasmids. We speculated that, because of the high copy number of the shuttle vector pEC-XK99E in *E. coli*, leaky expression of these target genes may cause cytotoxicity [37]. We then used a *pcnB*-deleted *E. coli* host that can lower the plasmid copy number for cloning [38], which successfully produced another ten correct plasmids. Finally, a total of 3049 plasmids each containing an IPTG-inducible gene of *C. glutamicum* were assembled, covering 99.7% of the total 3057 genes. The unsuccessful construction of the remaining eight genes (*cgl0060*, *cgl0069*, *cgl0253*, *cgl0537*, *cgl1130*, *cgl1043*, *cgl1777*, and *cgl2527*) may be due to the high cytotoxicity caused by their leaky expression in *E. coli*. *cgl0069* (*citB*, response regulator) together with *cgl0068* (*citA*, sensor histidine kinase) encode the CitAB two-component system that positively regulates the citrate transport genes involved in citrate utilization [39]. *cgl2527* encodes a transcriptional regulator, GntR1, which regulates gluconate catabolism and glucose uptake [40]. The biological functions of the remaining six genes have not yet been experimentally investigated. According to the annotation, *cgl0060* and *cgl1043* encode hypothetical proteins, *cgl0253* encodes a hypothetical membrane protein, *cgl0537* encodes the ribosomal protein S8, *cgl1130* encodes an ABC-type transporter, and *cgl1777* encodes a stress-sensitive restriction system protein. All the constructed plasmids were then successfully transformed into *C. glutamicum* competent cells by electroporation. After verifying the plasmid-borne gene copy in transformants by PCR, we numbered and deposited the 3049 strains in the Biobank of Tianjin Institute of Industrial Biotechnology, Chinese Academy of Sciences (Data S4 in the supplemental information online).

Glossary

Clustered regularly interspaced short palindromic repeats (CRISPR): technology for precise genome editing.

CRISPR interference (CRISPRi): method for sequence-specific repression of gene expression.

ChIP-Seq: powerful technique used in epigenetic studies to investigate protein-DNA interactions in the genome.

Clusters of orthologous groups (COG): database that categorizes proteins into groups based on their evolutionary relationships.

Gain-of-function (GOF): type of genetic mutation where a gene or protein gains a new or enhanced function that is not found in the wild-type allele.

Loss-of-function (LOF): type of genetic mutation where the function of a gene is diminished or eliminated.

Multiple cloning site (MCS): segment of DNA within a plasmid that contains a series of recognition sites for restriction enzymes.

Next-generation sequencing (NGS): also known as high-throughput sequencing or massively parallel sequencing; enables the sequencing of many DNA strands simultaneously, compared with one at a time as with traditional Sanger sequencing.

Scanning electron microscope (SEM): type of electron microscope that produces images of a sample by scanning it with a focused beam of high-energy electrons.

SNPs: type of genetic variation where a single nucleotide in the DNA sequence differs between individuals.

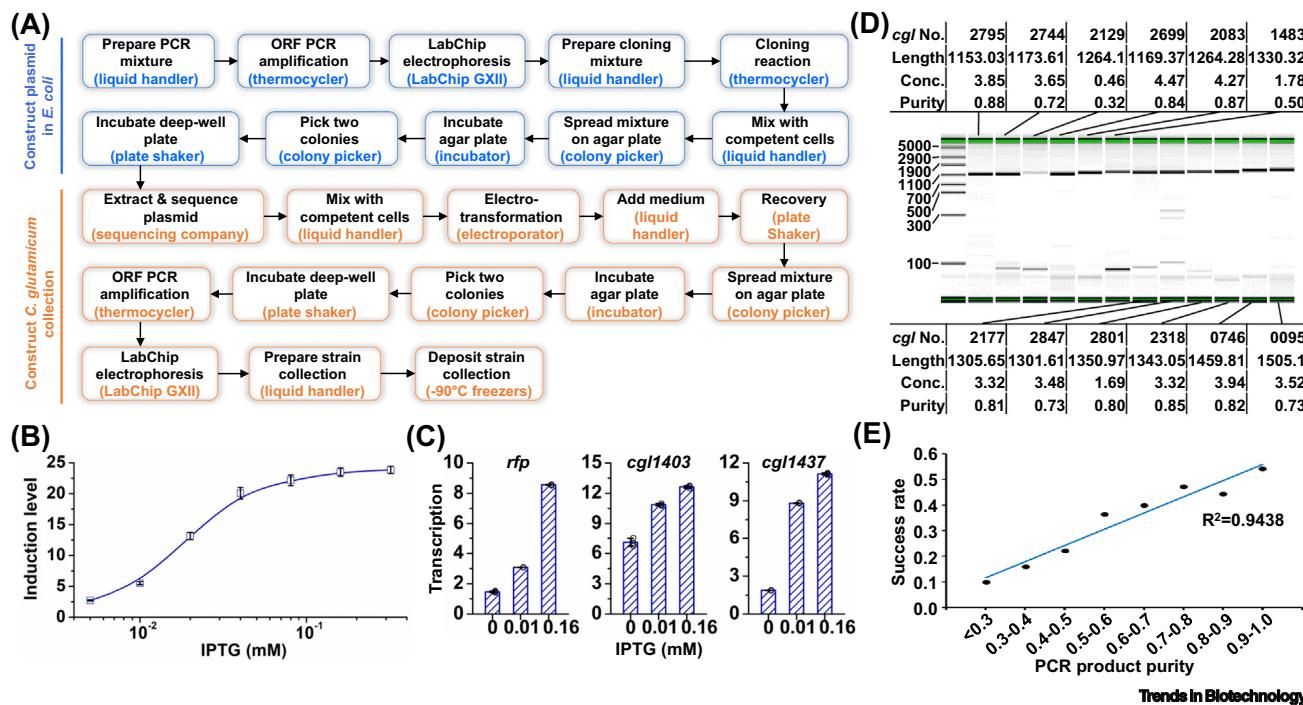


Figure 1. Automated construction of an array of *Corynebacterium glutamicum* strains with a single gene overexpressed. (A) Procedure for automated strain construction with the BioFoundry Tiangong-1. Equipment integrated in the BioFoundry is within parentheses. (B) Red fluorescent protein (RFP) fluorescence outputs of the isopropyl-β-D-thiogalactoside (IPTG)-inducible gene overexpression plasmid. Error bars are standard deviations ($n = 3$). (C) Transcription levels of *rfp*, *cg1403*, and *cg1437* in strains harboring the IPTG-inducible gene overexpression plasmid. Error bars are standard deviations ($n = 3$). Individual data points are shown by black circles. (D) LabChip electrophoresis and analysis of PCR products. Product length, DNA concentration (Conc., given in ng/μl), and product purity were analyzed using LabChip GX Touch Software v.1.11.144.0. (E) The correlation between the PCR product purity and the success rate of plasmid construction. The experimental data for constructing ~500 gene overexpression plasmids were used for the calculation. Abbreviation: ORF, open reading frame.

Comprehensive screening of functional components to enhance cellular tolerance to hyperosmotic stress

As an industrial workhorse for large-scale amino acid production, *C. glutamicum* is inevitably exposed to hyperosmotic stress during the mid- and late stages of fermentation due to the high concentration of extracellular products [41]. Therefore, functional components that can improve cellular tolerance to hyperosmotic stress are of significance for enhancing bioproduction. To discover beneficial functional components in a high-throughput manner, the arrayed strains were evaluated for their cell growth under hyperosmotic stress using the MicroScreen instrument, which enables simultaneous cultivation of four 48-deep-well microplates and real-time monitoring of the optical density at 600 nm (OD_{600}) of each well (Figure S3A in the supplemental information online). High concentrations of L-lysine (in the form of L-lysine sulfate, pH adjusted to 7.2), one of the most important products of the *C. glutamicum* fermentation industry, were added to exert hyperosmotic stress (Figure S3B). As a test, wild-type *C. glutamicum* was cultivated under different concentrations of L-lysine using the MicroScreen instrument. In the normal medium without L-lysine addition, a typical S-shaped growth curve and a final OD_{600} of ~7.0 [equal to ~11.3 g cell dry weight (CDW)/l] were obtained (Figure S3C). The OD_{600} values given by the MicroScreen instrument differed from those given by the standard spectrophotometer because of the different measuring instruments and pathlengths. A previous study demonstrated that the OD_{600} values measured by these two instruments have a fivefold difference and a linear relation [42]. With the increase in L-lysine concentration, the cell growth of *C. glutamicum* gradually decreased. When the L-lysine concentration was set as 1.2 M, the final biomass was approximately one-third of

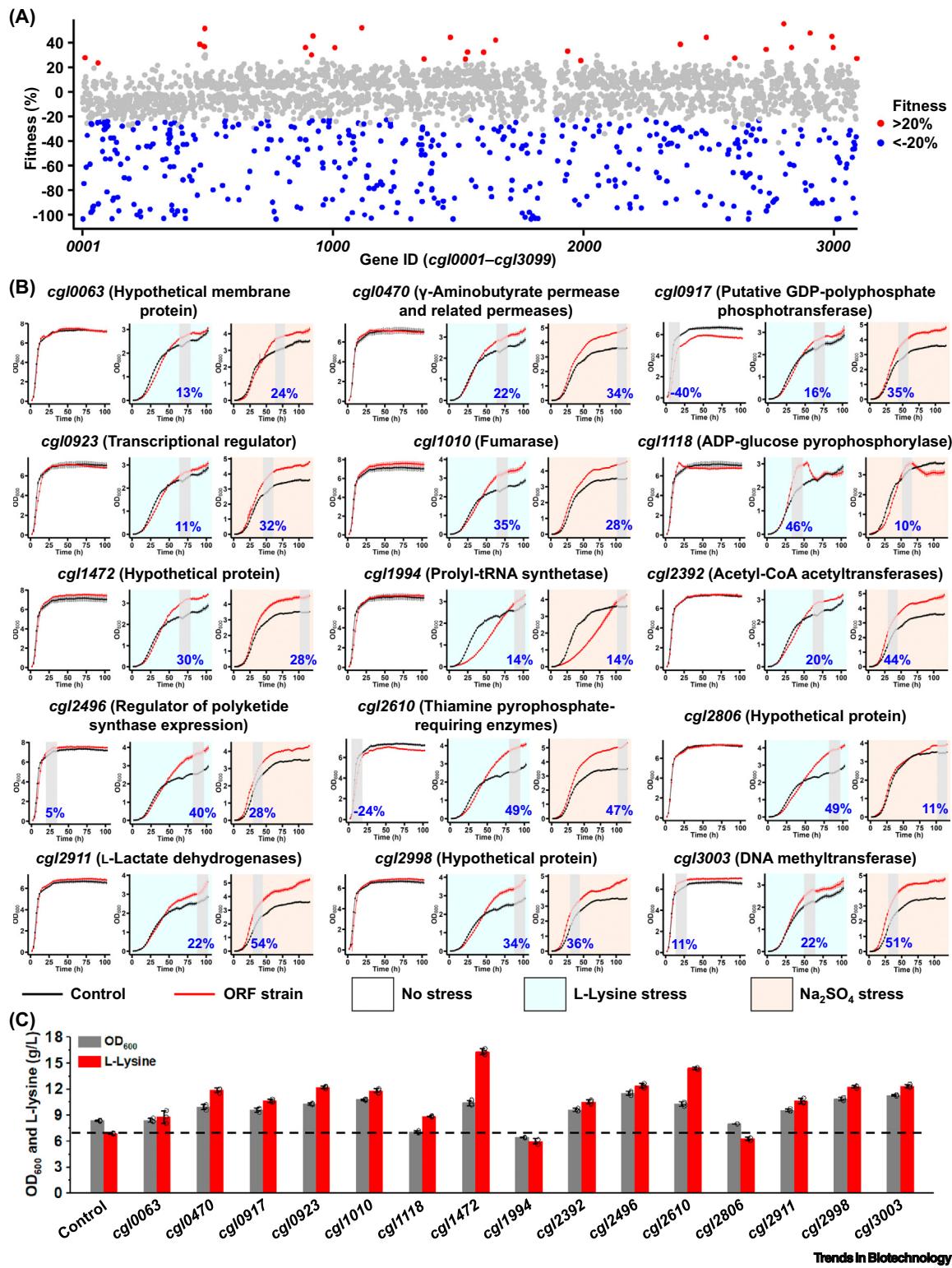
that without L-lysine stress. No cell growth was observed when the L-lysine concentration reached 2.0 M (Figure S3C). Before the growth test, the IPTG usage was optimized. Using RFP as a reporter protein, it was observed that addition of 0.04 mM IPTG produced a 20-fold induction in RFP overexpression, but a 46% decrease in cell growth rate. When the usage of IPTG increased further, RFP expression was not significantly enhanced, but growth inhibition was aggravated (up to 65% decrease) (Figure 1B and Figure S4 in the supplemental information online). Based on the gene expression level and cell growth rate, 0.04 mM IPTG was used for subsequent experiments.

Next, growth curves for all 3049 constructed gene overexpression strains were determined under the stress of 1.2 M L-lysine using the MicroScreen instrument to screen beneficial functional components. *C. glutamicum* harboring an empty plasmid was installed for each deep-well microplate as a control. An algorithm was developed to process the large amounts of growth data for rapid categorization of genes with different fitness tolerance of hyperosmotic stress. The fitness advantage of a gene represents improved tolerance and cell growth under stress conditions caused by overexpression of the gene. Continuous 15 data points that showed the biggest difference between the gene overexpression strain and the control strain were used to calculate the gene fitness (Figure S5 in the supplemental information online). Based on the analysis of the first round of growth characterization, 29 strains showed a >20% fitness advantage compared with the control, while 408 strains showed a >20% fitness defect (Figure 2A and Data S5 in the supplemental information online). The 29 strains entered a second round of growth characterization with three replicates under the stress of 1.2 M L-lysine. After two rounds of screening, 15 functional components the overexpression of which enhanced fitness to the high concentration of L-lysine were confirmed (Figure 2B; Table S3 and Data S6 in the supplemental information online). Interestingly, overexpression of these 15 genes also improved cell fitness under the stress of 1.2 M sodium sulfate, but caused no significant effect on growth without any stress, suggesting their positive function in enabling *C. glutamicum* to tolerate hyperosmotic stress (Figure 2B and Data S7 in the supplemental information online).

