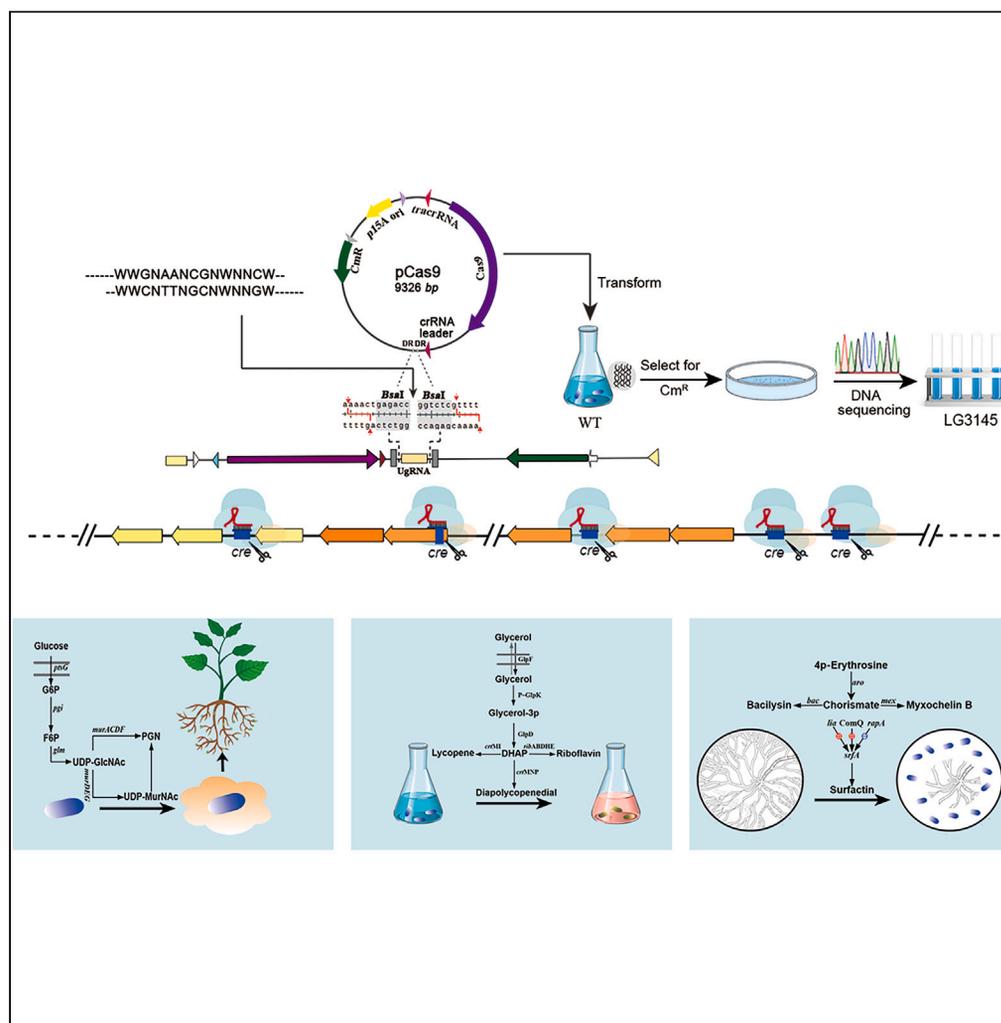


Article

Genome-scale *cis*-acting catabolite-responsive element editing confers *Bacillus pumilus* LG3145 plant-beneficial functions

Meiyang Bi,
Mingkun Li, Jiaxun
Wei, ..., Ming Ying,
Xiurong Yang, Lei
Huang

ym@tjut.edu.cn (M.Y.)
xiurongyang88@126.com (X.Y.)
huanglei@tjut.edu.cn (L.H.)

Highlights

B. pumilus acquired
disease resistance via non-
specific *cis*-acting element
editing

Cis-acting element editing
caused carbon source
utilization changes in *B.*
pumilus

Cis-acting element editing
allowed *B. pumilus* to
promote plant growth

Cre-like sequence
mutation induced *B.*
pumilus activity against
plant pathogens

Article

Genome-scale *cis*-acting catabolite-responsive element editing confers *Bacillus pumilus* LG3145 plant-beneficial functionsMeiying Bi,^{1,3} Mingkun Li,^{1,3} Jiaxun Wei,¹ Ziwen Meng,¹ Zhaoyang Wang,¹ Ming Ying,^{1,4,*} Xiurong Yang,^{2,*} and Lei Huang^{1,*}

SUMMARY

Rhizosphere dwelling microorganism such as *Bacillus* spp. are helpful for crop growth. However, these functions are adversely affected by long-term synthetic fertilizer application. We developed a modified CRISPR/Cas9 system using non-specific single-guide RNAs to disrupt the genome-wide *cis*-acting catabolite-responsive elements (*cres*) in a wild-type *Bacillus pumilus* strain, which conferred dual plant-benefit properties. Most of the mutations occurred around imperfectly matched *cis*-acting elements (*cre*-like sites) in genes that are mainly involved in carbon and secondary metabolism pathways. The comparative metabolomics and transcriptome results revealed that carbon is likely transferred to some pigments, such as riboflavin, carotenoid, and lycopene, or non-ribosomal peptides, such as siderophore, surfactin, myxochelin, and bacilysin, through the pentose phosphate and amino acid metabolism pathways. Collectively, these findings suggested that the mutation of global *cre*-like sequences in the genome might alter carbon flow, thereby allowing beneficial biological interactions between the rhizobacteria and plants.

INTRODUCTION

Bacillus strains play an essential role in maintaining soil health and promoting plant growth, and are considered as one of the plant growth-promoting rhizobacteria (PGPR). With accelerated agricultural development, long-term synthetic fertilizer application to soil always adversely affects PGPR functions, such as loss of antibiotic production, antifungal activity, and root colonization traits.¹ We isolated a wild-type *Bacillus pumilus* strain from intensively cultivated soil that had lost its PGPR potential. We aimed to develop a method to improve its PGPR ability, which would be important for sustainable agriculture with reduced use of chemical fertilizers.

Gene manipulation techniques have been used to restore PGPR potential. For example, the integration of desired genes, such as the isolation of the chitinase-encoding gene *chi113* from *Bacillus subtilis* Ap113 and its expression in *Burkholderia vietnamiensis*, which induced significant biocontrol effects on wheat sheath blight and tomato gray mold.² Genome reduction approaches have also been used, e.g., using double crossover homologous recombination to knock out the 2,3-butanediol dehydrogenase-encoding gene *bdh* in *Bacillus velezensis* GJ11, constructed in a mutant strain with a high level of acetoin to trigger induced systemic resistance against *Pseudomonas syringae* infection in plants.³ However, conventional methods based on homologous recombination to construct engineered *Bacillus* strains remain time-consuming, require a relatively high transformation efficiency, and only focus on a single PGPR attribute.⁴ Recent studies have achieved some breakthroughs using the clustered regularly interspaced short palindromic repeat (CRISPR)/CRISPR associated protein 9 (Cas9) system to edit and regulate multiple genes simultaneously.^{5,6}

However, there is a lack of studies reporting the success of CRISPR/Cas9-mediated multiple gene editing of *Bacillus* rhizobacteria for agriculture application. Rhizobacteria have advantages in terms of root colonization, which is dependent on their preference of rhizodeposits secreted by plants, such as organic acids, proteins, amino acids, and sugars.⁷ Moreover, plant growth, stress tolerance, and disease resistance can be enhanced by the secretion of certain secondary metabolites, such as siderophore, bacilysin, and surfactin. We speculated that gram-positive rhizobacteria in the soil, such as *Bacillus* species, might change their preferred carbon source because of their domestication under the effects of long-term chemical-fertilizer farming. Thus, they are subject to carbon catabolite repression (CCR) of many genes involved in secondary metabolism, which is a global mechanism of catabolite repression in the genus *Bacillus*. The presence of a rapidly metabolizable carbon source in the growth medium inhibits the synthesis of catabolic enzymes for other carbon sources. The mechanism of catabolite repression is mainly exerted through a “*cis*-acting catabolite-responsive element (*cres*)” in the promoter regions of the target genes.^{8,9}

¹Tianjin Key Laboratory of Organic Solar Cells and Photochemical Conversion, School of Chemistry and Chemical Engineering, Tianjin University of Technology, Tianjin 300384, People's Republic of China

²Institute of Plant Protection, Tianjin Academy of Agricultural Sciences, Tianjin 300384, People's Republic of China

³These authors contributed equally

⁴Lead contact

*Correspondence: ym@tjut.edu.cn (M.Y.), xiurongyang88@126.com (X.Y.), huanglei@tjut.edu.cn (L.H.)

<https://doi.org/10.1016/j.isci.2024.108983>



The *cre* site was first observed in *Bacillus subtilis* 168 in 1987 by Nicholson in the promoter region of the amylase gene (*amyE*) and has since been found in many operon promoters in the genome of the *Bacillus* genus, with the consensus sequence “WWGNAANCGNWNNCW”.^{10,11} Sometimes, the integration of only PGPR-associated genes into a heterogeneous host might not yield the desired results because of the original complex cascade regulatory system of their promoters. Mutation or replacement of the related part of the promoter might be an effective approach to enhance the function of these genes. Herein, we aimed to use the CRISPR/Cas9 technology to globally interfere with *cre*-like sites in the genome of a *Bacillus pumilus* strain, with the hope of restoring its agricultural function.

We searched for *cre*-like sequences using the previous *cre* site consensus sequence and identified 4465 potential sites in the genome. CRISPR/Cas9-mediated multiple gene editing using a non-specific single-guide RNA (UgRNA) was used to edit the genome-wide *cre*-like sites. Previously, we found seven target genes with mutations at their *cre* sites that resulted in the mutant strain exhibiting a carbon catabolite repression (CCR)-release phenotype, such as secreting a large number of secondary metabolites.^{12,13} Here, we continue to report the editing results occurring in other regions of the genome, and used omics analysis of the mutant strain to reveal the connection between the plant-beneficial functions and *cre*-like site editing.

RESULTS AND DISCUSSION

Differentially expressed genes (DEGs) resulting from global *cre*s perturbation

In our previous study, we designed three UgRNAs targeting the *cre* sites of seven genes, namely an acetate kinase gene (*ackA*), an amino-transferase gene (*ntdA*), a phosphocarrier protein HPr gene (*ptsH*), an acetolactate decarboxylase gene (*budA*), an acetyl-CoA synthetase gene (*acsA*), a succinate-CoA ligase gene (*sucC*), and a xylose isomerase gene (*xylA*), using 30–70% sequence similarity. These UgRNAs successfully induced mutations at their *cre* sites in *B. pumilus* LG3145. Compared with the wild-type (WT), the mutant strain exhibited a clear CCR release phenotype, as evidenced by the secretion of extracellular pigments, proteins, and the formation of polysaccharide capsules. Single nucleotide polymorphism (SNP) analysis revealed that the non-specific gRNA/Cas9 system (UgRNA/Cas9) edited 3370 genes, primarily enriched in the biosynthesis of secondary metabolites, regulatory and transporter systems, and carbon and amino acid metabolism pathways (Data S1). We isolated total RNA from the WT and LG3145 strains cultivated using different carbon sources for transcriptome analysis. The expression levels of approximately 3400 effective genes in *B. pumilus* were analyzed using a threshold Fragments Per Kilobase of transcript per Million mapped reads (FPKM) > 1. According to Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis, these genes were mainly enriched in carbohydrate, lipid, amino acid, and secondary metabolism pathways, and the membrane transport system. A comparison with the WT grown using glycerol as the carbon source (GLY) showed a greater number of upregulated genes than downregulated genes, especially in carbohydrate metabolism (Figure 1A). For example, there were 17 upregulated and 7 downregulated genes in the tricarboxylic acid (TCA) cycle (ko00020), 27 upregulated and 16 downregulated genes in glycolysis (ko00010), 27 upregulated and 19 downregulated genes in pyruvate metabolism (ko00620), and 24 upregulated and 9 downregulated genes in amino sugar metabolism (ko00520). These results suggested that the metabolic flux of carbohydrates and amino acid metabolism is likely to increase when strain LG3145 uses glycerol as a carbon source and produces more secondary metabolites. Interestingly, the expression levels of all the 15 genes in overflow metabolism (ko00640) declined, which might reduce acid production and redirect carbon sources to the TCA cycle or pentose phosphate pathway (PPP). In amino acid metabolism, 35 genes were upregulated and only 9 were downregulated in the cysteine and methionine synthesis pathway (ko00270), and 24 were upregulated and only 12 were downregulated in glycine, serine, and threonine metabolism (ko00260). In secondary metabolism, the unique biosynthesis pathway of siderophore group nonribosomal peptides (ko01053) had six upregulated genes, and the biosynthesis of various antibiotics pathway (ko00998) had seven upregulated genes, suggesting that strain LG3145 might produce more antibacterial substances (Data S2). Moreover, there was no significant difference in membrane transporter, carbohydrate metabolism, and amino acid metabolism while using glucose (GLU) or GLY as carbon sources (Figure 1B). These results indicated that the mutant strain LG3145 can grow normally and produce secondary metabolites using GLY as a carbon source. Furthermore, the fatty acid (FA) (ko00061) and carotenoid biosynthesis (ko00906) pathways uniquely increased, with 12 and 4 upregulated genes, respectively, which might cause strain LG3145 to produce pigments (Data S3). Finally, the transcriptome analysis data obtained in triplicate were aggregated and depicted in Figure 1C. Notably, the expression levels of 1926 medium-related genes in strain LG3145 cultivated in GLY were comparable to those of glucose-grown cells, but were significantly higher than those of the WT strain grown in GLY. Conversely, only 60 genes were not expressed in strain LG3145 grown in GLY. These results indicated that mutant strain LG3145 has a greater capacity to utilize GLY as a carbon source, and exhibited gene expression levels equivalent to those observed when GLU was used as the carbon source.

