Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Mechanisms of Enhanced Natamycin Biosynthesis in Streptomyces gilvosporeus 717: An Integrated Genomic and Transcriptomic Analysis

Renqiang Li
Linyi University
Caifeng Liu
Linyi University
Xu Zhang
Linyi University
Yinghua Zhang
Shandong Freda Biotechnology Co., Ltd
Haijun Li
Shandong Freda Biotechnology Co., Ltd
Cuijuan Gao
Linyi University
Jingjing Liu
liujingjing@lyu.edu.cn
Linyi University

Research Article

Keywords: Natamycin, Streptomyces gilvosporeus, ARTP mutagenesis, whole-genome resequencing, Transcriptomics

Posted Date: August 26th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10440944/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Mechanisms of Enhanced Natamycin Biosynthesis in *Streptomyces gilvosporeus* 717: An Integrated Genomic and Transcriptomic Analysis

Renqiang Li[1] | Caifeng Liu[1] | Xu Zhang[1] | Yinghua Zhang[2]| Haijun Li[2] | Cuijuan Gao[1] | Jingjing Liu[1]

[1]College of Life Sciences, Linyi University, Linyi 276000, China

2 Shandong Freda Biotechnology Co., Ltd, Linyi 276000, China

Correspondence: Jingjing Liu, liujingjing@lyu.edu.cn

Abstract

Natamycin is a polyene macrolide antifungal compound produced by *Streptomyces* species and is widely used in food preservation and antifungal therapy. Random mutagenesis is proved an effective strategy for improving natamycin production, but the genetic and transcriptional basis underlying high-yield phenotypes remains incompletely understood. In this study, we generated a high-producing mutant, *Streptomyces gilvosporeus* 717, from its parental strain 712 by atmospheric and room-temperature plasma (ARTP) mutagenesis and integrated whole-genome resequencing (WGS) with time-resolved transcriptome sequencing (RNA-seq) to investigate the mechanisms associated with enhanced natamycin biosynthesis. In 15 L fed-batch fermentation, strain 717 produced 21.15 g/L natamycin at 168 h, representing a 74.2% increase over strain 712. WGS identified 320 high-confidence variants in strain 717, including 241 SNPs and 79 InDels. Most variants were located in non-coding, intergenic, or upstream regulatory regions (298/320, 93.1%), whereas only 22 variants (6.9%) were located in coding sequences or were annotated as functionally relevant coding-region variants. RNA-seq at 72 h and 120 h revealed extensive stage-dependent transcriptional reprogramming, with 817 and 2097 differentially expressed genes (DEGs), respectively. The early response was characterized by upregulation of ABC transporter components, including ATP-binding proteins and permeases, whereas the later response involved increased expression of multiple substrate-binding proteins and genes associated with propionate metabolism and methylmalonyl-CoA supply. The natamycin biosynthetic gene cluster (BGC) was transcriptionally activated in strain 717, particularly at 120 h. Integrated WGS and RNA-seq analysis mined several candidate genes, including B1H19_RS30395, B1H19_RS33295, and B1H19_RS08875, which link regulatory variation to altered phosphate metabolism, redox balance, and competing secondary metabolic pathways. These results suggest that the high natamycin productivity of strain 717 is associated with regulatory-region variation and phase-specific transcriptional rewiring rather than extensive coding-sequence changes. Coordinated changes in membrane transport, nutrient acquisition, precursor

supply, redox metabolism, and natamycin BGC expression may collectively support enhanced production. The candidate genes and functional modules identified here provide potential targets for mechanism-guided improvement of natamycin-producing strains.

Key words

Natamycin; *Streptomyces gilvosporeus*; ARTP mutagenesis; whole-genome resequencing; Transcriptomics

1 Introduction

Natamycin, also known as pimaricin, is a polyene macrolide polyketide produced by *Streptomyces* species. Because of its broad antifungal activity against yeasts and filamentous fungi, high efficacy and low toxicity[1-3], natamycin has become an important natural antifungal agent for food preservation and has also been applied in the treatment of superficial fungal infections, including fungal keratitis[2,4]. As a typical type I polyketide synthase (PKS)-derived secondary metabolite, natamycin biosynthesis involves assembly by large modular PKS enzymes, multiple post-PKS tailoring steps, glycosylation, and transmembrane export. Its production is therefore controlled by multiple interconnected processes including precursor supply, NADPH availability, membrane transport, morphological differentiation, and nutrient-responsive regulatory networks, particularly phosphate homeostasis[5,6]. Elucidateing how these regulatory layers are rewired in the high-producing strain 717 is critical both for industrial strain improvement and for dissecting regulatory principles underlying secondary metabolism in actinomycetes.

Random mutagenesis remains a major strategy for developing high-yield natamycin producers. Ultraviolet mutagenesis, nitrosoguanidine treatment, ARTP mutagenesis, protoplast fusion, and genome shuffling have all been used to improve natamycin production[7-12]. For example, UV-ARTP-DES compound mutagenesis improved natamycin production to 1.64 g/L in *Streptomyces natalensis* HW-2, and transcriptome analysis implicated that gene transcript abundance of metabolism with relevant glycolysis, the pentose phosphate pathway, the tricarboxylic acid cycle, fatty acid metabolism, and natamycin biosynthesis underwent dramatic changes[8].

Genome shuffling generated *Streptomyces gilvosporeus* GS 4-21, which reached 4.69 g/L natamycin in shake-flask fermentation at 96 h [9]. Phosphate-tolerance screening and transcriptome-guided analysis further linked improved production to remodeling of carbon, nitrogen, phosphate metabolism, and natamycin biosynthesis [10]. In addition, ARTP combined with 2-deoxyglucose tolerance screening enriched carbon-metabolism-remodeled mutants with enhanced natamycin output [11]. These studies show that random mutagenesis can generate improved producers, but most analyses remain descriptive, linking yield changes to broad transcriptomic patterns without resolving which genetic changes may trigger stage-specific transcriptional responses and how these responses converge on final product accumulation.

Meanwhile, increasing understanding of the natamycin BGC and its regulatory mechanism has enabled more rational strain- and process-engineering strategies, including Strengthening precursor supply, enhancing product efflux capacity, intervening morphology-metabolism coupling and regulating phosphorus response[2,13]. Manipulating precursor pools can also improve production: propanol supplementation increased acetyl-CoA availability and raised the maximal natamycin titer to 10.38 g/L, 17% above the control[13]. In Streptomyces chattanoogensis L10, ScnA/ScnB were identified as primary natamycin exporters; deletion of both genes reduced efflux efficiency from approximately 86% in the wild type to approximately 54% in the double mutant, while total natamycin production remained similar[14]. Regulatory systems also contribute to natamycin production. The MacRS two-component system affects both morphological differentiation and natamycin biosynthesis in *S. gilvosporeus*[15]. In addition, a reporter-guided screening system using the expression of large PKS genes and the entire natamycin BGC as readouts increased natamycin production by 110%, 230%, and 340% after successive rounds of mutagenesis and screening, respectively, and these high-yield phenotypes were accompanied by coordinated transcriptional enhancement of the whole BGC[16]. Phosphate regulation is another important constraint, as the PhoR-PhoP system participates in the phosphate-responsive control of pimaricin or natamycin biosynthesis [5,10].

However, most of these targets were derived from single-gene functional validation. By contrast, high-yield phenotypes generated by physical mutagenesis approaches such as ARTP likely arise from coordinated, multi-locus remodeling across regulatory, metabolic, and biosynthetic layers[7-9]. In this regard, Single-target engineering often fails to recapitulate the overall advantages of mutagenized strains and may even be counterproductive because of metabolic imbalance. Therefore, an integrative strategy is urgently required to reliably identify genome-wide high-confidence variants, capture the variant-triggered transcriptional dynamics during critical fermentation phases, and integrate these two molecular layers into established metabolic and regulatory networks[17-19]. This multi-omics framework further enables the screening of core hub genes and functional modules responsible for governing high-yield phenotype. On this basis, we proposed and applied a tri-layer mechanistic framework linking genetic variation, stage-specific transcriptional responses, and the high-yield phenotype. Starting with *S. gilvosporeus* 712 as the parent strain, we generated a stably high-yield mutant designated 717 via ARTP mutagenesis. Whole-genome resequencing (WGS) was subsequently conducted to screen high-confidence variants within strain 717. The identified variants were markedly enriched in non-coding and upstream regulatory regions, implying that elevated product titers are mainly governed by remodeled global regulatory circuits instead of modified structural genes. In parallel, transcriptome sequencing (RNA-seq) at 72 h (corresponding to the biosynthetic initiation phase) and 120 h (corresponding to the high-production accumulation phase) revealed distinct stage-specific transcriptional features. The response of former stage was characterized by coordinated upregulation of ABC transporter components, including ATP-binding subunits and transmembrane permeases, whereas the later response shifted toward enhanced nutrient acquisition mediated by multiple substrate-binding proteins (SBPs), together with strong activation of propionate metabolism and methylmalonyl-CoA supply pathways. Most notably, gene-level integration of WGS and RNA-seq datasets, coupled with comprehensive functional annotation, enabled the prioritization of core candidate genes responsible for the high-yield phenotype. These genes did not act as isolated

loci but formed functional modules closely connected to the natamycin biosynthetic gene cluster (BGC) in the protein-protein interaction (PPI) network, thereby providing a molecular bridge between genetic variation and the final production phenotype[20,21].

