Journal of Applied Biology & Biotechnology Vol. 14(2), pp. 212-225, March-April, 2026
Available online at http://www.jabonline.in
DOI: 10.7324/JABB.2026.253744

# *In silico* analysis of pathogenic islands and metabolic pathways for understanding infection biology and lifestyle of *Clostridium chauvoei*

Keerthana Pasam, Surekha Challa, Nageswara Rao Reddy Neelapu*

Biochemistry and Bioinformatics Division, Department of Life Sciences, School of Science, Gandhi Institute of Technology and Management (Deemed to be University), Visakhapatnam, Andhra Pradesh, India.

**ARTICLE INFO**

*Article history*:
Received on: 22/09/2025
Accepted on: 28/12/2025
Available online: 25/01/2026

***Key words***:
*Clostridium chauvoei*,
Blackleg disease,
Pathogenic islands,
Gene prediction,
Functional annotation,
Metabolic pathway prediction.

**ABSTRACT**

*Clostridium chauvoei*, the causative agent of blackleg disease in cattle, presents significant economic and health challenges due to its high mortality rates and rapid disease progression. This study focuses on the *in silico* genomic characterization of *C. chauvoei* strain SBP 07/09 Swiss Bovine Pathogen, referring to a strain isolated in July 2009. The study identified and annotated pathogenic islands (PIs) contributing to the bacterium's virulence and adaptability. Using IslandViewer4, eight distinct PIs were identified, and 81 genes were detected using GeneMark. hmm-P across these PIs. The genes are categorized as 60 functional genes, 20 hypothetical proteins, and one gene with no assigned function. Functional annotation of genes using tools such as Basic Local Alignment Search Tool (BLASTp), InterPro, and BlastKOALA revealed that these genes are implicated in essential processes, including stress response, metabolism, genetic mobility, DNA repair, and anaerobic survival. Pathway analysis was performed using Kyoto Encyclopedia of Genes and Genomes (KEGG) Mapper and BioCyc Pathway Tools, utilizing KO identifiers assigned by BlastKOALA and whole genome data, respectively. This analysis revealed several key metabolic and regulatory pathways associated with the detected genes. These include nutrient transport, energy production, cofactor biosynthesis, and environmental adaptation. These pathways will likely contribute significantly to the organism's adaptability, anaerobic lifestyle, and survival within the host environment. Key findings include the identification of genes facilitating nutrient uptake, energy production, and genomic integrity maintenance. All of which enhance *C. chauvoei's* virulence and survival in hostile host environments. These insights offer valuable targets for developing preventative and therapeutic strategies to combat blackleg disease, reducing its economic burden on cattle farming.

## 1. INTRODUCTION

*Clostridium chauvoei*, a Gram-positive anaerobe, is the primary pathogen responsible for blackleg disease in cattle, a condition marked by acute myositis and high mortality. Blackleg primarily affects young cattle and is characterized by the sudden onset of lameness, fever, and rapid progression to fatality. The infection cycle of *C. chauvoei* in ruminant hosts begins with the entry of spores into the host through ingestion or contamination of wounds or mucosal abrasions. Ingested spores reach the intestines, traverse the bloodstream or lymphatic system, and migrate to muscle tissues. Alternatively, spores introduced through wounds follow a similar path to muscle tissues. In healthy tissues, the spores remain dormant until activation, which occurs under anaerobic conditions, such as those created by muscle injury or tissue damage that reduces oxygen levels. Once activated, the spores germinate into vegetative cells, marking the transition to the growth phase [1-4].

During vegetative growth, the bacteria proliferate rapidly in the anaerobic environment, releasing toxins and enzymes. Key virulence factors include beta pore-forming toxins (*C. chauvoei* cytotoxin A and Chauveolysin) [5,6], which lyse host cells and cause tissue necrosis. Clostridial exotoxins such as hyaluronidase, DNase, and neuraminidase, degrade the extracellular matrix to facilitate bacterial spread [7]. The toxin activity results in hemorrhagic necrosis of muscle tissues, gas production from bacterial metabolism, and the formation of characteristic dark, gas-filled lesions (blackleg), accompanied by severe inflammation and edema. As the infection progresses, toxins enter the bloodstream, causing toxemia and septicemia. This systemic dissemination leads to fever, lameness, swelling, and rapid deterioration of the host's condition, often culminating in shock and death if untreated [7]. While attempting to limit the infection through macrophages and antibody production, the host immune response is typically overwhelmed by the rapid bacterial proliferation and potent toxin activity. In severe cases, the infection cycle ends with the host's death, completing the bacterium's propagation in the environment through post-mortem spore release from decomposing tissues [8]

**Corresponding Author:*
*Nageswara Rao Reddy Neelapu,*
*Biochemistry and Bioinformatics Division,*
*Department of Life Sciences, School of Science, Gandhi Institute of Technology and Management (Deemed to be University), Visakhapatnam, Andhra Pradesh, India.*
*E-mail: nneelapu @ gitam.edu*

[Figure 1]. These spores in the environment and intestinal tracts of cattle can remain dormant until activated by anaerobic conditions, leading to tissue colonization and subsequent disease [5,8].

The pathogenicity of *C. chauvoei* stems from its ability to adapt to diverse environments within the host. This adaptability is primarily facilitated by genetic elements known as pathogenic islands (PIs), which harbor clusters of genes that enhance the bacterium's ability to evade the host immune system, utilize available nutrients, and survive under anaerobic conditions. Understanding the genes within these PIs is crucial for elucidating the mechanisms underlying *C. chauvoei's* virulence, which can inform effective disease management and prevention strategies [7,9] [Figure 1]. The objective of the present study is to employ a multi-tiered bioinformatic approach: (1) To identify PIs within the genome and to predict the genes and their functions within these PIs, and (2) to, consequently, perform the metabolic pathway analysis for the PIs and whole genome data of *C. chauvoei.* This integrative approach allows for identifying the genes that play a key role in processes and pathogenicity.

## 2. MATERIALS AND METHODS

### 2.1. Genome Data Sampling

The genome sampling considers the genome's completeness, strain history, and pathogenicity report. The whole genome sequence of *C. chauvoei* SBP 07/09, with the following accession number NZ_CP027286, is obtained from NCBI [10]. This whole genome sequence is used to predict the PIs.

### 2.2. Screening of PIs by *in silico* Genome Analysis

The modified protocol of Nammi *et al*. [11] is used to identify the PIs in *C. chauvoei* SBP 07/09. PIs are screened and identified using the tool Islandviewer4 [12]. IslandViewer4 integrates three methods for predicting genomic islands: sequence-based, comparative-based, and codon-based methods. The sequence-based method focuses on sequence composition, particularly dinucleotide bias, and detects the presence of mobility genes such as integrases and transposases to locate genomic islands. IslandPath-DIMOB [13] is an example of the sequence-based method. The comparative-based method uses comparative genomics to identify genomic islands by detecting regions present in the target genome but absent in closely related genomes, suggesting horizontal gene transfer. IslandPick [14] is an example of the comparative-based method. Finally, the codon-based method identifies genomic islands by analyzing codon usage bias, as these regions often exhibit codon patterns distinct from the rest of the genome. SIGI-Hidden Markov Model (HMM) [15] is an example of a codon-based method. HMMs pinpoint areas with atypical codon usage. The whole genome sequence of this strain was downloaded from NCBI and submitted in FASTA format as input to Islandviewer4 for predicting PIs. Islandviewer4 is freely available at https://www.pathogenomics.sfu.ca/islandviewer. These PIs are further used to predict the genes.

### 2.3. Prediction of Genes in PIs

GeneMark.hmm is a tool that uses HMM to predict genes in genomic sequences [16]. GeneMark.hmm family is widely used to identify protein-coding genes within DNA sequences. The prokaryotic version GeneMark.hmm-P identifies start and stop codons, ribosomal binding sites, and predicts operons in bacteria and archaea [17]. The tool provides gene prediction results in output formats, which include gene coordinates, coding sequences (CDS), and protein translations. The nucleotide sequences of each PI are downloaded from NCBI and are submitted in the FASTA format as an input to GeneMark.hmm-P for gene prediction. GeneMark.hmm-P is freely available at https://genemark.bme.gatech.edu/gmhmmp.cgi. These translated protein sequences are used to predict the function of the genes.

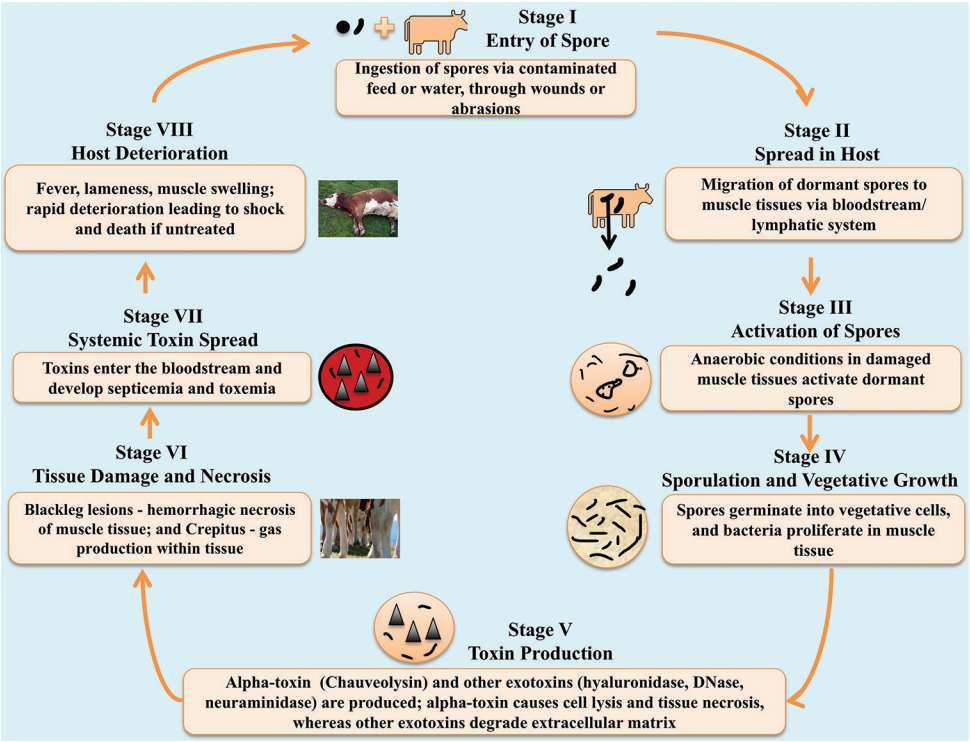


**Figure 1:** The lifecycle of *Clostridium chauvoei* in ruminants (cattle), causing blackleg disease.

### 2.4. Prediction of Gene Function Using Different Tools

The function of the genes was predicted using three different methods: (a) Homology-based method, (b) domain-based method, and (c) metabolic category-based method.