To understand the relationship between *cre* UgRNA/Cas9 editing and the phenotypic changes of strain LG3145, we selected some rate-limiting enzyme encoding genes or operons for DNA sequencing analysis (Figure S1). DNA sequencing of 35 genes or operons showed that more than 50 *cre*-like sites were edited by the UgRNAs/Cas9 system, including deletion, insertion, transversion, and transition mutations (Figure S2; Table 1). In this paper, the *cre*-like sequences with mismatches to the query sequence on the anti-sense and sense strands are given as the first and second numbers or with an underline in brackets, e.g., [0,0] means a perfect match; [0,0] means a perfect match to *cre* on the sense strand, [1,0] means the *cre* is same as the transcription direction of the target gene with only one mismatch; and [0,1] means the *cre* is opposite to the transcription direction with one mismatch. Most mutation sites were the [1,0], [2,0], and [0,2] types of *cre*-like sequences, and few of them occurred as *cre* [0,0], being perfectly matched with the consensus, except for *ackA* and *sucC* among the seven previous target genes and the *glpFKD* operon involved in GLY metabolism. These results suggested that *cre*-like sequences with one or two mismatch sites are the most edited targets by the UgRNA/Cas9 system.

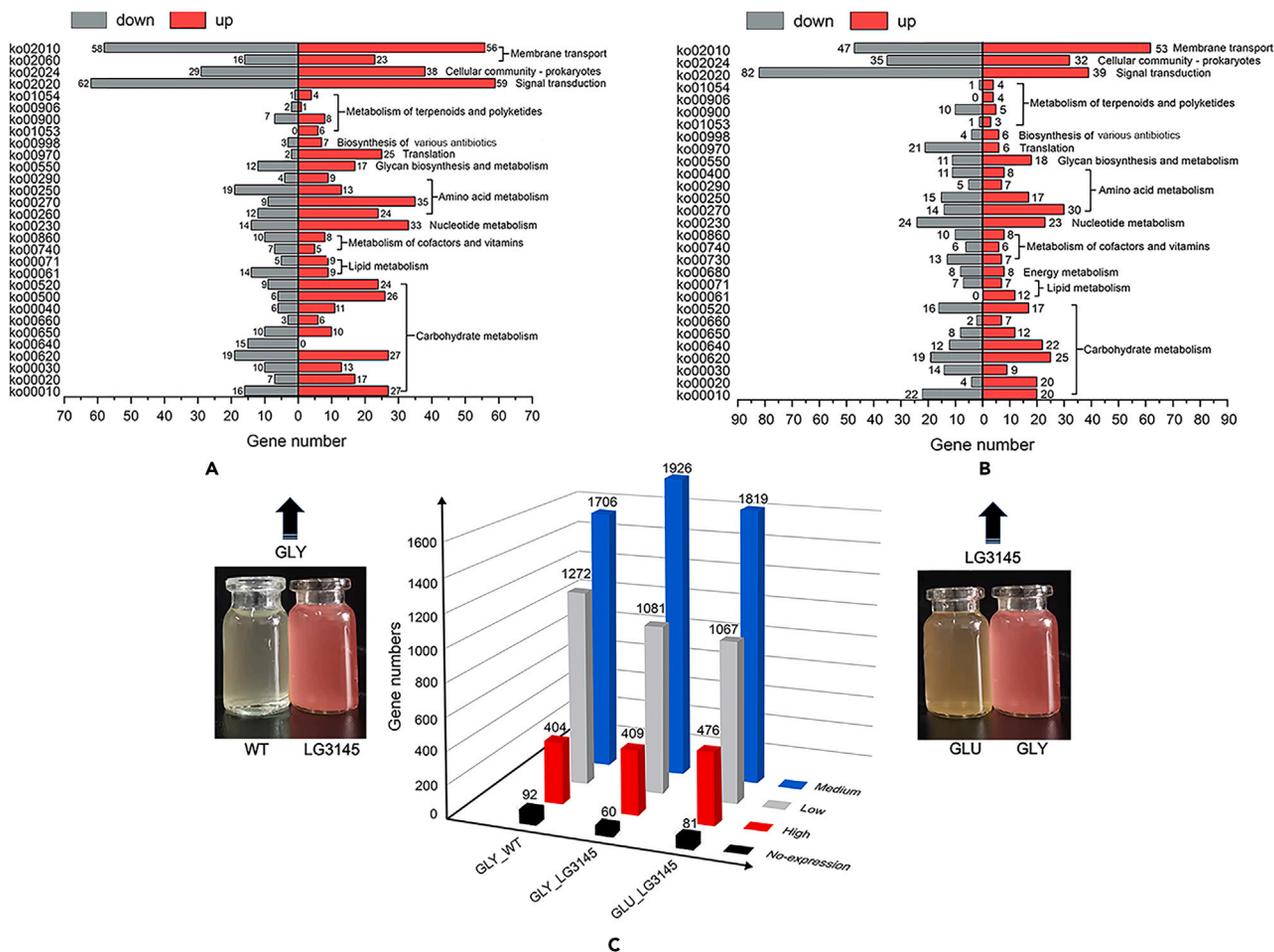


Figure 1. Transcriptomic and comparative transcriptomic feature of strain LG3145

(A) Major differences in metabolic pathways of strain LG3145 compared with the wild-type strain (WT) cultivated in glycerol minimal medium (GYMM) for 12 h. (B) Major differences in metabolic pathways of strain LG3145 in GYMM and in glucose minimal medium (GUMM) cultivated for 12 h. (C) Global abundance of the gene expression of the WT and LG3145 strains throughout the cultivation period in GYMM (GLY_WT, GLY_LG3145) and GUMM (GLU_LG3145). The transcripts were assessed based on Fragments Per Kilobase of transcript per Million mapped reads (FPKM) values: high expression (FPKM \geq 900), medium expression (500 \leq FPKM < 900), low expression (50 \leq FPKM < 500), and non-expression (1 < FPKM < 50). Down, downregulated; Up, upregulated.

Metabolic alterations caused by DEGs

As previously discussed, several *cre*-like sites located upstream of genes involved in key pathways related to metabolism were precisely edited using UgRNA/Cas9. To further investigate the underlying mechanism, we employed ultra-performance liquid chromatography-mass spectrometry (UPLC-MS) to analyze the metabolites of *B. pumilus* LG3145 cultured on GLY as a carbon source. The differential metabolites between the edited strain and the WT strain were identified and annotated for enrichment using KEGG, as depicted in Figure 2A and Data S4. Notably, the significantly altered metabolites (p value < 0.05, $|\log_2$ fold-change [FC]| \geq 1) were found to be concentrated in three metabolic pathways: biosynthesis of siderophore group nonribosomal peptides (ko01053), riboflavin metabolism (ko00740), and lysine biosynthesis (ko00300), which was consistent with the differential gene enrichment data depicted in Figure 2B and marked by the reverse red arrow. Thus, UgRNA/Cas9 editing significantly affected these three metabolic pathways. Additionally, as shown in Figure 2A, porphyrin metabolism (ko00860) and peptidoglycan biosynthesis (ko00550) were identified as significantly altered pathways. Notably, the gene numbers of their precursor synthesis pathways, aminoacyl-tRNA biosynthesis (ko00970) and vitamin B6 metabolism (ko00970; depicted by blue arrows in the same direction), and amino sugar metabolism (ko00520; depicted by green arrows), respectively, were also significantly altered (Figure 2B and Data S5).

The volcano plots in Figures 2C and 2D depict a comparison of metabolites and genes between the WT and mutant strain LG3145. Among them, there were 259 upregulated and 277 downregulated metabolites, as well as 505 upregulated genes and 465 downregulated genes (p value < 0.05, $|\log_2$ fold-change [FC]| \geq 1). Notably, we examined the names of the most significantly different metabolites and genes, and noted that some of them were linked. For example, the overexpression of a glycerol kinase encoding gene, *glpK*, shown in Figure 2D might explain the presence of (S)-3-hydroxytetradecanoyl-CoA (C05260) in Figure 2C, while the overexpression of a surfactin synthetase encoding gene, *srfAA*, shown in Figure 2D, might be responsible for the presence of surfactin (C12043) in Figure 2C. Expression of the operon

Table 1. The prediction of *cre*s in some operon in *B. pumilus* WT

Gene or operon	<i>cre</i> -like and flanking sequence ^{a,b}	TxSS ^c	Gene or operon	<i>cre</i> -like and flanking sequence ^{a,b}	TxSS ^c
<i>ackA</i>	TATCG ATGTAAGCGGTAACA ^[0,0] GTTCA	-52	<i>sucC</i>	TTAAA ATGAAAGCGCAGTCT ^[0,0] ATTTT	-6
<i>glpFKD</i>	AGAGG ATGAAAACGTTGTCA ^[0,0] ATAAA	-190	<i>glk</i>	CAATT TTTAAAGCTGAAACA ^[2,0] ATATA	-330
	ATTTA AAGAAATC [*] ACTCGCA ^[1,0] TATCA	81	<i>glmS</i>	TGTTG ATGCAAACGTAGCA [*] T ^[1,0] CTTCA	-259
<i>ptsGHI</i>	TAAct ATGAAAT [*] GATGTCA ^[1,0] GGGAG	-66	<i>glmM</i>	CTCTT TAGGG [*] ACT [*] GCTCTTT ^[3,0] TTCTA	-191
<i>ykgB</i>	AAATA T [*] G [*] CAAACGGTTTCT ^[0,2] CAAAT	-148	<i>glmU</i>	ACAGC ATTTAAA [*] GAAA [*] ACA ^[0,2] TCCAT	-157
<i>crr</i>	CGTGT ATGAAAC [*] G [*] ATACGCT ^[0,2] CATAc	-45	<i>murA</i>	CAAGG AAGAAAT [*] GCTGACT ^[0,1] GCAGA	-189
<i>fsa</i>	ATCTT [*] G [*] AAGCGCAATCA ^[1,0] AATTT	-59		TGCAa TAGCAATGGTAATGA ^[0,2] GCATC	-214
	<i>fbaA</i>	AGGTT TTGCAAAG [*] TGTGTCA ^[0,2] AAGCA	-261	<i>murC</i>	CAATT TTGGAAT [*] GGACGCT ^[1,0] TTTcG
	ATCCA AAGA [*] ACTCGGTGCTT ^[2,0] TGACA	-276	<i>mraY/murD</i>	CCATC AA [*] AAAAT [*] TGcATTCA ^[0,2] TTATT	-56
<i>eno/gpm/tpi/pgk</i>	AAATA AA [*] TAAACCGTTCACA ^[1,0] GTTTT	211	<i>murE</i>	AGCCA TTGTAG [*] C [*] AGAACACC ^[3,0] CTCTT	193
<i>gap/cggR</i>	AAGAG TTGAAAAT [*] GAAAG [*] AT ^[2,0] TATCC	-130	<i>murGB</i>	TTGTT TAGAAAACG [*] T [*] AT ^[2,0] TTGAC	-110
	CATGG ATGAAT [*] C [*] ATTCA ^[3,0] ATAGA	-255		TTATT TTGCAA [*] AGTATA ^[2,0] AAAAT	-161
<i>tkT</i>	GATAA TACA [*] AAACGGTCTCT ^[0,1] GTTAA	171		TGCTG AAGCTAA [*] TGAAACCA ^[2,0] TGTTG	-190
<i>pdhABCD</i>	ATGTT TCGACA [*] ACGATCCCA ^[2,0] GTGGC	-127	<i>pyk/pfkA/accAD</i>	GACAT TTGCAAAT [*] GAATGCA ^[1,0] AAACA	189
<i>pgi</i>	GGAAC TAGCTAT [*] T GAAATCA ^[0,2] AGTAA	-67	<i>srfA</i>	GGAGG ATGAAAAT [*] GAATACA ^[1,0] GTTTA	-6
	GGCTT TT [*] AAAATCGCTAAA ^[0,2] TGTCa	-155	<i>bacABCDF</i>	ACACA TT [*] GAAACGTAAGCT ^[1,0] CCATT	-161
<i>guaB</i>	AAATA AA [*] TAAACCGTTCACA ^[1,0] GTTTT	-259	<i>bacG</i>	AAACA AAGGAACCGGTGCT ^[0,0] CTTTT	189
<i>tenA</i>	GTCGT ATGTAAG [*] T [*] CCTCACT ^[0,2] AATCG	183		AGGGA TTGGAA [*] AGGCAATCG ^[2,0] CCATA	70
<i>ribA</i>	CGACT [*] CAGGAAGCGCAGTGA ^[0,2] TGGTG	-83		GTGTA [*] CGATGGCC [*] CATGAT ^[0,2] CCGCT	-235
<i>ribBD</i>	ATAAA AAGGCAA [*] AGAA [*] CGCT ^[0,2] CAAAG	58	<i>aroA</i>	CATTG TACCA [*] ACTCAT [*] ACA ^[2,0] TATAA	-68
<i>crtNIM</i>	AACGA TAC [*] GAAACGCAAAA ^[2,0] GGATC	389	<i>aroBC</i>	GAGAT TTGAAGCGTT [*] CG ^[2,0] GCTGC	126
	TCGTC TTGAAAAT [*] CAATCT ^[0,2] ATACA	273		GCGCA AATGAAACGAAG [*] ATT ^[2,0] AACGT	77
	AATGT AAGAAAG [*] CGGCCCT ^[2,0] TTATT	-145	<i>aroG</i>	GATCA ATGAAAG [*] AGGAAACA ^[1,0] TTGTA	73
<i>crtP</i>	AAGAA TTTAAAC [*] AGATCA ^[2,0] AGGAC	-71	<i>liaGFRS</i>	AAACT AAGAAAT [*] CAGAGACA ^[0,1] GGTCG	-113
	AAGGA TATGACCG [*] TATGCT ^[2,0] TGGCG	-165	<i>liaHI</i>	ATCCA AT [*] AAAA [*] AGCACTCT ^[0,2] TCAAT	-143
Consensus	WWGNAANCGNWNNCW (W:A /T, N:A/T/G/C)				

^aThe sequences shown in this table are *cre*-like sequences searched by SnapGene with the query sequence WWGNAANCGNWNNCW (W:A /T, N:A/T/G/C) in the genome of *B. pumilus* SH-B9, which are classified using the modified Ogiwara *cre*s search method.