This work not only establishes a systematic multi-omics evidence chain to decipher the molecular mechanisms governing high natamycin biosynthesis, but also constructs a general methodological framework applicable to the rational engineering of other polyketide antibiotic producers. The candidate genes and metabolic pathways uncovered in this work offer precise, experimentally verifiable targets for subsequent genome editing, enzyme modification, and AI-assisted combinatorial engineering. Collectively, these findings advance the shift of natamycin biosynthesis from empirical strain optimization to mechanism-informed next-generation intelligent biomanufacturing.

2 Materials and methods

2.1 Strains

S. gilvosporeus 712 was used as the parental strain and was provided by Shandong Freda Biotechnology Co., Ltd.

2.2 Media and culture conditions

Solid seed medium (SGM) was formulated as follows (per liter): 5 g peptone, 3 g yeast extract, 3 g NaCl, 10 g glucose, and 20 g agar. The medium was sterilized at 115 ºC for 30 min and poured into petri dishes before cooled to room temperature. Screening resistance plates (SGMR) were prepared by supplementing SGM with streptomycin sulfate at the indicated final concentration.

The liquid seed medium contained (per liter) 3 g peptone, 3 g yeast extract, 5 g NaCl, and 10 g glucose. The pH was adjusted to 7.4 with 15% NaOH before sterilization. Seed cultures were prepared in 250 mL shake flasks containing 30 mL medium and approximately 10 glass beads to disperse mycelia. The medium was sterilized at 115 ºC for 30 min as well. Cultures were inoculated with mycelia at 1% (v/v) and incubated at 29 ºC and 200 r/min for approximately 30 h before used as seed culture.

The fermentation medium contained (per liter) 70 g glucose, 20 g peptone, 4.5 g yeast extract, 2 g NaCl, 1 g citric acid, and 1 g magnesium sulfate. The pH was adjusted to 7.5 with 15% NaOH before sterilization. Each flask contained 10 glass beads and one drop of antifoam. Both medium and glucose stock solution were sterilized at 115 ^{o}C for 30 min.

Shake-flask fermentation was performed in 250 mL flasks containing 30 mL medium. Cultures were inoculated at 1% (v/v) and incubated at 29 ^{o}C and 200 r/min. Three biological replicates were used for each group and samples were collected every 12 h until 168 h.

For 15 L bioreactor fermentation, the initial working volume was 7.0 L, and fed-batch fermentation was used. During fermentation, the pH was maintained at 6.0 and residual glucose was maintained at 20-25 g/L. Dissolved oxygen was controlled at approximately 35% by coordinating agitation and aeration. The agitation speed was gradually increased from 200 r/min to a maximum of 700 r/min, and the aeration rate was increased from 10 L/min to a maximum of 20 L/min according to fermentation stage. A 50% glucose solution was fed according to residual glucose levels. Samples for transcriptome analysis were collected at 72 h and 120 h.

2.3 Natamycin sample processing and maximum absorption measurement

Natamycin concentration was determined using a UV spectrophotometric difference method. For routine measurement, 1.0 mL fermentation broth was transferred into a 10 mL volumetric flask, diluted to volume with methanol, and extracted by ultrasonication for 15 min. Then, 1.0 mL of the mixture was centrifuged at 8000 r/min for 2 min. A 100 microliter aliquot of the supernatant was transferred into a 10 mL volumetric flask and diluted to volume with 70% methanol. Absorbance was measured at 296, 303, 311, and 318 nm using 70% methanol as the blank. Natamycin in methanol shows characteristic absorption peaks at approximately 303 and 318 nm, with an absorption valley at approximately 311 nm.

2.4 Standard curve for natamycin quantification

Natamycin standard (98% purity, Shanghai Macklin Biochemical Co., Ltd.) was dissolved in methanol to prepare a stock solution and serially diluted with 70%

methanol to generate standard solutions. To reduce matrix effects, standard solutions were processed using the same pretreatment procedure as fermentation samples. Briefly, 1.0 mL standard solution or matrix-spiked standard was transferred into a 10 mL volumetric flask, diluted to volume with methanol, and extracted by ultrasonication for 15 min. Then, 1.0 mL of the mixture was centrifuged at 8000 r/min for 2 min. A 100 microliter aliquot of the supernatant was diluted to 10 mL with 70% methanol and mixed before measurement.

The net differential absorbance was calculated as:

$$\Delta A = \frac{A_{303} + A_{318}}{2} - \frac{A_{296} + A_{311}}{2}$$

The standard curve was established by linear regression using natamycin concentration C (g/L) as the independent variable and Delta A as the dependent variable:

$$\Delta \mathrm{A} = \mathrm{kC} + \mathrm{b}$$

The natamycin concentration in samples was calculated as:

$$\mathrm{C} = \frac{\Delta \mathrm{A} - \mathrm{b}}{\mathrm{k}}$$

All standards and samples were measured under the same instrumental conditions using a 1 cm quartz cuvette. Each standard concentration was measured in triplicate. A middle-concentration standard was included in each analytical batch as a quality-control sample. If sample absorbance exceeded the linear range of the instrument, for example 0.2-0.8, the sample was diluted and remeasured, and the final concentration was calculated according to the dilution factor.

2.5 Dry cell weight measurement

Dry cell weight (DCW) was determined gravimetrically. A 2 mL aliquot of fermentation broth was centrifuged at 8000 x g for 10 min to collect mycelial biomass. The pellet was washed with distilled water, dried at 105 ºC to constant weight, and weighed. Results were expressed as g/L or as dry weight corresponding to the sampled volume.

2.6 ARTP mutagenesis, lethality determination, and screening of high-producing mutants

ARTP mutagenesis was performed using *S. gilvosporeus* 712 as the parental strain. Before mutagenesis, colonies grown on solid medium at 28 ºC for 5-7 d, when spores had matured and turned white, were washed with sterile saline into a flask containing sterile glass beads. The spore suspension was dispersed by shaking and filtered through sterile absorbent cotton to remove mycelial fragments. Spores were counted using a hemocytometer and diluted with sterile saline to 10^8 CFU/mL.

Before ARTP treatment, the operating chamber was sterilized by ultraviolet irradiation for 30 min. High-purity helium (99.999%) was used as the working gas. The power was set to 120 W, the gas flow rate to 10 L/min, and the distance between the plasma source and the sample carrier to 2 mm. A 10 microliter aliquot of spore suspension was spread evenly on a sterile metal carrier. Treatment times of 0, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, and 120 s were tested. Immediately after treatment, the carrier was transferred into an EP tube containing 990 microliters sterile saline and vortexed for 2 min to release spores. The suspension was serially diluted, plated on solid medium, and incubated at 28 ºC for 5-7 d before colony counting. The lethality rate was calculated as:

Lethality (%) = (1 - viable cells after mutagenesis / viable cells before mutagenesis) x 100%.

The treatment condition close to 90% lethality was selected for subsequent mutant screening. After ARTP mutagenesis under the selected condition, spores were recovered, serially diluted, and plated on solid medium. Colonies were incubated at 28 ºC for 5-7 d. Well-separated colonies with clear morphology and good growth were purified to construct a mutant library. Mutants with slow growth, unstable morphology, or poor sporulation were removed during primary plate screening. Candidate mutants were then cultivated in seed medium and transferred to fermentation medium using the same inoculation ratio for shake-flask rescreening. Natamycin production was measured at predefined time points using the UV difference method and compared with that of strain 712. Mutants with significantly

higher natamycin production and stable growth phenotypes were retained. Candidate high-producing strains underwent several serial passages and re-evaluated via shake-flask fermentation to confirm genetic stability. The most stably high-yield mutant was chosen for subsequent analyses.

2.7 Whole-genome resequencing and variant analysis

Whole-genome resequencing was performed for *S. gilvosporeus* 712 and the ARTP-derived high-producing mutant 717. Genomic DNA was extracted from both strains and used to construct paired-end sequencing libraries. Raw reads were quality evaluated and filtered to obtain clean reads for bioinformatic analysis. Clean reads were mapped to the reference genome sequence, and SNPs and InDels were identified and annotated based on the mapping results. DNA-level differential genes and functional annotations were then obtained.

High-confidence variants were annotated using the reference genome annotation file. SnpEff was used to classify variants according to genomic location, including coding regions, non-coding regions, and intergenic regions, and to predict potential effects, including synonymous, missense, nonsense, and frameshift variants[17], The resulting set of functionally annotated variant genes was used for subsequent pathway-level interpretation and integrated multi-omics candidate screening.

2.8 Transcriptome sequencing and differential expression analysis

Biomass samples from strains 712 and 717 were collected at 72 h and 120 h, representing key fermentation stages. Three biological replicates were used for each group. Total RNA was extracted and assessed for concentration, purity, and integrity before library construction and Illumina sequencing. Raw reads were processed by adapter removal and low-quality read filtering to generate clean reads, following the Trimmomatic workflow[22].Clean reads were mapped to the *S. gilvosporeus* F607 reference genome (ASM208219v1, NZ_CP020569.1) using Bowtie2[23], Gene-level read counts were generated using HTSeq-count [24]. Gene expression levels were represented by read counts and FPKM values. Read counts were used for differential expression analysis, and FPKM values were used for expression visualization.

Differential expression analysis was performed using DESeq2[18]. Multiple testing was corrected using the Benjamini-Hochberg method. DEGs were defined using the thresholds |log2FC| >= 1 and adjusted P value (padj) < 0.05. GO and KEGG enrichment analyses were then performed for functional annotation and pathway interpretation[25,26].