#### *2.4.1. Homology-based method*

The homology-based method predicts the function of a gene or protein based on its homology (similarity) to other known sequences. The underlying assumption is that if two sequences share a significant level of similarity, they may have evolved from a common ancestor and are likely to retain similar functions. The query sequence (gene/protein of unknown function) is compared against a database of sequences with known functions using a tool based on a homology-based method. BLAST, is an essential tool for identifying sequences similar to a query sequence, which can provide insights into the sequence's function, structure, and evolutionary relationships. BLAST for proteins (BLASTp) predicts the protein's function by identifying similarities between a query protein sequence and known protein sequences in a database [18]. This method relies on the principle that proteins with similar sequences often share similar functions, as they tend to have conserved structural or functional domains. The translated protein sequences of the genes are used to predict the function of the genes. The amino acid sequences of each gene are submitted in the FASTA format as input to BLASTp to predict gene function. The sequence similarity and E-value are considered to interpret the results and predict the gene's function. BLASTp is freely available at https://blast.ncbi.nlm.nih.gov/Blast.cgi?PAGE=Proteins.

#### *2.4.2. Domain-based method*

The domain-based method predicts the function of a gene or a protein by identifying specific protein domains. Protein domains are conserved structural or functional units within proteins, often responsible for particular biochemical activities or interactions. Since these domains are evolutionarily conserved across different proteins, they can provide reliable clues about a protein's function even when the overall sequence similarity is low. InterPro is a tool built on a domain-based method to predict protein functions and provide annotations by combining information from several protein signature databases [19]. The translated protein sequences of each gene are submitted in the FASTA format as input to InterPro. The tool then scans these sequences against its databases and generates a report detailing identified protein families, domains, functional sites, and related GO terms. The generated report is used to interpret the results and predict the gene's function. The tool InterPro is freely available at https://www.ebi.ac.uk/interpro/.

#### *2.4.3. Metabolic category-based method*

The metabolic category-based method predicts a gene's function by identifying a protein's metabolic category. BLAST KEGG Orthology And Links Annotation (BlastKOALA) is a tool built on a metabolic category-based method that integrates homology-based and pathway-based function prediction within the context of metabolic and cellular processes. The tool BlastKOALA is part of the KEGG. BlastKOALA is an automatic annotation server for genome sequences, which performs KO (KEGG Orthology) assignments to characterize individual gene functions. BlastKOALA also reconstructs KEGG pathways, Biomolecular Relations in Information Transmission and Expression (BRITE) hierarchies, and KEGG modules to infer high-level functions of the organism or the ecosystem. In BlastKOALA, the KO assignment is performed by a modified version of the internally used KOALA algorithm. Initially, the BLAST search is against a non-redundant dataset of pangenome sequences at the species, genus, or family level, which are generated from the KEGG GENES database by retaining the KO content of each taxonomic category [20]. The translated protein sequences of each gene are submitted in the FASTA format to BlastKOALA. The output of these results includes gene function and KEGG Orthology identifiers (KO ID), which were further interpreted. The tool BlastKOALA is freely available at https:// www.kegg.jp/blastkoala/. The result files were downloaded and further used for KEGG Mapper analysis.

### 2.5. Metabolic Pathway Analysis for Identification of Genes and Pathways Involved in the Pathogenesis of *C. chauvoei* SBP 07/09

The metabolic pathways for *C. chauvoei* SBP 07/09 were analyzed to identify genes and pathways involved in pathogenesis. Two complementary approaches were used: Mapping pathways using KEGG mapper and constructing a pathway/genome database (PGDB) using BioCyc pathway tools. The analysis focused on genes located in PIs to explore their potential roles in metabolism and pathogenesis.

#### *2.5.1. Mapping of metabolic pathways using KEGG mapper based on KO IDs*

The genes identified during functional annotation through BlastKOALA, which predicts the KO IDs, were analyzed for their roles in metabolic pathways using KEGG Mapper. The tool KEGG Mapper is freely available at https://www.genome.jp/kegg/tool/map_pathway.html. KEGG Mapper [21,22] provided a pathway mapping of these annotated genes by integrating them into existing metabolic pathways within the KEGG framework. This tool highlighted pathways linked to metabolic processes and pathogenesis based on the input data.

#### *2.5.2. Building of metabolic pathways using BioCyc pathway tools based on whole genome sequence data*

KEGG Mapper assigned annotated genes to metabolic pathways based on KO identifiers, but its scope was limited to proteins represented in the KEGG database. To address this limitation, BioCyc pathway tools [23] were employed to reconstruct genome-wide pathways, including both characterized and uncharacterized proteins. A comprehensive metabolic map was generated by importing the annotated genome, offering a broader view of *C. chauvoei's* metabolic network. Pathway Tools covered KO-mapped proteins and inferred pathways for genes absent in KEGG, enabling deeper insight into metabolic potential. In addition, it supported pathway enrichment and experimental data integration, enhancing the resolution of genome-wide metabolic analysis.

#### *2.5.3. Correlation of pathways identified between KEGG mapper and BioCyc pathway tools*

A comparative manual correlation of results from both platforms enabled deeper insights into the genome's metabolic potential via KEGG Mapper and BioCyc Pathway Tools. By comparing the outputs, overlapping and unique pathways were identified. BioCyc has filled the gaps left by KEGG, while KEGG's standardized pathways helped validate BioCyc's results. This complementary approach leveraged the strengths of both tools: KEGG's precision in curated pathways and BioCyc's breadth in genome-wide reconstruction, yielding a more complete understanding of *C. chauvoei*'s metabolic network.

## 3. RESULTS

### 3.1. PIs in *C. chauvoei* strain SBP 07/09

The PIs of *C. chauvoei* strain SBP 07/09 were predicted using the IslandViewer4 tool. The different prediction tools of IslandViewer4

applied to predict PIs are IslandPath-DIMOB and SIGI-HMM. The IslandPath-DIMOB identified seven PIs, whereas SIGI-HMM detected one. The results of PIs are further refined by checking for any overlaps in island start and end positions to observe distinct PIs [Table 1]. The study identified eight distinct PIs in *C. chauvoei* strain SBP 07/09. The nomenclature for the PI is generated; for example, the first PI in the organism is *C. chauvoei* PI 1 labelled CCPI1 [Table 1].

### 3.2. Genes and their Functions

Gene prediction for the eight PIs of *C. chauvoei* strain SBP 07/09 using GeneMark.hmm.p identified 81 genes [Supplementary Table 1]. Functional annotation was performed using three complementary tools: BLASTp (homology-based), InterPro (domain-based), and BlastKOALA (metabolic category-based) [Supplementary Table 2]. Based on the combined results, genes were classified into three categories: (a) Genes with predicted functions, (b) hypothetical proteins, and (c) genes with no functional assignment. Genes with known homologues or functional annotations were assigned to the first category. Genes predicted with open reading frames but lacking homology were labeled hypothetical proteins. Genes that matched neither known functions nor hypothetical annotations were categorized as having no function, typically due to the absence of database matches. This classification yielded 60 functional genes, 20 hypothetical proteins, and one gene with no function [Figure 2 and Supplementary Table 3]. Each gene was assigned a unique identifier based on its position, for example, the first gene in a PI was labeled CCPI-G1 [Supplementary Table 1]. In addition, this organism's unique genes are summarized along with their roles in Table 2. The annotations present a key limitation: If genes such as hypothetical proteins and genes with no function reside within genomic islands, their unknown function prevents definitive inference of their role.

### 3.3. Metabolic Pathway Analysis for Identification of Genes and Pathways Involved in the Pathogenesis of *C. chauvoei* SBP 07/09

The metabolic pathway analysis of *C. chauvoei* SBP 07/09 was conducted using KEGG Mapper and Pathway Tools to identify genes and pathways associated with the organism's pathogenesis.

#### *3.3.1. KEGG mapper analysis*

The metabolic pathways of *C. chauvoei* SBP 07/09 were mapped using KEGG mapper based on KO IDs, which were assigned through functional annotation using the BLASTKOALA method. A total of 22 proteins were identified to be involved in 40 metabolic pathways within the KEGG database [Supplementary Table 4]. Further, after carefully evaluating the pathways, seven were assumed to be involved and play an essential role in the organism's pathogenicity.

**Table 1:** The number of pathogenic islands in *Clostridium chauvoei* strain SBP 07/09.

| S. No. | Island No. | Island start | Island end | Length |
|---|---|---|---|---|
| 1 | CCPI1 | 537993 | 545177 | 7184 |
| 2 | CCPI2 | 725827 | 738326 | 12499 |
| 3 | CCPI3 | 1322471 | 1328899 | 6428 |
| 4 | CCPI4 | 1811552 | 1819055 | 7503 |
| 5 | CCPI5 | 2258243 | 2268835 | 10592 |
| 6 | CCPI6 | 2610857 | 2631723 | 20866 |
| 7 | CCPI7 | 2786754 | 2790862 | 4108 |
| 8 | CCPI8 | 2830535 | 2845432 | 14897 |

#### *3.3.2. BioCyc pathway tools analysis*

The pathway analysis of *C. chauvoei* SBP 07/09 was conducted using BioCyc pathway tools, utilizing the whole genome sequence obtained from NCBI to reconstruct a metabolic PGDB. The resulting database provided a comprehensive overview of metabolic pathways, enzymatic reactions, transporter reactions, polypeptides, protein complexes, enzymes, transporters, compounds, transcription units, and tRNAs [Table 3, Figure 3 and Supplementary Figure 1]. To evaluate the accuracy and biological relevance of these predictions, the reconstructed pathways were further classified into six categories: Consistent, partially consistent, poor topological match, missing, not reported, and false positives [Table 4 and Supplementary Table 5]. In *C. chauvoei*, 155 out of 210 pathways were consistent, and 22 were missing [Table 4 and Supplementary Table 5]. Among these pathways,

**Table 2:** The summary of unique genes in the *Clostridium chauvoei* strain SBP 07/09.

| S. no | Gene. no | Gene name | Role |
|---|---|---|---|
| 1 | CCPI-G2 | Rpn family recombination-promoting nuclease/putative transposase | Recombination or phage defense |
| 2 | CCPI-G3 | Rpn family recombination-promoting nuclease/putative transposase | Recombination or phage defense |
| 3 | CCPI-G5 | Rpn family recombination-promoting nuclease/putative transposase | Recombination or phage defense |
| 4 | CCPI-G6 | Rpn family recombination-promoting nuclease/putative transposase | Recombination or phage defense |
| 5 | CCPI-G7 | Rpn family recombination-promoting nuclease/putative transposase | Recombination or phage defense |
| 6 | CCPI-G9 | Rpn family recombination-promoting nuclease/putative transposase | Recombination or phage defense |
| 7 | CCPI-G31 | ArpU family phage packaging/lysis regulator | Regulation of phage lysis and packaging genes |
| 8 | CCPI-G52 | PTS lactose/cellobiose transporter subunit IIA | PTS-mediated sugar uptake |
| 9 | CCPI-G53 | Lichenan-specific PTS enzyme IIB component | Transports β-glucans |
| 10 | CCPI-G54 | PEP phosphonomutase | Phosphonate metabolism |
| 11 | CCPI-G55 | PTS sugar transporter subunit IIC | Sugar transport via PTS system |
| 12 | CCPI-G58 | Class B sortase | Surface protein anchoring |
| 13 | CCPI-G59 | Alpha-L-fucosidase | Degradation of fucose-containing glycans |
| 14 | CCPI-G60 | Glycoside hydrolase family 16 protein | Environmental carbohydrate metabolism |
| 15 | CCPI-G61 | PTS sugar transporter subunit IIC (again) | Sugar transport via PTS system |
| 16 | CCPI-G77 | YjjG family noncanonical pyrimidine nucleotidase | Pyrimidine salvage |