^bThe *cre*-like sequences with the mismatch numbers to the query sequence on the anti-sense and sense strands are given as the first and second numbers or with an underline in brackets. e.g. [0,0] means perfect match; [0,0] means a perfect match to *cre* on the sense strand, [1,0] means the *cre* is same as the transcription direction of target gene with only one mismatch; and [0,1] means opposite to the transcription direction with one mismatch.

^cTxSS means the distance of each *cre* from the transcription start site (TSS).

liaFGHI, which belongs to the two-component system (ko02020), was significantly increased, which is associated with daptomycin resistance (a structural analog of surfactin). These results suggested that the bacterium has potential antagonistic functions against high surfactin concentrations.¹⁴ Additionally, the reduced expression of the operon *thiDFGO* and *tenAI*, which are thiamine biosynthesis genes (ko02020), indicates that mutant strain LG3145 produces less thiamine and other related substances after absorbing nutrients (Datas S6 and S7).

To verify the transcriptional profiling data obtained from RNA sequencing (RNA-Seq) (<https://edu.majorbio.com>), we performed quantitative real-time reverse transcription PCR (real-time qPCR) analysis on the original RNA extracts (Figure 2E and Data S8). Although there were some differences in the fold-changes for several DEGs as determined by real-time qPCR and RNA-Seq, the general trends were consistent, thus supporting the credibility of the RNA-Seq data. These differences likely resulted from the use of different methods, which has been observed previously.¹⁵ Among the upregulated genes confirmed by real-time qPCR, *ptsH* and *glpDK* are involved in phosphocarrier protein HPr and glycerol metabolism, respectively. Furthermore, *murA*, which is a critical gene in the generation of UDP-N-acetylmuramic Acid (MurNAc), the precursor of peptidoglycan, is associated with dihydrostreptomycin 6-phosphate (C01221) (Figure 2C). The genes *guaB* and *ribA* are involved in the synthesis of riboflavin, and *crtI* is responsible for generating lycopene. Finally, *srfA* is the surfactin synthase gene, which was significantly upregulated in bacteria grown in GLY minimal medium (GYMM) ($|\log_2FC| = 3.09$). Among the downregulated genes verified by real-time qPCR, *tenA*, which encodes thiaminase in the thiamine metabolism pathway (ko00730), was significantly downregulated (Figure 2D).

Based on the analysis of the volcano plots, we generated two heat maps to visualize the expression levels of some significant compounds and genes (p value <0.05) in strain LG3145 cultivated in different media, as shown in Figure 3 (Datas S9 and S10). The heatmap of genes in

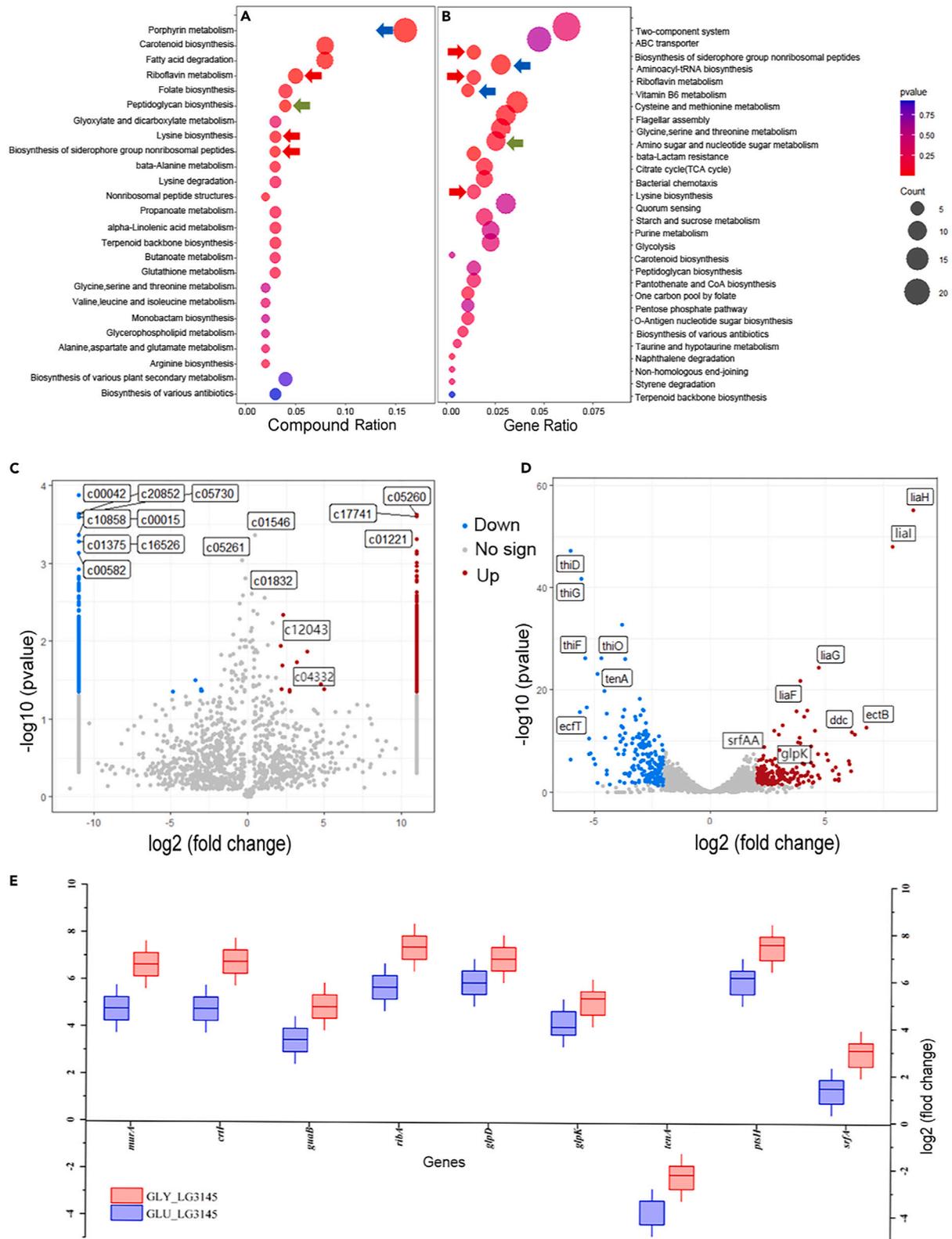


Figure 2. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment pathways analysis of altered metabolites and genes when strain LG3145 used glycerol as carbon source, and comparative analysis of differential metabolites and differentially expressed genes (DEGs) when strain LG3145 used glycerol and glucose as carbon sources

(A) Enriched metabolite pathways.
 (B) Enriched gene pathways. The bubble size represents the number of enrichment factors in the pathway. A bubble color change from purple to red indicates greater statistical significance. The red reverse arrow represents the same enrichment pathways. The blue and green arrows in the same direction represent associated metabolic pathways.
 (C) Differential metabolites analysis of strain LG3145 grown in glycerol (GLY) minimal medium (GYMM) (GLY_LG3145) and in glucose minimal medium (GUMM) (GLU_LG3145) using a volcano chart.
 (D) DEGs analysis of strain LG3145 grown in GYMM and GUMM using a volcano chart. Red dots represent upregulation. Blue dots represent downregulation.
 (E) Validation of RNA sequencing data (RNA)-Seq data for selected genes using quantitative real-time reverse transcription PCR (real-time qPCR). Data analysis was performed using the comparative $2^{-\Delta\Delta CT}$ method (CT, cycle threshold), with 16S rRNA serving as the internal control. Each group contains data for three independent samples. All data analysis was based on fold-changes in transcript levels and metabolite contents relative to the wild-type (WT) grown in GYMM (GLY_WT). Down, downregulated; Up, upregulated.

strain LG3145 cultivated in GYMM is more prominently red in Figure 3B, and the corresponding metabolites show similar patterns in Figure 3A. For example, the levels of precursors of riboflavin production, such as R-5-P (C00199), ribitol (C00474), and reduced riboflavin (C01007), were significantly higher in GYMM, which correspond to the upregulated genes *ribABDFH*, *ycsE*, and *nfsA*. Additionally, phytyfluene (C05414) and hydroxy-gamma-carotene (C15859) were also shown to be significantly increased in Figure 3A and correspond to the significantly upregulated genes *ispADEFHG*, *crtl*, *dxr*, and *dxs* in Figure 3B. These findings suggested that when strain LG3145 uses GLY as a carbon source, the expression levels of genes and corresponding metabolites are significantly higher than those of WT, and are equal to or higher than those when using GLU as a carbon source after UgRNA/Cas9 editing. These significant genes and metabolites are mainly involved in GLY transport and metabolism, purine metabolism, PPP, riboflavin, carotenoid, peptidoglycan (PGN), and surfactin synthesis pathways (p value <0.05), indicating that *B. pumilus* LG3145 has the potential to produce riboflavin, carotenoids, antibiotics, peptidoglycan, and other related compounds.

Acquiring wheat growth-driver ability

A pot experiment using wheat was conducted to investigate the agricultural functions of strain LG3145 (khaki powder) and WT (white powder) suspensions applied via root irrigation (RI)¹⁶ (Figure S3A). We observed a significant increase in plant height of 1.31 and 1.15-fold compared with the WT and non-inoculated wheat, respectively, when plants were subjected to strain LG3145-RI treatment after the seedling stage (Figure S3B). Moreover, strain LG3145 colonization significantly promoted vegetative growth during the three-leaf stage and enhanced the fresh weight of wheat shoots by 1.06-fold and the length of roots by 1.05-fold (Figures 4A and S3C; Table S3).¹⁷ These findings suggested that the metabolic changes in strain LG3145 after UgRNA/Cas9 editing are potentially related to restoration of its plant-beneficial functions.

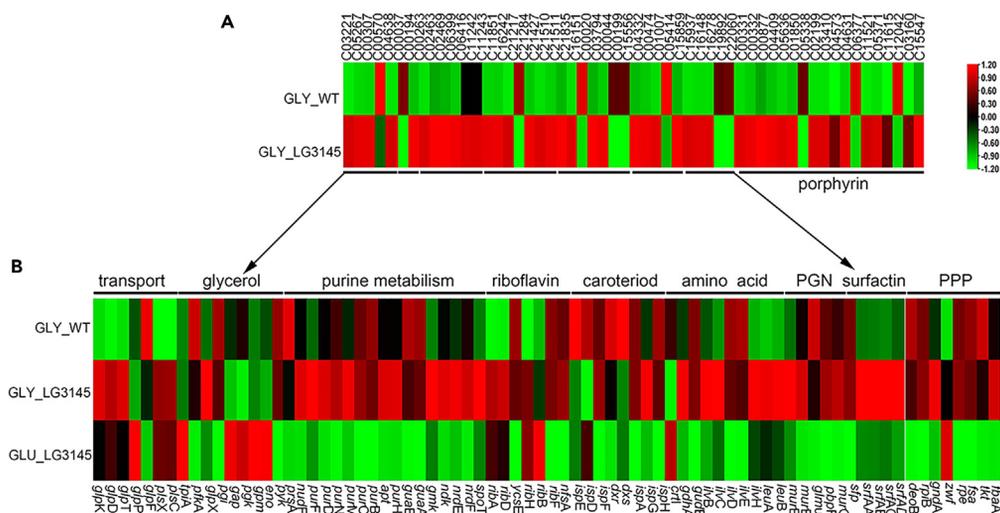


Figure 3. Schematic representation of the differentially expressed genes (DEGs) corresponding to differential metabolites when strain LG3145 was grown using glycerol as the carbon source

(A) Differential metabolite comparative analysis of strain LG3145 and the wild-type (WT) grown using glycerol (GLY) as the carbon source (Data S9).
 (B) DEGs between strain LG3145 grown in GLY minimal medium (GYMM), strain LG3145 grown in glucose minimal medium (GUMM), and the WT grown in GYMM (Data S10). The metabolites and genes with a p value <0.05 and $|\log_2 \text{fold-change [FC]}| \geq 1$ were determined as significantly different using Student's t test. GLY_LG3145 and GLU_LG3145 represent strain LG3145 using glycerol and glucose as a carbon source, respectively. GLY_WT represents the wild-type strain cultivated in GYMM. PGN, peptidoglycan; PPP, pentose phosphate pathway. Gene abbreviations are shown in Data S10.

