

Preplanned Studies

Identification and Genomic Characterization of *Corynebacterium belfantii* Strains — Fujian Province, China, 2019–2025

Wenwei Wu¹; Yuhang Deng¹; Rui Jin¹; Junjie Mao¹; Yadong Gao^{2,3}; Quwen Li^{1,2,3,#}

Summary

What is already known about this topic?

This is the first report of *Corynebacterium belfantii* (*C. belfantii*) in Fujian; genomic characterization confirms misidentification by routine methods and reveals regional phylogenetic characteristics.

What is added by this report?

This report adds the first local occurrence evidence, new genomic sequences, a distinct phylogenetic sub-branch, and a reinforced surveillance strategy for Fujian Province.

What are the implications for public health practice?

It emphasizes that laboratories should adopt molecular methods to accurately distinguish *C. belfantii*, strengthen surveillance of non-toxigenic strains, update MALDI-TOF databases, and monitor virulence and resistance genes despite the absence of diphtheria toxin.

ANI analysis showed 99.16%–99.22% identity with the reference genome GCA_041898185. This confirmed their classification as *C. belfantii*. Phylogenetic analysis indicated high conservation of the *dtxR* gene in *C. belfantii*. Pangenomic evolutionary trees showed that the Fujian strains clustered within the same major branch as global strains but formed a distinct sub-branch.

Conclusion: This study reports the first identification of *C. belfantii* in Fujian Province and describes its genomic characteristics and evolutionary position. These findings suggest that clinical and public health laboratories should use molecular methods to accurately distinguish members of the non-toxigenic *C. diphtheriae* complex and strengthen laboratory surveillance capacity.

ABSTRACT

Introduction: This study aimed to identify three non-toxigenic *Corynebacterium diphtheriae* (*C. diphtheriae*) complex strains isolated in Fujian Province between 2019 and 2025 and to determine their taxonomic classification through integrated phenotypic and genomic analyses.

Methods: Biochemical identification, matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), and polymerase chain reaction (PCR) were performed on three suspected *C. diphtheriae* isolates. Whole-genome sequencing was conducted using a nanopore sequencing platform. Average nucleotide identity (ANI) was calculated, and phylogenetic trees based on the 16S rRNA, the *dtxR* gene, and the pangenome were constructed, followed by comparative analysis against global NCBI sequences.

Results: All three strains were identified as *C. diphtheriae* by biochemical and mass spectrometric methods, with PCR detecting no *tox* gene carriage.

Over 160 species of the genus *Corynebacterium* have been identified to date, among which *Corynebacterium diphtheriae* (*C. diphtheriae*), *Corynebacterium ulcerans* (*C. ulcerans*), and *Corynebacterium pseudotuberculosis* (*C. pseudotuberculosis*) are recognized as pathogenic species capable of producing diphtheria toxin (DT). Subsequently, the *C. diphtheriae* complex was formally defined, encompassing additional potentially toxigenic species, namely, *C. belfantii*, *C. rouxii*, and *C. silvaticum* (1–2). These key pathogens have substantial implications for global public health, with broad host ranges encompassing both humans, including *C. diphtheriae*, *C. belfantii*, and *C. rouxii*, and wild and domestic mammals, including *C. ulcerans*, *C. pseudotuberculosis*, and *C. silvaticum* (1–2). With advances in genomic identification technologies and molecular phylogenetic analysis methods, the understanding of the ecological and host niches and pathogenic factors of *Corynebacterium* species previously regarded as commensals, particularly those isolated from human samples, has progressively deepened, highlighting their clinical importance.

This study analyzed three non-toxicogenic *C. diphtheriae* isolates obtained in Fujian Province from 2019–2025 through biochemical identification, mass spectrometry, polymerase chain reaction (PCR), and whole-genome sequencing. Comparative phylogenetic analysis with global National Center for Biotechnology Information (NCBI) sequences was performed to enhance epidemiological data on *Corynebacterium belfantii* and provide scientific evidence for public health laboratory diagnostics.

Three isolates, namely, FJ_201901, FJ_202501, and FJ_202502, from cases in Fujian Province were initially identified as *C. diphtheriae* by mass spectrometry and biochemical testing. Specimens were collected from the sputum, nasal secretions, and bronchoalveolar lavage fluid of female patients aged 25, 68, and 52 years, respectively. For comparative analysis, 33 *C. belfantii*, 15 *C. silvaticum*, six *C. rouxii*, three *C. diphtheriae* (strains NCTC11397, PW8, NCTC 13129), one *C. diphtheriae* bv. *gravis* (strain 5521/17), one *C. diphtheriae* bv. *mitis* (strain 230/19), and one *C. ulcerans* genome sequences were downloaded from NCBI, totaling 63 strains for analysis.