There were two special cases, *cgl1118* encoding ADP-glucose pyrophosphorylase and *cgl1994* (*proS*) encoding prolyl-tRNA synthetase. Overexpression of *cgl1118* increased exponential cell growth under stress conditions but the cell biomass quickly decreased after entering the stationary phase. For *cgl1994*, its overexpression inhibited cell growth under the stress conditions at the early exponential phase but led to a higher final biomass (Figure 2B).

To test whether deletion of the 15 genes affected the cellular tolerance to hyperosmotic stress, each gene was knocked out individually by CRISPR/Cas9-mediated gene editing of the wild-type strain [18]. Except for the essential gene *cgl1994* [43], the remaining 14 gene-deleted mutants were successfully constructed. Deletion of *cgl1010* (*fum*, fumarase), *cgl2911* (*ldh*, malate/lactate dehydrogenases), and *cgl2998* (hypothetical protein) largely decreased the cellular tolerance to hypermutation stress, whereas deletion of the other genes caused slight effects on cell growth under the same test conditions (Figure S6 in the supplemental information online). The results demonstrated that the relevance of these genes to hyperosmotic stress may be difficult to discover by **loss-of-function (LOF)** screening, such as gene knockout, transposon mutagenesis, or **CRISPR interference (CRISPRi)**. By contrast, GOF screening using gene overexpression collection is a useful strategy for discovering functional components.

In terms of the 408 genes the overexpression of which inhibited cell growth under hyperosmotic stress, we were interested in those encoding transcriptional regulators since their modifications may exert global effects on cellular metabolism. Based on the analysis of the first round of growth characterization, 46 transcriptional regulator-encoding genes were found to decrease cell growth under hyperosmotic stress by over 20%. Given their negative effects on cell fitness, we



(See figure legend at the bottom of the next page.)

hypothesized that deletion of these genes may increase cellular tolerance to hyperosmotic stress. We randomly selected 21 transcriptional regulator-encoding genes for further testing. Even without hyperosmotic stress, overexpression of 20 of the 21 genes caused different levels of growth inhibition (6–93% decreases in gene fitness), which may be due to the metabolic burden caused by gene overexpression (Figure S7 and Data S5 in the supplemental information online). *cgl0871* is an exception because its overexpression caused no growth inhibition under normal conditions but caused a 58% decrease in gene fitness under hyperosmotic stress. In addition, overexpression of *cgl2009*, *cgl2627*, and *cgl3089* showed slight effects on cell growth under normal conditions (6–12% decrease in gene fitness) but severely impaired cellular tolerance to hyperosmotic stress (40–66% decreases in gene fitness). The results suggest the relevance of these four genes to stress tolerance (Figure S7 and Data S8 in the supplemental information online). These 21 genes were then deleted using the CRISPR/Cas9 system. Unexpectedly, deletion of these 21 genes did not show the expected complementary effects (i.e., enhanced tolerance to hyperosmotic stress). Deletion of 14 genes (*cgl0016*, *cgl0020*, *cgl0445*, *cgl0627*, *cgl0662*, *cgl0871*, *cgl1314*, *cgl2009*, *cgl2627*, *cgl2711*, *cgl2894*, *cgl2988*, *cgl3024*, and *cgl3089*) showed no significant effects, while deletion of the remaining seven genes (*cgl0379*, *cgl1120*, *cgl1931*, *cgl2381*, *cgl2490*, *cgl2560*, and *cgl2941*) decreased cell growth (Figure S7). The results again suggest that the genotype–phenotype associations for many genes can only be established by GOF screening approaches, such as the genome-scale gene overexpression.

Identified functional components also enhance L-lysine production

The hyperosmotic stress caused by high concentrations of products cause cell damage and disturb cellular metabolism, consequently inhibiting the continuous formation of products in fermentation. Cellular tolerance to hyperosmotic stress is an important factor for the optimization of industrial strains [8]. Therefore, the effects of overexpression of the 15 identified functional components on L-lysine production were characterized in a L-lysine-producing *C. glutamicum* strain harboring three mutations (*lysC^{T311I}*, *pyc^{P458S}*, and *hom^{V59A}*) under hyperosmotic stress caused by high concentrations of sodium sulfate. Thirteen of the 15 identified functional components enhanced L-lysine production by 28–137%, except for *cgl1994* and *cgl2806* (hypothetical protein). Notably, overexpression of *cgl0470*, *cgl0923*, *cgl1010*, *cgl1472*, *cgl2496*, *cgl2610*, *cgl2998*, and *cgl3003* led to a >70% increase in L-lysine titer, with the highest increase of 137% achieved with *cgl1472* (Figure 2C). Different from the beneficial effects on the wild-type *C. glutamicum* strain, overexpression of *cgl1994* and *cgl2806* in the L-lysine-producing strain inhibited growth, glucose catabolism, and L-lysine biosynthesis under hyperosmotic stress (Figure 2C). Such difference may be due to the different genetic backgrounds between the wild-type strain and L-lysine producer.

Cgl3003 alleviates hyperosmotic stress by repairing elevated m6G methylation

Next, we investigated the functional mechanism of two identified components with good performance in improving stress tolerance and L-lysine biosynthesis. According to the annotation,

Figure 2. Comprehensive screening of functional compounds for hyperosmotic stress tolerance and L-lysine production in *Corynebacterium glutamicum*. (A) Effects of overexpressing the 3049 genes on cellular tolerance to 1.2 M L-lysine stress. Each point presents a strain from the collection. Strains with 20% higher and lower fitness than the control strain harboring an empty plasmid are highlighted in red and blue, respectively. For this first round of growth characterization, one replicate was conducted for each of the 3049 strains. (B) Growth curves of the 15 selected gene-overexpressing strains with no stress (white background), high concentration of L-lysine (blue background), and high concentration of Na₂SO₄ (yellow background). The strain with target gene overexpression and the control strain with an empty plasmid are shown by red and black lines, respectively. The gray-shaded area represents the continuous eight data points [optical density at 600 nm (OD₆₀₀) was measured every 2 h] used for calculating the gene fitness following the algorithm described in the STAR★METHODS. The gene fitness is shown in the growth curve as a percentage. Error bars are standard deviations (*n* = 3). (C) Enhanced L-lysine production by overexpression of the 15 selected genes under hyperosmotic stress (1.0 M Na₂SO₄). The plasmid for gene overexpression was transformed into a *C. glutamicum* strain with an engineered L-lysine biosynthetic pathway. Fermentation of the recombinant strains and the strain with an empty plasmid was conducted in 48-deep-well microplates and stopped after 120 h. Error bars are standard deviations (*n* = 3). Individual data points are shown by black circles. Abbreviation: ORF, open reading frame.

Cgl3003 is predicted to be a methylated DNA-protein L-cysteine methyltransferase involved in repairing DNA methylation [44]. However, its biological function has not been experimentally investigated, and the relevance of the balance between the formation and repair of DNA methylation and stress tolerance has not yet been characterized.

First, sequence alignment showed that Cgl3003 shares conserved residues, including the important L-cysteine essential for demethylation activity with the O^6 -methylguanine (m6G) methyltransferase (Ogt, EC: 2.1.1.63) from *Mycobacterium tuberculosis*, *E. coli*, and human (Figure 3A). While most types of nucleobase methylation participate in epigenetic regulation in bacteria and are not mutagenic, m6G incorrectly base-pairs with thymine and consequently generates G:C to A:T mutation [45] (Figure 3B). Second, to evaluate the level of m6G and the role of Cgl3003 under hyperosmotic

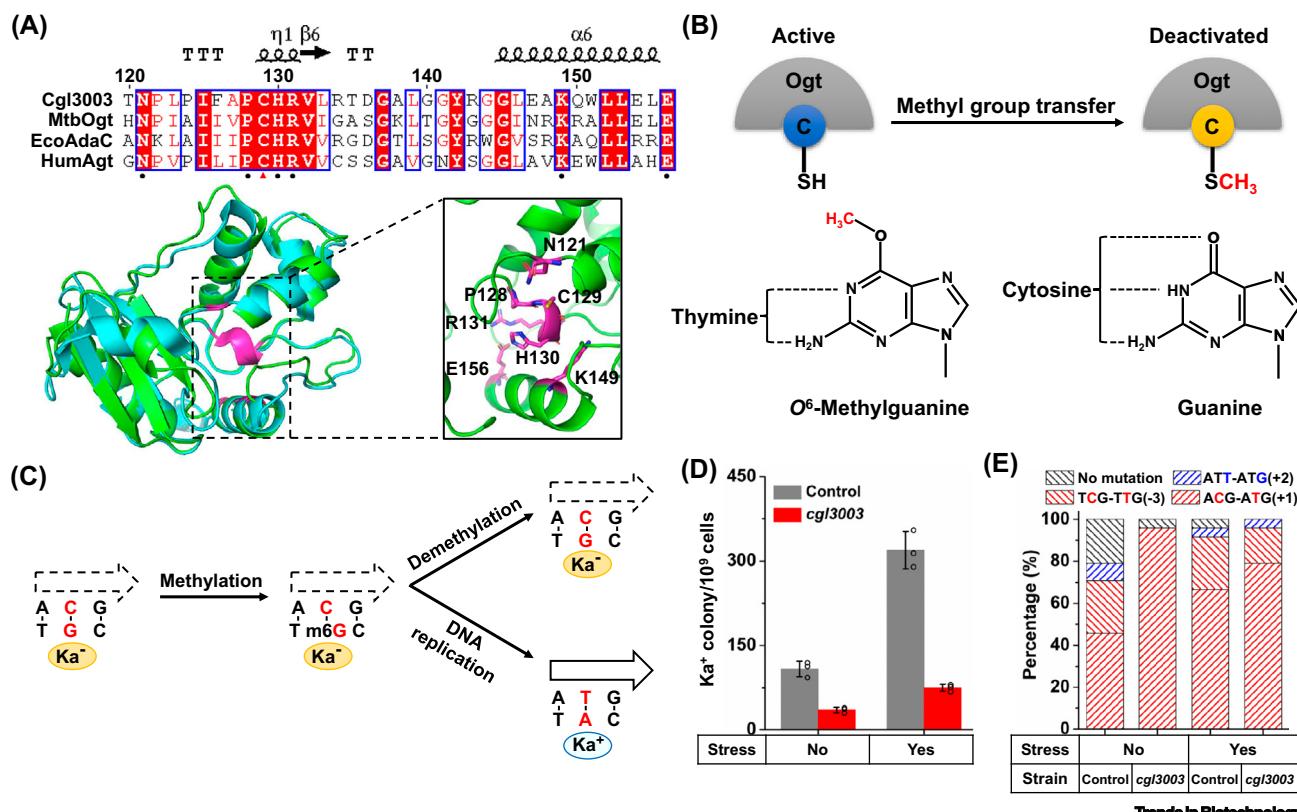


Figure 3. Functional analysis of Cgl3003 as an O^6 -methylguanine (m6G) methyltransferase (Ogt) in protecting cells from hyperosmotic stress. (A) Sequence and structure analysis of Cgl3003. Sequences of Ogs from different sources were aligned using CLUSTALW and analyzed and visualized using ENDscript [77]. Conserved amino acid residuals crucial for activity are marked by black circles and the reactive L-cysteine is marked by a red triangle. MtbOgt, *Mycobacterium tuberculosis* Ogt [Protein Data Bank (PDB) ID: 4WX9]; EcoAdaC, *Escherichia coli* AdaC (PDB ID: 1SFE); and HumAgt, human Agt (PDB ID: 1EH6). The structure of Cgl3003 (green) was predicted using AlphaFold2 [78] and aligned with that of *M. tuberculosis* Ogt (PDB ID: 4WX9) (cyan). Conserved amino acid residuals crucial for activity are highlighted in magenta. (B) Schematic of the methyl group transfer reaction catalyzed by Ogt. The methyl group is transferred from m6G, which incorrectly pairs with thymine to the reactive L-cysteine of Ogt through a suicidal mechanism. (C) Model for determining the frequency of m6G and the repair by Cgl3003. m6G results in G:C to A:T mutation and the recovery of kan^{ACG} to kan^{ATG}, leading to a kanamycin resistance phenotype. (D) Formation of kanamycin-resistant cells with or without hyperosmotic stress. A kan^{ACG} marker gene was integrated into the chromosome of *Corynebacterium glutamicum*. The control strain harboring an empty plasmid and the strain with cgl3003 overexpression were cultivated with or without hyperosmotic stress. The same amounts of cells were collected at the stationary phase and spread on plates with kanamycin. Colonies were counted after 48 h of cultivation. Error bars are standard deviations ($n = 3$). Individual data points are indicated by black circles. (E) kan^{ACG} mutation analysis by Sanger sequencing of kanamycin-resistant colonies. The kan^{ACG} genes from 24 colonies were sequenced for each test. Con. and 3003 represent the control and cgl3003-overexpressing strains, respectively. The designed start codon ATG produced by m6G-mediated G:C to A:T mutation is designated as position +1. Another two mutations leading to start codons TTG at position -3 and ATG at position +2 were also detected.

stress *in vivo*, a kanamycin-resistant marker gene with the start codon ATG mutated to ACG (*kan*^{ACG}) was integrated into the chromosome of *C. glutamicum*. m6G-mediated G:C to A:T mutation is expected to recover the start codon and result in kanamycin resistance (Figure 3C). The engineered *C. glutamicum* *kan*^{ACG} strains harboring an empty plasmid and a *cgl3003*-overexpressing plasmid were cultivated with or without hyperosmotic stress, and kanamycin-resistant colonies were counted and sequenced. Interestingly, for the control strain, hyperosmotic stress led to a greater than twofold increase in the number of kanamycin-resistant colonies. Overexpression of *cgl3003* decreased the emergence of kanamycin-resistant colonies under normal conditions by 2.8-fold, and a 4.3-fold decrease was observed for cultivation under hyperosmotic stress (Figure 3D). The *kan*^{ACG} genes in the kanamycin-resistant colonies were then sequenced to determine the genetic mutation. Besides the predicted ACG-ATG mutation at position +1, another two types of mutation resulting in translation initiation, TCG-TTG mutation at position -3 and ATT-ATG mutation at position +2, were also detected. The generation of a TTG start codon at position -3 may also be caused by the m6G-mediated G:C to A:T mutation, whereas the low-frequency ATT-ATG mutation at position +2 may be due to spontaneous mutation. The ACG-ATG and TCG-TTG mutations (both G:C to A:T mutations) accounted for 70–96% of kanamycin-resistant colonies (Figure 3E and Figure S8 in the supplemental information online). Therefore, the m6G-mediated G:C to A:T mutation was considered to be the main cause of the emergence of kanamycin resistance.