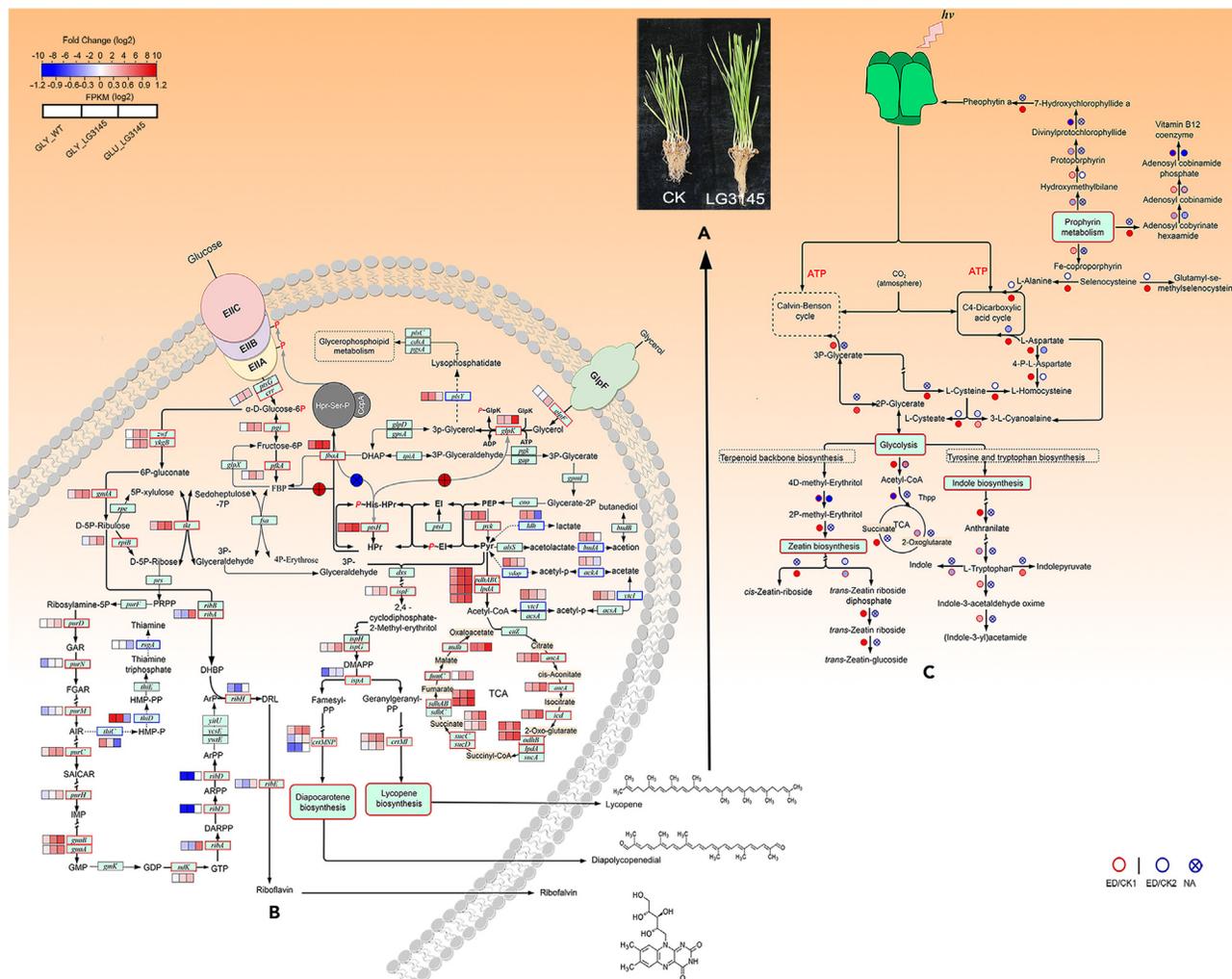


Figure 4. Metabolic pathway analysis of strain LG3145 in plant growth promotion

(A) Effect of strain LG3145-root irrigation (RI) treatment on wheat at the seedling stage with non-inoculated wheat as the control (CK).

(B) The metabolic network of pigments produced by strain LG3145 was reconstructed based on Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. Transcripts of strain LG3145 in glycerol minimal medium (GYMM) (GLY_LG3145) and glucose minimal medium (GUMM) (GLU_ LG3145) and WT in GYMM (GLY_WT) are shown near the pathway as a heatmap, based on normalized data of log₂ (Fragments Per Kilobase of transcript per Million mapped reads (FPKM)) using TBtools software.

(C) The metabolite pathway network of wheat growth under strain LG3145 stress shown as a compound heatmap, based on fold-changes in content relative to WT-RI and non-inoculated wheat. The color legend is shown at the bottom of the map. Gene abbreviations are shown in [Datas S11](#) and [S12](#).

We manually reconstructed a partial metabolic network integrating the transcriptomic analysis of strain LG3145 and the wheat metabolome, as shown in [Figures 4B](#) and [4C](#). Transcriptomic data of strain LG3145 grown in GYMM revealed that the FPKM values of the *glp* operon, which includes the glycerol transporter gene (*glpF*), the glycerol kinase gene (*glpK*), and the glycerol-3-phosphate dehydrogenase gene (*glpD*), increased from 217.8–5760.79 to 648.24–6213.75, suggesting enhanced GLY transport and metabolism. DNA sequencing results showed five mutation sites in the *glp* operon, including two transversions mutations, *T/A* and *A/C*, and three transition mutations with one *C/T* and two *G/A*, around the *cre* [0,0]–190 and *cre* [1,0]–81 sites. The expression level of the phosphotransferase system (PTS) operon, *ptsGHI*, encoding PTS transporter subunits, the HPR protein, and PTS system enzyme I, also increased from 44.10–1527.95 to 132.04–3160.40. Sequencing of the *ptsGHI* operon revealed two *cre*-like sites at –31 and –66, with some transitions of internal *C/T*, *T/G*, and *C/T* and some mutations surrounding them. Another PTS glucose transporter gene, *crr*, which showed an improved FPKM value from 199.19 to 556.27, contained *cre* [2,0]–330, with two transversions and two transitions. Overexpression of these three membrane transporter genes will enhance strain LG3145's ability to absorb non-favorable carbon sources, such as GLY. For example, the *cre* site mutation of *ptsH* resulted in overexpression of the Hpr protein, and the amount of Hpr-His-p and Hpr-Ser-p increased simultaneously. The *cre*-like site mutations in the *glp* operon increased the expression of transporter gene (*glpF*), the glycerol kinase gene (*glpD*), and the more active phospho (p)-GlpK was

induced by Hpr-P-His. Thus, more GLY enters the cell and is converted into glycerol-3p, which enters the glycolytic pathway after being converted into dihydroxyacetone phosphate (DHAP). The PTS transport system has a similar regulatory mechanism, enabling other non-favorable carbon sources to enter carbohydrate metabolism, where they provide carbon skeletons and energy for subsequent metabolism. There are abundant non-preferred carbons in soil, such as unsaturated FAs.¹⁸ After relieving CCR effect on the transport system, these carbon sources can be fully utilized and enter glycolysis, the TCA cycle, and the PPP.

In glycolysis, mutations were observed in the fructose bisphosphate aldolase gene (*fbaA*) with two C/T transitions at the *cre* [0,2]_–261 site and one T/A mutation at the *cre* [2,0]_–276 point. The *eno* operon, which includes the enolase gene (*eno*), the phosphoglycerate mutase gene (*gpm*), the triose-phosphate isomerase gene (*tpiA*), and the phosphoglycerate kinase gene (*pgk*), has a *cre* [3,0]_ 211 upstream of *pgk*, with a G/A transition inside and three mutations downstream: C/A, C/T, and T/C. Transcriptomic analysis revealed upregulation of most glycolysis genes, such as *fbaA*, *pgk*, and *pyk*, resulting in increased carbon flow into other carbon catabolism pathways. Specifically, the expression levels of *fbaA*, *pgk*, and *pyk* increased significantly from 1415.99 to 3092.21, 659.73 to 1413.97, and 281.54 to 754.26, respectively. In the TCA cycle, the pyruvate dehydrogenase operon (*pdhABC*), containing a *cre*-like site at *cre* [2,0]_127, has an internal C/T transition. The FPKM of *pdhABC* increased from 892.22–1223.90 to 3166.93–3933.69, providing more building precursors for cellular metabolism, such as pyruvate, 3P-glyceraldehyde, phosphoribosyl pyrophosphate (PRPP), and riboflavin-5-phosphate (R-5-P), which are used as construction precursors in purine metabolism and lycopene biosynthesis, respectively. In the PPP, fructose bisphosphate (FBP) is generated by DHAP through the FBP aldolase gene (*fbaA*), which has two *cre*-like sites *cre* [0,2]_–261 and *cre* [2,0]_–276 with three transversions and two transitions around them, resulting in an increased FPKM value of 1415.99–3092.21. FBP is converted to 4-P-erythrose by the transaldolase encoded by *fsa*, which has a point mutation at *cre* [1,0]_–59, resulting in an increased FPKM value of 1987.12–2999.40. Additionally, R-5-P is produced by the transketolase gene (*tkt*), which has a large number of mutations between or around two *cre*-like sites *cre* [2,0]_–212 and *cre* [0,1]_171, including five transitions, three transversions, and three insertions. The FPKM of *tkt* increased from 515.38 to 1764.62, leading to the production of more R-5-P and more PRPP as precursors for purine metabolism. Pyruvate and 3P-glyceraldehyde are used as substrates to generate (1S,2R)-D-erythro-2-(N-myrystoylamino)-1-phenyl-1-propanol (DMAPP) through the terpenoid backbone biosynthesis metabolic pathway under the action of the *ispFGA* gene, the FPKM of which was significantly increased 48.05–333.10 to 164.85–960.12, respectively. DMAPP enters the metabolic pathway of carotenoid biosynthesis and generates lycopene under the effect of the *crt* operon, in which the phytoene desaturase gene, *crtP*, has a large number of mutations between *cre* [2,0]_–71 and *cre* [2,0]_–165, including one transversion C/G and two transitions CG/TA inside the *cre*-like sites. The FPKM of *crtP* increased from 39.31 to 257.90. The *crtNIM* operon, with three *cre*-like sites *cre* [2,0]_–145, *cre* [0,2]_273, and *cre* [2,0]_–389, has many mutations, including ten transitions, six transversions, and one insertion mutation, resulting in an increased FPKM value of 104.31–449.95 to 363.05–1346.83. For the other metabolic pathway, PRPP enters purine metabolism under the regulation of the *pur* gene family, whose expression levels were significantly upregulated from 85.39 to 455.21 to 129.33–988.48. The *guaB* gene, responsible for the conversion of IMP to GMP, has a transition G/A mutation located upstream of the *cre*-like site *cre* [1,0]_–159, and its FPKM value increased from 443.89 to 2572.01. R-5-P and GTP, serving as precursors for riboflavin synthesis, enter the riboflavin metabolic pathway in a 1:2 ratio, and undergo seven reaction steps to generate riboflavin. The riboflavin synthesis pathway is regulated by the *ribABDHE* gene family, which plays a crucial role in controlling the rate of riboflavin production. DNA sequencing analysis revealed that the *ribBD* operon contains a *cre* [0,2]_58 with one transition G/A mutation and three transitions, four transversions, and one insertion mutation in its vicinity. In addition, the *ribA* gene harbors a *cre* [0,2]_–83 with a consecutive mutation of four bases AGTA/TCCT inside the *cre*-like site, as well as two deletion mutations nearby. The FPKM values of *ribA*, *ribB*, and *ribD* genes increased significantly from 130.01 to 1460.14, 122.27 to 281.50, and 16.26 to 236.30, respectively (Figure S2 and Data S11).

In the metabolic network of the wheat metabolism data presented in Figure 4C, we observed a significant increase in the components of porphyrin metabolism (ko00860), including hydroxymethylbilane (C01024), protoporphyrin (C02191), divinylprotoporphyrin (C11831), 7-hydroxychlorophyllide (C16540), and phytylphytyl a (C05797) (Data S12). Porphyrin is a fundamental molecular structural component of chlorophyll, which plays a pivotal role in plant photosynthesis. Therefore, as shown in Figure 4A, strain LG3145-RI-treated seedlings showed a higher biomass than that of the non-inoculated wheat. Furthermore, levels of metabolites such as 3P-glycerate, L-aspartate, and L-alanine in the Calvin-Benson and C4-dicarboxylic acid cycle were significantly increased, indicating that strain LG3145-RI-treated plants have an enhanced capacity for CO₂ absorption and fixation. In the strain LG3145-RI treatment group, levels of amino acids such as Cys and Asp were significantly higher than in the WT-RI treatment group and non-inoculated plants. These amino acids enter into tyrosine and tryptophan biosynthesis (ko00400), resulting in the production of indole complexes, such as indolepyruvate (C00331), (indol-3-yl) acetamide (C02693), and indole-3-acetaldehyde oxime (C02937), which are natural auxins in plants, promoting shoot tip formation.¹⁹ In addition, strain LG3145-RI treatment led to a significant improvement in zeatin biosynthesis (ko00908) and terpenoid backbone biosynthesis (ko00900), resulting in higher chlorophyll d levels of a variety of zeatin complexes, such as *cis*-zeatin riboside (C16449), *trans*-zeatin riboside (C16431), and *trans*-zeatin-glucoside (C16443). Zeatin is a member of the cytokinin family, which comprises a class of plant hormones involved in plant growth and development. The ability of strain LG3145 to promote plant growth might be derived from its capacity to produce more pigments and utilize non-preferred carbon in the soil, thereby stimulating plants to synthesize more porphyrins, zeatin, indole, and other compounds.