Preliminary identification was performed using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS), alongside biochemical identification using the VITEK CBC *Corynebacterium* identification card (bioMérieux China Ltd.). Strain identification was conducted using a *C. diphtheriae* nucleic acid detection kit (Shenzhen Bioyuan Life Science Co., Ltd.). Meanwhile, detection of the virulence gene *tox* in *C. diphtheriae* was performed using a duplex nucleic acid detection kit for *toxA* and *toxB* (Beijing Zhucheng Huisheng Biotechnology Co., Ltd.).

Genomic DNA was extracted using the DNeasy UltraClean Microbial Kit (QIAGEN, Venlo, The Netherlands). Library preparation was performed using the Rapid Genome Barcoding Kit (set)-24 (V3.0, Qitan Technology Co., Ltd., Chengdu), and sequencing was performed on the QNOME-3841 platform (Qitan Technology Co., Ltd., Chengdu) using QDS (V3.0, Qitan Technology Co., Ltd., Chengdu) sequencing reagents. Following quality control using FastQC (version 0.12.1, Babraham Bioinformatics, Cambridge, UK), genomes were assembled using SPAdes (version 4.0.0, Saint Petersburg State University, St. Petersburg, Russia). Average nucleotide identity (ANI) was calculated using fastANI (version 1.34, GitHub, San Francisco, CA,

USA). Single-nucleotide polymorphism (SNP) phylogenetic analysis was performed using snippy (version 4.6.0, Victorian Bioinformatics Consortium, Melbourne, Australia) for whole-genome alignment, followed by recombination removal using Gubbins (version 2.3.4, University of Oxford, Oxford, UK). FastTree (version 2.1.10, University of Washington, Seattle, WA, USA) was used for phylogenetic tree construction, with visualization in ChiPlot. Pangenome analysis was performed using Roary based on GFF3 files generated by Prokka (version 1.14.6, Victorian Bioinformatics Consortium, Melbourne, Australia).

All three strains yielded biochemical identification codes of 35431300405010, with a 95% probability of identification as *C. diphtheriae*. Mass spectrometry scores ranged from 2.25 to 2.31, indicating reliable identification as *C. diphtheriae*. Fluorescence PCR confirmed the absence of *toxA* and *toxB* genes in all three strains.

ANI analysis showed that Fujian strains FJ_201901, FJ_202501, and FJ_202502 shared 99.17%, 99.16%, and 99.22% identity, respectively, with the reference genome GCA_041898185. ANI values ranged from 94.85 to 100% compared with those of 33 other *C. belfantii* strains; from 79.28 to 79.94% with *C. silvaticum*; from 91.51 to 91.80% with *C. rouxii*; from 94.85 to 95.18% with *C. diphtheriae*; and from 79.61 to 79.82% with *C. ulcerans* (Figure 1). These results confirmed that FJ_201901, FJ_202501, and FJ_202502 were *C. belfantii*.

16S rRNA phylogenetic analysis indicated high conservation of the 16S rRNA gene throughout the genus *Corynebacterium*, establishing it as a reliable marker for genus-level classification (Figure 2). The clustering pattern resembled that of the *dtxR* gene tree (Figure 3), with strains clustering primarily by species or biovar. However, 16S rRNA could not distinguish extremely closely related strains or biovars, such as different virulence subtypes of *C. diphtheriae*. Meanwhile, the *dtxR* gene tree showed higher resolution. *C. belfantii* strains constituted the largest group, forming the main branch with relatively short branch lengths. This indicated high conservation of *dtxR* gene sequences within this species. Different biovars of *C. diphtheriae* (bv. *gravis*, bv. *mitis*) and related species (*C. ulcerans*, *C. silvaticum*, *C. rouxii*) formed independent small branches in the upper portion of the tree, consistent with their taxonomic positions. Independent clustering of *C. diphtheriae* biovars suggested that *dtxR* gene sequence