Hyperosmotic stress affects cells by driving the exit of water and consequently triggering cytoplasmic dehydration and a drop in turgor [46]. Based on the earlier results, it is suggested that exposure to hyperosmotic stress results in elevated m6G levels and mutagenesis frequency, which is also considered as a key factor for the inhibition caused by hyperosmotic stress. Our observations support the importance of Cgl3003 in protecting the GC-rich genome (53.8% GC-content for type strain ATCC 13032) of *C. glutamicum* from the promutagenic potential of m6G. *E. coli* has been reported to have two O⁶-methylguanine methyltransferases, the Cgl3003 homolog Ogt, and a bifunctional DNA repair protein/DNA-binding transcriptional regulator Ada. Expression of Ada is induced by DNA alkylation damage and, upon methylation, becomes a transcriptional activator of the genes involved in the adaptive response [47]. However, a blast analysis revealed that no Ada homolog exists in the *C. glutamicum* genome, making Cgl3003 the only m6G methylation repair enzyme in this species. It has been suggested that the Ogt from *M. tuberculosis* functions through a suicidal mechanism [44]. We hypothesize that the native expression level of Cgl3003 may be insufficient for m6G repair under hyperosmotic stress. Therefore, overexpression of Cgl3003 allows cells to alleviate hyperosmotic stress by repairing elevated m6G methylation.

Cgl2496 is a novel global regulator of stress tolerance

Cgl2496 was annotated as a PucR family transcriptional regulator. PucR was first identified as a transcriptional activator for genes involved in purine catabolism in *Bacillus subtilis* [48]. Another three PucR-type regulators, PrcR, AdeR, and GabR, have been identified in *B. subtilis* and *C. glutamicum*, which participate in the regulation of L-proline utilization, L-alanine dehydrogenase expression, and γ-aminobutyrate utilization, respectively [49–51]. However, Cgl2496 shares <20% amino acid sequence identities with the reported four PucR-type regulators. Homologs of Cgl2496 widely exist in the genomes of *Actinomycetia* (Figure S9 in the supplemental information online), although their biological function has not yet been characterized. Structure and function prediction using the Phyre2 bioinformatics tool [52] revealed the C-terminal helix-turn-helix (HTH) domain and implied its possible function as a transcriptional regulator.

To investigate the regulatory role of Cgl2496 in enhancing cellular tolerance to hyperosmotic stress, transcriptome analysis was performed to determine the genes controlled by Cgl2496.

According to the cell growth curves under hyperosmotic stress, the strain overexpressing Cgl2496 had similar growth with the control strain harboring an empty plasmid and showed growth advantage only after ~50 h of cultivation (Figure 2B). Therefore, cells grown under hyperosmotic stress were sampled at 36 h (with no growth advantage) and 82 h (with obvious growth advantage) for transcriptome analysis. The RNA sequencing data suggested that Cgl2496 overexpression led to the upregulation of 110 genes and downregulation of 73 genes (total 183 differentially expressed genes) at 36 h compared with the strain harboring an empty plasmid. However, the number of differentially expressed genes caused by Cgl2496 overexpression at 82 h increased to 538 genes (234 upregulated genes and 304 downregulated genes) (Figure 4A). A cluster analysis showed that the transcriptome profiles of the Cgl2496-overexpressing strain and control strain were more dissimilar at 82 h than at 36 h. According to the **clusters of orthologous groups (COG)** annotation [53], the differentially expressed genes were mainly involved in the biological processes of energy production and conversion, amino acid transport and metabolism, carbohydrate transport and metabolism, translation, ribosomal structure, and biogenesis (Figure 4B).

Cgl2496 overexpression significantly affected the transcription of over 500 genes. To identify the genes directly controlled by Cgl2496 and map the binding sites, we performed **ChIP-Seq**. N-terminal and C-terminal 3xFlag-tagged Cgl2496 proteins were episomally expressed under the control of an IPTG-inducible promoter. Since overexpression of the N-terminal tagged Cgl2496 maintained the growth advantage under hyperosmotic stress, it was selected for the ChIP-Seq experiment (Figure 4C). By sequencing and analyzing the DNA extracted from the DNA-Cgl2496 (3xFlag-tagged) complex, a total of 67 binding regions in the *C. glutamicum* chromosome were enriched (Figure S10 and Data S9 in the supplemental information online). Among these, 13 binding regions were located inside the gene and 54 were located in the intergenic region, which possibly facilitated transcriptional regulation of the downstream genes (Figure 4D). By using MEME [54], a DNA-binding motif shared by 37 binding sites was identified (Figure 4E and Data S10 in the supplemental information online). By integrating the transcriptome and ChIP-Seq data (peaks in the intergenic region), we found that six and 11 of the 67 genes were upregulated and downregulated, respectively due to the overexpression of *cgl2496* (Figure 4F).

Several of the downregulated genes are involved in oxidative phosphorylation, including *cgl2192* (*ctaE*), *cgl2195* (*ctaC*), and *cgl2523* (*ctaD*), which encode the Heme/copper-type cytochrome/quinol oxidase subunit 3, subunit 2, and subunit 1, respectively (Figure 5A). *cgl2192* and *cgl2195* belong to two oxidative phosphorylation-related gene operons, *cgl2192-cgl2191-cgl2190-cgl2189* and *cgl2195-cgl2194*, respectively. Given that *cgl2192* and *cgl2195* are the first genes of their operons, Cgl2496-mediated regulation also caused a two- to sixfold repression of other genes in these two operons. In addition, the *cgl2566* (*sucC*)-*cgl2565* (*sucD*) operon encoding succinyl-CoA synthetase β and α subunits were downregulated by 22- and 17-fold, respectively (Figure 5A). Considering the physiological role of these operons in oxidative phosphorylation and tricarboxylic acid (TCA) cycle, downregulation of their transcription may affect energy metabolism. We then determined the intracellular ATP level of the strain overexpressing *cgl2496*. The results suggest that *cgl2496* overexpression significantly decreases ATP level, which is consistent with the repression of genetic operons involved in energy metabolism (Figure 5B).

Besides these metabolic genes, Cgl2496 directly regulates several genes encoding transcriptional regulators, such as *cgl2167* (*mraZ*) [55], *cgl0768* (*whiB1*) [56], and *cgl0279* (*whiB4*) [57], suggesting the role of Cgl2496 as a master or global regulator. *cgl2167* is involved in the regulation of cell division [55,58]. *cgl0768* regulates the transcription of *cgl0279*, both of which are related to the stress response and cell morphology [59]. Environmental stress is suggested to

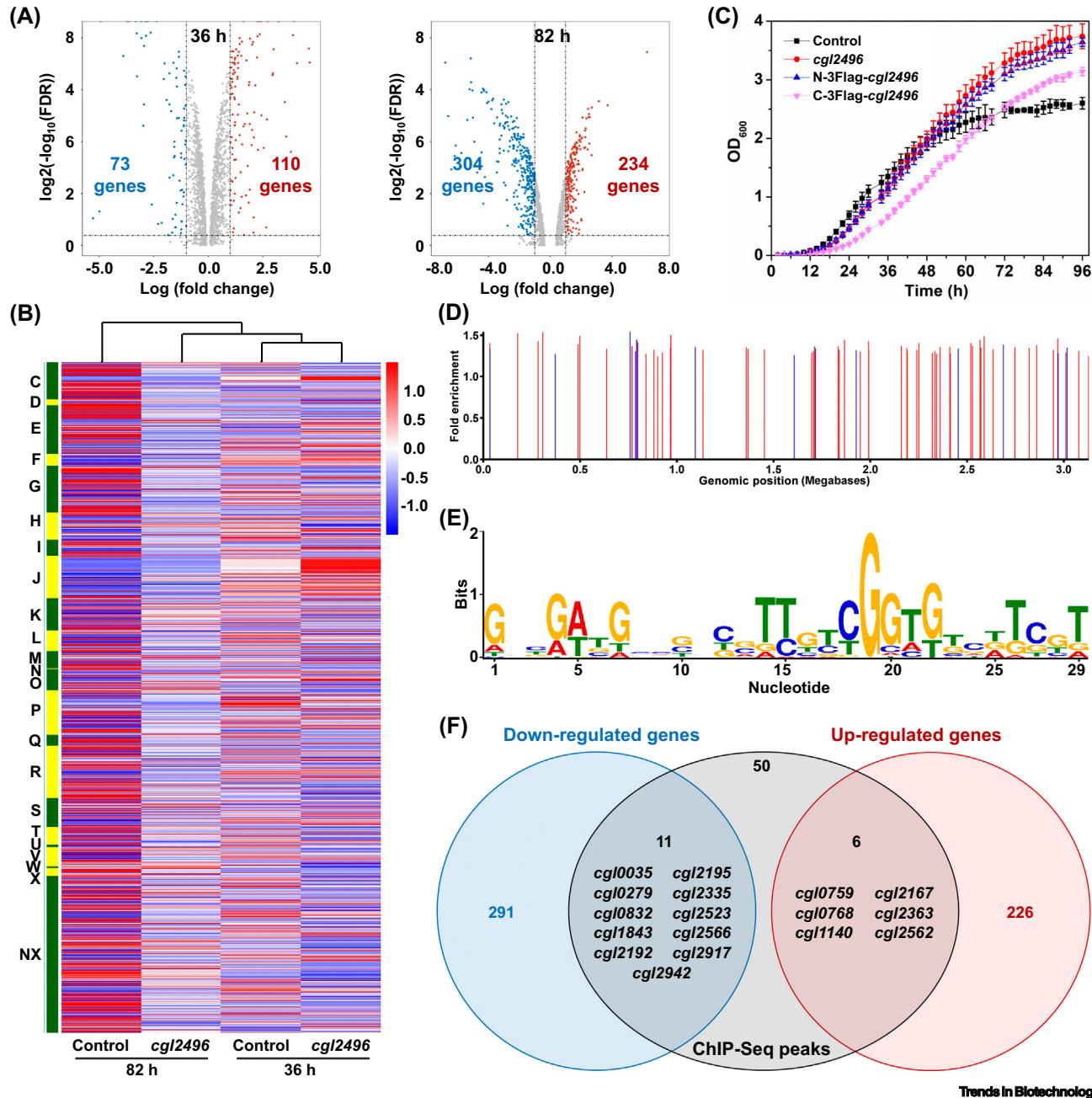
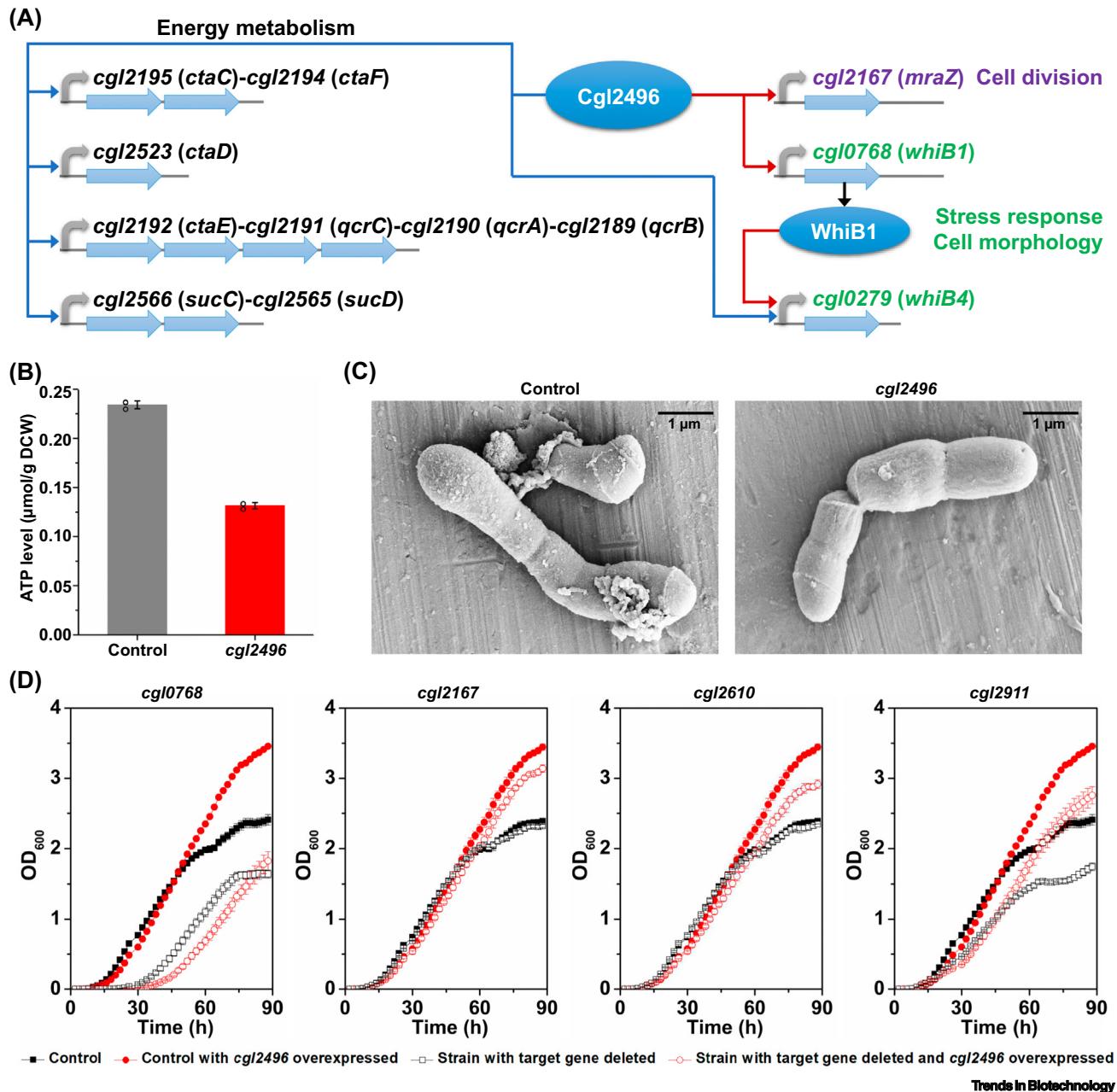


Figure 4. Effects of *cgl2496* overexpression on transcriptome profile. (A) Volcano plot of differential transcript levels caused by *cgl2496* overexpression as determined by RNA sequencing ($n = 3$). Cells grown under hyperosmotic stress were sampled at 36 h (with no growth advantage) and 82 h (with obvious growth advantage) for transcriptome analysis. (B) Heatmap of the transcriptional changes caused by *cgl2496* overexpression ($n = 3$). Genes were categorized according to the database of clusters of orthologous genes (COGs) [53]. (C) Cell growth of strains overexpressing N-terminal and C-terminal 3×Flag-tagged *cgl2496* under hyperosmotic stress. Error bars are standard deviations ($n = 2$). (D) The 67 enriched binding regions (peaks) in the *Corynebacterium glutamicum* chromosome by chromatin immunoprecipitation followed by sequencing (ChIP-Seq). The peaks inside the gene and the intergenic region are marked in blue and red, respectively. (E) DNA-binding motif of Cgl2496, predicted by using MEME [54]. Thirty-seven sites shared the predicted motif ($E\text{-value} = 4.6 \times 10^{-9}$). (F) Intersection of Cgl2496 ChIP-Seq peaks (in the intergenic region) and differentially expressed genes caused by *cgl2496* overexpression (82 h). Abbreviations: FDR, false discovery rate; OD₆₀₀, optical density at 600 nm.