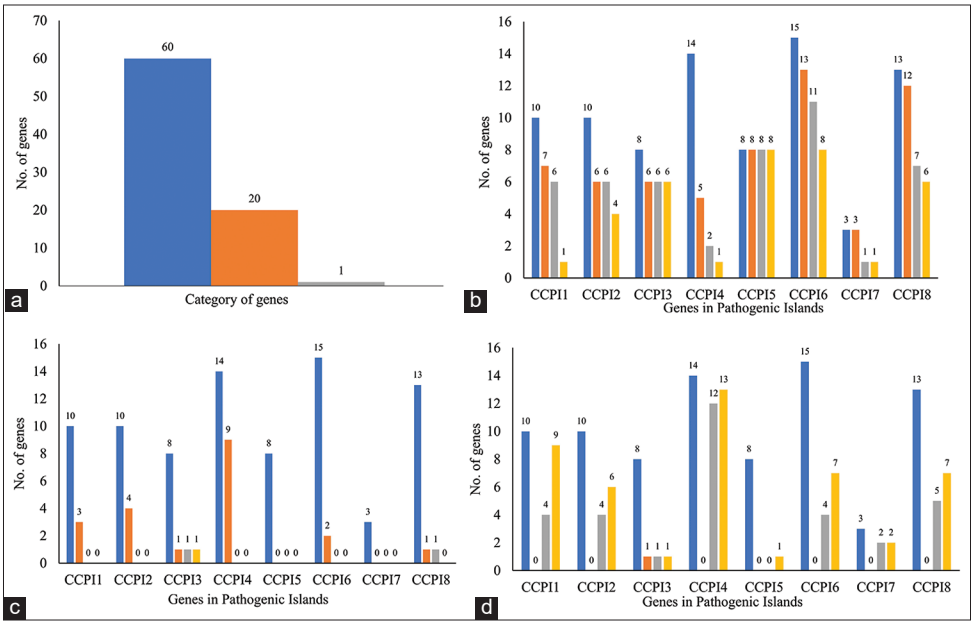


**Figure 2:** The figure shows the graphical representation of summary of gene category classification. (a) Gene categorized into genes with functions, genes with hypothetical proteins, and genes with no function. (b) Genes with Functions in each islands along with total number of proteins - total no. of genes, BLASTp, InterPro, and BLASTKOALA. (c) Genes with hypothetical proteins and (d) Genes with no function. The legend is the same for (c) and (d) as (b).

**Table 3:** The overview of reconstruction of metabolic pathways based on the whole genome of *Clostridium chauvoei* strain SBP 07/09.

| S. no | Description | Number |
|---|---|---|
| 1 | Pathways | 169 |
| 2 | Enzymatic reactions | 1052 |
| 3 | Transport reactions | 92 |
| 4 | Polypeptides | 2556 |
| 5 | Protein complexes | 118 |
| 6 | Enzymes | 558 |
| 7 | Transporters | 99 |
| 8 | Compounds | 798 |
| 9 | Transcription Units | 1754 |
| 10 | tRNAs | 87 |

**Table 4:** The different categories of metabolic pathways built using pathologic *Clostridium chauvoei* strain SBP 07/09.

| S. No. | Category of pathways | Number of pathways |
|---|---|---|
| 1 | Consistent | 155 |
| 2 | False positives | 8 |
| 3 | Not reported | 9 |
| 4 | Partially consistent | 12 |
| 5 | Poor topological match | 4 |
| 6 | Missing | 22 |
| | Total | 210 |

unique pathways and their roles that were not reported or studied are summarized in Table 5. The genes from PIs present within BioCyc pathways and that are assumed to be involved and play an essential role in the organism's pathogenicity are shown in Supplementary Table 6.

**Table 5:** The summary of unique metabolic pathways in *Clostridium chauvoei* SBP 07/09.

| S. no | Pathways | Role |
|---|---|---|
| 1 | Arsenate detoxification I | The pathway reduces arsenate to arsenite (detoxification) |
| 2 | Cytidylyl MoCo sulfurylation | The pathway is essential for sulfuration of MoCo, which is involved in crucial redox reactions in the global C-, N-, and S-cycles |
| 3 | Dipyrromethane cofactor biosynthesis | Cofactor for the formation of preuroporphyrinogen |
| 4 | Ethanol degradation I | The pathway oxidizes ethanol to acetaldehyde/acetate |
| 5 | Folate transformations III (*Escherichia coli*) | The pathway transforms folate into methyl and formyl derivatives. |
| 6 | Pyridoxal 5'-phosphate salvage I | The pathway salvages Pyridoxal 5'-phosphate (PLP) from pyridoxal, pyroxidine, and pyridoxamine. |
| 7 | Queuosine biosynthesis III (queuosine salvage) | The pathway is involved in the denovo biosynthesis of queuosine, and queuosine is a modified nucleoside in specific tRNAs in bacteria. |
| 8 | Tetrahydrofolate salvage from 5,10-methenyltetrahydrofolate | The pathway salvages tetrahydrofolate from 5,10-methenyltetrahydrofolate |
| 9 | Thiamine diphosphate salvage III | The pathway recycles thiamine when it's depleted and is essential for the bacteria's survival. |

MoCo: Molybdenum cofactor.

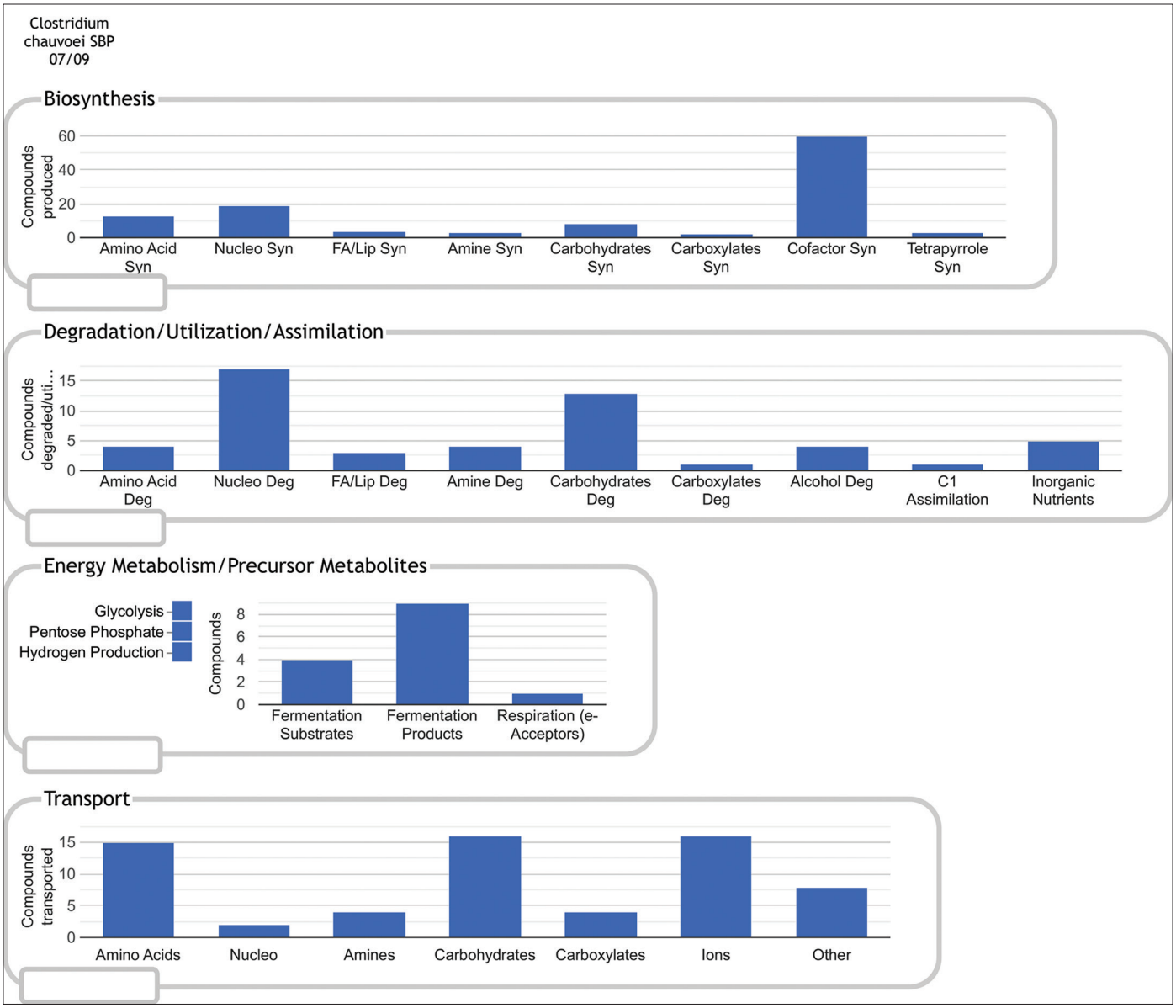


**Figure 3:** The figure shows the metabolic map or overview of classified pathways of pathway genome database of *Clostridium chauvoei* SBP 07/09 generated in pathway tools.

### *3.3.3. Correlation of KEGG mapper and BioCyc pathway tools*

The results obtained from both tools, KEGG Mapper and BioCyc pathway tools, were compared and correlated to assess the involvement of metabolic pathways in virulence and survival [Table 6]. A total of 36 proteins encoded within the PIs were mapped to 49 metabolic pathways across both KEGG mapper and BioCyc pathway tools analyses, highlighting their potential roles in the bacterium's pathogenic mechanisms. A comprehensive metabolic pathway analysis using KEGG Mapper and BioCyc pathway tools revealed a high similarity, demonstrating consistency in core metabolic predictions and highlighting unique pathways identified by each tool.