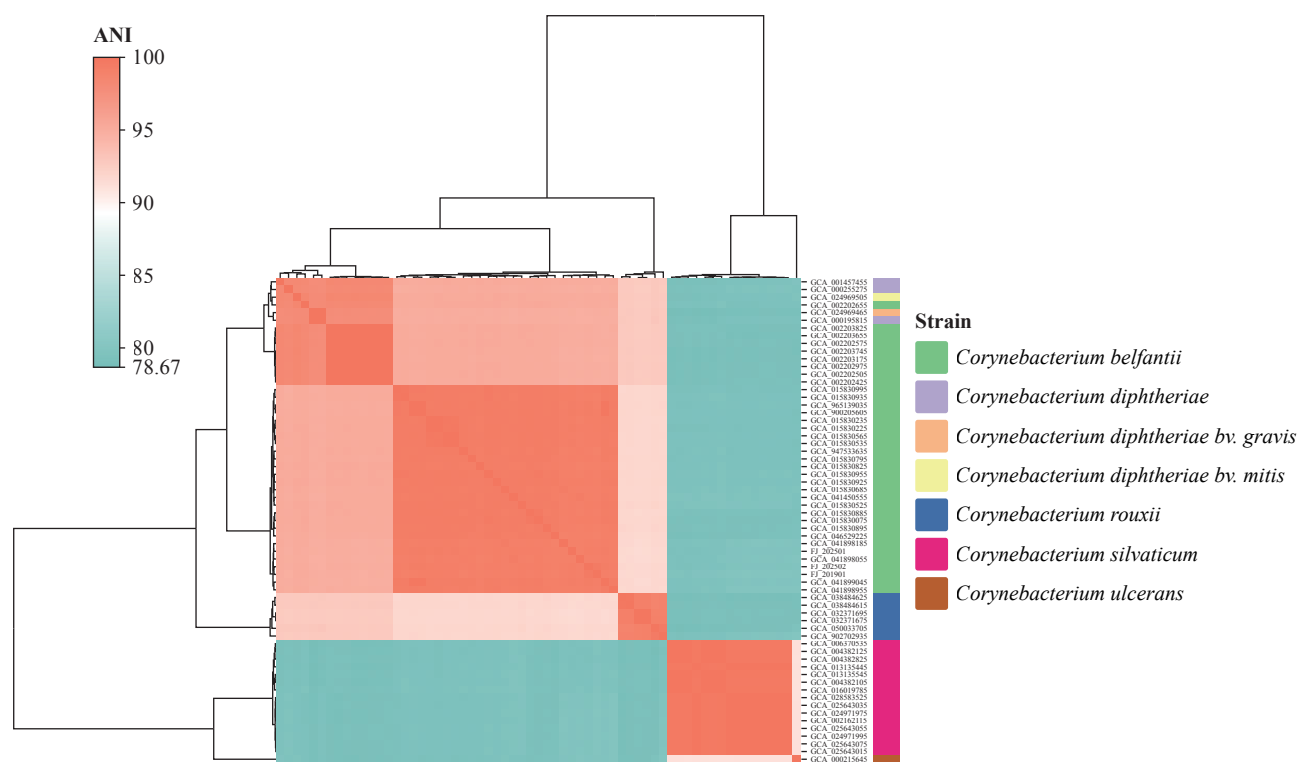


FIGURE 1. ANI analysis of 63 strains of the *Corynebacterium diphtheriae* complex. Abbreviation: ANI=average nucleotide identity.

differentiation may correlate with virulence subtype evolution. This phylogenetic tree validated species differentiation patterns within *Corynebacterium* and provided molecular evolutionary evidence for subsequent investigation of associations between *dtxR* gene variation and virulence phenotypes. 16S rRNA phylogeny primarily serves to verify genus- and species-level taxonomic status, whereas *dtxR* gene phylogeny proves more suitable for exploring fine-scale evolutionary relationships within species, such as among different virulence subtypes.