Trends In Biotechnology

Figure 5. Regulatory role of Cgl2496 in cell division and energy metabolism. (A) Schematic of the role of Cgl2496 in regulating genes involved in cell division and energy metabolism. The red and blue arrows represent upregulation and downregulation, respectively. The genes involved in energy metabolism, cell division, and stress response are in black, purple, and green, respectively. (B) Intracellular ATP levels of the control strain harboring an empty plasmid and the strain with *cgl2496* overexpressed under hyperosmotic stress. Error bars are standard deviations ($n = 3$). Individual data points are in black circles. (C) Scanning electron microscopy (SEM) of the *Corynebacterium glutamicum* control strain harboring an empty plasmid and the strain with *cgl2496* overexpressed under hyperosmotic stress. More SEM images are shown in Figure S11 in the supplemental information online. (D) Effects of target gene deletion on Cgl2496 overexpression-mediated tolerance to hyperosmotic stress. *cgl2496* was overexpressed in the control (wild-type) strain and the engineered strain with the target gene (*cgl0768*, *cgl2167*, *cgl2610*, and *cgl2911*) deleted, respectively. All strains were cultivated under the stress of 1.2 M L-lysine. Error bars are standard deviations ($n = 3$). Abbreviations: DCW, dry cell weight; OD₆₀₀, optical density at 600 nm.

cause morphological changes. To test whether Cgl2496-mediated regulation of *cgl2167*, *cgl0768*, and *cgl0279* changed the cell morphology under hyperosmotic stress, cells with and without *cgl2496* overexpression were checked under a **scanning electron microscope (SEM)**. Under hyperosmotic stress, the control strain suffered from disordered cell division and aggravated cell lysis. Cells with long and abnormal morphology were observed. By contrast, *cgl2496* overexpression counteracted the effects of hyperosmotic stress and allowed cells to maintain normal division and morphology (Figure 5C and Figure S11 in the supplemental information online), which may be the mechanism of *cgl2496*-mediated tolerance to hyperosmotic stress.

Since Cgl2496 directly bound the promoter region of *cgl0768* and *cgl2167* and upregulated their transcription, we hypothesized that deletion of *cgl0768* and *cgl2167* would break the regulatory chain of Cgl2496 and weaken the beneficial effects of *cgl2496* overexpression on hyperosmotic stress tolerance. We also analyzed the intersection of the upregulated genes caused by *cgl2496* overexpression (Figure 4A) and the 15 genes identified to enhance hyperosmotic stress tolerance (Figure 2B), and found two hits, *cgl2610* and *cgl2911*. To test the effects of these four genes (*cgl2167*, *cgl2610*, *cgl0768*, and *cgl2911*) on the regulatory function of Cgl2496, they were individually deleted in *C. glutamicum*. While deletion of *cgl2167* and *cgl2610* did not affect cell growth under hyperosmotic stress, deletion of *cgl0768* and *cgl2911* largely impaired cellular tolerance to hyperosmotic stress (Figure 5D). Next, *cgl2496* was overexpressed in the gene-deleted mutants; its overexpression in the *cgl0768*-deleted mutant did not improve cellular tolerance to hyperosmotic stress. Although *cgl2496* overexpression in the mutants with *cgl2167* and *cgl2610* deletion led to increased hyperosmotic stress tolerance, the beneficial effects were not as large as seen for *cgl2496* overexpression in the wild-type strain. These data suggest that *cgl0768*, *cgl2167*, and *cgl2610* are important for the *cgl2496*-mediated tolerance to hyperosmotic stress. However, it is difficult to decipher the role of *cgl2911* in the regulatory function of *cgl2496*, since *cgl2496* overexpression in the *cgl2911*-deleted mutant significantly increased cellular tolerance (Figure 5D).

Identification of novel L-threonine exporters by gene overexpression screening

Efficient export of intracellularly synthesized molecules to the culture medium is crucial for product hyper-production. Identification of exporters has largely depended on genome-scale random mutation combined with screening and deciphering target molecule (or structural analog)-resistant mutants [7,60,61]. However, poorly expressed or silent exporter-encoding genes have probably been neglected by previous strategies. For a comprehensive screening of novel transporters, all 389 strains overexpressing the annotated membrane transporter genes were reorganized into 48-deep-well microplates. The exporters of L-threonine were screened as a demonstration. Since construction of *cgl1130* (encoding an ABC-type transporter) was unsuccessful, it was not covered by the screening. By using the previously characterized L-threonine exporters Cgl2622 (ThrE) [62] and Cgl0605 (SerE) [63] as positive controls, a screening strategy based on cell growth under Thr–Thr dipeptide stress was developed. With the addition of Thr–Thr dipeptide, cells accumulate excessive L-threonine *in vivo* and cell growth is consequently inhibited. Overexpression of a functional L-threonine exporter can mediate efficient excretion of L-threonine and maintain homeostasis. As a result, compared with the control strain harboring an empty plasmid, strains overexpressing the L-threonine exporter showed a growth advantage upon the addition of Thr–Thr dipeptide (Figure 6A). Taking Cgl2622 as an example, its overexpression did not have significant effects on cell growth. The addition of Thr–Thr dipeptide inhibited cell growth, whereas the Cgl2622-overexpression strain showed an obvious growth advantage, suggesting the L-threonine export function of Cgl2622 (Figure S12 in the supplemental information online). In contrast to Cgl2622, overexpression of Cgl0605 caused obvious inhibition of

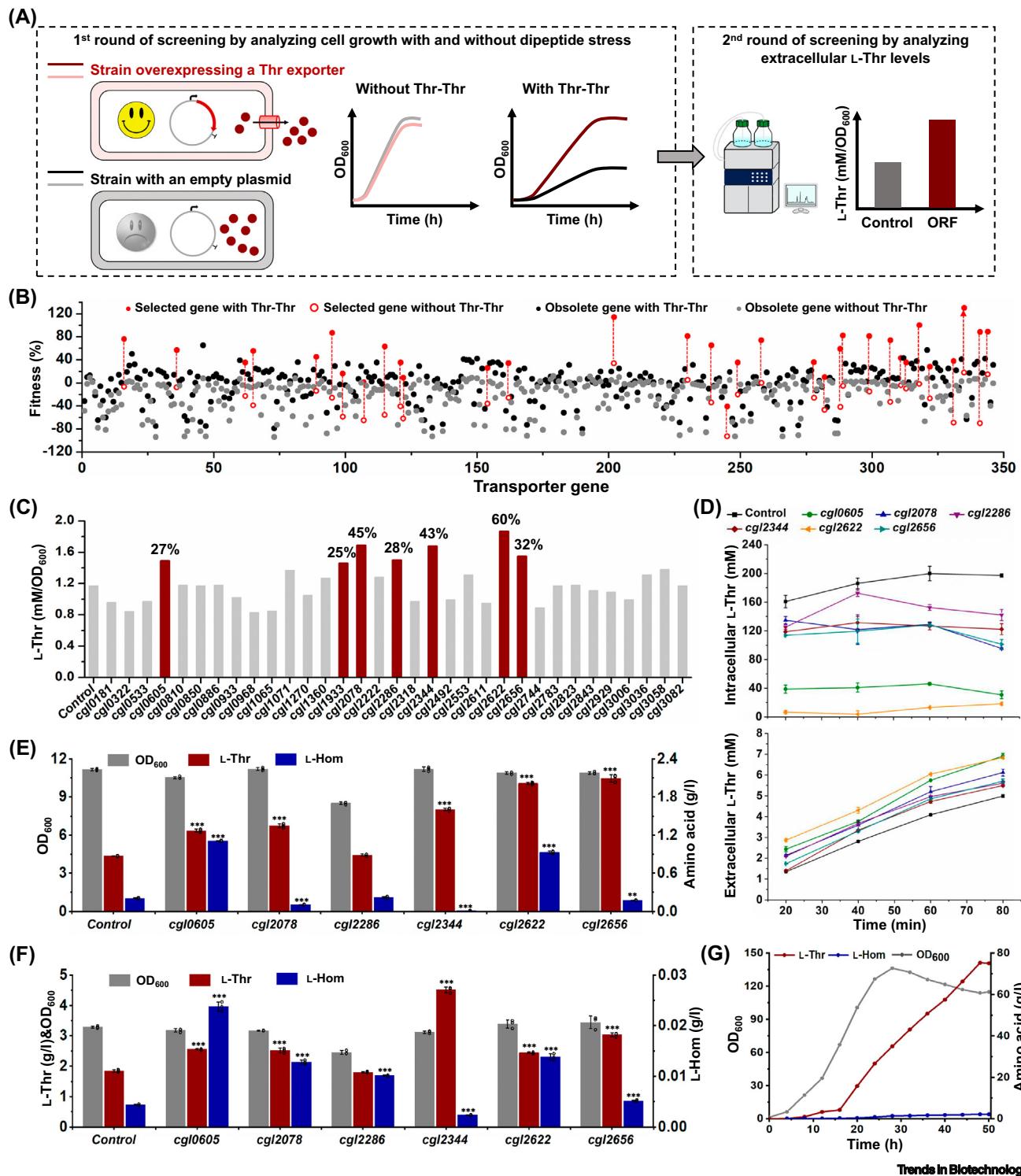


Figure 6. Identification and characterization of novel L-threonine exporters. (A) Schematic of the two-round screening procedure for identification of L-threonine exporters. For the first round of screening, all 389 strains overexpressing membrane transporter genes and a negative control with an empty plasmid were cultivated with or without Thr-Thr dipeptide stress. Those with relative growth advantage under Thr-Thr dipeptide stress entered the second round of screening to determine the (Figure legend continued at the bottom of the next page.)

cell growth, which was possibly due to the metabolic burden or unrestricted excretion of metabolites by Cgl0605. In the presence of Thr–Thr dipeptide, the difference in cell growth between the Cgl0605 overexpression strain and the control strain diminished (Figure S12). Therefore, Cgl0605 overexpression allowed cells to tolerate Thr–Thr dipeptide stress. In summary, it is feasible to identify potential L-threonine exporters by analyzing cell growth with and without dipeptide stress.

Two rounds of screening were conducted to identify novel L-threonine exporters. The first round of screening was performed by cultivating all 389 strains overexpressing membrane transporter genes and a negative control with an empty plasmid. By comparing the growth of gene-overexpressing strains with and without Thr–Thr dipeptide stress using a developed algorithm (Figure S13 in the supplemental information online), 33 transporters (including the previously reported L-threonine exporters, Cgl2622 [62] and Cgl0605 [63]) were found to enhance cell growth under Thr–Thr dipeptide stress compared with the negative control strain (Figure 6B; Figure S14 and Data S11 in the supplemental information online). These 33 strains were subjected to a second round of screening by quantifying the final concentration of extracellular L-threonine when they were cultivated with the addition of Thr–Thr dipeptide. Seven transporters (Cgl0605, Cgl1933, Cgl2078, Cgl2286, Cgl2344, Cgl2622, and Cgl2656) enhanced the accumulation of extracellular L-threonine by over 25% under Thr–Thr dipeptide stress (Figure 6C). However, the extracellular L-threonine levels of the remaining 26 transporters did not significantly increase, suggesting that the improvement in cell growth observed in the first round of screening was not due to accelerated L-threonine transport but to other unknown mechanisms. Cgl1933 (Ptsl) belongs to the phosphotransferase system (PTS) for sugar uptake [64] and, thus, it was not further investigated here. To verify the function of the remaining screened six transporters as functional L-threonine exporters, classic peptide uptake and amino acid export assays were conducted. Upon the addition of Thr–Thr dipeptide, the control strain harboring an empty plasmid showed very high intracellular L-threonine levels (>160 mM). Continuous accumulation of extracellular L-threonine was also observed because of the expression of chromosomally expressed L-threonine exporters. Overexpression of the screened six transporters led to different levels of decrease in intracellular L-threonine concentration and increase in L-threonine export (Figure 6D). These results confirmed the function of the six transporters as L-threonine exporters.