BioCyc pathway tools detected pathways such as L-threonine degradation I, S-propane-1,2-diol degradation, cardiolipin biosynthesis I, and inosine-5'-phosphate biosynthesis II, which were absent in KEGG mapper, likely due to its reliance on KO IDs. Conversely, the KEGG Mapper mapped pathways such as taurine and hypotaurine, pyruvate, and propanoate metabolism, which were

**Table 6:** The manual correlation of pathways identified by KEGG mapper and BioCyc pathway tools.

| S. no | Gene no. | Gene name | Gene synonym | KEGG pathways | Pathway tools |
|---|---|---|---|---|---|
| 1 | CCPI-G11 | K15024; putative phosphotransacetylase [EC: 2.3.1.8] | Phosphate propanoyltransferase | Pyruvate metabolism; Propanoate metabolism | L-threonine degradation I, *S*-propane-1,2-diol degradation |
| 2 | CCPI-G44 | gpmI; 2,3-bisphosphoglycerate-independent phosphoglycerate mutase [EC: 5.4.2.12] | Putative=2,3-bisphosphoglycerate-independent phosphoglycerate mutase | Glycolysis/ Gluconeogenesis | Gluconeogenesis I, glycolysis IV, glycolysis I (from glucose 6-phosphate) |
| 3 | CCPI-G45 | TPI, tpiA; triosephosphate isomerase (TIM) [EC: 5.3.1.1] | Putative Triosephosphate isomerase | Glycolysis/ Gluconeogenesis | Glycolysis I (from glucose 6-phosphate), gluconeogenesis I, glycolysis IV |
| 4 | CCPI-G46 | PGK, pgk; phosphoglycerate kinase [EC: 2.7.2.3] | Phosphoglycerate kinase | Glycolysis/ Gluconeogenesis | Glycolysis I (from glucose 6-phosphate), gluconeogenesis I, glycolysis IV |
| 5 | CCPI-G47 | GAPDH, gapA; glyceraldehyde 3-phosphate dehydrogenase (phosphorylating) [EC: 1.2.1.12] | Type I glyceraldehyde-3-phosphate dehydrogenase | Glycolysis/ Gluconeogenesis | Glycolysis I (from glucose 6-phosphate), gluconeogenesis I, glycolysis IV |

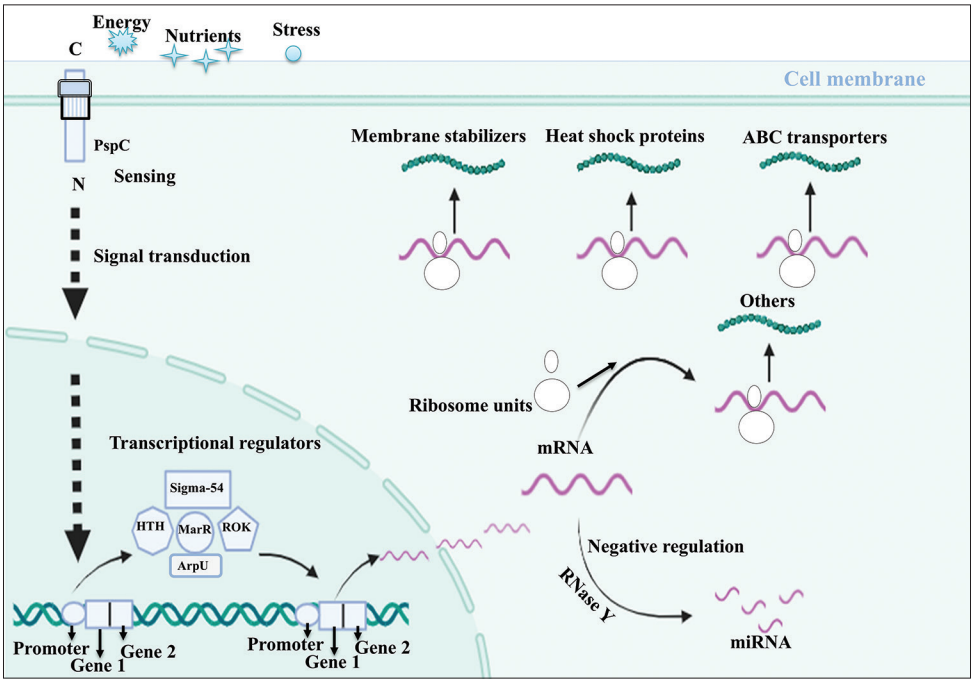


**Figure 4:** The genes identified as transcriptional regulators in the pathogenic islands of *Clostridium chauvoei* are involved in activating enzymes, transporters, energy-related genes, etc.

not explicitly detected in BioCyc Pathway Tools. BioCyc Pathway Tools primarily captured variations of core metabolic pathways, including glycolysis (e.g., glycolysis IV, glycolysis I from glucose-6-phosphate) and gluconeogenesis (e.g., gluconeogenesis I). At the same time, KEGG Mapper identified broader functional pathways such as microbial metabolism in diverse environments and biosynthesis of secondary metabolites. In addition, KEGG Mapper mapped several pathways linked to pathogenicity and stress response, including *Salmonella* infection, pathogenic *Escherichia coli* infection, biofilm formation in *Vibrio cholerae*, flagellar assembly, and two-component regulatory systems, suggesting a potential link between metabolism and virulence. Overall, KEGG Mapper provided a broad metabolic overview, whereas BioCyc Pathway Tools enabled a more detailed reconstruction of organism-specific variations, underscoring the complementary nature of both tools [Supplementary Table 4].

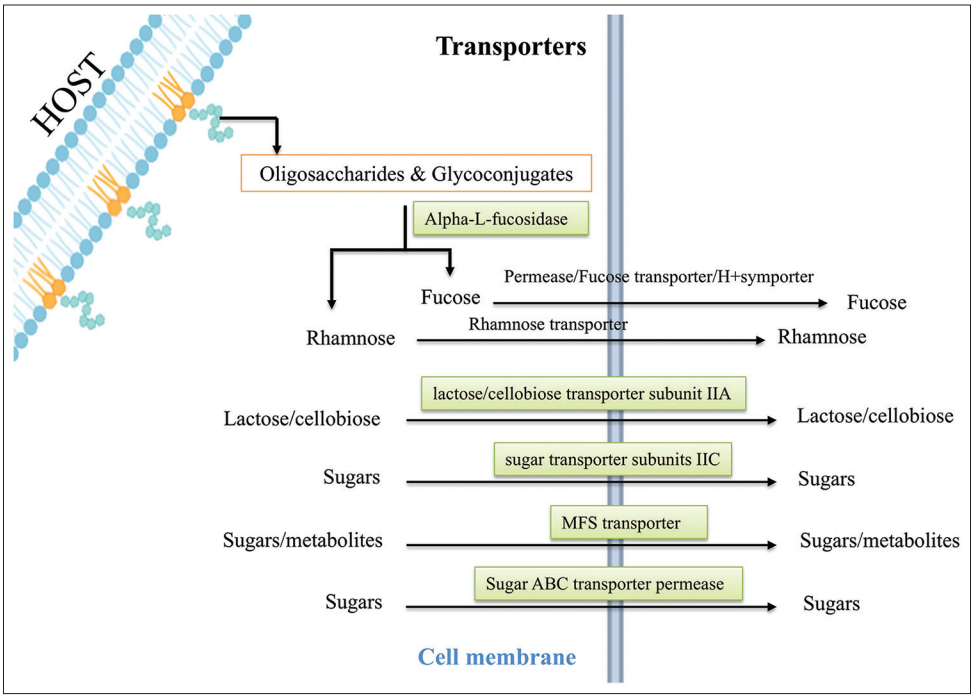


**Figure 5:** The genes that are transporters, which aid the uptake of sugars and metabolites from the environment. The highlighted genes are identified in the pathogenic islands of *Clostridium chauvoei*.

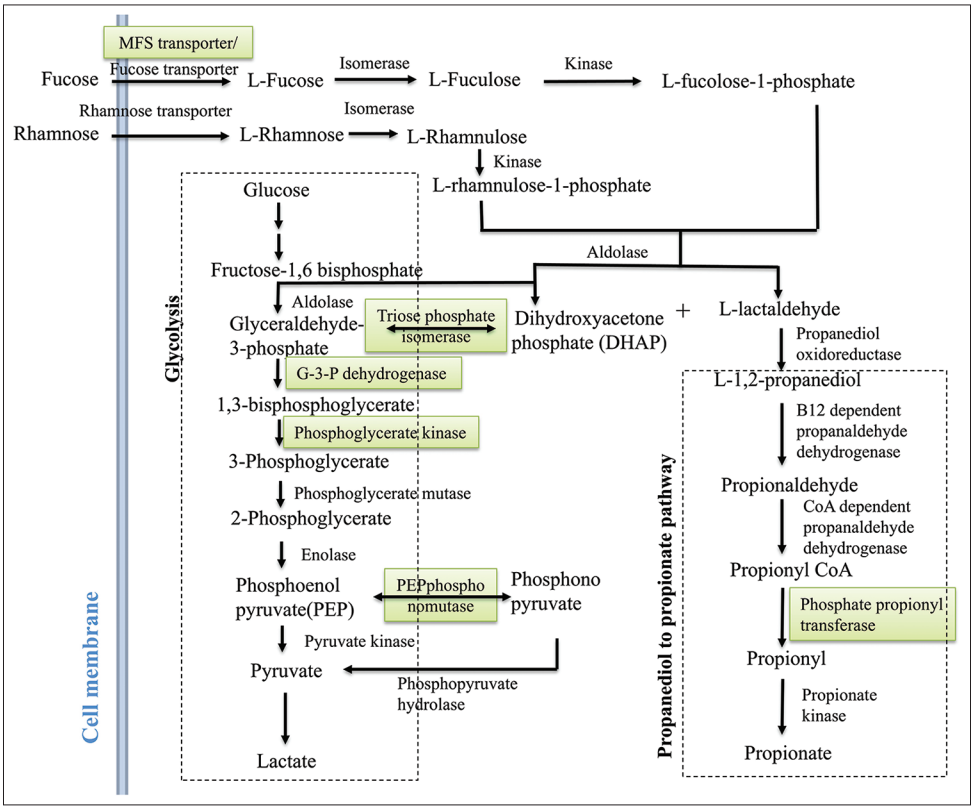


**Figure 6:** The pathways related to energy production which ensure efficient glucose utilization and others anaerobic metabolic pathways. The highlighted genes are present in the pathogenic islands of *Clostridium chauvoei*.