2.9 Multi-omics integration strategy

To integrate the genetic basis and transcriptional response of the high-producing mutant, high-confidence variant genes identified by WGS were compared with DEGs identified by RNA-seq. First, SNPs and InDels were annotated using SnpEff [17], and genes carrying non-synonymous, nonsense, frameshift, coding-region, or potentially regulatory-region variants were prioritized. DEGs were then used as transcriptional evidence and were evaluated according to fold change, adjusted P value, gene function, and GO/KEGG enrichment results. Genes supported by both genetic variation and differential expression were prioritized for further analysis. Genes supported by only one evidence layer but involved in carbon metabolism, nitrogen metabolism, sulfur metabolism, precursor supply, or secondary-metabolism regulation were also included as supplementary candidates[19]. Finally, GO, KEGG, and KO annotations were integrated to classify candidate genes and potential regulatory nodes associated with the high-yield phenotype.

3 Results

3.1 Isolation and phenotypic validation of a high-producing mutant

To determine the optimal ARTP mutagenesis dose, the lethality of *S. gilvosporeus* 712 was measured across different treatment times (Fig. 1A). The lethality rate increased monotonically with treatment duration. At 100 s, lethality reached 89.7%, close to the target threshold of 90%. Accordingly, this treatment condition was selected for subsequent mutant screening.

After ARTP mutagenesis, colony purification and primary screening, 50 candidate mutants were selected for shake-flask rescreening. Ten mutants produced elevated natamycin yields relative to the parental strain (Fig. 1B). After three consecutive passages and multiple fermentation validation, mutant 717 showed the

best performance. Its natamycin titer reached 9.06 g/L, representing a 159% increase relative to strain 712 (3.49 g/L; $P < 0.01$). The high-yield phenotype of strain 717 was stably maintained across subcultures with no noticeable drop in productivity, so this strain was chosen for further multi-omics characterization and subsequent analyses.

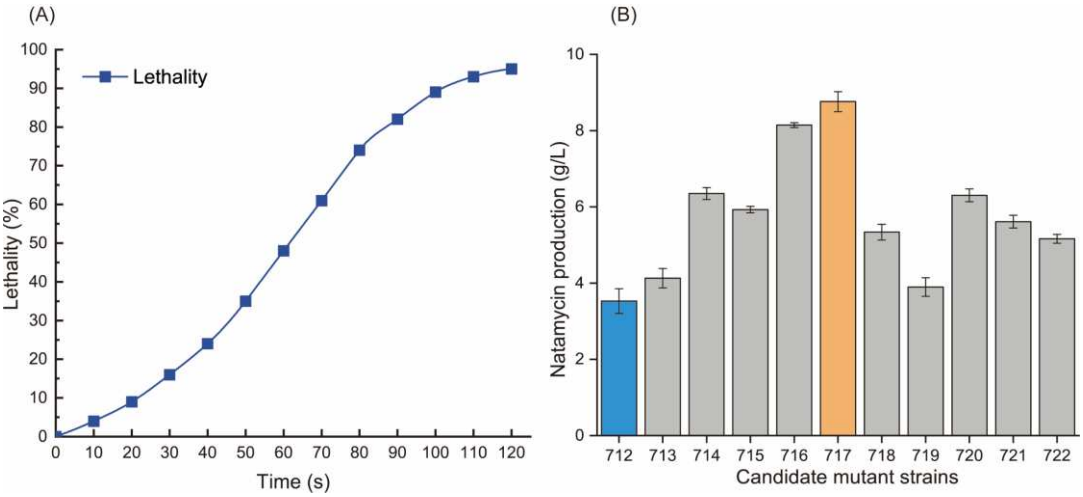


Figure 1. ARTP mutagenesis of *S. gilvosporeus* 712 and Shake-flask screening of candidate mutants. (A) Lethality curve under different ARTP treatment times. (B) Natamycin titers of different mutant strains after ARTP mutagenesis.

3.2 Morphological and physiological features of mutant 717

Compared with parental strain 712, mutant 717 showed clear colony morphology changes on solid medium. The mutant colonies were more wrinkled or granular, had more irregular margins, and showed increased morphological heterogeneity (Fig. 2A, B). During liquid shake-flask fermentation, mutant 717 growed faster. Its fermentation broth became dense and milky at an earlier cultivation stage compared with the parental strain 712 (Fig. 2C, D), and the DCW curve showed faster biomass accumulation in strain 717 (Fig. 2E).

Notably, this growth advantage was not obtained at the expense of secondary metabolism. For detail, at 72 h and 120 h, strain 717 produced 2.45 g/L and 8.57 g/L natamycin in shake flasks, respectively (Fig. 2F). And the specific production per biomass increased by 262.38% and 82.25% at these two time points, correspondingly. The above results demonstrated that strain 717 featured both rapid biomass buildup and improved natamycin productivity.

The production advantage of strain 717 was further validated in a 15 L

bioreactor (Fig. 2G). Strain 717 produced 12.09 g/L natamycin at 72 h, a 98.8% elevation above strain 712 and reached 16.26 g/L at 120 h, with an 80.1% increase. The natamycin titer peaked at 21.15 g/L at the terminal time point of 168 h, representing a 74.2% elevation relative to strain 712. These results confirmed that the stable retention of mutant 717’s superior productivity during fermentation scale-up, which establishes a reliable phenotypic framework to unravel the underlying molecular mechanisms.

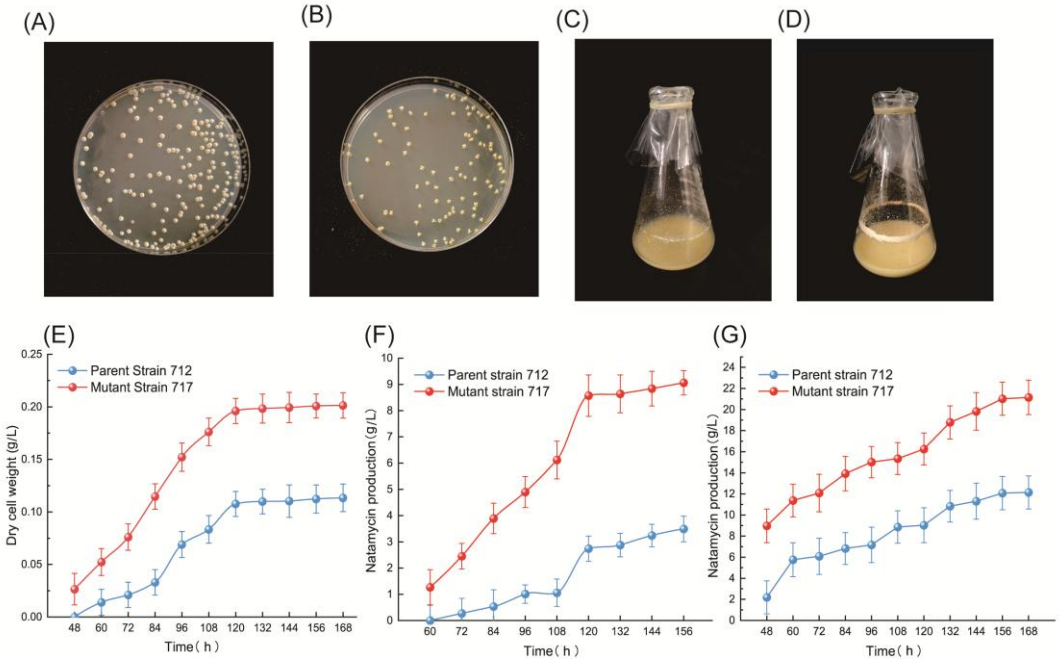


Figure 2. Growth and fermentation characteristics of high-producing mutant 717 compared with strain 712. (A) Colony morphology of strain 712 on solid medium. (B) Colony morphology of strain 717 on solid medium. (C) Shake-flask fermentation broth of strain 712 at 120 h. (D) Shake-flask fermentation broth of strain 717 at 120 h. (E) DCW changes during shake-flask fermentation. (F) Natamycin production during shake-flask fermentation. (G) Natamycin titer profiles in 15 L bioreactor fermentation.

3.3 Whole-genome re-sequencing revealed variant enrichment in non-coding and regulatory regions

To characterize the genetic basis of the high-yield phenotype, strains 712 and 717 were subjected to Illumina paired-end WGS and mapped to the *S. gilvosporeus* F607 reference genome (ASM208219v1). The sequencing data had an average depth of 116x, 99.64% coverage at >=30x, and a mapping rate of 96.06%, which met the requirements for high-confidence variant detection.

After quality control and SnpEff annotation (Fig. 3A), 320 high-confidence genetic variants were identified in strain 717 relative to strain 712. These included 241 SNPs (75.3%) and 79 InDels (24.7%). Variant distribution was strongly biased by genomic region. Only 22 variants (6.9%) mapped to coding sequences or were predicted to exert coding/functional impacts, while the remaining 298 variants (93.1%) resided within non-coding, intergenic, or upstream regulatory regions (Fig. 3B). Among coding-region or functionally annotated variants (Fig. 3C), the main predicted effects included synonymous, missense, frameshift, stop-gained, and disruptive in-frame insertion variants. The above results demonstrated that strain 717's superior productivity was unlinked to large-scale coding-sequence alterations, and instead correlated better with regulatory region polymorphisms and rewiring of transcriptional networks.