We next analyzed the classification of these L-threonine exporters. The two known L-threonine exporters, Cgl0605 and Cgl2622, belong to the DMT family transporter and ThrE family transporter, respectively. Interestingly, none of the four newly discovered L-threonine exporters (Cgl2078, Cgl2286, Cgl2344, and Cgl2656) belonged to either the DMT or ThrE family. Cgl2078 is classified as an amino acid ABC transporter and shares transmembrane structure similarity with the L-cysteine exporter CydDC of *E. coli* [65] (Figure S15 in the supplemental

extracellular L-threonine concentrations in the presence of Thr–Thr dipeptide stress. (B) First round of screening for identification of L-threonine exporters. The selected 33 potential L-threonine exporters that entered the second round of screening are in red. (C) Second round of screening for identification of L-threonine exporters. The selected six potential L-threonine exporters that caused a >25% increase in the extracellular L-threonine concentration upon Thr–Thr dipeptide addition are in red. The increase in the extracellular L-threonine concentration is marked as a percentage. (D) Thr–Thr dipeptide uptake and export assay for the selected potential L-threonine exporters. Intracellular and extracellular L-threonine concentrations were detected periodically after the addition of Thr–Thr dipeptide. Error bars are standard deviations ($n = 3$). (E) Effects of L-threonine exporter overexpression on L-threonine and L-homoserine production of an L-threonine-producing *Corynebacterium glutamicum* strain. Individual data points are shown as black circles. Error bars are standard deviations ($n = 3$). Student's two-tailed *t*-test was conducted to compare the L-threonine (or L-homoserine) titers of strains overexpressing L-threonine exporters with a control strain without L-threonine exporter overexpression (** $P < 0.01$, *** $P < 0.001$). (F) Application of L-threonine exporters in an L-threonine-producing *Escherichia coli* strain. Individual data points are shown in black circles. Error bars are standard deviations ($n = 3$). Student's two-tailed *t*-test was conducted to compare the L-threonine (or L-homoserine) titers of strains overexpressing L-threonine exporters with control strain without L-threonine exporter overexpression (** $P < 0.001$). (G) Fed-batch fermentation of the L-threonine-producing *C. glutamicum* strain with Cgl2344 overexpression in a 5-l bioreactor. Abbreviations: ORF, open reading frame; OD₆₀₀, optical density at 600 nm.

information online). Cgl2078 is the first confirmed L-threonine exporter of the amino acid ABC transporter family. Cgl2286 belongs to the hemolysin family and is the first member of this family with amino acid transport function. Both Cgl2344 and Cgl2656 belong to the LysE translocator family, and share transmembrane structure similarity with the L-threonine exporters RhtB and RhtC of *E. coli* [66,67] (Figure S15). An amino acid sequence alignment suggested that the four newly discovered L-threonine exporters (Cgl2078, Cgl2286, Cgl2344, and Cgl2656) all shared <30% similarity with the known L-threonine exporters (Figure S16 in the supplemental information online). Therefore, discovery of these exporters based on sequence blast would be difficult. Here, by using genome-scale screening, several L-threonine exporters with sequence, structure, and function diversities were identified, laying the foundation for investigating their physiological functions and applications in amino acid overproduction.

Application of the novel L-threonine exporters in L-threonine production in *C. glutamicum* and *E. coli*

One important application of transport proteins is to enhance the bioproduction of target molecules [7,60]. To test the application of the newly identified L-threonine exporters in L-threonine bioproduction, they were individually overexpressed by plasmids in an engineered L-threonine-overproducing *C. glutamicum* that harbored two chromosomal mutations (*hom*^{G378E} and *lysC*^{T311I}) and a *hom*^{G378E} and *thrB*-overexpressing plasmid. Overexpression of Cgl0605, Cgl2078, Cgl2344, Cgl2622, and Cgl2656 increased L-threonine production by 46–141% (Figure 6E). The accumulation of by-product L-homoserine was also largely increased when Cgl0605 and Cgl2622 were overexpressed, which was consistent with the reported substrate promiscuity of these two exporters [18,62,63]. Overexpression of the four newly identified L-threonine exporters did not increase L-homoserine accumulation, which was an advantage for improving conversion yield and simplifying product purification (Figure 6E). Therefore, we envision that these newly identified exporters will be valuable new targets for L-threonine transport engineering.

Next, the combinational effects of simultaneous overexpression of two exporters were tested. Since Cgl2656 showed the best performance in enhancing L-threonine production, it was overexpressed with the other five transporters. However, construction of the Cgl2656-Cgl2286 co-expressing plasmid in *E. coli* failed after several trials. The remaining five plasmids were successfully constructed, whereas transformation of the Cgl2656-Cgl2344 co-expressing plasmid into *C. glutamicum* did not produce any transformants after several trials, which may be due to the toxic effects on cells of the excessive excretion of intracellular amino acids. The resulting three engineered strains were cultivated and tested for L-threonine production. Although these strains produced 23–123% higher L-threonine compared with the control strain with an empty plasmid, the improvements were lower than that obtained by individually overexpressing Cgl2656 (Figure S17 in the supplemental information online).

In addition to application in their native host, transport proteins can usually be used in heterologous microbial hosts for enhancing bioproduction [7,60]. An L-threonine-producing *E. coli* strain was used as the host in this instance, and was constructed by expressing an L-threonine biosynthesis *thrA*^{E253K}*BC* operon-overexpression plasmid [68]. Individual overexpression of Cgl0605, Cgl2078, Cgl2344, Cgl2622, and Cgl2656 increased the L-threonine titer by 32–144%, with the newly identified Cgl2344 showing the largest increase (144%) (Figure 6F). Application of Cgl2344 also resulted in the lowest accumulation of the by-product L-homoserine, which was only 10% and 17% of those produced by strains overexpressing the known L-threonine exporters Cgl0605 and Cgl2622, respectively. However, Cgl2286 failed to improve L-threonine production

(Figure 6F). The results suggest that application of the identified transporters is not limited to the native host *C. glutamicum* but can be expanded to other microbial hosts, such as *E. coli*.

Finally, we demonstrated the application of the newly discovered L-threonine exporters in industry-level L-threonine production. We previously constructed a *C. glutamicum* strain ZcglT9 with the highest reported L-threonine production level (67.6 g/l and 1.2 g/l·h) by systems metabolic engineering. However, 4.6 g/l L-homoserine was detected in the fermentation broth as the main by-product [69]. In former tests, Cgl2344 led to the lowest by-product L-homoserine accumulation and Cgl2656 showed the highest L-threonine production (Figure 6E). Therefore, Cgl2344 and Cgl2656 were used to enhance L-threonine production. In 24-deep-well plate fermentation, the engineered strain with Cgl2344 overexpression produced 14.3 g/l L-threonine, which was 1.5-fold higher than that produced by the Cgl2656-overexpressing strain (Figure S18 in the supplemental information online). Next, the best-producing strain with Cgl2344 overexpression was subjected to fed-batch fermentation in a 5-l bioreactor. After 50 h fermentation, 75.1 g/l L-threonine was produced with a productivity of 1.50 g/l/h and a yield of 0.22 g/g glucose (Figure 6G). The titer and productivity were 11% and 25% higher, respectively, than that obtained in our previous study [69], representing the new record for L-threonine production in *C. glutamicum*. Given the substrate specificity of Cgl2344, only 2.2 g/l L-homoserine was produced, which was 48% of that produced by strain ZcglT9 [69]. These results demonstrate the application of the screened functional components in industry-level bioproduction.

Discussion

Industrial biomanufacturing requires microbial strains with not only enhanced biosynthesis, but also adaptability to industrially relevant environmental factors, particularly the high concentrations of intracellular and extracellular product molecules. Therefore, functional components capable of exporting product molecules outside cells and assisting cells to tolerate the stress caused by high-level product accumulation are always greatly valued [7,8].

In this study, phenotype profiling of a newly constructed gene overexpression collection of *C. glutamicum* was conducted for a comprehensive screening of novel targets for enhancing stress tolerance, amino acid export, and bioproduction. The screened functional components functioned in not only the wild-type *C. glutamicum*, but also industry-level amino acid-producing *C. glutamicum* strains, as well as heterologous microbial hosts, such as *E. coli*. In terms of the genes selected for hyperosmotic stress tolerance, we observed a large variety in their physiological functions involved in carbon and energy metabolism (*cgl0917*, *cgl1010*, *cgl1118*, *cgl2392*, *cgl2610*, and *cgl2911*), DNA repair (*cgl3003*), transportation (*cgl0470*), transcriptional regulation (*cgl0923* and *cgl2496*), protein synthesis (*cgl1994*), and unknown biological processes (*cgl0063*, *cgl1472*, *cgl2806*, and *cgl2998*). The correlation between these genes and hyperosmotic stress is difficult to predict based on existing knowledge. Indeed, complex phenotypes, such as cell growth and stress tolerance, are systematically controlled by a variety of genes and the genetic mechanisms are not yet fully characterized. Interestingly, when the selected genes were deleted to verify their functions, few led to phenotype changes, suggesting that many genes are natively expressed at very low levels or completely silent. Therefore, GOF screening with gene overexpression is an important complement to the widely used LOF screening approaches with CRISPRi [70], base editing [71], or transposon mutagenesis [72]. Combining these GOF and LOF screening approaches can be used for exploring the synergistic effects of overexpression and deletion of multiple genes. Although such combinations will generate large sets of data, booming machine learning and artificial intelligence technologies can be used for data processing and elucidation of the synergistic effects [73].

Besides the examples demonstrated in this study, the gene overexpression strain collection provides a valuable resource for functional genomics and systems biology research on *C. glutamicum*. Systematic phenotype profiling under selected conditions combined with multi-omics analyses [74] will generate a large data set for modeling and predicting *C. glutamicum* metabolism and regulation. Considering that the gene-overexpression plasmids are transferable to industrial *C. glutamicum* mutants, such as the L-lysine and L-glutamate-producing strains [5], the collection is useful for examining the properties of genes in industrial bioproduction and identifying novel genetic targets for strain improvement. The gene-overexpression plasmids can also be replicated in many other *Corynebacterium* strains [75], such as *Corynebacterium acetoacidophilam*, *Corynebacterium pekinense*, and *Corynebacterium crenatum*. The plasmid replication initiator (Rep) shares a high degree of similarity with the Rep protein encoded by the broad host-range plasmid pNG2 from the human pathogen *Corynebacterium diphtheriae* [76]. Therefore, the developed gene-overexpression plasmids may be used in other *Corynebacterium* strains. In addition, since the gene overexpression plasmids are based on an *E. coli*-*C. glutamicum* shuttle, they can also replicate in *E. coli*. Using RFP as a reporter, over tenfold gene induction by IPTG was observed (Figure S19 in the supplemental information online). The identified L-threonine exporters have also been proven effective in *E. coli*, suggesting the possibility of screening functional components in this species. By making this resource available to the research community, we hope to contribute to the worldwide efforts for a comprehensive understanding of *C. glutamicum* and accelerate the discovery of functional components that are applicable in various industrial microorganisms.

Concluding remarks

Development of industrial strains for biomanufacturing requires better understanding of the biology and identification of functional components associated with the production efficiency. This work provides a systematic strategy combining the automated construction of a genome-scale controllable gene overexpression strain collection using the state-of-the-art BioFoundry, and high-throughput phenotype screening for functional components based on the collection. As demonstrated, many new components enhancing cellular hyperosmotic tolerance and amino acid exports were identified and successfully used to generate amino acid hyperproducers. The collection should greatly promote the biological study and biotechnological development with the industrial workhorse *C. glutamicum*. With the rapid development of laboratory automation and artificial intelligence, this strategy should be readily extended to other well- or less well-studied industrial microorganisms (see [Outstanding questions](#)), greatly accelerating the development of efficient industrial microorganisms for biomanufacturing.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- METHOD DETAILS
 - Strains and culture conditions
 - Genome resequencing of *C. glutamicum* ATCC 13032
 - Plasmid construction
 - Automated construction of gene overexpression collection
 - High-throughput growth characterization of gene overexpression collection
 - Calculation of gene fitness
 - Chromosomal engineering of *C. glutamicum*
 - L-Lysine production by strains overexpressing selected genes
 - Detection of m6G-induced mutation and Cgl3003-mediated m6G repair
 - Transcriptome analysis
 - ChIP-Seq analysis

Outstanding questions

How can we integrate the genome-scale gene overexpression collection with other technologies, such as artificial intelligence and systems biology, to accelerate the discovery of functional components and the decipherment of the underlying mechanisms?

Can efficient high-throughput methods be developed other than growth-coupled phenotype screening to fully explore the potential of the arrayed library resources?

How can the identified functional components and knowledge further drive innovation in *C. glutamicum* and beyond?

Can the BioFoundry and the screening strategy developed in this study be extended to other less well-studied industrial microorganisms to understand the biology and upgrade their production efficiency?

- SEM of *C. glutamicum*
- Measurement of intracellular ATP levels
- Screening of L-threonine exporters
- Peptide uptake and amino acid export assay
- L-Threonine production by *C. glutamicum* overexpressing selected genes
- L-Threonine production by *E. coli* overexpressing selected genes

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to and will be fulfilled by the lead contact, Meng Wang (wangmeng@tib.cas.cn).

Materials availability

The authors agree to provide any materials and strains used in this study upon request. To access the collection information and request materials and strains used in this study, please visit our website <https://imrc.biodesign.ac.cn/>. For more details, please refer to the Strain Sharing Statement and Material Transfer Agreement on our website.

Data and code availability

All data supporting the findings of this work are available within the article and the supplemental information files online. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon request.

Author contributions

Y.W., P.Z., M.W., and J.S. conceived and designed this project. J.L., X.Z., H.C., Y.G., L.W., G.S., X.W., J.C., J.W., J.A., X.G., and Y.W. performed the experiments. J.L., X.Z., H.C., X.N., and Y.W. analyzed the data. Z.S., H.L., R.W., M.Z., and X.L. designed and built the website. Y.W., P.Z., M.W., and J.S. supervised the research and contributed reagents and analytic tools. Y.W. wrote the initial paper draft and all authors contributed to the discussion and writing of the final article.

Acknowledgments

This research was supported by the National Key R&D Program of China (2021YFC2100900), the National Natural Science Foundation of China (32270101 and 32222004), the Strategic Priority Research Program of the Chinese Academy of Sciences (XDC0110201), the Tianjin Synthetic Biotechnology Innovation Capacity Improvement Project (TSBICIP-CXRC-079, TSBICIP-PTJJ-007, and TSBICIP-PTJS-003), and the Youth Innovation Promotion Association of Chinese Academy of Sciences (2021177).

Declaration of interests

The authors have filed patent applications based on some of the technology described in this paper.

Supplemental information

Supplemental information associated with this article can be found online at <https://doi.org/10.1016/j.tibtech.2024.09.020>.