To analyze genome evolution, virulence gene distribution, and the epidemiological characteristics of *Corynebacterium* strains, a phylogenetic tree was constructed based on Roary pangenome analysis. This analysis integrated multilocus sequence typing (MLST) and presence/absence information for key virulence genes (*narK*, *dtxR*, *tox*) (Figure 4). The tree scale (0.3) reflects genetic distances at the pangenome level, indicating substantial whole-genome differences among strains. The results showed clear clustering of all strains primarily by species or biovar, with *C. belfantii* constituting the largest group and exhibiting branch structures highly consistent with traditional taxonomy. Strains within certain branches shared geographic origins, such as China or France, or host

types, such as human or swine, suggesting geographic or host specificity in strain transmission. The gene distribution matrix analysis indicated the conserved presence of *dtxR* across most strains. This finding is consistent with its role as a core regulatory gene, whereas *tox* genes were carried only by a few virulent *C. diphtheriae* biovars (bv. *gravis* and bv. *mitis*), correlating with pathogenic phenotypes. Carriage rates of other genes, such as *narK*, showed species specificity and may serve as auxiliary typing markers. Based on species annotation, non-pathogenic species, such as *C. belfantii*, generally lacked *tox* genes. Meanwhile, virulent biovars stably carried these genes, supporting the correlation between gene carriage and phenotype. The high resolution of the pangenome phylogenetic tree revealed fine-scale intra- and interspecies evolutionary relationships and provided whole-genome evidence. These findings, together with geographic and host information, provide a basis for tracing strain origins, transmission pathways, and cross-species transmission risks.

DISCUSSION

The *Corynebacterium diphtheriae* complex has undergone frequent taxonomic updates in recent years,

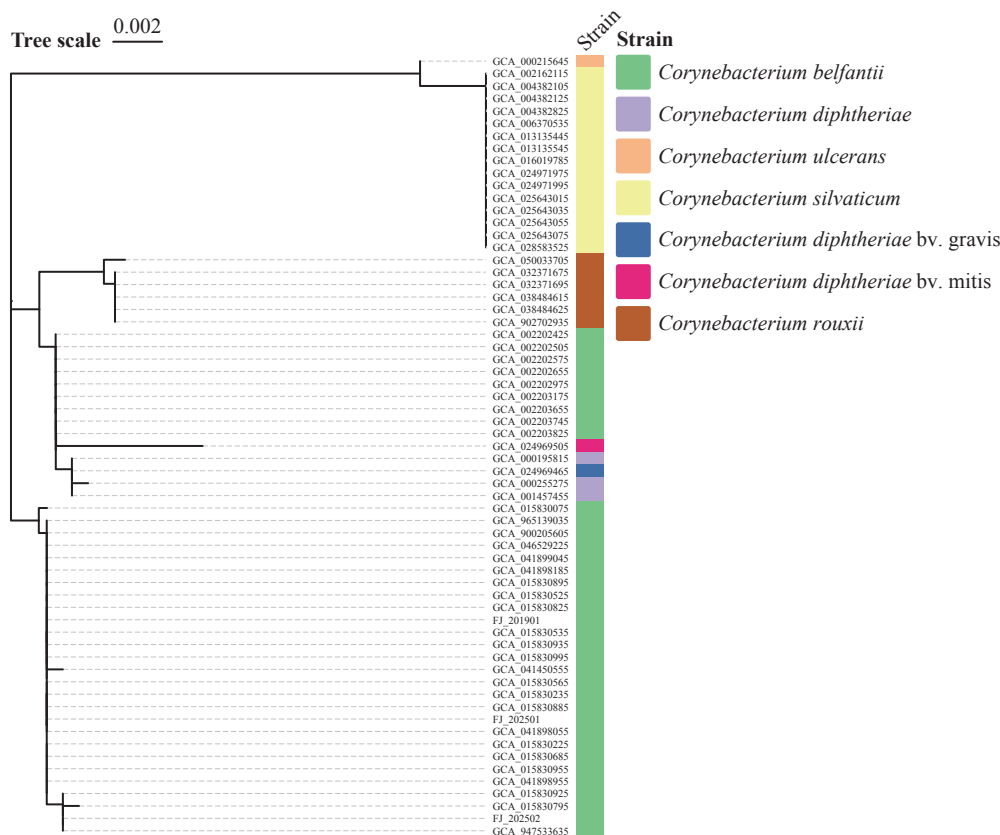


FIGURE 2. 16S rRNA gene-based phylogenetic tree of the *Corynebacterium diphtheriae* complex. Abbreviation: rRNA=ribosomal RNA; RNA=ribonucleic acid.

with *C. belfantii*, a newly described species sharing high phenotypic similarity with classical *C. diphtheriae* but typically lacking *tox* gene carriage (3–4). This presents challenges for routine microbiological identification and epidemiological surveillance. To our knowledge, this study represents the first biochemical, mass spectrometric, molecular, and genomic analysis of three non-toxigenic *C. diphtheriae* strains isolated in Fujian Province between 2019 and 2025, confirming their classification as *C. belfantii*. Comparative phylogenetic analysis with global sequences preliminarily characterized their molecular characteristics, virulence regulatory mechanisms, and evolutionary status.