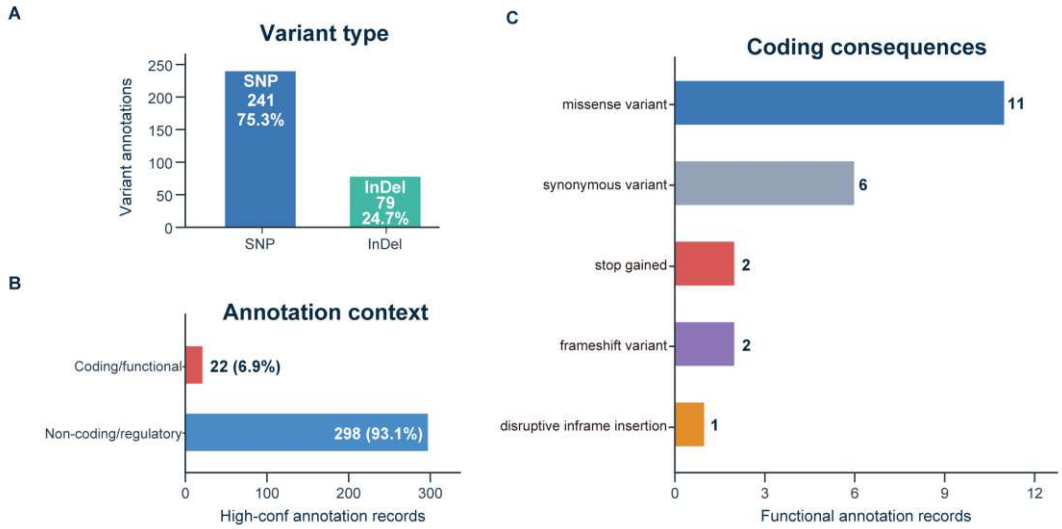


Figure 3. WGS variant features of mutant 717. (A) Distribution of high-confidence variants by variant type, including 241 SNPs and 79 InDels. (B) Genomic location distribution of high-confidence variants, showing that most variants were located in non-coding or regulatory regions. (C) Predicted consequences of coding-region or functionally annotated variants, including missense, synonymous, stop-gained, frameshift, and disruptive in-frame insertion variants. High-confidence variants were defined as SnpEff-annotated records satisfying PASS, DP >= 10, GQ >= 30, and AF >= 0.8.

3.4 RNA-seq revealed stage-dependent transcriptional reprogramming

To explore the transcriptional basis for elevated natamycin titers, RNA-seq was carried out for strains 712 and 717 at 72 h and 120 h, corresponding to the biosynthesis initiation stage and high-yield accumulation stages respectively. Each

group contained three independent biological replicates. DESeq2 analysis using |log2FC| >= 1 and padj < 0.05 showed extensive transcriptional differences between strains 717 and 712 at both time points (Fig. 4A, B).

A total of 817 DEGs were screened at 72 h, among which 428 were upregulated and 389 were downregulated. At 120 h, the number of DEGs rose substantially to 2097, consisting of 826 upregulated and 1271 downregulated genes. Downregulated genes showed a more prominent increment compared with upregulated ones, suggesting widespread transcriptional remodeling during the later fermentation stage.

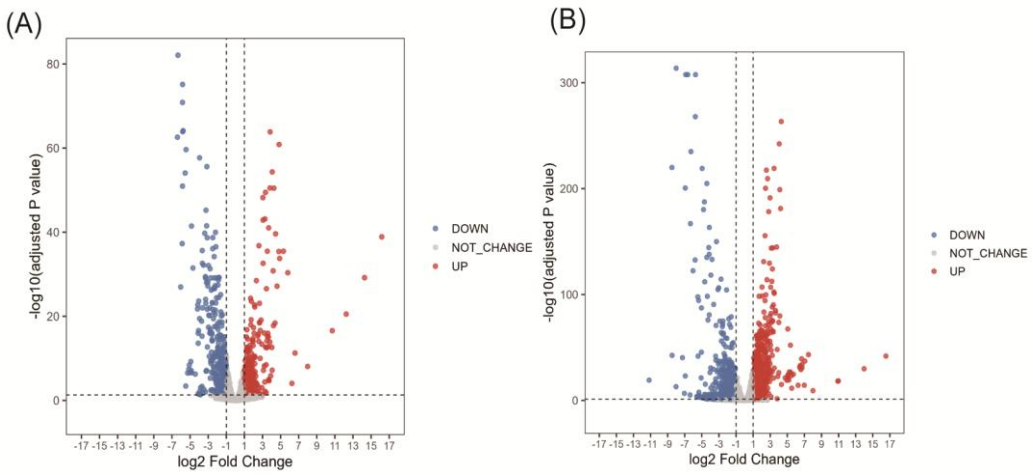


Figure 4. Volcano plot of DEGs between the mutant and parental strains at different fermentation stages of 72 h (A) and120 h (B) .

3.4.1 Functional enrichment features of DEGs

To summarize the key transcriptional responses in strain 717, representative protein-coding DEGs with significant expression alterations and explicit functional annotations were chosen at 72 h and 120 h (Table 1). At the two fermentation stages, genes linked to transcriptional regulation, material transport, tolerance, and genetic rearrangement displayed prominent expression shifts. An ABC transporter ATP-binding protein gene, B1H19_RS06390, was strongly upregulated (log2FC = 6.60) at 72 h, indicating early activation of energy-dependent membrane transport. At the 120 h fermentation time point, numerous genes linked to redox processes, metal binding and cellular metabolism were differentially transcribed, with a gene for cupin domain-containing protein exhibiting profound downregulation.At the 120 h fermentation time point, multiple genes linked to redox processes, metal binding and

cellular metabolism were differentially transcribed, with a gene for cupin domain-containing protein exhibiting profound downregulation (B1H19_RS08440, log2FC = -8.50). These representative DEGs indicated that strain 717 underwent stage-dependent transcriptional remodeling during fermentation.

Table 1. Top annotated DEGs between strain 717 and strain 712 at 72 h and 120 h.

Time	GeneID	log2FC	P-adj	Description	Basemean
72h	B1H19_RS20885	16.1741	7.77E-39	aminoglycoside N-acetyltransferase AAC(3)-IVa	6350
	B1H19_RS20875	14.2721	1.63E-28	LuxR C-terminal-related transcriptional regulator	1699
	B1H19_RS20895	12.2463	1.36E-19	conjugal transfer transcriptional regulator TraJ	417
	B1H19_RS20865	10.7085	5.62E-17	recombinase family protein	143
	B1H19_RS06390	6.5952	1.95E-16	ABC transporter ATP-binding protein	835
120h	B1H19_RS20885	16.5395	3.12E-42	aminoglycoside N-acetyltransferase AAC(3)-IVa	8083
	B1H19_RS20875	13.9864	2.89E-30	LuxR C-terminal-related	1377

			transcriptional regulator	
B1H19_RS20895	10.9825	4.35E-19	conjugal transfer transcriptional regulator TraJ	171
B1H19_RS20865	10.9108	6.15E-19	recombinase family protein	163
B1H19_RS08440	-8.4993	5.65E-217	cupin domain-containing protein	2265

Note: Padj indicates Benjamini-Hochberg-adjusted P value, and BaseMean indicates normalized mean expression. Entries lacking clear functional annotation, including novel transcript or antisense transcript entries, were excluded from this representative list.

Subsequent GO and KEGG enrichment results demonstrated that strain 717's DEGs correlated with different functional groups across the two fermentation stages (Fig. 5).

At the 72 h fermentation time point, DEGs were primarily enriched in ABC transporter functional terms, including ATP-binding proteins, permeases, and substrate-binding proteins. By 120 h, substrate-binding protein genes showed more pronounced differential expression, alongside visible expression changes in genes linked to propionate metabolism. The above results demonstrated that strain 717 displayed distinct transcriptional signatures at the early versus late fermentation stages.

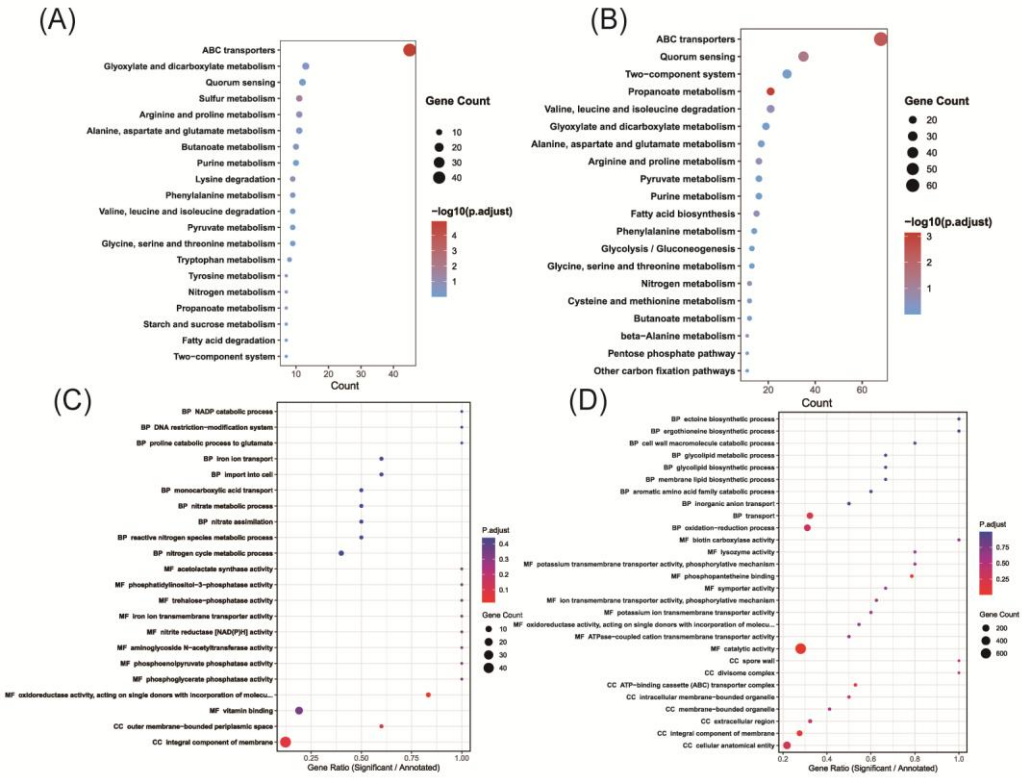


Figure 5. Functional enrichment analysis of DEGs between mutant 717 and parental strain 712 at different fermentation stages. (A) KEGG pathway enrichment at 72 h. (B) KEGG pathway enrichment at 120 h. (C) GO enrichment at 72 h. (D) GO enrichment at 120 h.