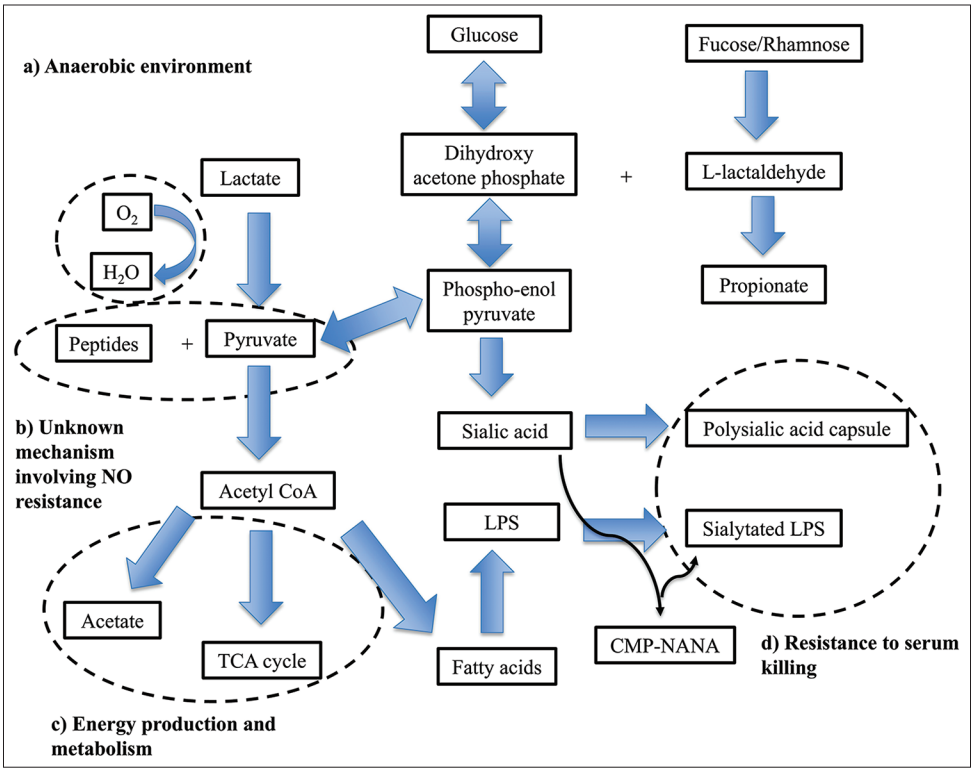


**Figure 7:** The figure was derived from Jiang *et al*., [77] and modified. The summary of pathways and flow describing lactate utilization in an anaerobic environment, unknown mechanisms involving NO resistance, energy production and metabolism, and resistance to serum killing.

## 4. DISCUSSION

This study aimed to enhance the understanding of *C. chauvoei* by identifying its PIs and predicting genes potentially linked to virulence and survival. Functional annotation of these genes was performed using homology-based, domain-based, and metabolic category-based approaches to ensure a comprehensive classification. To further investigate the metabolic potential of *C. chauvoei*, a genome-scale metabolic pathway reconstruction was carried out, and KO IDs derived from metabolic category-based function prediction were mapped onto metabolic pathways. A comparative analysis was then conducted by manually correlating pathways obtained from KEGG and BioCyc Pathway Tools. This allowed for an in-depth evaluation of pathway organization, enzyme annotations, and metabolic variations. This comparison provided insights into key metabolic processes that may contribute to the organism's adaptation and pathogenicity. Finally, the reconstructed pathways were examined in the context of *C. chauvoei's* lifecycle, linking its metabolic capabilities to survival strategies and infection mechanisms. The steps based on the standard tools provided us with the expected results. At the same time, the combinatorial use of these tools in this study linked the genomic islands of the genomes to metabolic pathways. The present section comprehends the understanding of the current study, provides insights, and discusses how the proteins encoded by the PIs help *C. chauvoei's* adapt to the situation during pathogenicity.

The genes encoded by these PIs are the genomic data generated in this method, which is helpful for disease management. Disease management is an essential aspect for understanding and controlling a disease. The first step in understanding a disease is to identify the organism causing the disease. However, the bacterial pathogen rapidly evolves and generates highly variable genotypes or isolates. Therefore, identifying the isolate among the group of isolates mainly associated with the disease is essential. Thus, the genomic data can be used as molecular markers that allow us to discriminate different strains within a species and can be applied to disease management.

The proteins encoded by the PIs enhance *C. chauvoei's* ability to survive and adapt to diverse environments. They also manage stress responses, effectively sustain metabolic processes, manage energy production, respond to environmental cues, maintain genomic integrity, and acquire beneficial genetic traits. The bacterium's flexible genetic toolkit, robust energy production, and membrane stabilization systems collectively allow it to persist and evade host defenses. If conditions within the host become unfavorable, some bacterial cells initiate sporulation, forming resilient spores. These spores can withstand environmental stress, ensuring the bacterium's survival and potential for future transmission.

### 4.1. Genes Coding for Stress Sensor Proteins Responding to Environmental Cues

*C. chauvoei* spores enter a ruminant host, often through ingestion or wound contamination; they encounter a nutrient-rich but hostile environment that initiates spore germination into active bacterial cells. The active bacterial cells are known for having proteins that detect stress due to the acidic and low-oxygen conditions in host tissues. Gram-negative bacteria harbor a highly conserved stress response system known as the envelope stress response (Esr) system, formerly known as phage shock protein (Psp) response system [24] [Figure 4]. The response system senses the signal from the environment and transduces it to the cytoplasm [25]. The Psp systems of *E. coli* have six proteins, PspA, B, C, D, F, and G. In general, stress mislocalizes protein secretin from the cell envelope due to its dissociation from the chaperone-like pilot protein and also reduces proton motive force. These events help proteins PspB and PspC (CCPI-G10) sense the signals in the extracytoplasmic space and help them bind to PspA, thereby releasing PspF from PspA. The

protein PspF activates the promoters of pspG and pspA and subsequently turns on the pspABCDE operon [26]. A rapid sensing of environmental changes marks this transition, and this detection acts as an alert. This initiates a bacterial response to activate RNA polymerase and a number of transcriptional regulators that trigger defense mechanisms, coordinating numerous proteins across several functions. The environment's hostility becomes apparent, and transcriptional regulators play a critical role in adjusting gene expression to maximize bacterial survival.

### 4.2. Genes Coding for Transcriptional Regulators to Sustain Metabolic Process

The sigma factor of RNA polymerase (CCPI-G49) is a transcription initiation factor that enables specific binding of RNA polymerase (RNAP) to gene promoters needed to initiate transcription in bacteria. The specific sigma factor used to activate transcription of a given gene will vary, depending on the gene and the environmental signals. RNAP factor sigma-54 is needed to initiate transcription in bacteria in a nitrogen-limited environment [27,28]. The protein PspF also activates the σ54-dependent transcription of the pspABCDE operon [26]. The sigma-54 transcriptional regulator and its interacting counterpart, the Sigma-54-interacting transcriptional regulator (CCPI-G51), add specificity to this response. This helps to focus on genes that help the bacterium function optimally in low-oxygen conditions [Figure 4] [29]. The sigma-54 transcriptional regulator is known for flagellar biosynthesis, motility [30-34], amino acid metabolism [35-37], quorum sensing, biofilm formation, virulence [38-40], and bacterial natural product genes [41].

The helix-turn-helix transcriptional regulator (CCPI-G40, G41) and multiple antibiotic resistance regulator (MarR) family regulators (CCPI-G75), belonging to the family winged helix-turn-helix, activate genes necessary for pathogenicity and stress response [Figure 4]. This helps the bacterium to adapt its metabolism and defense mechanisms to the nutrient-limited, anaerobic conditions within the host. The MarR protein is essential for diverse biological functions, which are crucial to the survival of pathogenic bacteria. The functions include resistance to multiple antibiotics, regulation of virulence-associated traits, virulence genes, hemolytic activity, extracellular protease activity, and motility [42-44]. The signaling molecules or ligands that activate the transcriptional regulator MarR are small phenolic compounds, metal ions, small peptides, and oxidative stress [45]. The repressor, open reading frame, kinase (ROK) family transcriptional regulators (CCPI-G56) are characterized by carbohydrate-sensing domains shared with sugar kinases, and the ROK family transcriptional regulator modulates pathways involved in carbohydrate metabolism, fine-tuning energy production to the available substrates in host tissues [Figure 4].

Bacteria extensively use autolysins to remodel, recycle, and even destroy their cell walls. The autolysin regulatory protein (ArpU) family transcriptional regulator (CCPI-G31) is linked to the regulation of muramidase-2 (peptidoglycan hydrolase-2 or autolysin) [Figure 4]. Bacteria would use ArpU to remodel the cell wall affected by stress. Autolysins may also degrade peptidoglycan to avoid their own recognition by the host's innate immune system. Collectively, these transcriptional regulators control a network of genes that help *C. chauvoei* respond to host conditions and establish infection. During this process, the equilibrium between the synthesis and degradation of mRNA in the pathogen is needed. The processing and degradation of mRNAs are initiated by RNase Y (CCPI-G24), an endoribonuclease anchored to the cell membrane [Figure 4 and Supplementary Figure 2]. Exoribonucleases degrade the cleaved products. In many bacteria, these RNases, RNA helicases, and other proteins are organized in a protein complex called the RNA degradosome, which plays an essential role in virulence and pathogenicity [46,47].

### 4.3. Genes for Managing Energy Production

The study identified transcription regulators that activate enzymes, transporters, and energy-related genes that manage energy production. The enzymes such as alpha-L-fucosidase (CCPI-G59) and glycoside hydrolase family 16 protein (CCPI-G60) break down polysaccharides [Figure 5]. Gut microbes produce fucosidases [48-51], cleaving fucose from host glycans (free oligosaccharides and glycoconjugates) to maintain intestinal homeostasis [52,53]. This outcome also ensures additional carbon sources to microbes for energy production, even under oxygen-limited conditions in anaerobic respiration. Transporters are crucial for nutrient uptake, which sustains the bacterium during infection. The major facilitator superfamily (MFS) transporter (CCPI-G76) facilitates the uptake of sugars and possibly other metabolites from the host [Figures 4 and 6, Supplementary Figure 3]. FucP and its homologues belonging to the MFS family transported L-fucose across cell membranes in a pH-dependent manner [54-56].

In general, pathogens form communities of microorganisms known as biofilms, and biofilms are protected by extracellular polymeric substance (EPS) made of polysaccharides, proteins, lipids, and extracellular DNA [57]. The mature biofilms undergo dispersal, which can be divided into two types: Active and passive dispersal, where active dispersal plays a vital role in the life cycle of a biofilm that contributes to bacterial survival and disease progression. Passive dispersal refers to dispersal by external forces. Active dispersal refers to dispersal triggered by microbes in the biofilm in response to environmental changes such as nutrient starvation, phagocyte challenge, and unfavorable oxygen levels [58]. The enzyme glycoside hydrolase family 16 protein produced by the bacterium may degrade the polysaccharide, poly(1,6)- N-acetyl-d-glucosamine (PNAG), by hydrolyzing β(1,6) glycosidic linkages, forming N-acetylglucosamine (GlcNAc) [59]. GlcNAc then enters the cell and is deacetylated into acetate and GlcN-6-P by GlcNAc-6-phosphate deacetylase, which belongs to the amidohydrolase superfamily (CCPI-G62) [60,61]. Then, GlcN-6-P is used in two main pathways: PG recycling pathway and the glycolysis pathway. Thus, the two enzymes glycoside hydrolase family 16 protein and GlcNAc-6-phosphate deacetylase play a role in bacterial survival and disease progression.