Traditional biochemical identification methods could not accurately identify the three strains as *C. belfantii* because of their highly similar biochemical metabolic profiles to those of *C. diphtheriae*. This highlights the limitations of phenotype-based identification systems in distinguishing closely related species within the complex. MALDI-TOF mass spectrometry similarly could not differentiate *C. belfantii* from closely related species within the

complex because of the absence of *C. belfantii* in reference databases, highlighting the need for database updates or custom library construction for accurate identification. Through PCR, genome sequencing, and ANI analysis, the species classification of the three strains was definitively determined. This highlights the need for clinical and public health laboratories to establish molecular biology- or genomics-based confirmatory protocols as the gold standard for non-toxigenic *C. diphtheriae* isolates to prevent underdetection or misreporting.

Genomic analysis showed that all three *C. belfantii* strains lacked the *tox* gene cassette and lysogenic phage elements carrying this gene, consistent with the typically non-toxigenic nature of this species. Although *C. belfantii* does not produce diphtheria toxin, it retains the capacity for colonization and localized infection. Its public health relevance as an opportunistic pathogen should not be overlooked (5–8). The *dtxR* gene phylogenetic tree elucidated species clustering patterns and gene evolutionary characteristics within the genus *Corynebacterium*. The high conservation among *C. belfantii* strains in the

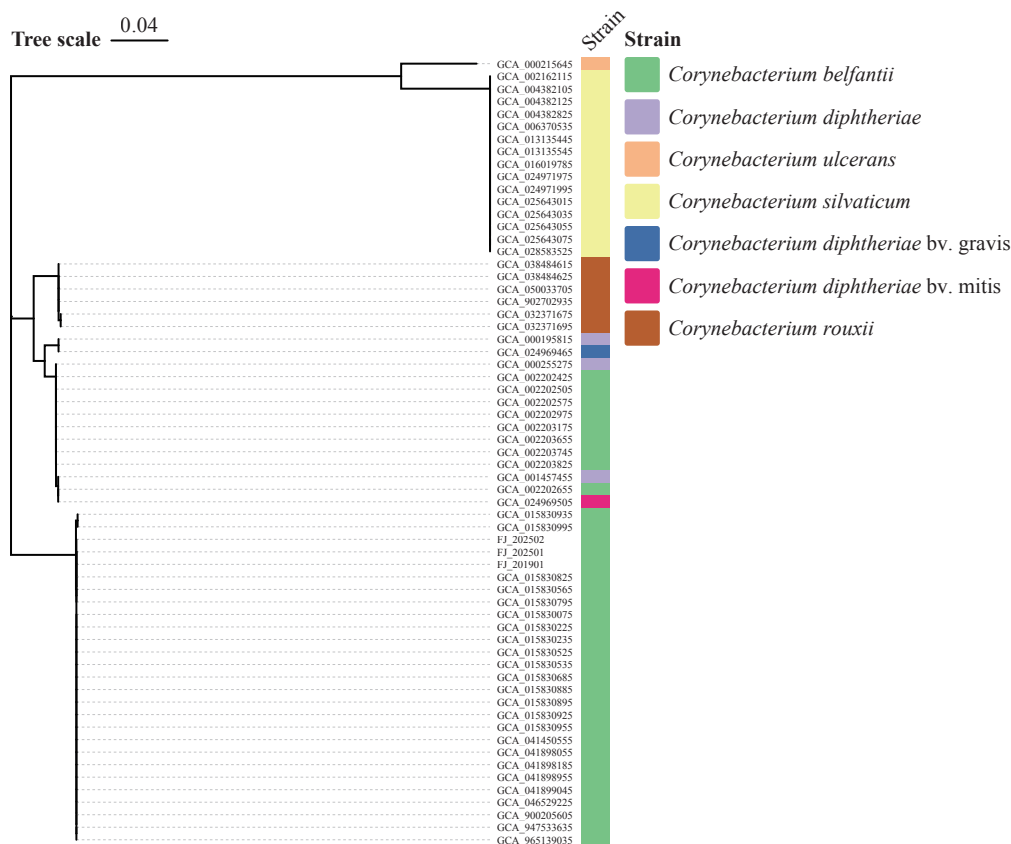


FIGURE 3. *dtxR* gene-based phylogenetic tree of the *Corynebacterium diphtheriae* complex..
Abbreviation: *dtxR*=diphtheria toxin repressor.