A detailed analysis of enriched pathways revealed that genes associated with ABC transporters and propionate metabolism constituted the primary functional groups with significant expression changes in strain 717, as shown in table 2. The KEGG pathway "ABC transporters" (map02010) was significantly enriched at both 72 h and 120 h, indicating sustained changes in transmembrane transport and environmental adaptation.

Table 2. Differentially expressed genes in functional modules between high-producing strain 717 and parental strain 712 at different fermentation stages.

GeneID	log2FC	P-adj	Description	Basemean
ABC transporter 72h				
B1H19_RS06395	5.79	1.74E-19	ABC transporter family substrate-binding protein	1711
B1H19_RS06390	6.60	1.95E-16	ABC transporter ATP-binding protein	835
B1H19_RS06385	3.19	4.24E-07	ABC transporter permease	374
B1H19_RS31475	3	0.009164	ABC transporter permease	619
B1H19_RS31480	2.5	0.033164	ABC transporter ATP-binding protein	422
B1H19_RS31485	2.2988	0.047914	aliphatic sulfonate ABC transporter substrate-binding protein	1142
ABC transporter 120h				
B1H19_RS06395	5.06	4.3×10^{-47}	ABC transporter family substrate-binding protein	327
B1H19_RS15915	3.62	3.0×10^{-70}	2-aminoethylphosphonate ABC transporter substrate-binding protein	395

B1H19_RS23935	2.23	5.4×10^{-154}	phosphate ABC transporter substrate-binding protein PstS	3.8×10^{4}
B1H19_RS29515	1.83	2.0×10^{-30}	Glutamate ABC transporter substrate-binding protein	683
B1H19_RS30480	1.58	6.3×10^{-30}	sugar ABC transporter substrate-binding protein	1488
Propionate metabolism				
B1H19_RS30825	6.59	2.66E-36	acyl-CoA carboxylase subunit β	2543
B1H19_RS30810	6.266967	8.37E-19	biotin carboxylase N-terminal domain-containing protein	323
B1H19_RS27660	1.12	2.44E-25	methylmalonyl-CoA epimerase	750
B1H19_RS15465	1.052608	1.09E-07	CoA-acylating methylmalonate-semialdehyde dehydrogenase	1135
B1H19_RS03755	-2.37104	1.34E-07	acetate--CoA ligase	16
Quorum sensing				
B1H19_RS31640	3.45	2.88E-197	phosphocholine-specific phospholipase C	42752
B1H19_RS06555	2.40	9.96E-188	phospholipase C	42204

B1H19_RS30795	6.50	4.79E-40	class I adenylate-forming enzyme family protein	3973
B1H19_RS19520		9.60E-96	alkaline phosphatase family protein	898
B1H19_RS35805		5.91E-30	M4 family metallopeptidase	1119
		Biosynthesis of various other secondary metabolites		
B1H19_RS08420	-4.70	8.53E-135	2-amino-3,7-dideoxy-D-threo-hept-6-ulosonate synthase	1358
B1H19_RS08425	-3.61	4.71E-61	3-dehydroquinate synthase II family protein	826
B1H19_RS19875	2.43	2.04E-51	2,3-diaminopropionate biosynthesis protein SbnA	867
B1H19_RS08970	2.88	7.41E-19	2,3-diaminopropionate biosynthesis protein SbnA	292
B1H19_RS09140	-1.45	1.90E-27	lysine N(6)-hydroxylase/L-ornithine N(5)-oxygenase family protein	500

At 72 h, ABC transporter-related DEGs included genes encoding ATP-binding proteins, permeases, and substrate-binding proteins. B1H19_RS06390, encoding an

ABC transporter ATP-binding protein, was upregulated with a log2FC of 6.60, and B1H19_RS06385, encoding a permease, was upregulated with a log2FC of 3.19. Another group of transporters, namely B1H19_RS31475, B1H19_RS31480 and B1H19_RS31485, exhibited significant upregulation. Collectively, these findings indicate that strain 717 activates energy-coupled membrane transport at the early stage of fermentation.

At 120 h, ABC transporter-related DEGs were mainly concentrated in substrate-binding proteins (SBPs), including B1H19_RS06395, B1H19_RS15915, B1H19_RS23935, B1H19_RS29515, and B1H19_RS30480. These genes are annotated as SBPs responsible for general ABC transport, 2-aminoethylphosphonate transport, phosphate transport, glutamate transport, and sugar transport. Their upregulation suggests an improved capacity for substrate recognition and translocation during the late fermentation stage.

3.4.2 Activation of the natamycin biosynthetic gene cluster

Natamycin biosynthesis in *Streptomyces* is generally mediated by large PKS genes, such as *pimS0-pimS4*, together with post-PKS tailoring enzymes and cluster-associated regulators, including LuxR-family regulatory proteins such as PimR/PimM. In *S. gilvosporeus*, extra regulatory modules such as two-component systems have also been linked to natamycin biosynthesis[15,16]. Expression profiles of core structural, tailoring, regulatory and transporter genes in the natamycin BGC were extracted and visualized to link transcriptional variation to natamycin formation.

Comparative analysis of natamycin biosynthetic genes between strains 712 and 717 revealed that fifteen genes within the sgn cluster exhibited significantly higher transcript abundances in strain 717 at both 72 h and 120 h (Fig. 6). Time-point internal comparison via heatmap further revealed that a greater number of genes possessed positive $\log_2$FC values at 120 h, demonstrating that the transcriptional superiority of strain 717 was intensified at the late fermentation stage. The upregulated gene set covered core PKS components, tailoring enzyme coding sequences, regulatory factors, and transporter genes. Taken together, these observations reveal widespread transcriptional activation of the natamycin BGC in

strain 717, consistent with its robust natamycin overproduction trait.

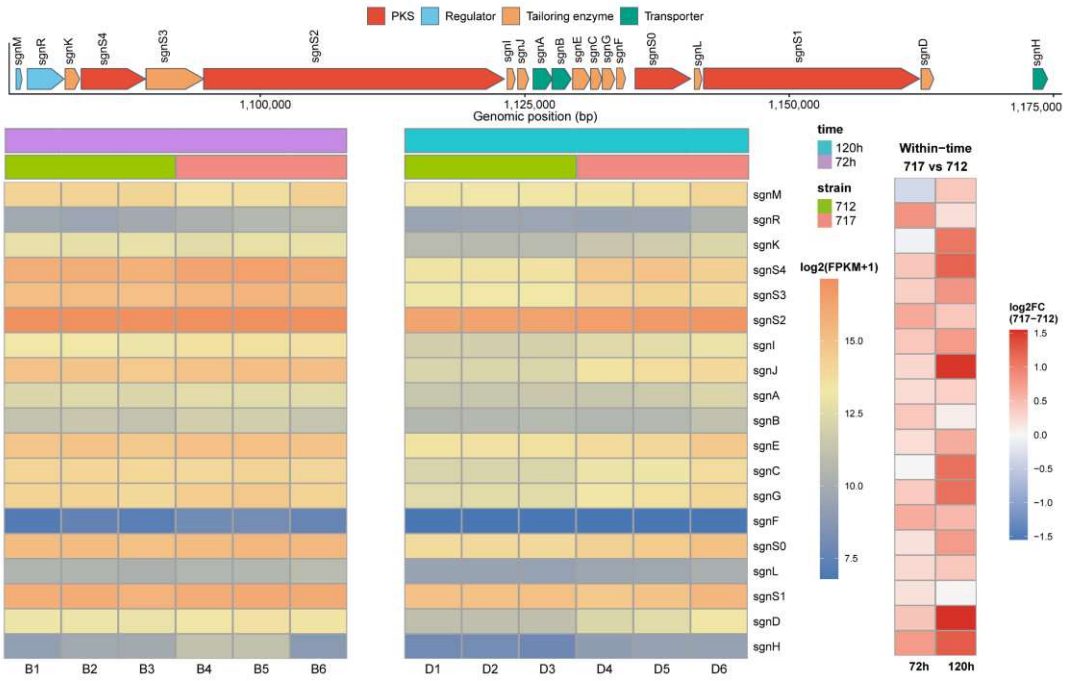


Figure 6. Expression profile of the natamycin BGC. The upper panel shows gene arrangement, transcriptional orientation, and functional classification of the natamycin BGC. The central heatmap shows gene expression levels across samples, with color indicating log2(FPKM + 1). The right heatmap shows expression differences between strain 717 and strain 712 at the same time point.