References

1. Clomburg, J.M. *et al.* (2017) Industrial biomanufacturing: The future of chemical production. *Science* 355, eaag0804
2. Lee, S.Y. and Kim, H.U. (2015) Systems strategies for developing industrial microbial strains. *Nat. Biotechnol.* 33, 1061–1072
3. Teng, Y. *et al.* (2024) The expanded CRISPR toolbox for constructing microbial cell factories. *Trends Biotechnol.* 42, 104–118
4. Wang, Y. *et al.* (2018) MACBETH: multiplex automated *Corynebacterium glutamicum* base editing method. *Metab. Eng.* 47, 200–210
5. Liu, J. *et al.* (2017) Development of a CRISPR/Cas9 genome editing toolbox for *Corynebacterium glutamicum*. *Microb. Cell Factories* 16, 205
6. Wang, Y. *et al.* (2021) *In-situ* generation of large numbers of genetic combinations for metabolic reprogramming via CRISPR-guided base editing. *Nat. Commun.* 12, 678
7. Perez-Garcia, F. and Wendisch, V.F. (2018) Transport and metabolic engineering of the cell factory *Corynebacterium glutamicum*. *FEMS Microbiol. Lett.* 365, fry166
8. Guan, N. *et al.* (2017) Microbial response to environmental stresses: from fundamental mechanisms to practical applications. *Appl. Microbiol. Biotechnol.* 101, 3991–4008
9. Choi, K.R. *et al.* (2019) Systems metabolic engineering strategies: integrating systems and synthetic biology with metabolic engineering. *Trends Biotechnol.* 37, 817–837

10. Wendisch, V.F. *et al.* (2016) Updates on industrial production of amino acids using *Corynebacterium glutamicum*. *World J. Microbiol. Biotechnol.* 32, 105
11. Ikeda, M. and Takeno, S. (2020) Recent advances in amino acid production. In *Corynebacterium glutamicum: Biology and Biotechnology* (Inui, M. and Toyoda, K., eds), pp. 175–226, Springer International Publishing
12. Wang, Y. *et al.* (2020) Eliminating the capsule-like layer to promote glucose uptake for hyaluronan production by engineered *Corynebacterium glutamicum*. *Nat. Commun.* 11, 3120
13. Zhan, C. *et al.* (2023) Improved polyketide production in *C. glutamicum* by preventing propionate-induced growth inhibition. *Nat. Metab.* 5, 1127–1140
14. Li, Z. *et al.* (2022) Systems metabolic engineering of *Corynebacterium glutamicum* for high-level production of 1,3-propanediol from glucose and xylose. *Metab. Eng.* 70, 79–88
15. Hoffmann, S.L. *et al.* (2021) Cascaded valorization of brown seaweed to produce L-lysine and value-added products using *Corynebacterium glutamicum* streamlined by systems metabolic engineering. *Metab. Eng.* 67, 293–307
16. Tuyishime, P. *et al.* (2018) Engineering *Corynebacterium glutamicum* for methanol-dependent growth and glutamate production. *Metab. Eng.* 49, 220–231
17. Liu, B. *et al.* (2022) Increased NADPH supply enhances glycolysis metabolic flux and L-methionine production in *Corynebacterium glutamicum*. *Foods* 11, 1031
18. Liu, J. *et al.* (2022) CRISPR-assisted rational flux-tuning and arrayed CRISPRi screening of an L-proline exporter for L-proline hyperproduction. *Nat. Commun.* 13, 891
19. Sheng, Q. *et al.* (2021) Highly efficient biosynthesis of L-ornithine from mannitol by using recombinant *Corynebacterium glutamicum*. *Biosour. Technol.* 327, 124799
20. Hao, Y. *et al.* (2020) High-yield production of L-valine in engineered *Escherichia coli* by a novel two-stage fermentation. *Metab. Eng.* 62, 198–206
21. Wu, H. *et al.* (2020) Highly efficient production of L-histidine from glucose by metabolically engineered *Escherichia coli*. *ACS Synth. Biol.* 9, 1813–1822
22. Ding, X. *et al.* (2023) High-level and -yield production of L-leucine in engineered *Escherichia coli* by multistep metabolic engineering. *Metab. Eng.* 78, 128–136
23. Jiang, S. *et al.* (2023) Metabolic reprogramming and biosensor-assisted mutagenesis screening for high-level production of L-arginine in *Escherichia coli*. *Metab. Eng.* 76, 146–157
24. Cao, M. *et al.* (2024) Efficient L-valine production using systematically metabolic engineered *Klebsiella oxytoca*. *Biosour. Technol.* 395, 130403
25. Ahn, J.H. *et al.* (2020) Enhanced succinic acid production by *Mannheimia* employing optimal malate dehydrogenase. *Nat. Commun.* 11, 1970
26. Sayers, E.W. *et al.* (2023) Database resources of the National Center for Biotechnology Information in 2023. *Nucleic Acids Res.* 51, D29–D38
27. Pfeifer-Sancar, K. *et al.* (2013) Comprehensive analysis of the *Corynebacterium glutamicum* transcriptome using an improved RNAseq technique. *BMC Genomics* 14, 888
28. Fan, L. *et al.* (2021) Transcriptome analysis reveals the roles of nitrogen metabolism and sedoheptulose bisphosphatase pathway in methanol-dependent growth of *Corynebacterium glutamicum*. *Microb. Biotechnol.* 14, 1797–1808
29. Kitagawa, M. *et al.* (2005) Complete set of ORF clones of *Escherichia coli* ASKA library (a complete set of *E. coli* K-12 ORF archive): unique resources for biological research. *DNA Res.* 12, 291–299
30. Gelperin, D.M. *et al.* (2005) Biochemical and genetic analysis of the yeast proteome with a movable ORF collection. *Genes Dev.* 19, 2816–2826
31. Sopko, R. *et al.* (2006) Mapping pathways and phenotypes by systematic gene overexpression. *Mol. Cell* 21, 319–330
32. Matsuyama, A. *et al.* (2006) ORFeome cloning and global analysis of protein localization in the fission yeast *Schizosaccharomyces pombe*. *Nat. Biotechnol.* 24, 841–847
33. Palanisamy, N. *et al.* (2020) C-terminal eYFP fusion impairs *Escherichia coli* MinE function. *Open Biol.* 10, 200010
34. Ferenc-Mrozek, A. *et al.* (2020) Effect of the His-tag location on decapping scavenger enzymes and their hydrolytic activity toward Cap analogs. *ACS Omega* 5, 10759–10766
35. Weill, U. *et al.* (2019) Assessment of GFP tag position on protein localization and growth fitness in yeast. *J. Mol. Biol.* 431, 636–641
36. Wolf, S. *et al.* (2021) Advances in metabolic engineering of *Corynebacterium glutamicum* to produce high-value active ingredients for food, feed, human health, and well-being. *Essays Biochem.* 65, 197–212
37. Wen, X. *et al.* (2021) A CRISPR/dCas9-assisted system to clone toxic genes in *Escherichia coli*. *Biochim. Biophys. Acta Gen. Subj.* 1865, 129994
38. Lopilato, J. *et al.* (1986) Mutations in a new chromosomal gene of *Escherichia coli* K-12, *pcnB*, reduce plasmid copy number of pBR322 and its derivatives. *Mol. Gen. Genet.* 205, 285–290
39. Brocker, M. *et al.* (2009) Citrate utilization by *Corynebacterium glutamicum* is controlled by the CitAB two-component system through positive regulation of the citrate transport genes *citH* and *tctCBA*. *J. Bacteriol.* 191, 3869–3880
40. Frunzke, J. *et al.* (2008) Co-ordinated regulation of gluconate catabolism and glucose uptake in *Corynebacterium glutamicum* by two functionally equivalent transcriptional regulators, GntR1 and GntR2. *Mol. Microbiol.* 67, 305–322
41. Huang, J. *et al.* (2021) Development of a hyperosmotic stress inducible gene expression system by engineering the MtrA/MtrB-dependent NCgl1418 promoter in *Corynebacterium glutamicum*. *Front. Microbiol.* 12, 718511
42. Cai, N. *et al.* (2023) Engineering of the DNA replication and repair machinery to develop binary mutators for rapid genome evolution of *Corynebacterium glutamicum*. *Nucleic Acids Res.* 51, 8623–8642
43. Suzuki, N. *et al.* (2006) High-throughput transposon mutagenesis of *Corynebacterium glutamicum* and construction of a single-gene disruptant mutant library. *Appl. Environ. Microbiol.* 72, 3750–3755
44. Mattosovich, R. *et al.* (2020) O⁶-alkylguanine-DNA alkyltransferases in microbes living on the edge: From stability to applicability. *Int. J. Mol. Sci.* 21, 2878
45. Miggiano, R. *et al.* (2013) Biochemical and structural studies of the *Mycobacterium tuberculosis* O⁶-methylguanine methyltransferase and mutated variants. *J. Bacteriol.* 195, 2728–2736
46. Bremer, E. and Krämer, R. (2019) Responses of microorganisms to osmotic stress. *Ann. Rev. Microbiol.* 73, 313–334
47. Rebeck, G.W. and Samson, L. (1991) Increased spontaneous mutation and alkylation sensitivity of *Escherichia coli* strains lacking the *ogt* O⁶-methylguanine DNA repair methyltransferase. *J. Bacteriol.* 173, 2068–2076
48. Schultz, A.C. *et al.* (2001) Functional analysis of 14 genes that constitute the purine catabolic pathway in *Bacillus subtilis* and evidence for a novel regulon controlled by the PucR transcription activator. *J. Bacteriol.* 183, 3293–3302
49. Huang, S.C. *et al.* (2011) PrcR, a PucR-type transcriptional activator, is essential for proline utilization and mediates proline-responsive expression of the proline utilization operon *putBCP* in *Bacillus subtilis*. *Microbiology* 157, 3370–3377
50. Lin, T.H. *et al.* (2012) AdeR, a PucR-type transcription factor, activates expression of L-alanine dehydrogenase and is required for sporulation of *Bacillus subtilis*. *J. Bacteriol.* 194, 4995–5001
51. Zhu, L. *et al.* (2020) Regulation of γ-aminobutyrate (GABA) utilization in *Corynebacterium glutamicum* by the PucR-type transcriptional regulator GabR and by alternative nitrogen and carbon sources. *Front. Microbiol.* 11, 544045
52. Kelley, L.A. *et al.* (2015) The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* 10, 845–858
53. Galperin, M.Y. *et al.* (2015) Expanded microbial genome coverage and improved protein family annotation in the COG database. *Nucleic Acids Res.* 43, D261–D269
54. Bailey, T.L. and Elkan, C. (1994) Fitting a mixture model by expectation maximization to discover motifs in biopolymers. *Proc. Int. Conf. Intell. Syst. Mol. Biol.* 2, 28–36
55. Maeda, T. *et al.* (2016) RNase III mediated cleavage of the coding region of *mraZ* mRNA is required for efficient cell division in *Corynebacterium glutamicum*. *Mol. Microbiol.* 99, 1149–1166
56. Kim, T.H. *et al.* (2005) The *whcE* gene of *Corynebacterium glutamicum* is important for survival following heat and oxidative stress. *Biochim. Biophys. Res. Commun.* 337, 757–764

57. Choi, W.W. *et al.* (2009) The *whcA* gene plays a negative role in oxidative stress response of *Corynebacterium glutamicum*. *FEMS Microbiol. Lett.* 290, 32–38
58. Valbuena, N. *et al.* (2006) Morphological changes and proteome response of *Corynebacterium glutamicum* to a partial depletion of FtsI. *Microbiology* 152, 2491–2503
59. Bush, M.J. (2018) The actinobacterial WhiB-like (Wbl) family of transcription factors. *Mol. Microbiol.* 110, 663–676
60. Eggeling, L. (2017) Exporters for production of amino acids and other small molecules. *Adv. Biochem. Eng. Biotechnol.* 159, 199–225
61. Radi, M.S. *et al.* (2022) Membrane transporter identification and modulation via adaptive laboratory evolution. *Metab. Eng.* 72, 376–390
62. Simic, P. *et al.* (2001) L-Threonine export: use of peptides to identify a new translocator from *Corynebacterium glutamicum*. *J. Bacteriol.* 183, 5317–5324
63. Zhang, X. *et al.* (2020) High-yield production of L-serine through a novel identified exporter combined with synthetic pathway in *Corynebacterium glutamicum*. *Microb. Cell Factories* 19, 115
64. Ikeda, M. (2012) Sugar transport systems in *Corynebacterium glutamicum*: features and applications to strain development. *Appl. Microbiol. Biotechnol.* 96, 1191–1200
65. Pittman, M.S. *et al.* (2002) Cysteine is exported from the *Escherichia coli* cytoplasm by CydDC, an ATP-binding cassette-type transporter required for cytochrome assembly. *J. Biol. Chem.* 277, 49841–49849
66. Zakataeva, N.P. *et al.* (1999) The novel transmembrane *Escherichia coli* proteins involved in the amino acid efflux. *FEBS Lett.* 452, 228–232
67. Kruse, D. *et al.* (2002) Influence of threonine exporters on threonine production in *Escherichia coli*. *Appl. Microbiol. Biotechnol.* 59, 205–210
68. Liu, Y. *et al.* (2015) Developing a high-throughput screening method for threonine overproduction based on an artificial promoter. *Microb. Cell Factories* 14, 121
69. Liu, J. *et al.* (2024) Reconstruction of the feedback regulation of amino acid metabolism to develop a non-auxotrophic L-threonine producing *Corynebacterium glutamicum*. *Bioresour. Bioprocess.* 11, 43
70. Wang, T. *et al.* (2018) Pooled CRISPR interference screening enables genome-scale functional genomics study in bacteria with superior performance. *Nat. Commun.* 9, 2475
71. Liu, Y. *et al.* (2022) Base editor enables rational genome-scale functional screening for enhanced industrial phenotypes in *Corynebacterium glutamicum*. *Sci. Adv.* 8, eabq2157
72. Mormann, S. *et al.* (2006) Random mutagenesis in *Corynebacterium glutamicum* ATCC 13032 using an IS6100-based transposon vector identified the last unknown gene in the histidine biosynthesis pathway. *BMC Genomics* 7, 205
73. Kim, G.B. *et al.* (2023) Metabolic engineering for sustainability and health. *Trends Biotechnol.* 41, 425–451
74. Reboul, J. *et al.* (2003) *C. elegans* ORFeome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression. *Nat. Genet.* 34, 35–41
75. Jiang, Y. *et al.* (2017) CRISPR-Cpf1 assisted genome editing of *Corynebacterium glutamicum*. *Nat. Commun.* 8, 15179
76. Nesvera, J. *et al.* (1997) Plasmid pGA1 from *Corynebacterium glutamicum* codes for a gene product that positively influences plasmid copy number. *J. Bacteriol.* 179, 1525–1532
77. Robert, X. and Gouet, P. (2014) Deciphering key features in protein structures with the new ENDscript server. *Nucleic Acids Res.* 42, W320–W324
78. Jumper, J. *et al.* (2021) Highly accurate protein structure prediction with AlphaFold. *Nature* 596, 583–589
79. Keilhauer, C. *et al.* (1993) Isoleucine synthesis in *Corynebacterium glutamicum*: molecular analysis of the *ilvB*-*ilvN*-*ilvC* operon. *J. Bacteriol.* 175, 5595–5603
80. Ruan, Y. *et al.* (2015) Improving the electro-transformation efficiency of *Corynebacterium glutamicum* by weakening its cell wall and increasing the cytoplasmic membrane fluidity. *Biotechnol. Lett.* 37, 2445–2452
81. Bae, S. *et al.* (2014) Cas-OFFinder: a fast and versatile algorithm that searches for potential off-target sites of Cas9 RNA-guided endonucleases. *Bioinformatics* 30, 1473–1475
82. Zhang, C. *et al.* (2020) Metabolic engineering of an auto-regulated *Corynebacterium glutamicum* chassis for biosynthesis of 5-aminolevulinic acid. *Bioresour. Technol.* 318, 124064
83. Ruffert, S. *et al.* (1997) Efflux of compatible solutes in *Corynebacterium glutamicum* mediated by osmoregulated channel activity. *Eur. J. Biochem.* 247, 572–580