evolutionary tree suggests strong purifying selection pressure on the *dtxR* gene in this species. This is consistent with its non-pathogenic or weakly pathogenic phenotype. Independent clustering of different *C. diphtheriae* biovars (*gravis* and *mitis*) suggests that *dtxR* gene sequence variation may co-occur with virulence subtype differentiation. This finding aligns with previous reports associating diphtheria toxin expression levels with biovars, supporting the central role of *dtxR* in virulence regulation (9–12). Branching relationships among closely related species *C. ulcerans*, *C. silvaticum*, and *C. diphtheriae* reflect the balance between conservation and variation of the *dtxR* gene during species evolution. This validates the taxonomic framework of *Corynebacterium* and provides molecular evidence for tracing the horizontal transfer and evolutionary pathways of virulence genes.

Further pangenome-based phylogenetic tree construction, combined with MLST typing and virulence gene analysis, revealed evolutionary patterns and virulence characteristics of *Corynebacterium* at the whole-genome level. The high resolution of the

pangenome phylogenetic tree revealed genomic differences among strains, with clustering patterns highly consistent with traditional taxonomy. This supports the reliability of the species classification. Whole-genome SNP-based phylogenetic trees showed that the three Fujian *C. belfantii* strains (FJ_201901, FJ_202501, and FJ_202502) clustered within the same major branch as strains from the NCBI database (GCA_015830685, GCA_041450555, GCA_041898055, GCA_041898955, and GCA_041898185), indicating close genetic relationships. Strain FJ_201901 and FJ_202501 clustered in the same sub-branch, suggesting a possible cross-regional transmission or a shared evolutionary origin. However, the Fujian strains also formed an independent small cluster, implying potential regional microevolutionary characteristics.

The conserved presence of the virulence regulatory gene *dtxR* across all 63 strains indicates its fundamental function as a core regulatory gene in this bacterial genus. Meanwhile, *tox* gene carriage exclusively in a few virulent *C. diphtheriae* biovars suggests that acquisition of diphtheria toxin represents a critical