3.5 Integrated WGS and RNA-seq analysis identified candidate genes and functional modules associated with high production

For the identification of genes underlying the high-yield trait, high-reliability variant loci from WGS were cross-compared with RNA-seq DEGs on a gene-wide basis, followed by comprehensive interpretation combining functional annotation and metabolic pathway profiles. WGS identified 320 high-confidence variants in strain 717 relative to strain 712, and these variants were strongly biased toward non-coding or upstream regulatory regions (298/320). In parallel, RNA-seq identified 817 DEGs at 72 h and 2097 DEGs at 120 h. The coexistence of regulatory region mutations and widespread transcriptional divergence suggests that strain 717 undergoes systematic transcriptomic reprogramming at both the biosynthesis initiation and metabolite accumulation phases.

Subsequently, protein-protein interaction (PPI) analysis was conducted on the candidate genes screened via multi-omics integration, as well as key genes involved

in natamycin biosynthesis (Fig. 7). The interaction network revealed that these candidate genes did not function independently; instead, they assembled into multiple functional modules that physically link to the natamycin biosynthetic gene cluster (BGC). Core natamycin biosynthetic genes occupied the central hub of the network, while peripheral nodes corresponded to genes governing transport and export, transcriptional regulation, redox homeostasis, and precursor supply metabolism.

Transport-related proteins, including RS08945, RS03070, RS30775, and RS05960, belonged to the ABC or MFS transporter families. These genes potentially facilitate substrate uptake, transmembrane translocation, and product efflux throughout the onset of natamycin biosynthesis and sustained high-yield production. Regulatory proteins encoded by RS08955, RS08860, RS32430 and RS20875 constituted a putative transcriptional regulatory layer. This observation aligns with the hypothesis that random mutagenesis boosts high-yield biosynthesis via rewiring of the global regulatory network. RS30395, RS33295, RS34740, and multiple FAD- or NAD(P)-dependent oxidoreductase nodes formed a subnetwork associated with precursor supply and redox metabolism, indicating coordinated changes in phosphate homeostasis, amino acid metabolism, acyl-CoA turnover, and redox balance.

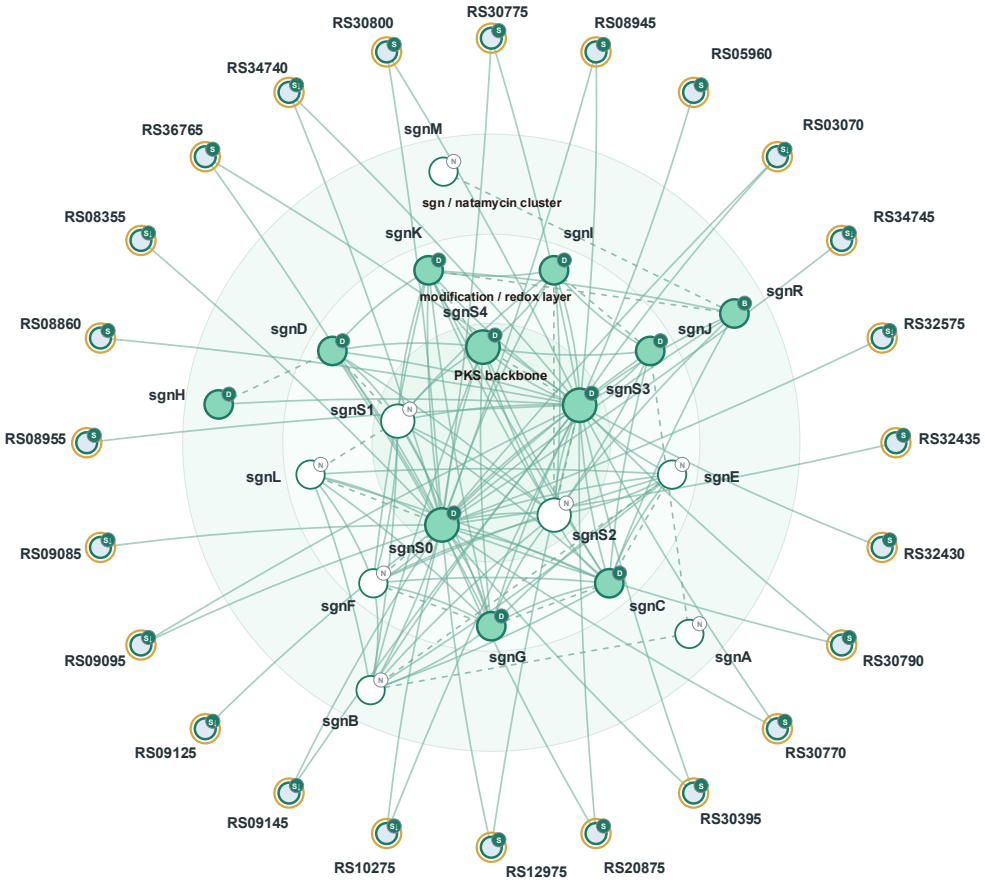


Figure 7. PPI network centered on the natamycin *sgn* gene cluster. The network was constructed using the natamycin *sgn* BGC as the central module. The central concentric region contains *sgn* genes, classified into PKS backbone, modification/redox, regulatory, and transport-related genes. Peripheral nodes represent non-*sgn* genes connected to the DEG-related PPI subnetworks at the early and late fermentation stages. Green nodes indicate *sgn* genes, blue-gray nodes indicate peripheral connector genes, and gold outlines indicate genes carrying WGS-annotated variants in mutant 717. Expression labels indicate RNA-seq differential expression categories: S, significantly upregulated at both comparisons; S down, significantly downregulated at both comparisons; B, significant only at the early stage; D, significant only at the late stage; and N, not significant but retained as a core *sgn* gene. Solid lines represent PPI edges, and dashed lines represent the genomic order of *sgn* genes.

Three candidates were prioritized from the gene set jointly validated by both WGS and RNA-seq evidence (Table 3). *B1H19_RS30395* encodes a haloacid dehalogenase (HAD) family hydrolase and carried a stop-gained SNP. This gene was highly expressed in strain 717 ($log_2FC = 2.55$; BaseMean approximately 3.77 x 10^4). HAD superfamily proteins often act as metal-dependent phosphatases or phosphotransfer-related hydrolases and can participate in sugar-phosphate balance, metabolite detoxification, and inorganic phosphate homeostasis[27]. This expression pattern reveals the co-occurrence of a truncation mutation and elevated transcript

abundance, suggesting that this gene participates in phosphorylated metabolite turnover and intracellular inorganic phosphate (Pi) homeostasis, which in turn modulates carbon resource partitioning toward secondary metabolic pathways[19]. For Streptomyces species, phosphate metabolic pathways and phosphate deprivation signal transduction exert profound control over the initiation and overall output of secondary metabolite production, such as antibiotic and polyketide synthesis, which is well illustrated via PhoP-mediated transcriptional regulation[28,29]. Therefore, the characteristic RS30395 profile combining truncating mutation with high expression likely drives aberrant turnover of phosphorylated metabolites and disrupted phosphate signal transduction in high-producing strains at key fermentation stages, which in turn modulates intracellular resource allocation and the activation threshold of secondary metabolism.

B1H19_RS33295, which encodes a protein harboring an NAD(P)-binding domain, harbors a disruptive in-frame insertion and exhibits transcriptional upregulation in strain 717 (log2FC = 1.43; BaseMean approximately 242). Since ketoreductase domains within type I PKS assembly lines necessitate NADPH for catalysis, alterations to NAD(P)-dependent enzymes could interfere with redox balance and reduce overall biosynthetic capacity[30]. Previous studies into *S. gilvosporeus* have likewise demonstrated that boosting sugar metabolism or elevating pentose phosphate pathway flux augments natamycin biosynthesis via enhanced NAD(P) pool turnover[6,11]. Therefore, B1H19_RS33295 is a promising candidate that bridges redox metabolism and improved natamycin biosynthesis.

B1H19_RS08875 encodes a nonribosomal peptide synthetase (NRPS) and bears a frameshift InDel; such mutations typically trigger premature translational termination or complete ablation of biosynthetic function. However, the gene was transcriptionally upregulated in strain 717 ($log_2FC = 2.41$; BaseMean approximately 1736), suggesting a state in which transcriptional activation coincides with structural disruption of a competing secondary-metabolite pathway. Since the genomes of *Streptomyces* species generally encode multiple secondary-metabolite BGCs, NRPS and PKS pathways may compete with one another for precursors, ATP, reducing

equivalents, and regulatory resources [31,32]. Hence, , B1H19_RS08875 serves as a candidate locus linked to perturbed resource allocation during secondary metabolism, though its biological role needs further targeted verification.B1H19_RS08875 serves as a candidate locus linked to perturbed resource allocation during secondary metabolism, although its biological role needs further targeted validation.