STAR★METHODS

KEY RESOURCES TABLE

Reagent or resource	Source	Identifier
Bacterial and virus strains		
<i>Escherichia coli</i> DH5α	TaKaRa	D9057
<i>E. coli</i> DB3.1	TransGen Biotech	CD531
<i>Corynebacterium glutamicum</i> ATCC 13032	ATCC	13032
Chemicals, peptides, and recombinant proteins		
Thr-Thr dipeptide	Hefei-specific peptide organisms	N/A
BacTiter-Glo™ Microbial Cell Viability Assay	Promega	G8230
Wizard Genomic DNA Purification Kit	Promega	A1125
ClonExpress MultiS One-Step Cloning Kit	Vazyme	C112
T4 DNA ligase	New England Biolabs	M0202S
Restriction endonucleases Bsal	New England Biolabs	R3733S
2×TransStart FastPfu SuperMix	TransGen Biotech	AS221-02
HT DNA Extender Range LabChip	Perkin Elmer	760517
DNA 12K reagent kit	Perkin Elmer	760569
ClonExpress II One Step Cloning kit	Vazyme	C112
MagBeads Plasmid DNA Mini Extraction Kit	Biomiga	BW-MPD1211
2 × Es Taq Mix	CWBIO	CW0718M

METHOD DETAILS

Strains and culture conditions

Strains used in this study, excluding the 3,049 *C. glutamicum* strains of the gene overexpression collection, are listed in Data S1 in the supplemental information online. *E. coli* strains DH5α and DB3.1 used for plasmid cloning were cultivated aerobically at 37°C in Luria–Bertani (LB) broth. Ampicillin (100 µg/ml), kanamycin (50 µg/ml), or chloramphenicol (20 µg/ml) was added to the medium as required. *C. glutamicum* ATCC 13032 and the derivatives were cultivated aerobically at 30°C in TSB medium or CGXII-YE medium. The TBS medium [18] contains 5 g/l glucose, 9 g/l soya peptone, 5 g/l yeast extract, 1 g/l K₂HPO₄·3H₂O, 0.1 g/l MgSO₄·7H₂O, 3 g/l urea, 0.5 g/l succinic acid, 10 µg/l biotin, 100 µg/l vitamin B1, and 20 g/l MOPS (pH 7.2). The CGXII-YE medium is the CGXII minimal medium [79] supplemented with 2 g/l yeast extract. Kanamycin (25 µg/ml), chloramphenicol (5 µg/ml), or IPTG (0.005 to 0.32 mM) was added as required.

Genome resequencing of *C. glutamicum* ATCC 13032

To identify the differences in the genome sequence of the *C. glutamicum* ATCC 13032 copy stored in our laboratory and the one deposited in NCBI, the genomic DNAs were extracted using Wizard Genomic DNA Purification Kit [Promega (Beijing) Biotech Co., China]. Library construction and genome sequencing were performed by Biomarker Technologies (Beijing, China) by using Illumina and Oxford Nanopore sequencing platforms. Quality assurance of the output was analyzed by using FastQC software (v.0.10.1) and NGSQC Toolkit software (v.2.3.3). BWA alignment software (v.0.7.17), and SAM tools software (v.1.9) were used for alignment and variant calling, respectively. Variations were annotated by using the SnpEff software (v.4.3i).

Plasmid construction

Plasmids used in this study, excluding the 3049 plasmids of the gene overexpression collection, are listed in Data S2 in the supplemental information online. Plasmids were constructed via recombination or Golden Gate assembly. Recombination was conducted

using the ClonExpress MultiS One-Step Cloning Kit (Vazyme, Nanjing, China). Restriction endonucleases and T4 DNA ligase used for Golden Gate assembly were purchased from New England Biolabs (Beijing) (Beijing, China). DNA polymerase and reagents used for PCR were purchased from TransGen Biotech (Beijing, China). Services of primer and gene synthesis and DNA sequencing were provided by GENEWIZ Inc. (Suzhou, China). Primers and details for constructing plasmids are described in Data S3 in the supplemental information online.

Automated construction of gene overexpression collection

Automated construction of *C. glutamicum* gene overexpression collection was performed on the BioFoundry TianGong I. The procedure consists of two parts, plasmid assembly in *E. coli* DH5α and plasmid transformation to *C. glutamicum*. BioFoundry TianGong I integrates various instruments to facilitate the automation of PCR reaction, DNA fragment analysis, DNA purification, plasmid ligation, plasmid transformation, plating, colony picking, cell cultivation, plasmid extraction, etc. (Table S1 in the supplemental information online). The PCR system was reduced to 5 µl per well, consisting of 2×TransStart FastPfu SuperMix (TransGen Biotech, China) 2.5 µl, genomic DNA (~50 ng/µl) 0.05 µl, nuclease-free water 2.25 µl, and 0.1 µl each of 10 µM forward and reverse primers. First, all components, except for the primers, were prepared as a PCR master mix in a reservoir and dispensed into a 384-well PCR plate (ThermoFisher Scientific, USA) by Biomek i7 Automated Liquid Handler (Beckman Coulter, USA). Next, the 384-well PCR plate was transferred to Echo Acoustic Liquid Handler (Beckman Coulter, USA) via selective compliance assembly robot arm (SCARA), where the forward and reverse primers synthesized and stored in the 384-well polypropylene plate (Beckman Coulter, USA) were added to the 384-well PCR plate. The 384-well PCR plate was then sent to the sealer followed by an automated thermal cycler (ATC) using SCARA. After 30 cycles of PCR reaction, 3 µl PCR product per well was aspirated into a new 384-well PCR plate, and the reaction mix for plasmid ligation stored in the reservoir was dispensed into the new 384-well PCR plate by Biomek i7 Automated Liquid Handler. The total volume of the ligation reaction was reduced to 5 µl, containing Exnase (Vazyme, China) 0.5 µl, 5×reaction buffer 1 µl, plasmid backbone (~80 ng/µl) 0.5 µl, PCR product 3 µl. All the components except for the PCR product were pre-made as a reaction mix. After incubating at 37°C for 30 min, the plate was sent back to Biomek i7 Automated Liquid Handler, and the entire ligation product was transferred to a pre-prepared 96-well microplate containing 50 µl *E. coli* DH5α competent cells per well. After incubation at 4°C for 5 min, all the cells were plated on twelve 8-well QTrays (Molecular Devices, USA) containing LB agar medium with ampicillin (100 µg/ml) by Microbial Colony Picker QPix420 (Molecular Devices, USA). All QTrays were then moved to Cytomat 2 C450-LiN Automated Incubator (ThermoFisher Scientific, USA) by SCARA for overnight cultivation at 37°C. On the second day, two colonies were picked from each well of the plates using Microbial Colony Picker QPix420 and inoculated in 96-deep-well plates containing 1 ml LB medium with ampicillin (100 µg/ml) in each well. After overnight incubation at 37°C, 100 µl culture was aspirated out for Sanger sequencing. When sequencing results were incorrect, 100 µl culture from the backup plate was aspirated out for Sanger sequencing. If more than two Sanger sequencing reactions were required to go through the target gene, the plasmids were mixed and sent for NGS. After verification by plasmid sequencing, *E. coli* strains with the correct plasmid were stored in an Arktic Automated Sample Storage System (SPT Labtech, United Kingdom) with 15% glycerol.

The *E. coli* strains with the correct plasmid were cultivated for extracting plasmid for transformation into *C. glutamicum*. Plasmid extraction was carried out using the MagBeads Plasmid DNA Mini Extraction Kit (Biomiga, USA) with the Biomek i7 Automated Liquid Handler, centrifuge, Kingfisher Presto (ThermoFisher Scientific, USA), and SCARA. In the second and third rounds of plasmid construction, PCR products were analyzed by electrophoresis using the LabChip GXII with the HT DNA Extender Range LabChip 760517 (Perkin Elmer, USA) and the DNA 12K reagent kit (Perkin Elmer, USA). PCR products were diluted 20-fold in 384-well PCR plates and then loaded into the LabChip GXII for analysis according to the instructions. Product length, DNA concentration (Conc., given in ng/µl), and product purity were analyzed using LabChip GX Touch Software v.1.11.144.0.

All successfully constructed plasmids were transformed into *C. glutamicum* by electroporation with the BioFoundry. After electroporation, cells were recovered at 30°C for 1 h and then spread on 8-well QTrays containing LBHIS agar medium [80] with kanamycin (25 µg/ml) by Microbial Colony Picker QPix420. All QTrays were then moved to Cytomat 2 C450-LiN Automated Incubator by SCARA for overnight cultivation at 30°C. After 2 days of incubation, two colonies were picked from each well of the plates using Microbial Colony Picker QPix420 and inoculated in a 96-deep-well plate containing 1 ml LBHIS medium with kanamycin (25 µg/ml) in each well. After 4 hours of cultivation, PCR verification was performed with a 5 µl reaction system consisting of 2.5 µl 2 × Es Taq Mix (CWBIOD, China), 1 µl culture, 0.25 µl forward primer, 0.25 µl reverse primer, and 1 µl nuclease-free water. The reaction mix containing 2 × Es Taq Mix and nuclease-free water was prepared as a PCR mix in a reagent reservoir and was dispensed into a 384-well

PCR plate by Biomek i7 Automated Liquid Handler. Cultures and primers were then added to the PCR plates by Biomek i7 Automated Liquid Handler and Echo Acoustic Liquid Handler, respectively. Once verified, the recombinant *C. glutamicum* strains were stored in an Arktic Automated Sample Storage System with 15% glycerol.

High-throughput growth characterization of gene overexpression collection

MicroScreen HT automatic microbial growth curve analyzer (Gering Instrument Ltd., Tianjin, China) and 48-deep-well plates with a capacity of 4 ml for each well were used for the high-throughput growth characterization of gene overexpression collection. Preparation of the seed culture and cultivation were performed with the abovementioned equipment. OD₆₀₀ values of each well were measured online every one or two hours as required. The strains with gene overexpression were cultivated in 400 µl TSB medium supplemented with 25 µg/ml kanamycin. After cultivation at 30°C and 800 rpm for 8 h, 10 µl of the culture was used as the seed to inoculate 400 µl CGXII-YE medium supplemented with 0.04 mM IPTG, 25 µg/ml kanamycin, and 1.2 M L-lysine sulfate (or 1.2 M Na₂SO₄). The pH of the medium was adjusted to 7.2. The strains were cultivated at 30°C and 800 rpm, and OD₆₀₀ values were monitored every hour. Cultivation of the control strain *C. glutamicum* ATCC13032 (pEC) was conducted for every 48-deep-well plate. For verifying the effects of selected genes on cell growth and tolerance to hyperosmotic stress, strains were cultivated under the same conditions as the first round of screening. At least two biological replicates were performed and OD₆₀₀ values were monitored every two hours.

Calculation of gene fitness

Gene fitness was used to describe the difference in cell growth with or without stress conditions caused by gene overexpression. Gene fitness was calculated according to the procedure and equations shown in Figure S5a. For the first round of screening of gene overexpression collection under hyperosmotic stress, no replicates were performed and OD₆₀₀ values were monitored every hour. Continuous fifteen data points (≥ 30 h) that showed the biggest difference between the gene overexpression strain and the control strain were used for calculating the gene fitness. For verifying the effects of selected genes on cell growth and tolerance to hyperosmotic stress, at least two biological replicates were performed and OD₆₀₀ values were monitored every two hours. To calculate the gene fitness for hyperosmotic stress condition, continuous eight data points (≥ 30 h, mean of biological replicates, $P < 0.05$, Student's two-tailed *t*-test) that showed the biggest difference between the gene overexpression strain and the control strain were used. Without hyperosmotic stress, cells grew much faster compared with those with hyperosmotic stress and commonly entered the stationary phase after twenty hours of cultivation. Therefore, to calculate the gene fitness for normal conditions, continuous eight data points (≥ 4 h, mean of biological replicates, $P < 0.05$, Student's two-tailed *t*-test) that showed the biggest difference between the gene overexpression strain and the control strain were used.

Chromosomal engineering of *C. glutamicum*

Gene deletion and insertion in *C. glutamicum* were performed using the developed CRISPR/Cas9 system [18]. The plasmids used for gene deletion and insertion are listed in Data S10. The plasmid was transformed into *C. glutamicum* by electroporation. Preparation of competent cells and electroporation were conducted following the reported procedure [80]. After electroporation, 1 ml TSB medium was added immediately and the suspension was quickly transferred into a 1.5 ml tube and incubated at 46°C for 6 min. After recovery at 30°C for 3 h, cells were spread on a TSB solid medium supplemented with 0.05 mM IPTG and 5 µg/ml chloramphenicol for selection. After 48 h of cultivation at 30°C, colonies were verified by PCR. Correctly edited strains were cultivated at 37°C with antibiotic-free TSB medium to cure the plasmid. The gRNA design and off-target analysis were performed using the online tool CRISPR RGEN Tools [81]. Primers used for plasmid construction and strain verification are listed in Data S3.