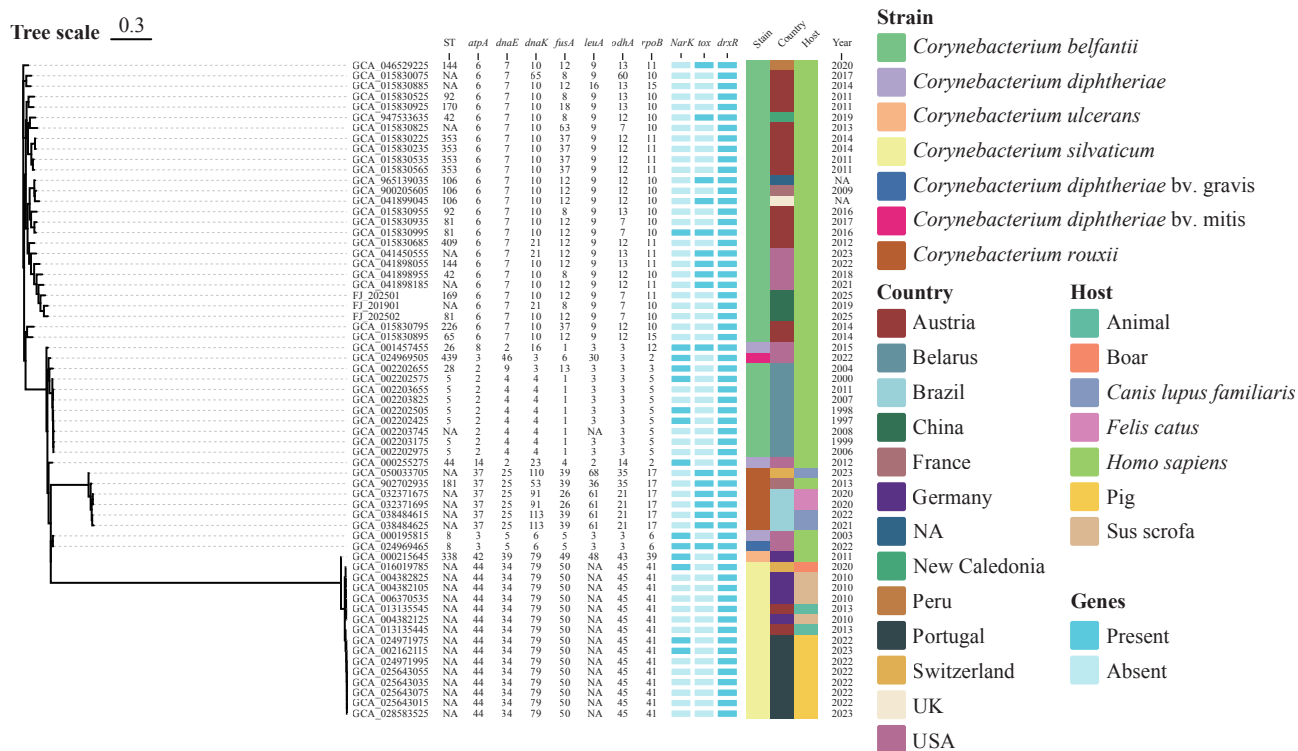


FIGURE 4. Roary-based core-genome phylogenetic tree of the *Corynebacterium diphtheriae* complex.

Abbreviation: MLST=multilocus sequence typing; narK=nitrate reductase; dtxR=diphtheria toxin repressor; tox=diphtheria toxin; NA=Not applicable; UK=United Kingdom; USA=United States of America.

event in the evolution of high pathogenicity, with stable inheritance restricted to specific lineages. This distribution pattern aligns with clinical observations that diphtheria cases are caused only by specific virulent strains. This also provides molecular targets for the rapid screening of virulent strains. Comparison of the pangenome analysis with single-gene phylogenetic trees, such as 16S rRNA and *dtxR* phylogenetic trees, showed that pangenome phylogeny enables higher-resolution analysis of intraspecies and interbiovar evolutionary relationships. Combined with the MLST typing results, this facilitates tracing of strain origins and transmission pathways, providing robust supporting data for tracking strain sources and understanding global dissemination patterns. Future studies should incorporate additional Asian strain sequences for comparison to more clearly delineate regional epidemiological characteristics of *C. beltantii*.

The findings of this study have implications for laboratory surveillance of diphtheria and diphtheria-like diseases. First, the carriage rate of non-toxigenic *C. diphtheriae* complex strains in human populations may be underestimated, particularly under immune pressure, where these strains might fill ecological niches. Second, laboratories should enhance

identification capabilities by avoiding the misidentification of *C. belfantii* as non-toxigenic *C. diphtheriae* and instead using accurate species names to accumulate precise epidemiological data. Third, the conservation of *dtxR* as a core virulence regulatory element suggests its potential as a molecular detection target for assessing strain virulence potential. Although *C. belfantii* does not produce diphtheria toxin, carriage of other virulence and resistance genes warrants continued monitoring.

However, this study has certain limitations. First, the small sample size, that is, only three strains from specific regions and time periods, may not fully reflect the prevalence patterns and genetic diversity of *C. belfantii* in Fujian Province or throughout China. Second, the lack of detailed clinical data, such as patient symptoms, vaccination history, and underlying diseases, limits the assessment of the actual pathogenic risk and clinical relevance of these strains. Future studies should expand the sample size, conduct case-control studies incorporating clinical data, and perform in vitro cellular and in vivo animal model validation of pathogenic potential. Genome-wide association studies (GWAS) should be used to further elucidate associations between genomic variations and virulence

phenotypes or host adaptation, particularly the impact of *dtxR* gene variants on strain virulence phenotypes. This can provide more precise molecular targets for diphtheria and related disease prevention and control.

Conflicts of interest: No conflicts of interest.

Funding: Supported in part by a grant from the Fujian Provincial Science and Technology Plan Project (2026J0011260).

doi: [10.46234/ccdcw2026.148](https://doi.org/10.46234/ccdcw2026.148)

Corresponding author: Quwen Li, lqw0043@163.com.

¹ School of Public Health, Fujian Medical University, Fuzhou City, Fujian Province, China; ² Fujian Provincial Center for Disease Control and Prevention, Fuzhou City, Fujian Province, China; ³ Fujian Provincial Key Laboratory of Zoonoses Research, Fujian Provincial Institute of Preventive Medicine, Fuzhou City, Fujian Province, China.