Transporters are crucial for nutrient uptake, which sustains the bacterium during infection. The MFS transporter works alongside the phosphotransferase system (PTS) transporters, facilitating the uptake of sugars and possibly other metabolites from the host [Supplementary Figure 4]. The various PTS transporters include PTS lactose/cellobiose transporter subunit IIA (CCPI-G52), and PTS sugar transporter subunit IIC (CCPI-G55) [Figure 5 and Supplementary Figure 4]. In addition, the putative Lichenan-specific phosphotransferase enzyme IIB component (CCPI-G53) is also identified in the study. Histidine-phosphorylatable phosphocarrier protein (HPr) family phosphocarrier protein (CCPI-G26) acts as a regulator and is essential in the PTS, which transfers sugar molecules across the cell membrane for energy. HPr is present in the cell in the phosphorylated (HPr-P) or non-phosphorylated (HPr) form, depending on the presence or absence of a sugar substrate of the PTS [62,63]. Such sugars, when present, give rise to the dephospho form of the protein due to sugar phosphorylation, but when exogenous PTS sugars are absent, HPr-P should predominate. The concentration of the phosphorylated form of HPr decreases in the

presence of a PTS substrate [64,65]. Sugar ABC transporter permease (CCPI-G67) proteins enhance this by efficiently channeling nutrients across the bacterial membrane [Figure 5]. Together, these transporters maximize nutrient intake, helping *C. chauvoei* thrive in the host's nutrient-rich but competitive environment.

The energy production of *C. chauvoei* depends on an array of proteins that work collectively to ensure efficient glucose utilization and other anaerobic metabolic pathways. The genes related to the fucose pathway, 1,2-propanediol to propionate pathway, and DHAP to lactate pathway are identified in the study. The fucose pathway starts with fucose and ends with DHAP and L-acetaldehyde [Figure 6]. L-acetaldehyde enters the 1,2-propanediol to propionate pathway, whereas DHAP enters the glycolysis pathway. In the propanediol pathway, the enzyme phosphate propanoyl transferase (CCPI-G11) is identified in the study, and this enzyme catalyzes the reaction of propanoyl-CoA to propanoyl [Figures 6 and 7, Supplementary Figures 5-7]. The study identifies enzymes such as triosephosphate isomerase (CCPI-G45), type I glyceraldehyde-3-phosphate dehydrogenase (CCPI-G47), phosphoglycerate kinase (CCPI-G46), and 2,3-bisphosphoglycerate-independent phosphoglycerate mutase (CCPI-G44) in the glycolysis pathway. These enzymes catalyze the reactions from DHAP to lactate [Figure 6 and Supplementary Figures 8-12]. In addition, enzyme PEP phosphonomutase (CCPI-G54) catalyzes the conversion of PEP to phosphonopyruvate, and the enzyme phosphonopyruvate hydrolase catalyzes the conversion from phosphonopyruvate to pyruvate [Figure 6]. In contrast, the lactate utilization protein (CCPI-G80) enables *C. chauvoei* to metabolize lactate, a common byproduct in anaerobic environments within host tissue [Figure 7 and Supplementary Figure 13]. Anaerobic sulfite reductase subunits AsrA (CCPI-G69) and AsrB (CCPI-G70), alongside sulfite reductase subunit C (CCPI-G71), ensure energy production even under oxygen-limited conditions by reducing sulfite to a form usable as hydrogen sulfide in anaerobic respiration [66].

### 4.4. Genes for Acquiring Beneficial Genetic Traits and Maintaining Genomic Integrity

DNA replication and repair systems are critical as the infection progresses to ensure accurate replication and to counteract DNA damage from host immune defenses. Proteins such as replicative DNA helicase (CCPI-G35) facilitate DNA replication, ensuring continuous bacterial cell division and colony expansion. The recombinase RecA (CCPI-G23) plays an important role in DNA repair through homologous recombination, protecting the bacterium from genotoxic stress imposed by the host immune system [Supplementary Figure 14]. Exonuclease SbcCD subunit D (CCPI-G14) and AAA family ATPase (CCPI-G15) assist in repairing double-strand breaks in DNA and removing damaged nucleotides, respectively, maintaining genome integrity. Meanwhile, the YjjG family nucleotidase (CCPI-G77) helps to remove and recycle damaged or noncanonical pyrimidine bases, preserving DNA and RNA fidelity.

The genetic adaptability of *C. chauvoei* is based on proteins that promote genome plasticity. The IS256 family transposase (CCPI-G12, G64) and the Rpn family recombination-promoting nuclease (CCPI-G2, G3, G5-7, G9) and tyrosine-type recombinase/integrase (CCPI-G29) enable horizontal gene transfer by allowing DNA segments to integrate into the genome. This capability may enhance virulence or antibiotic resistance, aiding survival under immune pressures. The DDE domain protein (CCPI-G50, G57, G72), DDE-type integrase (CCPI-G66, G68), and other recombinases further promote genetic diversity, allowing the bacterium to adapt to varying conditions within the host and potentially evade immune responses.

During the above reactions or in the TCA cycle, electron transport, DNA repair; flavin adenine dinucleotide (FAD) exists in four redox states flavin-N(5)-oxide, hydroquinone, quinone, and semiquinone; and is converted between these states by accepting or donating electrons [67]. FAD (Quinone or oxidized form) accepts two electrons and two protons to become $FADH_2$ (hydroquinone form). The oxidation of $FADH_2$ or reduction of FAD by donating or accepting one electron and one proton, respectively, to form semiquinone (FADH·) [68].

### 4.5. Genes for Managing Stress Responses

The cell structure and membrane integrity of *C. chauvoei* rely on membrane-stabilizing proteins, which are vital as the bacterium proliferates. The genes coding for CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyl transferase, Class B sortase, DegV family proteins, stage V sporulation protein S, pseudouridine synthase, and the 23S rRNA pseudouridine methyltransferase RlmH are known for managing stress responses. CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyl transferase (CCPI-G21) is involved in phospholipid biosynthesis, such as phosphatidylglycerol, the most abundant acidic phospholipid, which is essential in the outer membrane. The acidic phospholipids strengthen the cell membrane, providing more stability to bacterial cell walls. The mutants were able to grow at normal temperatures but showed slow growth at higher temperatures (>40°C) [69]. Thus, CDP-diacylglycerol-glycerol-3-phosphate 3-phosphatidyl transferase is important in stabilizing the bacterial cell walls.

Class B sortase (CCPI-G58) aids by anchoring surface proteins to the bacterial cell wall, which may enhance adhesion to host tissues and provide a structural advantage. DegV family proteins (CCPI-G43) and stage V sporulation protein S (CCPI-G25) are involved in structural maintenance and sporulation, with the latter ensuring that some cells can form spores for persistence or future transmission.

Pseudouridine is a ubiquitous constituent of structural RNA (transfer, ribosomal, small nuclear (snRNA), and small nucleolar), and is the most abundant RNA modification [70]. Pseudouridine in rRNA and tRNA has been shown to stabilize and maintain structure and functions in mRNA decoding, ribosome assembly, processing, and translation [71]. Pseudouridine synthase (CCPI-G17) is an enzyme that modifies RNA by converting uridine to pseudouridine. The absence of pseudouridine synthase demonstrated temperature sensitivity and decreased virulence in *Candida albicans* [72]. Stem–loop 69 of 23S rRNA contains three highly conserved pseudouridines (C) at positions 1911, 1915, and 1917, synthesized by the pseudouridine synthase RluD [73]. One of the pseudouridines in stem-loop 69 (position 1915) is further methylated by RlmH to take part in the final steps of ribosome biosynthesis [74,75]. Cells lacking the *rlmH* gene have a clear growth disadvantage [76], and the methyltransferase confers a fitness advantage under stress conditions. Pseudouridine synthase and the 23S rRNA pseudouridine methyltransferase RlmH (CCPI-G78) modify RNA structures, potentially increasing the stability of ribosomal RNA and enhancing protein synthesis under stress. The Hsp20 family protein (CCPI-G79) acts as a molecular chaperone, protecting bacterial proteins from denaturation under heat or immune attack.

## 5. CONCLUSION

This study offers a comprehensive genomic analysis of *C. chauvoei* strain SBP 07/09, identifying eight PIs that play crucial roles in the

bacterium's ability to cause blackleg disease in cattle. Advanced bioinformatic tools allowed for detailed functional prediction of genes, functions, and pathways within these PIs. The combinatorial use of this study linked the genomic islands or genes of the genomes to metabolic pathways, revealing critical insights into processes such as stress response, nutrient acquisition, and anaerobic survival. Key findings include the identification of genes associated with DNA recombination, glycolysis, and cellular membrane stability, which enhance the bacterium's adaptability and virulence. The study underscores the importance of these genetic elements in facilitating *C. chauvoei's* survival under anaerobic conditions typical of infected tissues, enabling efficient colonization and persistence. Future studies will include developing a bioinformatics method for linking the genomic islands of the genomes to the phenotypes via metabolic pathways and (2) using experimental validation strategies such as microarray studies to validate the proof of concept. Understanding these mechanisms provides a foundation for future research on developing vaccines or therapeutic targets of *Clostridium* species, leading to potential vaccines or drugs, ultimately helping to mitigate its economic impact on cattle farming.

## 6. ACKNOWLEDGMENTS

The author PK is thankful to GITAM (Deemed to be University) for the PhD fellowship. PK, CS and NNRR are also grateful to the management of GITAM (Deemed to be University) for providing the necessary facilities to carry out the research work and extending constant support.

## 7. AUTHORS' CONTRIBUTIONS

All authors made substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; took part in drafting the article or revising it critically for important intellectual content; agreed to submit to the current journal; gave final approval of the version to be published; and agree to be accountable for all aspects of the work. All the authors are eligible to be author as per the International Committee of Medical Journal Editors (ICMJE) requirements/guidelines.

## 8. CONFLICTS OF INTEREST

The authors report no financial or any other conflicts of interest in this work.

## 9. ETHICAL APPROVALS

This study does not involve experiments on animals or human subjects.

## 10. DATA AVAILABILITY

All the data is available with the authors and shall be provided upon request.

## 11. PUBLISHER'S NOTE

All claims expressed in this article are solely those of the authors and do not necessarily represent those of the publisher, the editors and the reviewers. This journal remains neutral with regard to jurisdictional claims in published institutional affiliation.

## 12. USE OF ARTIFICIAL INTELLIGENCE (AI)-ASSISTED TECHNOLOGY

The authors declare that they have not used artificial intelligence (AI)-tools for writing and editing of the manuscript, and no images were manipulated using AI.