B. *pumilus* LG3145 confers plant pathogen resistance

In soilless wheat cultivation experiments, *B. pumilus* LG3145-RI treatment significantly reduced the damage to the wheat shoot and root architecture, and showed high biocontrol efficacy against *Fusarium graminearum*, as demonstrated in Figure 5A (Data S13).²⁰ The root and shoot lengths of strain LG3145-treated plants were longer and healthier than those of non-inoculated wheat, measuring approximately

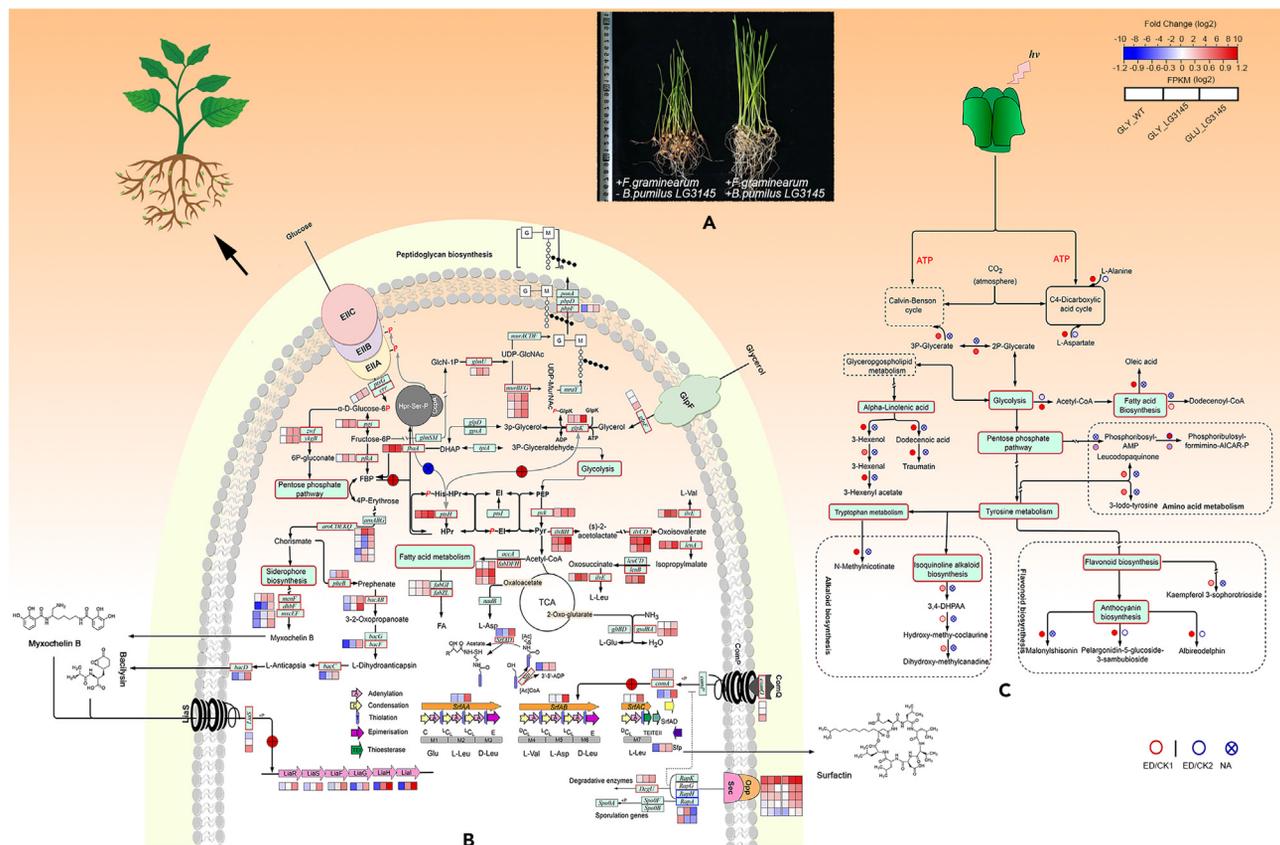


Figure 5. Metabolic pathway analysis of how strain LG3145 assists wheat to resist plant pathogens

(A) Effect of strain LG3145-root irrigation (RI) treatment on wheat infected by *F. graminearum*, with non-inoculated wheat infected by *F. graminearum* as the control (CK).

(B) The metabolic network of antibiotic substances produced by strain LG3145 was reconstructed based on Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway analysis. Transcripts of strain LG3145 in glycerol minimal medium (GYMM) (GLY_LG3145) and glucose minimal medium (GUMM) (GLU_ LG3145) and wild-type (WT) in GYMM (GLY_WT) and GUMM (GLU_ WT) are shown near the pathway as a heatmap, based on normalized data of log₂ (Fragments Per Kilobase of transcript per Million mapped reads (FPKM)) using TBtools software.

(C) The metabolite pathway network of wheat against *F. graminearum* under strain LG3145 stress is shown with a compounds heatmap, based on fold-changes in content relative to non-inoculated wheat infected by *F. graminearum*. The color legend is shown at the bottom of the map. The gene abbreviations are shown in [Datas S13 and S14](#).

~13.4 cm and ~25.4 cm, respectively. We found that strain LG3145 induced resistance against other fungal pathogens, such as *Rhizoctonia solani* Kühn, *Botrytis cinerea* Pers, *Fusarium graminearum*, *Fusarium moniliforme* Sheld, *Helminthosporium turcicum* Pass, and *Stemphylium solani* Weber (Figure S4; Table S4). The average inhibition reached 67%. The photomicrographs shown in Figure 6A indicate that the pathogenic fungal hyphae were hydrolyzed by strain LG3145.

We further investigated the antibiotic synthesis pathway of *B. pumilus* LG3145 through heat-mapping of individual gene expression levels, as shown in Figure 5B.²¹ As previously discussed, overexpression of PTS transport system-related genes enhanced strain LG3145's absorption of carbon sources to produce more fructose-6-phosphate (F6P) in glycolysis, after which F6P enters amino sugar metabolism (ko00520) to synthesize UDP-GlcNAc via the *glm* gene family and UDP-MurNAc by the *mur* and *pbp* gene clusters. The expression levels of *glm*, *mur*, and *pbp* were significantly upregulated, with *glmU* increasing from 206.16 to 429.41, *murA* increasing from 1692.01 to 1925.89, and *pbp* increasing from 108.84 to 174.50. Several mutations were identified in the *cre*-like sites of these key genes involved in peptidoglycan biosynthesis in the LG3145 mutant strain. Specifically, *glmM* with *cre* [3,0]_191 exhibited a T deletion and a G/A transition mutation, while *glmS* has a T/C transition in *cre* [1,0]_175, and *glmU* with *cre* [0,2]_157 had four transitions, one transversion, and a deletion of A around *cre* [0,2]_157. The *murGB* operon also had *cre*-like sites, including, *cre* [2,0]_190, *cre* [2,0]_161, and *cre* [2,0]_110, which exhibited three transversions and three transitions. The FPKM values of *murGB* increased significantly from 77.09 and 139.92 to 276.13 and 295.23. Similarly, *murE*, with a G/A transition inside and four transversions around the *cre* [3,0]_193 site, showed a significant increase in FPKM from 81.80 to 227.39. Mutations were also identified at the *cre*-like sites of *murA*, *murD*, and *murC*, resulting in an upregulation of their FPKM values from 1692.01 to

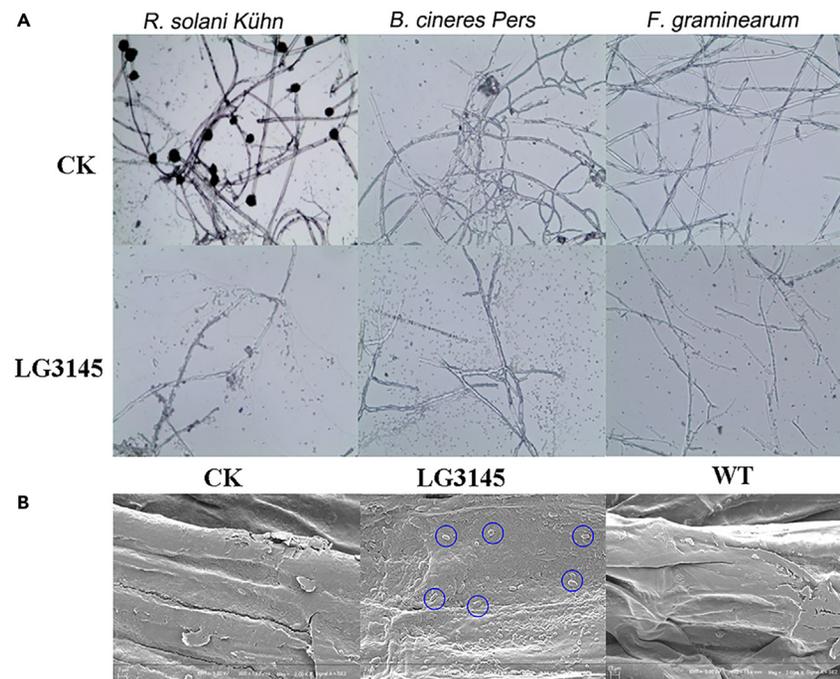


Figure 6. Micromorphology analysis of the interaction of strain LG3145 with hyphae and its colonization of wheat roots

(A) Photomicrographs of strain LG3145's effect on pathogenic fungi hyphae, *R. solani* Kühn, *B. Pers*, and *F. graminearum*. Images of the corresponding healthy fungal hyphae (control (CK)).

(B) Scanning electron microscopy (SEM) images of the colonization pattern of strain LG3145 and wild-type (WT) in wheat roots. Blue circles indicate strain LG3145 cells. SEM images of non-inoculated wheat (CK).

1925.89, 117.60 to 194.32, and 410.62 to 438.15, respectively. These mutations ultimately led to a substantial increase in peptidoglycan production in strain LG3145 and facilitated the formation of polysaccharide capsules, which aided the attachment of strain LG3145 to wheat roots. The scanning electron microscopy (SEM) images of the root cross-section in [Figure 6B](#) confirmed the abundance of attached strain LG3145 cells.

In the alternative branch of the PPP, the 4p-Erythrosine produced in PPP enters the physalalanine, tyrosine, and tryptophan biosynthesis pathway (ko00400). The FPKM values of the *aro* gene family for chorismate synthesis were upregulated, including *aroG* from 2692.28 to 3876.75, *aroA* from 524.11 to 848.06, and *aroBC* from 157.82 to 445.17, which led to an increased amount of chorismate entering the siderophore biosynthetic pathway (ko01053). Furthermore, the bacilysin and myxochelin synthase gene family, the *bac* operon, and the *pheB* gene in the bacilysin biosynthesis pathway (ko00998) also exhibited higher expression, with *bacA* increasing from 34.50 to 97.23, *bacB* from 25.80 to 186.91, and *pheB* from 73.68 to 277.86.²² The expression of the *mxo* operon, comprising myxochelin precursor synthase genes, also increased from 1.8 to 21.92.²³ These operons or genes possess some point mutations at their *cre*-like sites in their upstream sequences, which correlated with their increased expression ([Figure S2](#)).

The metabolites of strain LG3145 in GYMM exhibited distinct absorption peaks at 4.85, 12.04, and 12.15 min, which were not observed in the WT. MS analysis of the 12.04 min peak indicated the presence of surfactin homologs with $m/z = 1034.6840$ and 1035.6748 , as shown in [Figure S5](#). Furthermore, the other two peaks were identified using the mzCloud MS database as myxochelin B and lycopene. Further analysis of the UPLC-MS raw data also revealed the presence of bacilysin and riboflavin. ([Data S6](#)). In addition, the retention times of surfactin from strain LG3145 were identical to those of the standard surfactin from TargetMol ([Figure S6](#)). The maximum concentration of surfactin at 72 h was nearly 1.75 g/L. The MS fragments of surfactin produced by strain LG3145 agreed with those of previous reports ([Figure S7](#)).²⁴

The utilization of glycerol as a carbon source significantly increased most of the primary metabolites in glycolysis and the TCA cycle, providing more precursors for the synthesis of secondary metabolites, such as L-Asp and L-Glu. In addition, the gene expression levels in the leucine and valine metabolism pathway, including *ilvBCDEH* and *leuAB*, were significantly upregulated (p value <0.01). Most of the genes in the FA metabolism pathway were also upregulated (p value <0.01), such as *fabD* (increased from 148.07 to 503.71) and *fabL* (increased from 169.68 to 383.71) ([Data S6](#)). As a result, all the precursors required for surfactin biosynthesis in GYMM, namely L-Val, L-Leu, L-Asp, L-Glu, and FA, were increased compared with those in the WT strain, and were even higher than in strain LG3145 when using glucose as the carbon source.

Transcriptome analysis revealed that the FPKM values of the surfactin synthesis operon were also significantly increased (p value <0.01), such as *srfAA* (increased to 379.20), *srfAB* (increased to 547.11), *srfAC* (increased from 135.76 to 662.83), and the thioesterase gene (*srfAD*) (increased from 30.81 to 394.44). The FPKM value of the 4'-phosphopantetheinyl transferase gene (*sfp*) increased from 15.95 to 46.60.²⁵ Moreover, we identified some point mutations at the upstream region of the *srfA* operon, and DNA sequencing indicated three transitional

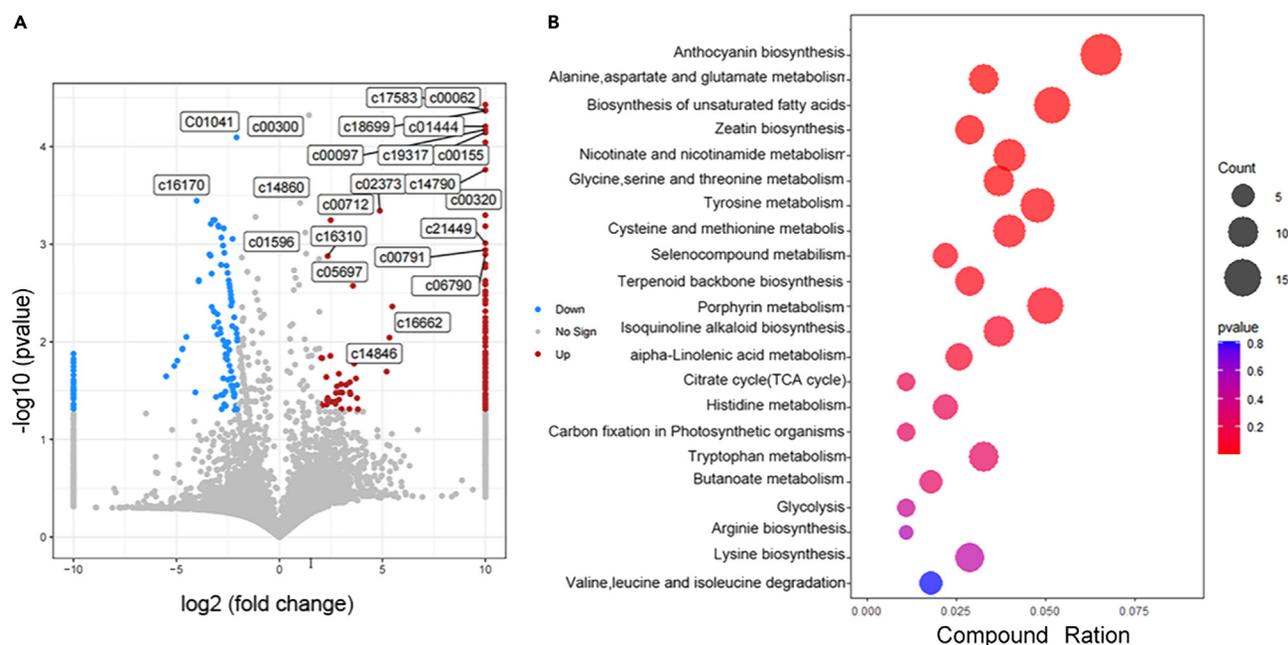


Figure 7. Metabolic changes in wheat shoots infected by *F. graminearum* under strain LG3145 stress at the seedling stage

(A) Differential metabolites analysis of wheat shoots under strain LG3145 stress using a volcano chart. Red dots represent upregulation. Blue dots represent downregulation.