Copyright © 2026 by Chinese Center for Disease Control and Prevention & Chinese Academy of Preventive Medicine. All content is distributed under a Creative Commons Attribution Non Commercial License 4.0 (CC BY-NC).

Submitted: May 09, 2026

Accepted: June 26, 2026

Issued: July 17, 2026

REFERENCES

- Mitchell BI, Markantonis JE. An underestimated pathogen: *Corynebacterium* species. *J Clin Microbiol* 2025;63(10):e01552 – 24. <https://doi.org/10.1128/jcm.01552-24>.
- Dazas M, Badell E, Carmi-Leroy A, Criscuolo A, Brisse S. Taxonomic status of *Corynebacterium diphtheriae* biovar Belfanti and proposal of *Corynebacterium belfantii* sp. nov. *Int J Syst Evol Microbiol* 2018;68(12):3826 – 31. <https://doi.org/10.1099/ijsem.0.003069>.
- Peng YH, Fueston H, Irfan M, Hammond J, Morales D, Ju H, et al. Complete genome sequences of four representative *Corynebacterium belfantii* strains. *Microbiol Resour Announc* 2024;13(11):e00755 – 24. <https://doi.org/10.1128/MRA.00755-24>.
- Pivot D, Fanton A, Badell-Ocando E, Benouachkou M, Astruc K, Huet F, et al. Carriage of a single strain of nontoxigenic *Corynebacterium diphtheriae* bv. Belfanti (*Corynebacterium belfantii*) in four patients with cystic fibrosis. *J Clin Microbiol* 2019;57(5):e00042 – 19. <https://doi.org/10.1128/JCM.00042-19>.
- Burkovski A. Proteomics of toxigenic corynebacteria. *Proteomes* 2024;12(1):2. <https://doi.org/10.3390/PROTEOMES12010002>.
- Ramos JN, Araújo MRB, Baio PVP, Sant'Anna LO, Veras JFC, Vieira ÉMD, et al. Molecular characterization and phylogenetic analysis of the first *Corynebacterium rouxii* strains isolated in Brazil: a recent member of *Corynebacterium diphtheriae* complex. *BMC Genom Data* 2023;24(1): 65. <https://doi.org/10.1186/S12863-023-01167-W>.
- Viana MVC, Profeta R, Cerqueira JC, Vattam AR, Barh D, Silva A, et al. Evidence of episodic positive selection in *Corynebacterium diphtheriae* complex of species and its implementations in identification of drug and vaccine targets. *PeerJ* 2022;10:e12662. <https://doi.org/10.7717/PEERJ.12662>.
- Hoefer A, Pampaka D, Herrera-León S, Peiró S, Varona S, López-Perea N, et al. Molecular and epidemiological characterization of toxigenic and nontoxigenic *Corynebacterium diphtheriae*, *Corynebacterium belfantii*, *Corynebacterium rouxii*, and *Corynebacterium ulcerans* isolates identified in Spain from 2014 to 2019. *J Clin Microbiol* 2021;59(3): e02410 – 20. <https://doi.org/10.1128/JCM.02410-20>.
- Li QW, Liang XC, Gao YD, Weng ST, Gao Y, Shao ZJ. Epidemiological and pathogenic analyses of three cases of imported *Corynebacterium diphtheriae* infection in Fujian Province. *Chin J Zoonoses* 2024;40(6):590 – 4,598. <https://doi.org/10.3969/j.issn.1002-2694.2024.00.081>.
- Li ZC, Liu MZ, Fang YM, Gong ZJ, Wang XL, Chen JD, et al. Identification and analysis of 2 *Corynebacterium diphtheriae* strains in Guangdong Province. *Chin J Prev Med* 2022;56(4):427 – 32. <https://doi.org/10.3760/cma.j.cn112150-20210517-00478>.
- Lyman LR, Peng ED, Schmitt MP. DtxR is a global iron-dependent regulatory protein with multiple roles in the control of gene expression in *Corynebacterium diphtheriae*. *J Bacteriol* 2026;208(3):e00530 – 25. <https://doi.org/10.1128/JB.00530-25>.
- Valenzuela-Cardenas MB, Goodfellow SM. Microbe snapshot: corynebacterium diphtheriae. *Clin Microbiol Newsl* 2026;54:1 – 3. <https://doi.org/10.1016/J.CLINMICNEWS.2025.11.001>.