## 13. SUPPLEMENTARY MATERIAL

The supplementary material can be accessed at the journal's website: Link here [https://jabonline.in/admin/php/uploadss/1442_pdf.pdf].

## REFERENCES

1. Jubb K, Kennedy P, Palmer N. Pathology of Domestic Animals; 2012. Available from: https://books.google.com/books?hl=en&lr=&id=qjFfJ8pcttMC&oi=fnd&pg=PP1&ots=oC24mbMSGt&sig=KHw3xtP2z5E1H-DfWZB43J1VGaM [Last accessed on 2024 Dec 24].
2. Lange M, Neubauer H, Seyboldt C. Development and validation of a multiplex real-time PCR for detection of *Clostridium chauvoei* and *Clostridium septicum*. Mol Cell Probes. 2010;24(4):204-10. https://doi.org/10.1016/j.mcp.2010.03.003
3. Bagge E, Lewerin SS, Johansson KE. Detection and identification by PCR of *Clostridium chauvoei* in clinical isolates, bovine faeces and substrates from biogas plant. Acta Vet Scand. 2009;51(1):8. https://doi.org/10.1186/1751-0147-51-8.
4. Rychener L, In-Albon S, Djordjevic SP, Chowdhury PR, Nicholson P, Ziech RE, *et al*. *Clostridium chauvoei*, an evolutionary dead-end pathogen. Front Microbiol. 2017;8:1054. https://doi.org/10.3389/fmicb.2017.01054
5. Uzal FA, Songer JG, Prescott JF, Popoff MR. Clostridial Diseases in Animals; 2016. p. 1-336. Available from: https://onlinelibrary.wiley.com/doi/book/10.1002/9781118728291 [Last accessed on 2024 Dec 24].
6. Popoff MR. Clostridial pore-forming toxins: Powerful virulence factors. Anaerobe. 2014;30:220-38. https://doi.org/10.1016/j.anaerobe.2014.05.014
7. Useh NM, Nok AJ, Esievo KAN. Pathogenesis and pathology of blackleg in ruminants: The role of toxins and neuraminidase. A short review. Vet Q. 2003;25(4):155-9. https://doi.org/10.1080/01652176.2003.9695158
8. Blackleg in Animals - Infectious Diseases - MSD Veterinary Manual. Available from: https://www.msdvetmanual.com/infectious-diseases/clostridial-diseases/blackleg-in-animals [Last accessed on 2024 Dec 24].
9. Frey J, Johansson A, Bürki S, Vilei EM, Redhead K. Cytotoxin CctA, a major virulence factor of *Clostridium chauvoei* conferring protective immunity against myonecrosis. Vaccine. 2012;30(37):5500-5. https://doi.org/10.1016/j.vaccine.2012.06.050
10. Sayers EW, Bolton EE, Brister JR, Canese K, Chan J, Comeau DC, *et al*. Database resources of the National Center for Biotechnology Information in 2023. Nucleic Acids Res. 2023;51(D1):D29-38. https://doi.org/10.1093/nar/gkac1032
11. Nammi D, Yarla NS, Chubarev VN, Tarasov VV, Barreto GE, Corolina Pasupulati AM, *et al*. A systematic *in-silico* analysis of *Helicobacter pylori* pathogenic islands for identification of novel drug target candidates. Curr Genomics. 2017;18(5):450-65. https://doi.org/10.2174/1389202918666170705160615
12. Bertelli C, Laird MR, Williams KP, Simon Fraser University Research Computing Group, Lau BY, Hoad G, *et al*. IslandViewer 4: Expanded prediction of genomic islands for larger-scale datasets. Nucleic Acids Res. 2017;45(W1):W30-5. https://doi.org/10.1093/nar/gkx343
13. Hsiao W, Wan I, Jones SJ, Brinkman FSL. IslandPath: Aiding detection of genomic islands in prokaryotes. Bioinformatics. 2003;19(3):418-20. https://doi.org/10.1093/bioinformatics/btg004
14. Langille MGI, Hsiao WWL, Brinkman FSL. Evaluation of genomic island predictors using a comparative genomics approach. BMC Bioinformatics. 2008;9(1):329. https://doi.org/10.1186/1471-2105-9-329
15. Waack S, Keller O, Asper R, Brodag T, Damm C, Fricke WF, *et al*.

Score-based prediction of genomic islands in prokaryotic genomes using hidden Markov models. BMC Bioinformatics. 2006;7(1):142. https://doi.org/10.1186/1471-2105-7-142
16. Besemer J, Borodovsky M. GeneMark: Web software for gene finding in prokaryotes, eukaryotes and viruses. Nucleic Acids Res. 2005;33(Web Server issue):W451-4. https://doi.org/10.1093/nar/gki487
17. Lukashin A, Borodovsky M. GeneMark.hmm: New solutions for gene finding. Nucleic Acids Res. 1998;26(4):1107-15. https://doi.org/10.1093/nar/26.4.1107
18. McGinnis S, Madden TL. BLAST: At the core of a powerful and diverse set of sequence analysis tools. Nucleic Acids Res. 2004;32(Web Server issue):W20-5. https://doi.org/10.1093/nar/gkh435
19. Paysan-Lafosse T, Blum M, Chuguransky S, Grego T, Pinto BL, Salazar GA, *et al*. InterPro in 2022. Nucleic Acids Res. 2023;51(D1):D418-27. https://doi.org/10.1093/nar/gkac993
20. Kanehisa M, Sato Y, Morishima K. BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. J Mol Biol. 2016;428(4):726-31. https://doi.org/10.1016/j.jmb.2015.11.006
21. Kanehisa M, Furumichi M, Sato Y, Ishiguro-Watanabe M, Tanabe M. KEGG: Biological systems database as a model of the real world. Nucleic Acids Res. 2023;51(D1):D672-80. https://doi.org/10.1093/nar/gkae909
22. Kanehisa M, Sato Y. KEGG Mapper for inferring cellular functions from protein sequences. Protein Sci. 2020;29(1):28-35. https://doi.org/10.1002/pro.3711
23. Karp PD, Paley SM, Krummenacker M, Kothari A, Midford PE, Subhraveti P, *et al*. Pathway Tools version 28.0: Integrated software for pathway/genome informatics and systems biology. Brief Bioinform. 2021;22(1):109-26. https://doi.org/10.1093/bib/bbz104
24. Kleine B, Chattopadhyay A, Polen T, Pinto D, Mascher T, Bott M, *et al*. The three-component system EsrISR regulates a cell envelope stress response in *Corynebacterium glutamicum*. Mol Microbiol. 2017;106(5):719-41. https://doi.org/10.1111/mmi.13839
25. Flores-Kim J, Darwin AJ. Activity of a bacterial cell envelope stress response is controlled by the interaction of a protein binding domain with different partners. J Biol Chem. 2015;290(18):11417-30. https://doi.org/10.1074/jbc.M114.614107
26. Darwin AJ. The phage-shock-protein response. Mol Microbiol. 2005;57(3):621-8. https://doi.org/10.1111/j.1365-2958.2005.04694.x
27. Gruber TM, Gross CA. Multiple sigma subunits and the partitioning of bacterial transcription space. Annu Rev Microbiol. 2003;57(1):441-66. https://doi.org/10.1146/annurev.micro.57.030502.090913
28. Kang JG, Hahn MY, Ishihama A, Roe JH. Identification of sigma factors for growth phase-related promoter selectivity of RNA polymerases from *Streptomyces coelicolor A3(2)*. Nucleic Acids Res. 1997;25(13):2566-73. https://doi.org/10.1093/nar/25.13.2566
29. Yang B, Nie X, Gu Y, Jiang W, Yang C. Control of solvent production by sigma-54 factor and the transcriptional activator AdhR in *Clostridium beijerinckii*. Microb Biotechnol. 2020;13(2):328-38. https://doi.org/10.1111/1751-7915.13505
30. Anderson DK, Ohta N, Wu J, Newton A. Regulation of the *Caulobacter crescentus* rpoN gene and function of the purified sigma 54 in flagellar gene transcription. Mol Gen Genet. 1995;246(6):697-706. https://doi.org/10.1007/BF00290715
31. Arora SK, Ritchings BW, Almira EC, Lory S, Ramphal R. A transcriptional activator, FleQ, regulates mucin adhesion and flagellar gene expression in *Pseudomonas aeruginosa* in a cascade manner. J Bacteriol. 1997;179(17):5574-81. https://doi.org/10.1128/jb.179.17.5574-5581.1997
32. Jiménez-Fernández A, López-Sánchez A, Jiménez-Díaz L, Navarrete B, Calero P, Platero AI, *et al*. Complex interplay between FleQ, cyclic diguanylate and multiple σ factors coordinately regulates flagellar motility and biofilm development in *Pseudomonas putida*. PLoS One. 2016;11(9):e0163142. https://doi.org/10.1371/journal.pone.0163142
33. Correa NE, Peng F, Klose KE. Roles of the regulatory proteins FlhF and FlhG in the *Vibrio cholerae* flagellar transcription hierarchy. J Bacteriol. 2005;187(18):6324-32. https://doi.org/10.1128/JB.187.18.6324-6332.2005
34. Martinez RM, Dharmasena MN, Kirn TJ, Taylor RK. Characterization of two outer membrane proteins, FlgO and FlgP, that influence *Vibrio cholerae* motility. J Bacteriol. 2009;191(18):5669-79. https://doi.org/10.1128/JB.00632-09
35. Shao X, Zhang X, Zhang Y, Zhu M, Yang P, Yuan J, *et al*. RpoN-dependent direct regulation of quorum sensing and the type VI secretion system in *Pseudomonas aeruginosa PAO1*. J Bacteriol. 2018;200(16):e00205-18. https://doi.org/10.1128/JB.00205-18
36. Matsuyama BY, Krasteva PV, Baraquet C, Harwood CS, Sondermann H, Navarro MVAS. Mechanistic insights into c-di-GMP-dependent control of the biofilm regulator FleQ from *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A. 2016;113(2):E209-18. https://doi.org/10.1073/pnas.1523148113
37. Lloyd MG, Lundgren BR, Hall CW, Gagnon LBP, Mah TF, Moffat JF, *et al*. Targeting the alternative sigma factor RpoN to combat virulence in *Pseudomonas aeruginosa*. Sci Rep. 2017;7(1):12615. https://doi.org/10.1038/s41598-017-12667-y
38. Hayrapetyan H, Tempelaars M, Nierop Groot M, Abee T. *Bacillus cereus ATCC 14579* RpoN (Sigma 54) is a pleiotropic regulator of growth, carbohydrate metabolism, motility, biofilm formation and toxin production. PLoS One. 2015;10(8):e0134872. https://doi.org/10.1371/journal.pone.0134872
39. Arai H, Hayashi M, Kuroi A, Ishii M, Igarashi Y. Transcriptional regulation of the flavohemoglobin gene for aerobic nitric oxide detoxification by the second nitric oxide-responsive regulator of *Pseudomonas aeruginosa*. J Bacteriol. 2005;187(12):3960-8. https://doi.org/10.1128/JB.187.12.3960-3968.2005
40. Weiner L, Brissette JL, Model P. Stress-induced expression of the *Escherichia coli* phage shock protein operon is dependent on sigma 54 and modulated by positive and negative feedback mechanisms. Genes Dev. 1991;5(10):1912-23. https://doi.org/10.1101/gad.5.10.1912
41. Ma M, Welch RD, Garza AG. The σ54 system directly regulates bacterial natural product genes. Sci Rep. 2021;11(1):4771. https://doi.org/10.1038/s41598-021-84057-4
42. Grove A. Regulation of metabolic pathways by MarR family transcription factors. Comput Struct Biotechnol J. 2017;15:366-71. https://doi.org/10.1016/j.csbj.2017.06.001
43. Weatherspoon-Griffin N, Wing HJ. Characterization of SlyA in *Shigella flexneri* identifies a novel role in virulence. Infect Immun. 2016;84(4):1073-82. https://doi.org/10.1128/IAI.00806-15
44. Li Z, Li W, Lu J, Liu Z, Lin X, Liu Y. Quantitative proteomics analysis reveals the effect of a MarR family transcriptional regulator AHA_2124 on *Aeromonas hydrophila*. Biology (Basel). 2023;12(12):1473. https://doi.org/10.3390/biology12121473
45. Grove A. MarR family transcription factors. Curr Biol. 2013;23(4):R142-3. https://doi.org/10.1016/j.cub.2013.01.013
46. Matos RG, Casinhas J, Bárria C, Dos Santos RF, Silva IJ, Arraiano CM. The role of ribonucleases and sRNAs in the virulence of foodborne pathogens. Front Microbiol. 2017;8:910. https://doi.org/10.3389/fmicb.2017.00910
47. Jester BC, Romby P, Lioliou E. When ribonucleases come into play in pathogens: A survey of Gram-positive bacteria. Int J Microbiol. 2012;2012:592196. https://doi.org/10.1155/2012/592196
48. Wu H, Rebello O, Crost EH, Owen CD, Walpole S, Bennati-Granier C, *et al*. Fucosidases from the human gut symbiont