(B) Kyoto Encyclopedia of Genes and Genomes (KEGG) pathways enrichment analysis of differential metabolites of wheat shoots. The bubble size represents the number of enriched factors in the pathway. A bubble color change from purple to red indicates greater statistical significance.

mutations around the *cre* [1,0]₋₆, in which G/A and A/G occurred inside the site. Furthermore, transcriptome profiling showed changes in the expression of some genes involved in two-component signal transduction systems (ko02020), such as members of the *lia* family, which are responsible for self-resistance and were all significantly upregulated (p value < 0.01). DNA sequencing revealed many mutations around *cre* [0,2]₋₁₄₃ and *cre* [0,1]₋₁₁₃. Hence, surfactin or bacilysin produced by strain LG3145 stimulated higher levels of histidine kinase (*liaS*) and transcriptional regulator (*liaR*), resulting in increased expression of the *lia* family.²⁶ Additionally, in the quorum sensing system (ko02024), the positive regulator genes, *comQ*, *opp*, and *sop* of the *srfA* operon were upregulated,²⁷ while the negative regulator gene, *rapA*, decreased significantly from 133.92 to 20.06.^{28–30} These changes are beneficial for high surfactin production. To investigate the mechanism of plant pathogen resistance induced by *B. pumilus* LG3145, we reconstructed the metabolic network using a compounds heatmap based on the comparative metabolomics data of strain LG3145-RI-treated and non-treated plants infected with *F. graminearum*,³¹ as depicted in Figure 5C (Data S14).

Notably, significant differentially abundant compounds (p value < 0.05 , $|\log_2 FC| \geq 1$) were observed for most primary metabolites in the unsaturated FA biosynthesis pathway (ko01040), such as oleic acid (C00712) and 3-hexenal (C16310), marked in the volcano chart of Figure 7A. This pathway was also significantly enriched according to Figure 7B (p value < 0.05), indicating that these FAs might not be preferred carbon sources for strain LG3145, which instead utilizes them to produce antibacterial substances, surfactants, bacilysin, and myxohelin B, which assist wheat to resist *F. graminearum* infection. Additionally, the amino acid metabolites in wheat treated by *B. pumilus* LG3145 showed a significant increase, as observed in the enriched pathway of tyrosine metabolism (ko00350) in Figure 7B, which can provide precursors for isoquinoline alkaloid and anthocyanin biosynthesis.³² Anthocyanin biosynthesis (ko00942) was the most significantly enriched pathway in Figure 7B (p value = 0.00), in which kaempferol 3-sophorotropioside (C12635), pelargonidin-5-glucoside-3-sambubioside (C20492), malonylshisonin (C16299), and albireodelphin (C16354), which are biological flavonoids, have the ability to scavenge free radicals and provide antioxidation functions.³³ Isoquinoline alkaloid is a plant alkaloid derived from tyrosine metabolism (ko00350, p value = 0.01), which is involved in various biological activities.³⁴ N-methylnicotinate (C01004) and dihydroxy-methylcanadine (C21587) are examples of isoquinoline skeleton components upregulated in strain LG3145,^{35,36} which also possess strong pharmacological and antibacterial properties (Data S15 and S16).

Conclusions

In this study, we conferred dual-plant-benefit properties on wild-type *B. pumilus*, which was isolated from intensively cultivated soil, by perturbation of its genome-scale *cre*-like sequences. The presence of *cre*-mediated global regulation in gram-positive bacteria has been known since 1990. However, the universal distribution and sequence diversity of this *cis*-acting carbon repression element pose a significant obstacle to utilizing site-directed mutagenesis to reconstruct metabolic networks. DNA sequencing demonstrated that the UgRNA/Cas9 gene editing system can effectively edit multiple *cre*-like sites simultaneously, most of which occurred around [1,0], [2,0], and [0,2] *cre*-like sites, especially in

srfa, *ptsG*, *glpFKD*, *bacABCDF*, *murABCDEG*, and *ribABD*, which contained genes encoding surfactin, myxochelin B, and bacilysin, with good antibacterial performance. When colonizing the wheat rhizosphere, strain LG3145 significantly increased the plant height and fresh weight, while improving stem rot resistance.

Transcriptomic profiling of 3505 genes and metabolomic data assayed using ultra UPLC-MS revealed that the mutations of *cre*-like sequences in strain LG3145 relieved the CCR effect for most genes, allowing the bacterium to convert rhizodeposits, e.g., FAs produced by plants, into secondary metabolites, such as pigments, PG, surfactant, bacilysin, and myxochelin B through the pentose phosphate and amino acid metabolism pathways. In turn, this enabled strain LG3145 to adhere to the root more easily, and the plant provides more FA to the bacterium, ultimately promoting plant growth and enhancing resistance against pathogens. Our findings suggested that the mutations of *cre*-like sequences play a crucial role in the overall transcription and metabolism of *B. pumilus* LG3145, and restoration of its plant-beneficial function is closely related to *cre*-mediated global regulation.

Limitations of the study

The non-specific gRNA/Cas9 system used here successfully edited the *cre*s sites of multiple genes simultaneously, but there were still many details that need to be considered in the design of UgRNA. We still use the traditional method combined with manual adjustment of similarity to design the UgRNA which is time-consuming and laborious. Future users can develop computer-aided algorithm based on their edit target specificities for making more precise similarity ratios of UgRNA. We also recognize that identifying potential causal genes involved in CCR was still very limited. The SNP-based summary statistics showed many mutation loci in the promoters, which were not discussed due to the large amount of data. We believe these mutation loci also contributed to the LG3145 strain to restore agricultural functions. In addition, field experiments need to be conducted in future studies to further verify the results of soilless and indoor experiments. The method used in this study to restore PGPR's function is more suitable for gram-positive bacteria. We need to explore more editing targets and methods for gram-negative PGPR bacteria and fungi.

STAR★METHODS

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

- KEY RESOURCES TABLE
- RESOURCE AVAILABILITY
 - Lead contact
 - Materials availability
 - Data and code availability
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
- METHOD DETAILS
 - Identifying *cre*-like sites in the *B. pumilus* genome
 - Analysis of next-generation sequencing data
 - DNA sequencing
 - RNA isolation and transcriptomic analysis
 - Metabolomics analysis
 - Fungal pathogen resistance studies
 - Plant and *B. pumilus* LG3145 interaction
 - Surfactin production assays
 - Quantitative real-time reverse transcription PCR
- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.isci.2024.108983>.

ACKNOWLEDGMENTS

We are grateful to Tianjin Academy of Agricultural Sciences of China for kindly providing the wild-type strain of *B. pumilus*.

Funding: This work was supported by the National Natural Science Foundation of China, China. [grant number 42077212]; the Tianjin Technical Expert Project [grant number 14JCTPJ00479]; and the Training Project of Innovation Team of Colleges and Universities in Tianjin, China [grant number 202210060077].

AUTHOR CONTRIBUTIONS

M.B. and M.L. designed and performed experiments, and wrote the first draft of the manuscript. J.W., Z.M., and Z.W. participated in the experimental design and the analysis of the experiment results. M.Y., X.Y., and L.H. designed and supervised the project, and guided the experimental design, data analysis, manuscript writing, and revision. All authors read and approved the final manuscript.

DECLARATION OF INTERESTS

The authors declare no conflict of interest.

Received: July 4, 2023

Revised: October 26, 2023

Accepted: January 17, 2024

Published: January 20, 2024

REFERENCES

- Kumar, A., Prakash, A., and Johri, B. (2011). *Bacillus* as PGPR in crop ecosystem. In *Bacteria in Agrobiolgy: Crop Ecosystems*, D.K. Maheshwari, ed. (Springer), pp. 37–59. https://doi.org/10.1007/978-3-642-18357-7_2.
- Zhang, X., Huang, Y., Harvey, P.R., Ren, Y., Zhang, G., Zhou, H., and Yang, H. (2012). Enhancing plant disease suppression by *Burkholderia vietnamiensis* through chromosomal integration of *Bacillus subtilis* chitinase gene *chi113*. *Biotechnol. Lett.* 34, 287–293. <https://doi.org/10.1007/s10529-011-0760-z>.
- Peng, G., Zhao, X., Li, Y., Wang, R., Huang, Y., and Qi, G. (2019). Engineering *Bacillus velezensis* with high production of acetoin primes strong induced systemic resistance in *Arabidopsis thaliana*. *Microbiol. Res.* 227, 126297. <https://doi.org/10.1016/j.micres.2019.126297>.
- Yi, Y., Li, Z., Song, C., and Kuipers, O.P. (2018). Exploring plant-microbe interactions of the rhizobacteria *Bacillus subtilis* and *Bacillus mycoides* by use of the CRISPR-Cas9 system. *Environ. Microbiol.* 20, 4245–4260. <https://doi.org/10.1111/1462-2920.14305>.
- Watzlawick, H., and Altenbuchner, J. (2019). Multiple integration of the gene *ganA* into the *Bacillus subtilis* chromosome for enhanced β -galactosidase production using the CRISPR/Cas9 system. *AMB Expr* 9, 158. <https://doi.org/10.1186/s13568-019-0884-4>.
- Ferrando, J., Filluelo, O., Zeigler, D.R., and Picart, P. (2023). Barriers to simultaneous multilocus integration in *Bacillus subtilis* tumble down: development of a straightforward screening method for the colorimetric detection of one-step multiple gene insertion using the CRISPR-Cas9 system. *Microb. Cell Fact.* 22, 21. <https://doi.org/10.1186/s12934-023-02032-2>.
- Zheng, W., Zeng, S., Bais, H., LaManna, J.M., Hussey, D.S., Jacobson, D.L., and Jin, Y. (2018). Plant growth-promoting rhizobacteria (PGPR) reduce evaporation and increase soil water retention. *Water Resour. Res.* 54, 3673–3687.
- Weickert, M.J., and Chambliss, G.H. (1990). Site-directed mutagenesis of a catabolite repression operator sequence in *Bacillus subtilis*. *Proc. Natl. Acad. Sci. USA* 87, 6238–6242. <https://doi.org/10.1073/pnas.87.16.6238>.
- Hueck, C.J., Hillen, W., and Saier, M.H. (1994). Analysis of a *Cis*-active sequence mediating catabolite repression in Gram-positive bacteria. *Res. Microbiol.* 145, 503–518. [https://doi.org/10.1016/0923-2508\(94\)90028-0](https://doi.org/10.1016/0923-2508(94)90028-0).
- Nicholson, W.L., Park, Y.K., Henkin, T.M., Won, M., Weickert, M.J., Gaskell, J.A., and Chambliss, G.H. (1987). Catabolite repression-resistant mutations of the *Bacillus subtilis* alpha-amylase promoter affect transcription levels and are in an operator-like sequence. *J. Mol. Biol.* 198, 609–618. [https://doi.org/10.1016/0022-2836\(87\)90204-X](https://doi.org/10.1016/0022-2836(87)90204-X).
- Miwa, Y., Nakata, A., Ogiwara, A., Yamamoto, M., and Fujita, Y. (2000). Evaluation and characterization of catabolite-responsive elements (*cre*) of *Bacillus subtilis*. *Nucleic Acids Res.* 28, 1206–1210. <https://doi.org/10.1093/nar/28.5.1206>.
- Sun, S., Wang, X., Xu, J., Fu, H., and Sun, Z. (2021). Study of tomato growth weight-distribution model based on real-time plant weight in a solar greenhouse. *J. Taibah Univ. Sci.* 15, 1027–1037. <https://doi.org/10.1080/16583655.2021.2014690>.
- Wang, Y., Cao, L., Bi, M., Wang, S., Chen, M., Chen, X., Ying, M., and Huang, L. (2021). Wobble editing of *cre*-box by unspecific CRISPR/Cas9 causes CCR release and phenotypic changes in *Bacillus pumilus*. *Front. Chem.* 9, 717609. <https://doi.org/10.3389/fchem.2021.717609>.
- Jacob-Dubuisson, F., Mechaly, A., Betton, J.M., and Antoine, R. (2018). Structural insights into the signalling mechanisms of two-component systems. *Nat. Rev. Microbiol.* 16, 585–593. <https://doi.org/10.1038/s41579-018-0055-7>.
- Costa-Silva, J., Domingues, D., and Lopes, F.M. (2017). RNA-Seq differential expression analysis: An extended review and a software tool. *PLoS One* 12, e0190152. <https://doi.org/10.1371/journal.pone.0190152>.
- Zhang, J., Abdelraheem, A., Zhu, Y., Elkins-Arce, H., Dever, J., Whitelock, D., Hake, K., Wedegaertner, T., and Wheeler, T.A. (2022). Studies of evaluation methods for resistance to fusarium wilt race 4 (*Fusarium oxysporum* f. sp. *vasinfectum*) in cotton: effects of cultivar, planting date, and inoculum density on disease progression. *Front. Plant Sci.* 13, 900131. <https://doi.org/10.3389/fpls.2022.900131>.
- Xu, Y.L., Zhang, Y., Li, J.M., Gao, T., Zhang, L.N., Si, L., Li, Q., Li, G.Y., and Yang, Y.L. (2021). Comparison of antioxidant enzyme activity and gene expression in two new spring wheat cultivars treated with salinity. *Biol. Plant. (Prague)* 65, 131–144. <https://doi.org/10.32615/BP.2020.171>.
- Yang, J., Zhu, Y., Li, J., Men, Y., Sun, Y., and Ma, Y. (2015). Biosynthesis of rare ketoses through constructing a recombination pathway in an engineered *Corynebacterium glutamicum*. *Biotechnol. Bioeng.* 112, 168–180. <https://doi.org/10.1002/bit.25345>.
- Ponomarenko, P.M., and Ponomarenko, M.P. (2015). Sequence-based prediction of transcription upregulation by auxin in plants. *J. Bioinform. Comput. Biol.* 13, 1540009. <https://doi.org/10.1142/S0219720015400090>.
- Brennan, J.M., Egan, D., Cooke, B.M., and Doohan, F.M. (2005). Effect of temperature on head blight of wheat caused by *Fusarium culmorum* and *F. graminearum*. *Plant Pathol.* 54, 156–160. <https://doi.org/10.1111/j.1365-3059.2005.01157.x>.
- Hu, X., Roberts, D.P., Maul, J.E., Emche, S.E., Liao, X., Guo, X., Liu, Y., McKenna, L.F., Buyer, J.S., and Liu, S. (2011). Formulations of the endophytic bacterium *Bacillus subtilis* Tu-100 suppress *Sclerotinia sclerotiorum* on oilseed rape and improve plant vigor in field trials conducted at separate locations. *Can. J. Microbiol.* 57, 539–546. <https://doi.org/10.1139/W11-041>.
- Silakowski, B., Kunze, B., Nordsiek, G., Blöcker, H., Höfle, G., and Müller, R. (2000). The myxochelin iron transport regulon of the myxobacterium *Stigmatella aurantiaca* Sg a₁₅. *Eur. J. Biochem.* 267, 6476–6485.
- Nakano, M.M., Magnuson, R., Myers, A., Curry, J., Grossman, A.D., and Zuber, P. (1991). *srfA* is an operon required for surfactin production, competence development, and efficient sporulation in *Bacillus subtilis*. *J. Bacteriol.* 173, 1770–1778. <https://doi.org/10.1128/jb.173.5.1770-1778.1991>.
- Chen, Z., Wu, Q., Wang, L., Chen, S., Lin, L., Wang, H., and Xu, Y. (2020). Identification and quantification of surfactin, a nonvolatile lipopeptide in Moutai liquor. *Int. J. Food Prop.* 23, 189–198. <https://doi.org/10.1080/10942912.2020.1716791>.
- Kim, H.S., Kim, S.B., Park, S.H., Oh, H.M., Park, Y.I., Kim, C.K., Katsuragi, T., Tani, Y., and Yoon, B.D. (2000). Expression of *sfp* gene and hydrocarbon degradation by *Bacillus subtilis*. *Biotechnol. Lett.* 22, 1431–1436. <https://doi.org/10.1023/A:1005675231949>.
- Mascher, T., Zimmer, S.L., Smith, T.A., and Helmann, J.D. (2004). Antibiotic-inducible promoter regulated by the cell envelope stress-sensing two-component system LiARs of *Bacillus subtilis*. *Antimicrob. Agents Chemother.* 48, 2888–2896. <https://doi.org/10.1128/AAC.48.8.2888-2896.2004>.
- Bacon Schneider, K., Palmer, T.M., and Grossman, A.D. (2002). Characterization of *comQ* and *comX*, two genes required for production of ComX pheromone in *Bacillus subtilis*. *J. Bacteriol.* 184, 410–419. <https://doi.org/10.1128/JB.184.2.410-419.2002>.
- Hayashi, K., Kensuke, T., Kobayashi, K., Ogasawara, N., and Ogura, M. (2006). *Bacillus subtilis* RghR (YvaN) represses *rapG* and *rapH*, which encode inhibitors of expression of the *srfA* operon. *Mol. Microbiol.* 59, 1714–1729. <https://doi.org/10.1111/j.1365-2958.2006.05059.x>.
- Singh, P.K., Schaefer, A.L., Parsek, M.R., Moninger, T.O., Welsh, M.J., and Greenberg, E.P. (2000). Quorum-sensing signals indicate that cystic fibrosis lungs are infected with bacterial biofilms. *Nature* 407, 762–764. <https://doi.org/10.1038/35037627>.
- Tjalsma, H., Koetje, E.J., Kiewiet, R., Kuipers, O.P., Kolkman, M., van der Laan, J., Daskin, R., Ferrari, E., and Bron, S. (2004). Engineering of quorum-sensing systems for improved