Table 3. Differentially expressed genes carrying mutations

GeneID	log2FC	P-adj	Description	Variant Class	Consequence	Basemean
B1H19_RS30395	2.55	5.58E-25	HAD family hydrolase	SNP	stop_gained	37,715.21
B1H19_RS33295	1.425	7.78E-24	NAD(P)-binding domain-containing protein	INDEL	disruptive_inframe_insertion	241.92
B1H19_RS08875	2.414	4.20E-17	non-ribosomal peptide synthetase	INDEL	frameshift_variant	1,735.62

Taken together, ARTP mutagenesis introduced a spectrum of SNPs and InDels into the genome of strain 717, accompanied by genome-wide transcriptional rewiring (Fig. 8). At the 72 h time point, upregulation of ATP-binding proteins, permeases, and related transporter components suggests strengthened energy-dependent membrane transport and material exchange. At 120 h, multiple SBPs were significantly upregulated, which points to strengthened uptake and translocation capabilities for phosphate, nitrogen, and carbon sources. Meanwhile, the upregulated propionate metabolism and methylmalonyl-CoA-associated precursor supply pathways can boost the abundance of extension units and reducing equivalents to polyketide assembly. Such alterations occurred alongside transcriptional activation of the natamycin BGC, which together substantiate a model: nutrient uptake, precursor supply, membrane transport and BGC activation synergistically facilitate natamycin accumulation.

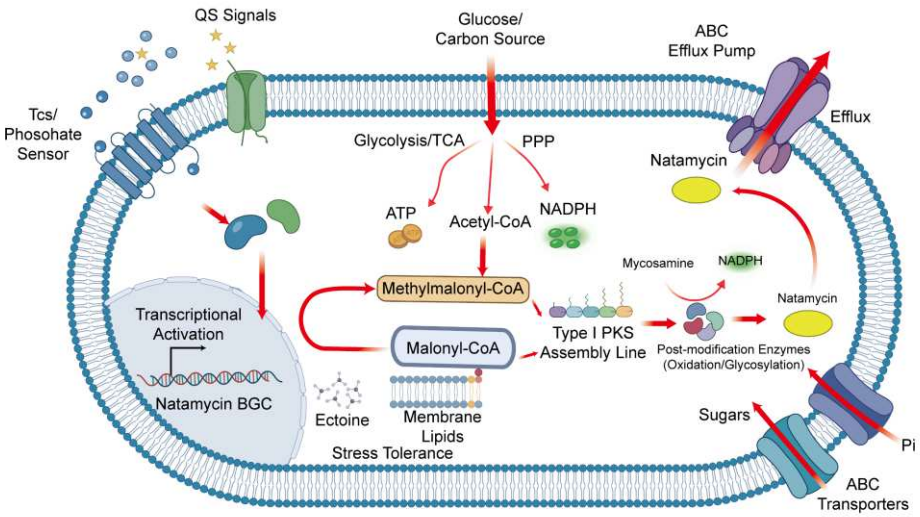


Figure 8. Proposed mechanism of enhanced natamycin production based on integrated WGS and RNA-seq analysis. Red arrows indicate key pathways or processes enriched in the high-producing mutant.

4 Conclusions and discussion

4.1 Conclusions

In this work, we generated a genetically stable high-natamycin-yielding *S. gilvosporeus* mutant (strain 717) via ARTP mutagenesis, and dissected its overproduction phenotype through combined whole-genome sequencing (WGS) and time-series RNA-seq analyses.. Strain 717 reached 21.15 g/L natamycin in 15 L fed-batch fermentation, representing a 74.2% increase over the parental strain 712. WGS analysis demonstrated that most high-confidence variants mapped to non-coding or upstream regulatory regions, while RNA-seq profiling uncovered stage-dependent transcriptional reprogramming involving ABC transporters, SBPs, propionate metabolic pathways, methylmalonyl-CoA supply routes, redox-relevant genes, and the natamycin BGC. Integrative multi-omics analysis pinpointed candidate genes and functional modules that bridge cis-regulatory variants to elevated natamycin titers. Collectively, these results yield tractable targets for subsequent genetic validation and establish a multi-omics roadmap for rational, mechanism-driven engineering of high-yield natamycin strains.

4.2 Discussion

This work constructs a multi-omics evidence basis for unraveling natamycin overproduction mechanisms and advances our understanding of dynamic secondary

metabolic regulation in *Streptomyces*. Firstly, the stage-specific functional differentiation of transport machineries reveals an underlying spatiotemporal regulatory logic governing secondary metabolism. While prior work largely frames ABC transporters solely as product efflux pumps, our transcriptomic data demonstrate that these transport modules undergo functional uncoupling and temporal compartmentalization throughout natamycin biosynthetic fermentation. At 72 h, coordinated upregulation of ATP-binding proteins, permeases, and SBPs indicates energy-dependent substrate uptake and preserved membrane homeostasis, potentially constructing an intracellular microenvironment conducive to PKS activation. By contrast, enhanced expression of nutrient acquisition systems at 120 h, represented by PstS (a phosphate SBP), glutamate SBPs, and carbohydrate SBPs, facilitates the reallocation of core phosphorus, nitrogen, and carbon nutrients amid the high-titer production phase. Consistent with this interpretation, deletion of the primary exporter genes *scnA/scnB* markedly reduced natamycin efflux efficiency without altering total production[14], indicating that these exporters primarily affect product export and self-resistance rather than biosynthetic capacity. In comparison, strain 717 possesses comprehensively augmented transport machinery covering the entire metabolic cycle, which coordinates nutrient uptake, substrate acquisition, membrane stability and product secretion simultaneously. Secondly, the coupled mutation-expression profiles of candidate genes indicate multi-layered regulatory nodes that modulate secondary metabolism. Integrated analysis screened out three core candidate genes, each bearing unique molecular characteristics determined by mutation type and transcriptional shift. *B1H19_RS30395*, which codes a HAD family phosphatase, carried a stop-gained mutation while presenting markedly elevated expression levels ($log_2FC = 2.55$). This transcriptional pattern implies that the resultant truncated protein may undergo altered catalytic activity or subcellular localization. Such changes likely perturb the turnover of phosphorylated metabolites and interfere with PhoP-dependent phosphate repression, ultimately reshaping the transcriptional activation cascade of secondary metabolic gene clusters[5,28]. *B1H19_RS33295* encodes an NAD(P)-binding protein bearing an in-frame insertion

and exhibits concomitant transcriptional upregulation (log_2FC = 1.43). This structural lesion likely remodels the cofactor-binding pocket and disturbs intracellular NADPH regeneration and redox cycling, consistent with the obligate requirement of ketoreductase (KR) domains within type I PKS machineries for NADPH as a reducing cofactor[30], which provides molecular corroboration for the interplay of reductive power supply and biosynthetic flux. *B1H19_RS08875*, encoding a non-ribosomal peptide synthetase (NRPS), harbors a frameshift mutation yet exhibits elevated transcriptional expression. This phenotype implies that a competing secondary metabolic branch undergoes transcriptional activation accompanied by functional impairment. Such a regulatory shift may reroute scarce biosynthetic substrates—including acyl carrier proteins (ACPs), ATP, and amino acid precursors—to the natamycin polyketide synthase (PKS) biosynthetic pathway [31,32]. Taken together, the three candidate loci delineate a presumptive multi-tiered regulatory cascade: RS30395 mediates upstream signal uncoupling, RS33295 strengthens the supply of cellular reducing equivalents at the midstream node, and RS08875 facilitates downstream shunting of biosynthetic resources. Lastly, the "variant-transcription-phenotype" framework constructed herein delivers well-defined targets for follow-up reverse genetic verification. Subsequent research can be carried out across three hierarchical tiers. At the single-gene level, CRISPR-Cas9-mediated gene knockout and genetic complementation of RS30395, RS33295 and RS08875 will enable quantification of their respective individual contributions to natamycin biosynthesis in strain 717. Lastly, the "variant-transcription-phenotype" framework established herein delivers well-defined targets for follow-up reverse genetic validation. Subsequent research can be carried out across three hierarchical tiers. At the single-gene level, CRISPR-Cas9-mediated gene deletion and genetic complementation of *RS30395*, *RS33295*, and *RS08875* will enable quantification of their respective individual contributions to natamycin production in strain 717. At the module level, strains co-overexpressing ABC transporter genes (*RS06390/RS06385/RS06395*) and propionate-metabolism genes (*RS30825/RS30810*) could be constructed to test potential inter-module synergy. At

the systems level, machine learning algorithms combined with mechanistic modeling can be implemented to forecast optimal multi-gene combinatorial engineering strategies[20,21]. Most importantly, the transcriptional shift observed between 72 h and 120 h in this work can serve as a valuable indicator for precise fermentation regulation in subsequent cultivation processes. Specifically, implementing phase-dependent feeding of propionate precursors or tuning phosphate levels beyond 72 h exhibits superior performance relative to invariant culture conditions for tapping into the strain’s maximum biosynthesis capacity.

Concurrently, artificial intelligence (AI) and machine learning are emerging as important methodological tools bridging multi-omics profiling, metabolic-network modeling, rational strain engineering, and fermentation process optimization[21,33,34]. For the “genomic variant-transcriptional response-high-yield phenotype” framework established in this study, hybrid modeling strategies integrating machine learning and mechanistic models can simulate how gene overexpression, deletion, and combinatorial genetic perturbation reshape metabolic flux partitioning and target metabolite biosynthesis. This enables the prioritization of candidate engineering targets and the prediction of optimal combinatorial modification schemes[20,21]. For key enzyme engineering, structural modeling, protein language models, and machine learning-assisted directed evolution could be further integrated to screen favorable mutations, evaluate enzyme activity and stability, and optimize substrate specificity of precursor supply enzymes, BGC catalytic enzymes and transporters[35-37]. In parallel, machine-learning models could also be applied to optimize medium formulation, feeding strategies, dissolved oxygen and pH regulation, soft sensor construction, and overall fermentation performance. This provides robust providing methodological support for further improving natamycin titer, yield, and productivity[34,38].