L-Lysine production by strains overexpressing selected genes

An L-lysine-producing *C. glutamicum* strain harboring three mutations (*lysC*^{T311I}, *pyc*^{P458S}, and *hom*^{V59A}) was used as the host to evaluate the effects of gene overexpression on L-lysine production. The strains transformed with the gene overexpression plasmid and the control harboring an empty plasmid were cultivated in 24-deep-well plates containing 800 µl TSB medium each well, which was supplemented with 25 µg/ml kanamycin. After cultivation at 30°C and 800 rpm for 8 h, the culture was used as a seed to inoculate 800 µl fermentation medium in a 24-deep-well plate with an initial OD₆₀₀ of 0.1. The fermentation medium contains 60 g/l glucose, 1 g/l yeast extract, 1 g/l soya peptone, 1 g/l NaCl, 1 g/l (NH₄)₂SO₄, 8 g/l urea, 1 g/l K₂HPO₄·3H₂O, 0.45 g/l MgSO₄·7H₂O, 0.05 g/l FeSO₄·7H₂O, 0.4 mg/l biotin, 0.1 mg/l vitamin B1, and 40 g/l MOPS (pH7.2). The fermentation medium was supplemented with 0.04 mM IPTG, 25 µg/ml kanamycin, and 1.0 M Na₂SO₄. The strains were cultivated at 30°C and

800 rpm for 120 h. Extracellular L-lysine levels were measured by HPLC [18]. Glucose was quantified using an SBA-40D biosensor analyzer equipped with glucose oxidase (Institute of Biology of Shandong Province Academy of Sciences, China).

Detection of m6G-induced mutation and Cgl3003-mediated m6G repair

A reporter gene *kan*^{ACG} (kanamycin resistance gene with the first codon changed from ATG to ACG) was inserted into the chromosome between *cgl2830* and *cgl2831* using the CRISPR/Cas9 system. Since the kanamycin resistance was used as a reporting phenotype, the *kan* gene in the plasmid pEC-*cgl3003* was replaced with a chloramphenicol resistance gene *cm*. The resulting plasmid pEC-*cm-cgl3003* and the control plasmid pEC-*cm* were transformed into *C. glutamicum* ATCC 13032 *kan*^{ACG}. The engineered strains were cultivated in 100 ml shake flasks containing 5 ml TSB medium supplemented with 5 µg/ml chloramphenicol. After cultivation at 30°C and 220 rpm for 8 h, the cultures were used as seeds to inoculate 30 ml CGXII-YE medium in 500 ml shake flasks with an initial OD₆₀₀ of 0.1. The medium was supplemented with 0.04 mM IPTG and 5 µg/ml chloramphenicol. To provide a hyperosmotic condition, 1.2 M L-lysine sulfate was added. The strains were cultivated at 30°C and 220 rpm for 96 h. The same amount of cells were collected and spread on plates with 25 µg/ml kanamycin. Colonies were counted after 48 hours of cultivation. The *kan*^{ACG} genes of the kanamycin-resistant colonies were amplified by PCR and sequenced to analyze mutations.

Transcriptome analysis

C. glutamicum ATCC 13032 (pEC-*cgl2496*) and *C. glutamicum* ATCC 13032 (pEC) were cultivated in a TSB medium supplemented with 25 µg/ml kanamycin. The cultures were used as seeds to inoculate 48-deep-well plates containing 400 µl CGXII-YE medium in each well with an initial OD₆₀₀ of 0.1. The medium was supplemented with 0.04 mM IPTG, 25 µg/ml kanamycin, and 1.2 M L-lysine sulfate. Strains were cultivated at 30°C and with shaking at 800 rpm. Cells at 36 h and 82 h were collected for RNA isolation. RNA preparation, library construction, sequencing on Illumina HiSeq, and data processing were performed by Novogene (Tianjin, China). Genes with a false discovery rate (FDR)<0.05 and log₂(Fold change)>1 or ≤1 were considered to be differentially expressed.

ChIP-Seq analysis

3xFlag was fused to the N-terminal or C-terminal of Cgl2496. The recombinant plasmids pEC-N-3Flag-*cgl2496* and pEC-C-3Flag-*cgl2496* were transformed into *C. glutamicum* ATCC 13032. The transformants were characterized by their tolerance to hyperosmotic stress. Since the strain *C. glutamicum* ATCC 13032 (pEC-N-3Flag-*cgl2496*) showed similar tolerance to hyperosmotic stress to *C. glutamicum* ATCC 13032 (pEC-*cgl2496*), the fusion protein with 3xFlag fused to the N-terminal of Cgl2496 was used for ChIP-Seq analysis to identify the binding motif of Cgl2496. *C. glutamicum* ATCC 13032 (pEC-N-3Flag-*cgl2496*) and *C. glutamicum* ATCC 13032 (pEC) grown under the hyperosmotic stress were collected at the logarithmic phase. ChIP-Seq analysis and data processing were conducted by Wuhan Igenebook Biotechnology (Wuhan, China). The antibody used for immunoprecipitation was anti-3xFlag (Huabio, Hangzhou, China), and ChIP-enriched DNAs were sequenced using the Illumina HiSeq 2000 sequencing system. Trimmomatic (version 0.36) was used to filter out low-quality reads. Clean reads were mapped to the genome of *C. glutamicum* ATCC 13032 by Bwa (version 0.7.15). Samtools (version 1.3.1) was used to remove potential PCR duplicates. MACS2 software (version 2.1.1.20160309) was used to call peaks by default parameters (bandwidth, 300 bp; model fold, 5, 50; *q* value, 0.05). If the midpoint of a peak is located closest to the transcriptional start site of one gene, the peak will be assigned to that gene. HOMER (version 3) was used to predict motif occurrence within peaks with default settings.

SEM of *C. glutamicum*

C. glutamicum ATCC 13032 (pEC-*cgl2496*) and *C. glutamicum* ATCC 13032 (pEC) strains grown with 1.2 M Na₂SO₄ were collected at the logarithmic phase. Cells were washed with Na₂SO₄ isotonic medium containing 1 g/l KH₂PO₄, 1.3 g/l K₂HPO₄·3H₂O, 42 g/l MOPS, and 1.2 M Na₂SO₄ and then fixed in 2.5% glutaraldehyde tissue fixative (prepared with the Na₂SO₄ isotonic medium) and 1% osmium tetroxide. Then, cells were washed with sodium sulfate hypertonic buffer, dehydrated in ethanol, dried at a critical point, and plated by the SU8010 Ultra-High Resolution Scanning Electron Microscope FE-SEM (Hitachi, Japan).

Measurement of intracellular ATP levels

Cells of *C. glutamicum* ATCC 13032 (pEC-*cgl2496*) and *C. glutamicum* ATCC 13032 (pEC) were collected at the logarithmic phase. The intracellular ATP levels were measured using the BacTiter-Glo™ Microbial Cell Viability Assay Kit, which has been used for the determination of intracellular ATP levels of *C. glutamicum* [82]. The measurement was performed following the manufacturer's instructions.

Screening of L-threonine exporters

From the *C. glutamicum* ATCC 13032 gene overexpression collection, 390 strains overexpressing membrane transporter genes were selected for screening of L-threonine exporters. The first round of screening was performed based on the principle that strains overexpressing functional L-threonine exporters could tolerate the high intracellular L-threonine concentration caused by the addition of Thr–Thr dipeptide. The seeds were cultivated in 48-deep-well plates with 400 µl TSB medium supplemented with 25 µg/ml kanamycin. After cultivation at 30°C and 800 rpm for 8 h, the seed cultures were used to inoculate 48-deep-well plates containing 400 µL TSB medium supplemented with 0.04 mM IPTG and 25 µg/ml kanamycin. After cultivation at 30°C and 800 rpm for 7 h, the cultures were used to inoculate 48-deep-well plates containing 400 µL modified CGXII minimal medium (20 g/l glucose, 4 g/l (NH₄)₂SO₄, and 1 g/l urea) without Thr–Thr or with 6 mM Thr–Thr dipeptide. The modified CGXII medium was supplemented with 25 µg/ml kanamycin and 0.4 mM IPTG. The strains were incubated at 30°C and with shaking at 800 rpm using the MicroScreen HT automatic microbial growth curve analyzer (Gering Instrument Ltd., Tianjin, China), and the OD₆₀₀ values were monitored online every 1 h. Gene fitness was used to describe the difference in cell growth under dipeptide stress conditions caused by transporter overexpression. Gene fitness was calculated according to the procedure and equations shown in Figure S13 in the supplemental information online. The data of the fifth to ninth-time points that showed the biggest difference between the gene overexpression strain and the control strain were used for calculating the gene fitness.

After the first round of screening, 33 transporters that enhanced cell growth under Thr–Thr dipeptide stress compared with the negative control strain were subjected to a second round of screening by quantifying the final concentration of extracellular L-threonine when they were cultivated with the addition of Thr–Thr dipeptide. The 33 strains with transporter overexpression were cultivated with the same conditions as the first round of screening. After cultivation at 30°C and 800 rpm for 7 h, cells were harvested by centrifugation at 5,000 × g for 10 min and resuspended in modified CGXII minimal medium (20 g/l glucose, 4 g/l (NH₄)₂SO₄, and 1 g/l urea) supplemented with 0.4 mM IPTG, 25 µg/ml kanamycin, and 6 mM Thr–Thr dipeptide. After incubation at 30°C and with shaking at 800 rpm for 1 h, the cultures were sampled and extracellular L-threonine levels were measured by HPLC [18].

Peptide uptake and amino acid export assay

After two rounds of screening, six transporters were selected for further characterization. *C. glutamicum* ATCC 13032 strains overexpressing the transporters and harboring the pEC empty plasmid (the negative control) were cultivated in 100 ml shake flasks containing 10 ml TSB medium supplemented with 25 µg/ml kanamycin. After cultivation at 30°C and 220 rpm for 8 h, the seed cultures were used to inoculate 10 ml TSB medium supplemented with 0.04 mM IPTG and 25 µg/ml kanamycin. After cultivation at 30°C and 800 rpm for 7 h, cells were collected by centrifugation at 5,000 × g for 10 min and resuspended in modified CGXII minimal medium (20 g/l glucose, 4 g/l (NH₄)₂SO₄, and 1 g/l urea) supplemented with 0.4 mM IPTG, 25 µg/ml kanamycin, and 6 mM Thr–Thr dipeptide. The OD₆₀₀ values were set at 10. Cells were incubated at 30°C and 220 rpm for 2 h and samples were taken every 20 min. Intracellular and extracellular L-threonine concentrations were quantified using HPLC [18]. The intracellular volume used to calculate the internal amino acid concentration was 1.7 µl/mg DCW [83].

L-Threonine production by *C. glutamicum* overexpressing selected genes

An L-threonine-producing *C. glutamicum* ATCC 13032 strain harboring two mutations (*hom*^{G378E} and *lysC*^{T311I}) and overexpressing *hom*^{G378E} and *thrB* via plasmid was used as the host to evaluate the effects of transporter overexpression on L-threonine production. The transporter overexpression plasmids were transformed into the L-threonine-producing strain. The recombinant strains were cultivated in 24-deep-well plates containing 800 µL TSB medium supplemented with 25 µg/ml kanamycin and 5 µg/ml chloramphenicol. After cultivation at 30°C and 800 rpm for 8 h, the seed cultures were used to inoculate 24-deep-well plates containing 800 µL fermentation medium with an initial OD₆₀₀ of 0.06. The fermentation medium contains 80 g/l glucose, 1 g/l yeast extract, 1 g/l soya peptone, 1 g/l NaCl, 1 g/l (NH₄)₂SO₄, 8 g/l urea, 1 g/l K₂HPO₄·3H₂O, 0.45 g/l MgSO₄·7H₂O, 0.05 g/l FeSO₄·7H₂O, 0.4 mg/l biotin, 0.1 mg/l vitamin B1, and 40 g/l MOPS (pH7.2). The fermentation medium was supplemented with 0.4 mM IPTG, 25 µg/ml kanamycin, and 5 µg/ml chloramphenicol. The 24-deep-well plates were cultivated at 30°C and 800 rpm for 20 h. L-Threonine and L-homoserine in the fermentation broth were quantified using HPLC [18].

To test the application of Cgl2344 and Cgl2656 in industry-level L-threonine production, a *C. glutamicum* strain ZKCgW3 was used as the host. To improve the expression levels of *cgl2344* and *cgl2656*, their initiation codon was changed from GTG to ATG. *cgl2344*^{ATG} and *cgl2656*^{ATG} were used to replace the *rhtC* on the previously reported plasmid overexpressing L-threonine transporter and *thrB* [69]. The recombinant strains were cultivated in 24-deep-well plates and fermentation products were determined as described in

the last paragraph, except that only 25 µg/ml kanamycin was added and chloramphenicol was not supplemented. Fed-batch fermentation in a 5 L bioreactor was performed using the same procedure described previously for a direct comparison between the previous and present study [69].

L-Threonine production by *E. coli* overexpressing selected genes

An overexpression plasmid pS-*thrA*^{E253K}BC previously reported was introduced into the wide-type *E. coli* MG1655 strain [68], and the resulting strain EcThr was used as the host to evaluate the effects of transporter overexpression on L-threonine production. The transporter overexpression plasmids were transformed into strain EcThr. The recombinant strains were cultivated in 24-deep-well plates containing 800 µL LB medium supplemented with 25 µg/ml kanamycin and 25 µg/ml streptomycin. After overnight incubation at 37°C and 800 rpm, the seed cultures were used to inoculate 24-deep-well plates containing 800 µL fermentation medium with an initial OD₆₀₀ of 0.06. The fermentation medium contains 30 g/l glucose, 3 g/l yeast extract, 10 g/l (NH₄)₂SO₄, 2 g/l KH₂PO₄·3H₂O, 0.5 g/l MgSO₄·7H₂O, 0.01 g/l FeSO₄·7H₂O, 0.01 g/l MnSO₄·5H₂O and 83.6 g/l MOPS (pH 7.0). The fermentation medium was supplemented with 0.4 mM IPTG, 25 µg/ml kanamycin, and 25 µg/ml streptomycin. The 24-deep-well plates were cultivated at 37°C and 800 rpm for 24 h. L-Threonine and L-homoserine in the fermentation broth were quantified using HPLC [18].