- production of alkaline protease by *Bacillus subtilis*. *J. Appl. Microbiol.* 96, 569–578. <https://doi.org/10.1111/j.1365-2672.2004.02179.x>.
31. Hsieh, F.C., Li, M.C., Lin, T.C., and Kao, S.S. (2004). Rapid detection and characterization of surfactin-producing *Bacillus subtilis* and closely related species based on PCR. *Curr. Microbiol.* 49, 186–191. <https://doi.org/10.1007/s00284-004-4314-7>.
 32. Goldbach, R., Bucher, E., and Prins, M. (2003). Resistance mechanisms to plant viruses: an overview. *Virus Res.* 92, 207–212. [https://doi.org/10.1016/S0168-1702\(02\)00353-2](https://doi.org/10.1016/S0168-1702(02)00353-2).
 33. Koirala, M., Karimzadegan, V., Liyanage, N.S., Mérindol, N., and Desgagné-Penix, I. (2022). Biotechnological approaches to optimize the production of amaryllidaceae alkaloids. *Biomolecules* 12, 893. <https://doi.org/10.3390/biom12070893>.
 34. Kogawa, K., Kazuma, K., Kato, N., Noda, N., and Suzuki, M. (2007). Biosynthesis of malonylated flavonoid glycosides on the basis of malonyltransferase activity in the petals of *Clitoria ternatea*. *J. Plant Physiol.* 164, 886–894. <https://doi.org/10.1016/j.jplph.2006.05.006>.
 35. Rozwadowska, M.D., Matecka, D., and Brózda, D. (1991). Synthesis of isoquinoline alkaloids. Total synthesis of (±)-tetrahydroescholamine and its(±)-threo- and(±)-erythro- α -hydroxy analogs. *Liebigs Ann. Chem.* 1991, 73–75. <https://doi.org/10.1002/jlac.199119910113>.
 36. Chrzanowska, M. (1995). Synthesis of isoquinoline alkaloids. total synthesis of (±)-stylopine. *J. Nat. Prod.* 58, 6669–6672. <https://doi.org/10.1021/np50117a008>.
 37. Chen, T., Chen, X., Zhang, S., Zhu, J., Tang, B., Wang, A., Dong, L., Zhang, Z., Yu, C., Sun, Y., et al. (2021). Toward Explosive Data Growth and Diverse Data Types. *Dev. Reprod. Biol.* 19, 578–583. <https://doi.org/10.1016/j.gpb.2021.08.001>.
 38. CNCB-NGDC Members and Partners (2023). Database Resources of the National Genomics Data Center, China National Center for Bioinformation in 2023. *Nucleic Acids Res.* 51, D18–D28. <https://doi.org/10.1093/nar/gkac1073>.
 39. Dao, V.L., Chan, S., Zhang, J., Ngo, R.K.J., and Poh, C.L. (2022). Single 3'-exonuclease-based multifragment DNA assembly method (SENAX). *Sci. Rep.* 12, 4004. <https://doi.org/10.1038/s41598-022-07878-x>.
 40. Bolger, A.M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. <https://doi.org/10.1093/bioinformatics/btu170>.
 41. Li, D., Liu, C.M., Luo, R., Sadakane, K., and Lam, T.W. (2015). MEGAHIT: an ultra-fast single-node solution for large and complex metagenomics assembly via succinct de Bruijn graph. *Bioinformatics* 31, 1674–1676. <https://doi.org/10.1093/bioinformatics/btv033>.
 42. Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUAST: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. <https://doi.org/10.1093/bioinformatics/btt086>.
 43. Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. <https://doi.org/10.1093/bioinformatics/btu153>.
 44. Xia, H., Zhang, Z., Luo, C., Wei, K., Li, X., Mu, X., Duan, M., Zhu, C., Jin, L., He, X., et al. (2023). MultiPrime: A reliable and efficient tool for targeted next-generation sequencing. *iMeta* 2, e143. <https://doi.org/10.1002/imt2.143>.
 45. Cao, L., Meng, Z., Tan, J., Ying, M., Bi, M., Liu, Y., Tong, X., Wei, J., and Huang, L. (2022). Self-assembled endogenous DNA nanoparticles for auto-release and expression of the eGFP gene in *Bacillus subtilis*. *Commun. Biol.* 5, 1373. <https://doi.org/10.1038/s42003-022-04233-8>.
 46. Langmead, B., and Salzberg, S.L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. <https://doi.org/10.1038/nmeth.1923>.
 47. Li, B., and Dewey, C.N. (2011). RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform* 12, 323. <https://doi.org/10.1186/1471-2105-12-323>.
 48. Kanehisa, M., and Goto, S. (2000). KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 28, 27–30.
 49. Wu, T., Hu, E., Xu, S., Chen, M., Guo, P., Dai, Z., Feng, T., Zhou, L., Tang, W., Zhan, L., et al. (2021). clusterProfiler 4.0: A universal enrichment tool for interpreting omics data. *Innov. J. (Camb)* 2, 100141. <https://doi.org/10.1016/j.xinn.2021.100141>.
 50. Hageman, J.H., Shankweiler, G.W., Wall, P.R., Franich, K., McCowan, G.W., Cauble, S.M., Grajeda, J., and Quinones, C. (1984). Single, chemically defined sporulation medium for *Bacillus subtilis*: growth, sporulation, and extracellular protease production. *J. Bacteriol.* 160, 438–441. <https://doi.org/10.1111/j.1365-2672.1984.tb01383.x>.
 51. Zhi, Y., Wu, Q., and Xu, Y. (2017). Genome and transcriptome analysis of surfactin biosynthesis in *Bacillus amyloliquefaciens* MT45. *Sci. Rep.* 7, 40976. <https://doi.org/10.1038/srep40976>.
 52. Livak, K.J., and Schmittgen, T.D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2^{- $\Delta\Delta$ CT} Method. *Methods* 25, 402–408.
 53. R Core Team (2020). R: A Language and Environment for Statistical Computing (R Foundation for Statistical Computing). <https://www.R-project.org/>.

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Chemicals, peptides, and recombinant proteins		
Yeast powder	AOBOX	Cat# 01-008A
Tryptone	Thermo Scientific™	Cat# CM0595B
AGAR	Sangon Biotech	Cat# A505255
Trichloroacetic acid	Sigma	Cat# T9159
0.1% formic acid aqueous solution	Thermo Scientific™	Cat# 85170
Sodium carbonate	Sangon Biotech	Cat# A610615
Copper sulfate	Sangon Biotech	Cat# A603008
Tris	Sangon Biotech	Cat# A610195
SDS	Thermo Scientific™	Cat# 28365
Taq DNA polymerase	Ruibio Biotech	Cat# RBPM06
DNA Marker	Ruibio Biotech	Cat# RB-MK8
Fluorescent dyes for DNA electrophoresis gel	Ruibio Biotech	Lot# S11494
Methanol	Merck	Cat# 106035
Glucose	Thermo Scientific™	Cat# A2494001
Glycerol	Thermo Scientific™	Cat# 17904
Peptone	Thermo Scientific™	Cat# 211677-4
EDTA	Sangon Biotech	Cat# B540625
Lysozyme	Solarbio	Cat# L8120
RNAiso Plus reagent	Takara	Cat# 9108
PrimeScript RT reagent Kit	Takara	Cat#RR047A
Chloroform	Merck	Cat# 102444
TB Green Premix Ex Taq II	Takara	Lot# AL92672A
ROX Reference Dye	Sangon Biotech	Cat# B541010
Critical commercial assays		
TIANamp bacteria DNA kit	Tiangen Biotech	Cat# DP302-02
Transcriptome analysis	Meiji Biomedical Technology	N/A
DNA sequencing	Ruibio Biotech	N/A
Experimental models: Organisms/strains		
<i>Bacillus. pumilus</i> WT	Laboratory stock	N/A
<i>Bacillus. pumilus</i> LG3145	Laboratory stock	N/A
Zhongmai 100	Beijing, China	N/A
<i>Fusarium graminearum</i>	Tianjin, China	N/A
<i>Fusarium moniliforme</i> Sheldon	Tianjin, China	N/A
<i>Rhizoctonia solani</i> Kühn	Tianjin, China	N/A
<i>Helminthosporium turcicum</i> Pass	Tianjin, China	N/A
<i>Botrytis cinerea</i> Pers	Tianjin, China	N/A
<i>Stemphylium solani</i> Weber	Tianjin, China	N/A
Deposited data		
RNA-seq data	SRA database	PRJNA997144
Supplementary file data	This paper; Mendeley Data	https://doi.org/10.17632/zd6x83mpv.2

(Continued on next page)

Continued

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Raw sequence reads	GSA database	CRA011959
Software and algorithms		
RSEM	Bo Li and Colin Dewey et al., 2020	https://www.deweylab.github.io/RSEM/
Bowtie2	Langmead and Salzberg, 2012	https://github.com/BenLangmead/bowtie2
Kallisto	Nicolas L Bray and Harold Pimentel et al., 2016	https://pachterlab.github.io/kallisto
GraphPad Prism 9.4.0	Harvey Motulsky.2022	https://github.com/Jaycadox/Graphpad-Patcher.git
DESeq2	Michael Love and Constantin Ahlmann-Eltze et al.,2014	https://github.com/thevelab/DESeq2.git
SnapGene	Drummond, E.and Kavanagh, T. et al.,2022	https://github.com/topics/snapgene
R statistical software 4.1.3	R Foundation for Statistical Computing	https://www.r-project.org/
KEGG	Kanehisa Laboratories, 1996	https://github.com/FlyPythons/KEGGTools.git
Oligonucleotides		
Primers used for DNA sequencing in this study listed in Table S1.	This study	N/A
Primers used in the real-time qPCR assay listed in Table S2	This study	N/A

RESOURCE AVAILABILITY

Lead contact

Further information and requests should be directed to and will be fulfilled by the lead contact, Ming Ying (ym@tjut.edu.cn).