*Ruminococcus gnavu*s. Cell Mol Life Sci. 2021;78(2):675-93. https://doi.org/10.1007/s00018-020-03514-x
49. Sakurama H, Tsutsumi E, Ashida H, Katayama T, Yamamoto K, Kumagai H. Differences in the substrate specificities and active-site structures of two α-L- fucosidases (glycoside hydrolase family 29) from *Bacteroides thetaiotaomicron*. Biosci Biotechnol Biochem. 2012;76(5):1022-4. https://doi.org/10.1271/bbb.111004
50. Sela DA, Garrido D, Lerno L, Wu S, Tan K, Eom HJ, *et al*. *Bifidobacterium longum* subsp. infantis ATCC 15697 α-Fucosidases are active on fucosylated human milk oligosaccharides. Appl Environ Microbiol. 2012;78(3):795-803. https://doi.org/10.1128/AEM.06762-11
51. Kostopoulos I, Elzinga J, Ottman N, Klievink JT, Blijenberg B, Aalvink S, *et al*. *Akkermansia muciniphila* uses human milk oligosaccharides to thrive in the early life conditions *in vitro*. Sci Rep. 2020;10(1):14330. https://doi.org/10.1038/s41598-020-71113-8
52. Pickard JM, Maurice CF, Kinnebrew MA, Abt MC, Schenten D, Golovkina TV, *et al*. Rapid fucosylation of intestinal epithelium sustains host-commensal symbiosis in sickness. Nature. 2014;514(7524):638-41. https://doi.org/10.1038/nature13823
53. Lei C, Sun R, Xu G, Tan Y, Feng W, McClain CJ, *et al*. Enteric VIP-producing neurons maintain gut microbiota homeostasis through regulating epithelium fucosylation. Cell Host Microbe. 2022;30(10):1417-34.e8. https://doi.org/10.1016/j.chom.2022.09.001
54. Dang S, Sun L, Huang Y, Lu F, Liu Y, Gong H, *et al*. Structure of a fucose transporter in an outward-open conformation. Nature. 2010;467(7316):734-8. https://doi.org/10.1038/nature09406
55. Pao SS, Paulsen IT, Saier MH. Major facilitator superfamily. Microbiol Mol Biol Rev. 1998;62(1):1-34. https://doi.org/10.1128/MMBR.62.1.1-34.1998
56. Law CJ, Maloney PC, Wang DN. Ins and outs of major facilitator superfamily antiporters. Annu Rev Microbiol. 2008;62:289-305. https://doi.org/10.1146/annurev.micro.61.080706.093329
57. Flemming HC, Wingender J. The biofilm matrix. Nat Rev Microbiol. 2010;8(9):623-33. https://doi.org/10.1038/nrmicro2415
58. Fleming D, Rumbaugh KP. Approaches to dispersing medical biofilms. Microorganisms. 2017;5(2):15. https://doi.org/10.3390/microorganisms5020015.
59. Baker P, Hill PJ, Snarr BD, Alnabelseya N, Pestrak MJ, Lee MJ, *et al*. Exopolysaccharide biosynthetic glycoside hydrolases can be utilized to disrupt and prevent *Pseudomonas aeruginosa* biofilms. Sci Adv. 2016;2(5):e1501632. https://doi.org/10.1126/sciadv.1501632
60. Plumbridge J. An alternative route for recycling of N-acetylglucosamine from peptidoglycan involves the N-acetylglucosamine phosphotransferase system in *Escherichia coli.* J Bacteriol. 2009;191(18):5641-7. https://doi.org/10.1128/JB.00448-09.
61. Amidohydrolase family (PF01979) - Pfam entry - InterPro. Available from: https://www.ebi.ac.uk/interpro/entry/pfam/PF01979 [Last accessed on 2024 Dec 24].
62. Kim HM, Park YH, Yoon CK, Seok YJ. Histidine phosphocarrier protein regulates pyruvate kinase A activity in response to glucose in *Vibrio vulnificus*. Mol Microbiol. 2015;96(2):293-305. https://doi.org/10.1111/mmi.12936
63. Deutscher J, Aké FMD, Derkaoui M, Zébré AC, Cao TN, Bouraoui H, *et al*. The bacterial phosphoenolpyruvate: carbohydrate phosphotransferase system: Regulation by protein phosphorylation and phosphorylation-dependent protein-protein interactions. Microbiol Mol Biol Rev. 2014;78(2):231-56. https://doi.org/10.1128/MMBR.00001-14
64. Postma PW, Lengeler JW, Jacobson GR. Phosphoenolpyruvate: Carbohydrate phosphotransferase systems of bacteria. Microbiol Rev. 1993;57(3):543-94. https://doi.org/10.1128/mr.57.3.543-594.1993
65. Hogema BM, Arents JC, Bader R, Eijkemans K, Yoshida H, Takahashi H, *et al*. Inducer exclusion in *Escherichia coli* by non-PTS substrates: The role of the PEP to pyruvate ratio in determining the phosphorylation state of enzyme IIAGlc. Mol Microbiol. 1998;30(3):487-98. https://doi.org/10.1046/j.1365-2958.1998.01053.x
66. Parey K, Warkentin E, Kroneck PMH, Ermler U. Reaction cycle of the dissimilatory sulfite reductase from *Archaeoglobus fulgidus.* Biochemistry. 2010;49(41):8912-21. https//doi.org/10.1021/bi100781f
67. Teufel R, Agarwal V, Moore BS. Unusual flavoenzyme catalysis in marine bacteria. Curr Opin Chem Biol. 2016;31:31-9. https://doi.org/10.1016/j.cbpa.2016.01.001
68. Teufel R, Miyanaga A, Michaudel Q, Stull F, Louie G, Noel JP, *et al*. Flavin-mediated dual oxidation controls an enzymatic Favorskii-type rearrangement. Nature. 2013;503(7477):552-6. https://doi.org/10.1038/nature12643
69. Kikuchi S, Shibuya I, Matsumoto K. Viability of an *Escherichia coli* pgsA null mutant lacking detectable phosphatidylglycerol and cardiolipin. J Bacteriol. 2000;182(2):371-6. https://doi.org/10.1128/JB.182.2.371-376.2000
70. Penzo M, Guerrieri A, Zacchini F, Treré D, Montanaro L. RNA pseudouridylation in Physiology and medicine: For better and for worse. Genes (Basel). 2017;8(11):301. https://doi.org/10.3390/genes8110301
71. Davis FF, Allen FW. Ribonucleic acids from yeast which contain a fifth nucleotide. J Biol Chem. 1957;227(2):907-15.
72. Pickerill ES, Kurtz RP, Tharp A, Guerrero Sanz P, Begum M, Bernstein DA. Pseudouridine synthase 7 impacts *Candida albicans* rRNA processing and morphological plasticity. Yeast. 2019;36(11):669-77. https://doi.org/10.1002/yea.3436
73. Huang L, Ku J, Pookanjanatavip M, Gu X, Wang D, Greene PJ, *et al*. Identification of two *Escherichia coli* pseudouridine synthases that show multisite specificity for 23S RNA. Biochemistry. 1998;37(45):15951-7. https://doi.org/10.1021/bi981002n
74. Kowalak JA, Bruenger E, Hashizume T, Peltier JM, Ofengand J, McCloskey JA. Structural characterization of U*-1915 in domain IV from *Escherichia coli* 23S ribosomal RNA as 3-methylpseudouridine. Nucleic Acids Res. 1996;24(4):688-93. https://doi.org/10.1093/nar/24.4.688
75. Ero R, Peil L, Liiv A, Remme J. Identification of pseudouridine methyltransferase in *Escherichia coli.* RNA. 2008;14(10):2223-33. https://doi.org/10.1261/rna.1186608
76. Purta E, Kaminska KH, Kasprzak JM, Bujnicki JM, Douthwaite S. YbeA is the m3Psi methyltransferase RlmH that targets nucleotide 1915 in 23S rRNA. RNA. 2008;14(10):2234-44. https://doi.org/10.1261/rna.1198108
77. Jiang T, Gao C, Ma C, Xu P. Microbial lactate utilization: Enzymes, pathogenesis, and regulation. Trends Microbiol. 2014;22(10):589-99. https://doi.org/10.1016/j.tim.2014.05.008

**How to cite this article**: Pasam K, Challa S, Neelapu R. *In silico* analysis of pathogenic islands and metabolic pathways for understanding infection biology and lifestyle of *Clostridium chauvoei.* J Appl Biol Biotech 2026;14(2):212-225. DOI: 10.7324/JABB.2026.253744