Taken together, this work not only unravels the molecular underpinnings underlying robust natamycin biosynthesis in the ARTP mutant strain 717, but also establishes a secondary metabolism engineering framework guided by time-resolved transcriptomics. This framework centers on coordinating functional modules and

implementing multi-tiered regulatory interventions. Candidate variant sites, DEGs and enriched pathways identified herein provide valuable references for exploring natamycin overproduction mechanisms and guiding rational strain modification. Future efforts will be directed toward causal verification and modular reconstitution of the aforementioned candidate genes, advancing natamycin strain development from empirical breeding toward rational design. Particular emphasis should be placed on functionally validating the causal roles of genes associated with precursor provision, metabolite efflux transport, and cellular responses to phosphate and nitrogen nutrients.

Authors contributions

RL performed the experiments, analysed the data, and drafted the manuscript. CL, XZ, YZ, HL, and CG contributed to strain cultivation, fermentation experiments, data analysis, and manuscript revision. JL conceived and supervised the study, acquired funding, and revised the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by the National Natural Science Foundation of China (NSFC) Youth Project (22408145) and the Shandong Provincial Higher Education Institutions' Development Plan for Youth Innovation Teams in 2025 (2025KJH184).

Availability of data and materials

The raw whole-genome resequencing and RNA-sequencing data generated in this study are publicly available in the NCBI Sequence Read Archive (SRA) under BioProject accession PRJNA1504471 and SRA Study accession SRP72 2582 (https://www.ncbi.nlm.nih.gov/sra/?term=SRP722582). The whole-genome resequencing data for *Streptomyces gilvosporeus* 717 are available under SRA Run accession SRR39910663 (BioSample SAMN62059381). The 12 RNA-se quencing datasets for strains 712 and 717, collected at 72 and 120 h with t hree biological replicates per strain at each time point, are available under S RA Run accessions SRR39913424–SRR39913435 (BioSamples SAMN6206256 6–SAMN62062577).

Declarations

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

References

1. Meena M, Prajapati P, Ravichandran C, et al. Natamycin: a natural preservative for food applications - a review. Food Sci Biotechnol. 2021;30:1481-1496.

2. Aparicio JF, Barreales EG, Payero TD, et al. Biotechnological production and application of the antibiotic pimaricin: biosynthesis and its regulation. Appl Microbiol Biotechnol. 2016;100:61-78.

3. Robertsen HL, Musiol-Kroll EM. Actinomycete-derived polyketides as a source of antibiotics and lead structures for the development of new antimicrobial drugs. Antibiotics (Basel). 2019;8:157.

4. Qiu S, Zhao GQ, Lin J, et al. Natamycin in the treatment of fungal keratitis: a systematic review and meta-analysis. Int J Ophthalmol. 2015;8:597-602.

5. Mendes MV, Tunca S, Antón N, et al. The two-component PhoR-PhoP system of *Streptomyces natalensis*: inactivation or deletion of *phoP* reduces the negative phosphate regulation of pimaricin biosynthesis. Metab Eng. 2007;9:217-227.

6. Zong G, Meng Z, Yu M, et al. Abundant organic nitrogen enhances natamycin biosynthesis by increasing NAD(P) metabolic pathway activity in *Streptomyces gilvosporeus* F607. Front Microbiol. 2025;16:1684019.

7. Zhu Z, Ding X, Rang J, et al. Application and research progress of ARTP mutagenesis in actinomycetes breeding. Gene. 2024;929:148837.

8. Sun J, Li J, Yao L, et al. UV-ARTP-DES compound mutagenesis breeding improves natamycin production of *Streptomyces natalensis* HW-2 and reveals transcriptional changes by RNA-seq. Food Sci Biotechnol. 2023;32:341-352.

9. Luo JM, Li JS, Liu D, et al. Genome shuffling of *Streptomyces gilvosporeus* for

improving natamycin production. J Agric Food Chem. 2012;60:6026-6036.

10. Wang L, Xiao W, Qiu T, et al. Enhanced natamycin production in *Streptomyces gilvosporeus* through phosphate tolerance screening and transcriptome-based analysis of high-yielding mechanisms. Microb Cell Fact. 2025;24:79.

11. Wang L, Xiao W, Zhang H, et al. Improved natamycin production in *Streptomyces gilvosporeus* through mutagenesis and enhanced nitrogen metabolism. Microorganisms. 2025;13:390.

12. Xue J, Xiao W, Xu Y, et al. Enhancement of natamycin production by combining ARTP mutagenesis with temperature control strategy development in *Streptomyces gilvosporeus*. Bioprocess Biosyst Eng. 2025;48:817-827.

13. Li M, Chen S, Li J, et al. Propanol addition improves natamycin biosynthesis of *Streptomyces natalensis*. Appl Biochem Biotechnol. 2014;172:3424-3432.

14. Wang TJ, Shan YM, Li H, et al. Multiple transporters are involved in natamycin efflux in *Streptomyces chattanoogensis* L10. Mol Microbiol. 2017;103:713-728.

15. Zong G, Cao G, Fu J, et al. MacRS controls morphological differentiation and natamycin biosynthesis in *Streptomyces gilvosporeus* F607. Microbiol Res. 2022;262:127077.

16. Wang Y, Tao Z, Zheng H, et al. Iteratively improving natamycin production in *Streptomyces gilvosporeus* by a large operon-reporter based strategy. Metab Eng. 2016;38:418-426.

17. Cingolani P. Variant annotation and functional prediction: SnpEff. In: Ng C, Piscuoglio S, editors. Variant calling: methods and protocols. New York: Springer US; 2022. p. 289-314.

18. Love MI, Huber W, Anders S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. Genome Biol. 2014;15:550.

19. Krysenko S, Wohlleben W. Role of carbon, nitrogen, phosphate and sulfur metabolism in secondary metabolism precursor supply in *Streptomyces* spp. Microorganisms. 2024;12:1571.

20. Li X, Guo Z, Li Y, et al. Leveraging large language models for metabolic engineering design. Trends Biotechnol. 2026. doi:10.1016/j.tibtech.2026.03.026.

21. Lawson CE, Martí JM, Radivojevic T, et al. Machine learning for metabolic engineering: a review. Metab Eng. 2021;63:34-60.

22. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics. 2014;30:2114-2120.

23. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods. 2012;9:357-359.

24. Anders S, Pyl PT, Huber W. HTSeq - a Python framework to work with high-throughput sequencing data. Bioinformatics. 2015;31:166-169.

25. Kanehisa M, Furumichi M, Sato Y, et al. KEGG for taxonomy-based analysis of pathways and genomes. Nucleic Acids Res. 2023;51:D587-D592.

26. Gene Ontology Consortium. The Gene Ontology (GO) database and informatics resource. Nucleic Acids Res. 2004;32:D258-D261.

27. Martín JF, Liras P. Molecular mechanisms of phosphate sensing, transport and signalling in *Streptomyces* and related actinobacteria. Int J Mol Sci. 2021;22:1129.

28. Martín JF, Rodríguez-García A, Liras P. The master regulator PhoP coordinates phosphate and nitrogen metabolism, respiration, cell differentiation and antibiotic biosynthesis: comparison in *Streptomyces coelicolor* and *Streptomyces avermitilis*. J Antibiot (Tokyo). 2017;70:534-541.

29. Barreiro C, Martínez-Castro M. Regulation of the phosphate metabolism in the *Streptomyces* genus: impact on the secondary metabolites. Appl Microbiol Biotechnol. 2019;103:1643-1658.

30. Wang J, Deng Z, Liang J, et al. Structural enzymology of iterative type I polyketide synthases: various routes to catalytic programming. Nat Prod Rep. 2023;40:1498-1520.

31. Xu Z, Ji L, Tang W, et al. Metabolic engineering of *Streptomyces* to enhance the synthesis of valuable natural products. Eng Microbiol. 2022;2:100022.

32. Bibb MJ. Regulation of secondary metabolism in streptomycetes. Curr Opin Microbiol. 2005;8:208-215.

33. Cheng Y, Bi X, Xu Y, et al. Machine learning for metabolic pathway optimization: a review. Comput Struct Biotechnol J. 2023;21:2381-2393.

34. Helleckes LM, Hemmerich J, Wiechert W, et al. Machine learning in bioprocess development: from promise to practice. Trends Biotechnol. 2023;41:817-835.
35. Zhou J, Huang M. Navigating the landscape of enzyme design: from molecular simulations to machine learning. Chem Soc Rev. 2024;53:8202-8239.
36. Notin P, Rollins N, Gal Y, et al. Machine learning for functional protein design. Nat Biotechnol. 2024;42:216-228.
37. Singh N, Lane S, Yu T, et al. A generalized platform for artificial intelligence-powered autonomous enzyme engineering. Nat Commun. 2025;16:5648.
38. Cheng Y, Bi X, Xu Y, et al. Artificial intelligence technologies in bioprocess: opportunities and challenges. Bioresour Technol. 2023;369:128451.