Materials availability

This study did not generate new materials.

Data and code availability

- Data: The raw sequence data reported in this paper have been deposited in the Genome Sequence Archive³⁷ in National Genomics Data Center,³⁸ China National Center for Bioinformation/Beijing Institute of Genomics, Chinese Academy of Sciences (GSA: CRA011959), which are publicly accessible at <https://ngdc.cncb.ac.cn/gsa>. RNA-seq data have been deposited in the SRA database (Accession number: PRJNA997144). This paper reports the original data in Mendeley Data (<https://doi.org/10.17632/zd6x83rmpv.2>).
- Code: This paper does not report original code.
- Any additional information required to reanalyze the data reported in this paper is available from the [lead contact](#) upon request.

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Wild-type *Bacillus pumilus* (*B. pumilus* WT) was isolated from intensively cultivated soil and shares the same genotype as *B. pumilus* SH-B9 (NCBI Accession no: NZ_CP011007.1). *Bacillus pumilus* LG3145 was derived from the *B. pumilus* WT, as previously described by Wang et al. (2021). *B. pumilus* LG3145 and the WT strain of *B. pumilus* were cultured in MM broth (3.48 g/L KH₂PO₄, 1.5 g/L Na₂HPO₄•12H₂O, 3.96 g/L (NH₄)₂SO₄, 0.7 g/L MgSO₄•7H₂O) with either 60% (w/v) glucose (GUMM) or glycerol (GYMM) and 0.01 g/L yeast powder at 35°C and 150 rpm. Fungal pathogens were cultivated on potato dextrose agar (PDA) plates containing 20 g/L glucose, 200 g/L potato, 15–20 g/L agar, and pH 7.3–7.5 at 25°C. Wheat seeds were immersed in a 5% (w/w) NaClO solution for 4 min and then washed three times with ddH₂O. The seeds were sown into organic soil or into solid nutrients containing 0.034 g/L KH₂PO₄, 0.120 g/L MgSO₄•7H₂O, 1 g/L KCl, 1 g/L NaCl, 1 g/L C₆H₁₂O₆, 0.5 mL of 1 mM Ca (NO₃)₂ stock solution, and 1 mL of 1 mM FeSO₄ stock solution, and cultivated at 25°C or room temperature.

METHOD DETAILS

Identifying cre-like sites in the *B. pumilus* genome

The cre-like sites present in the genome of the wild-type strain of *B. pumilus* were detected using SnapGene software.³⁹ The query sequence utilized for the search was “WWGNAANCGNWNCCW”.

Analysis of next-generation sequencing data

Adapters and low-quality reads were removed using Trimmomatic, employing parameters "LEADING: 3 TRAILING: 3 SLIDINGWINDOW: 4:15 MINLEN: 7".⁴⁰ The quality of the resulting clean reads was evaluated using FastQC (<https://github.com/s-andrews/FastQC>) before undergoing assembly by Megahit with default parameters.⁴¹ Quast⁴² was utilized for assembly quality control. Gene annotation was carried out using Prokka with the parameter "-kingdom Bacteria".⁴³ All the primers were designed by multiPrime.⁴⁴

DNA sequencing

B. pumilus LG3145 was cultured in Luria-Bertani (LB) broth (10 g/L peptone, 5 g/L yeast extract, g/L NaCl, pH 7.5). The genome of strain LG3145 was extracted and purified using a TIANamp bacteria DNA kit (TIANGEN Biotech, Beijing, China) and used as the PCR template for the genes under investigation. The *cre*-like sites regions of 35 genes and operons were amplified using the primers detailed in Table S1. The resulting PCR products were detected using agarose gel electrophoresis and subsequently sent to Ruibio Biotech (Beijing, China) for DNA sequencing using high-throughput DNA sequencing on an Illumina HiSeq 4000 instrument (Illumina, San Diego, CA, USA). The obtained results were aligned with the genome of *B. pumilus* SH-B9 using Snappgene software.⁴⁵ The DNA extraction kit was obtained from TIANGEN Biotech, while the Taq DNA polymerase and oligonucleotides utilized for the PCR experiments were purchased from Ruibio Biotech.

RNA isolation and transcriptomic analysis

The cells of strain LG3145 were cultured in glycerol minimal medium (GYMM) and glucose minimal medium (GUMM) broth at 35°C for 24 h. *B. pumilus* WT was used as the control and was grown in GYMM. The cells were harvested by centrifugation at 10,000 × *g* for 10 min. Subsequently, the pellets were treated with liquid nitrogen and subjected to genome-wide RNA sequencing via the Illumina HiSeq Novaseq 6000 system at a 2 × 150 bp read length by Majorbio Co., Ltd. (Shanghai, China). The high-quality data of each sample obtained from the Illumina platform were mapped to the reference genome GCF_001578205.1 using Bowtie2.⁴⁶ The expression level of each gene was analyzed using RSEM⁴⁷ and measured using fragments per kilobases per million reads (FPKM). DEGs between the groups were analyzed using DESeq2, DEGseq, and edgeR in the R software. Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis was performed online.⁴⁸

Metabolomics analysis

Strain LG3145 was cultured in GYMM at 35°C for 72 h, and 300 mL shaking cultures (OD₆₂₀ = 12) were harvested by centrifugation (10,000 × *g*, 5 min) to remove the supernatant. The resulting cells were scraped onto a clean Petri dish and sealed with parafilm, with some small holes pricked on it. The dish was stored at -20°C for 5 h and then freeze-dried for 8 h. The freeze-dried bacterial powder was ground and immersed in 3 mL of 2 M NaOH. The cells were then crushed by shaking in an ultrasonic ice bath for 5 min and harvested again by centrifugation at 10,000 × *g* for 5 min. The cells were washed twice with ddH₂O and freeze-dried for 6 h before UPLC-MS detection. The freeze-dried powder of *B. pumilus* WT, prepared from GYMM cultures, was used as the control (CK). For the plant samples, small pieces of roots from three replicates of each treatment were fully ground with quartz sand and soaked in 1 mL of methanol solution (CH₃OH: H₂O = 1:1, v/v) for 30 min. The extracts were separated by centrifugation at 10,000 × *g* for 2 min and then subjected to UPLC-MS detection. The WT-RI treated and non-inoculated wheat were used as controls. All samples (5 μL) were loaded onto a Waters ACQUITY UPLC BEH C₁₈ column (5 μm, 150 × 4.6 mm; Waters Corp., Milford, MA, USA) and analyzed using a Waters UPLC I-Class/Xevo G2 XS-Quadrupole Time-of-flight system (Waters Corp) at 4°C with a flow rate of 0.3 mL/min under isocratic elution with water (0.1% formic acid) and acetonitrile (0.1% formic acid). Mass spectrometry detection was performed in the positive ESI-MS mode with the following conditions: capillary voltage: 0.5 kV; cone voltage: 35 V; extractor voltage: 4.0 V; source temperature: 11°C; desolvation temperature: 550°C. UPLC-MS data were collected and processed using MassLynx NT 4.1 software with QuanLynx program (Waters Corp) and converted to the CVS file format raw data for bioinformatic analysis using R version 4.1.0. Enrichment pathway analysis was performed to identify significant differential metabolites (p value <0.05) using the clusterProfile package⁴⁹ and the KEGG pathway database.

Fungal pathogen resistance studies

B. pumilus LG3145 and the WT strain were cultured in 5 mL of LB broth for 12 h at 37°C and 150 rpm to prepare bacterial suspensions (OD₆₂₀ = 12). The bacterial suspensions were then spread on potato dextrose agar (PDA) plates using a sterilized cotton swab. The plates were incubated for 3 h at 37°C. Thereafter, 1-cm fungal cakes of the pathogens mentioned above were inoculated into the center of the strain LG3145-PDA plates and WT-PDA plates simultaneously. The plates were then cultivated for 72 h at 25°C. The diameter of the fungal growth was measured every 24 h and images were taken after 72 h. The WT-PDA plates were used as the control.

Plant and *B. pumilus* LG3145 interaction

Fresh cells of *B. pumilus* LG3145 were cultivated in Chemically Defined Sporulation medium (CDSM)⁵⁰ for 24 h when the spores were about 10⁵–10⁹/mL. The spore cultures were collected by centrifugation and freeze-dried for 8 h to make a powder. The freeze-dried powder was diluted with ddH₂O to prepare the spore suspension (OD₆₂₀ = 4). The *B. pumilus* WT spore suspension was prepared using the same method as the control. Wheat seeds were sown into 15 organic soil pots (250 g per pot) with five seeds in each at a depth of 1 cm, or into solid nutrient medium with 36 seeds in each vessel. The organic soil was a Universal nutrient soil purchased from Jurong Kaihong Plastic Industry Co., Ltd

(Jurong, Jiangsu, China). In the indoor experiment, the organic soil was moistened using 100 mL of ddH₂O and kept at approximately 60% wetness at a temperature of approximately 25°C each day. After seed germination, the strain LG3145 spore suspension was sprayed around the germinated plants once every two days. The plant height was recorded every day and the fresh weight and the root lengths were recorded after nine days of cultivation. The non-inoculation (CK1) and WT spore suspension treatment (CK2) were used as the controls. In the disease resistance experiment, when the wheat seeds had germinated for approximately one week at 25°C, an equal number of hyphae from *F. graminearum* were inoculated to infect the wheat embryos. The next day, strain LG3145 spore suspensions were sprayed around the embryos as the experimental group (EG), and an equal amount of water and the WT strain were used as controls CK1 and CK2, respectively. The growth of wheat was observed and recorded, and images of wheat shoots were acquired 10 days later.

Surfactin production assays

A modified method was utilized to purify and detect crude surfactin from *B. pumilus* LG3145⁵¹. *B. pumilus* LG3145 was cultivated in GYMM at 35°C and 125 rpm for 72 h. Then, 60 mL of cell culture broth was taken and adjusted to pH 2.0 using 6M HCl, followed by centrifugation to remove the supernatant (10,000 × g, 30 min). The pellet was then re-dissolved in 1 mL of 100% CH₃OH and centrifuged (10,000 × g, 30 min) to remove the cell residues. A high-performance liquid chromatography (HPLC) assay was performed using an SHIMADZU-CTO-20A HPLC with photo diode array (PDA) detector (Shanghai, China) and an XAqua C₁₈ column (250 mm × 4.6 mm, 5 μm 100 Å). The mobile phase consisted of solvent A (ddH₂O) and solvent B (100% CH₃OH), and a binary solvent delivery system was used. A linear biphasic gradient of 85% solvent B over 10 min at a flow rate of 1 mL/min and a column temperature of 40°C was used. The injection volume was 25 μL, monitored at 210 nm. The surfactin concentration was quantified using a calibration curve constructed using surfactin standards (CAS: 24730-31-2 TargetMol Chemicals, Boston, MA, USA). The experiment was performed in triplicate. The same samples (5 μL) were analyzed using an SHIMADZU-LCMS-8045 with a Q3. detector (Shanghai, China). Mass spectrometry detection was performed on a Shim-pack GIST C₁₈ column (2 μm, 2.1 mm × 50 mm (High Strength Silica (HSS)); SHIMADZU, Kyoto, Japan) using a linear biphasic gradient elution with 90% CH₃OH (100%) and ddH₂O at a flow rate of 0.3 mL/min. The detection time was 20 min, and the m/z range was from 400 to 1300.

Quantitative real-time reverse transcription PCR

The same cell samples used for transcriptomic analysis were re-suspended in TE buffer (1 mM EDTA and 10 mM Tris-HCl, pH 8.0) supplemented with 1.0 mg/mL lysozyme. Following incubation at 37°C for 10 min, RNA extraction was carried out by adding 1 mL of RNAiso Plus reagent (cat No.9108, Takara, Dalian, China), incubating at room temperature for 5 min, and then removed by centrifugation (12,000 × g, 5 min). The precipitate was washed three times with chloroform, isopropanol, and 75% ethanol, consecutively, and then dissolved in RNase free water. The extracted RNA was reverse transcribed into cDNA using a PrimeScript RT reagent Kit (cat No. RR047A, Takara). Quantitative real-time PCR (real-time qPCR) was then used to determine the gene expression levels of *crtI*, *glpD*, *glpK*, *guaB*, *murA*, *ptsH*, *ribA*, *srfA*, and *tenA*, using the primers listed in Table S2 together with the QuantStudio 3 Real-Time PCR System (Applied Biosystems, Foster City, CA, USA) and the prepared cDNA as the template. The real-time qPCR reaction was performed in a 20 μL volume containing 10 μL of TB Green Premix Ex Taq II, 0.8 μL of each primer, 0.4 μL of ROX Reference Dye, 2 μL of template cDNA, and 6 μL of RNase-free H₂O. The amplification protocol was as follows: preheating at 95°C for 30 s; followed by 40 cycles of 95°C for 5 s and 60°C for 34 s. The relative expression levels of target genes were calculated using the $2^{-\Delta\Delta CT}$ method with 16SrRNA gene as the internal control.⁵² Three replicate assays were performed for each target gene.

QUANTIFICATION AND STATISTICAL ANALYSIS

The transcriptome and metabolome analysis in this study were performed in triplicate. Pearson's correlation coefficients were used to explore the correlations between variables. If the variables fitted a normal distribution, Student's *t* test was used to analyze the differences between groups. There were three biological replicates in each experiment. Differences between categorical variables were determined using Pearson's chi-squared test or Fisher's exact test. Statistical analysis of raw counts was conducted using DESeq2 software (version 1.24.0) included in the R software⁵³ based on a negative binomial distribution. *p* values obtained from statistical tests were corrected, and BH (false discovery rate (FDR) correction with Benjamin/Hochberg) multiple tests were performed to obtain the basis for comparing expression differences between the groups based on certain screening conditions. Default parameters: *p* value <0.05 & |log₂FC| ≥ 1. Analyses were conducted using the R software (version 4.1.3) and GraphPad Prism (version 9.4.0; GraphPad inc. La Jolla, CA, USA).