BMC Infectious Diseases

RESEARCH

Open Access

# Genomic characterisation and traceability analysis of a *Clostridium botulinum* strain involved in a food poisoning incident

Guiman Li[1†], Huiwen Wang[1†], Yan Zhang[1], Jun Yan[1], Zhongwen Duan[1], Lei Pang[1], Xianhui Xiong[1], Qiyan Cha[1], Siwen Zhao[1], Min Hou[1*] and Yanyan Liu[1*]

## Abstract

**Objective** To identify the source of *Clostridium botulinum* (*C. botulinum*) involved in a food poisoning case in Kunming and analyze its molecular characteristics.

**Results** This study examined samples from a clustered food poisoning incident, including pickled pig liver consumed by three patients, hospital vomit from a deceased patient, and household waste mixed with vomit. Enrichment culture, microscopic examination, and real-time quantitative PCR confirmed the presence of *C. botulinum* (designated KM001) with type E botulinum toxin. The full-length genomic sequence of this strain was obtained through second and third-generation sequencing, revealing a genome size of 3,713,470 bp. KEGG annotation indicated that 1,840 single genes were assigned to 44 KEGG pathways. Whole-genome sequencing revealed genetic diversity among toxin gene clusters, with 94.76% homology to *C. botulinum* E3 strain (Alaska E43) and 93.65% homology to *C. botulinum* B strain (Eklund 17B, NRP). Evolutionary analysis, incorporating complete genomes of foodborne and environmental *C. botulinum* strains worldwide along with KM001, showed stronger phylogenetic affinity of KM001 to other type E strains.

**Conclusion** Overall, this study identified KM001 as the causative agent in a food poisoning incident, marking the first report of type E botulinum toxin poisoning in the region. Genomic analysis revealed the serotype and genetic diversity among toxin gene clusters, providing insights into its gene function, virulence, and evolutionary relationships. Understanding the genetic relationships and evolutionary pathways of different *C. botulinum* strains is crucial for predicting infection risks and implementing effective control measures.The findings contribute to the documentation of botulinum toxin incidents in Yunnan, China, but do not examine the antimicrobial resistance of *C. botulinum* or its interactions with other microorganisms; further research is needed to address these aspects.

**Keywords** *C. botulinum*, Real-time PCR, Strain isolation, Whole genome sequencing, Trace the source

[†]Guiman Li and Huiwen Wang contributed equally to this work.

*Correspondence:
Min Hou
houmin7250@163.com
Yanyan Liu
liuyuxiu07@sina.com
[1]Microbiology Testing Department, Kunming Center for Disease Control and Prevention, Kunming, Yunna 650228, China

## Introduction

*Clostridium botulinum (C. botulinum)* is an anaerobic bacterium commonly found in soil and the intestines of animals, typically existing in a spore form. Under suitable environmental conditions, it can proliferate and produce various toxins. Its pathogenicity primarily arises from the production of neurotoxins such as botulinum toxin and tetanus toxin, while the bacterium itself lacks invasive properties [1].Foodborne illness caused by botulinum toxins remains a significant public health concern, often linked to improperly processed or stored foods. Different types of *C. botulinum* produce various neurotoxins, which are the primary cause of foodborne botulism [2].*C. botulinum* is classified into seven serotypes (A-G), with types A, B, E, and F primarily associated with human botulism, particularly A and B in foodborne cases [3]. Recent studies have reported food poisoning incidents linked to botulinum toxin in food products. PCR detection of type A neurotoxin in two commercial canned food samples from Turkey underscores the critical importance of stringent quality control in food production [4].Furthermore, genomic sequencing in Northern Italy linked type B botulism to industrial asparagus cream, demonstrating the effectiveness of genomic tools in tracing foodborne pathogens [5].In a case report from Shenzhen, China, type B1 *C. botulinum* was isolated from the patient's intestinal wash [6], highlighting the necessity for rapid identification in clinical settings. Abroad, type E botulinum toxin is predominantly found in marine environments and coastal soils of high-latitude regions [7].Studies in Finland indicate a high prevalence of type E in sediments, particularly in Arctic communities consuming traditional fermented foods [8].Type E botulinum toxin remains stable and active at temperatures below 0 °C [9].In mainland China, type E botulism cases are primarily linked to undercooked pork, beef, and fermented soy products [10].Therefore, a deeper understanding of *C. botulinum* and its toxins is crucial for assessing their impact on food safety and related foodborne illness outbreaks.

This study aimed to investigate the sources and characteristics of *C. botulinum* implicated in foodborne illness outbreaks. Our team collected samples of leftover pickled pig liver, household waste containing vomit, and hospital vomitus. We isolated and identified pathogenic strains and their toxins as type E using enrichment culture media for nucleic acid extraction. Type E botulism was rare in low-latitude inland areas, prompting comprehensive whole-genome sequencing and sequence analysis to elucidate the genetic features and evolutionary relationships of the strain. These methods determined the cause of the outbreak, revealing genomic characteristics, functional annotations, and virulence factors of the pathogenic strains, while comparative genomic analysis explored relationships among the strains. This research provided valuable insights into the characteristics of *C. botulinum* and its role in foodborne illness events.

## Materials and methods

### Brief case description and sample collection

The three samples analyzed in this study were collected from a single food poisoning incident involving three patients who consumed home-prepared pickled pig liver stored in a refrigerator for approximately one week. Within 24 h of consumption, the patients exhibited symptoms including dizziness, weakness, numbness of the tongue, slurred speech, blurred vision, double vision, and vomiting. Upon hospitalization, they developed clinical signs of pharyngeal and respiratory muscle paralysis. Despite the use of ventilators, one patient succumbed to the illness due to ineffective treatment.

The samples collected included one portion of the remaining pickled pig liver, one sample of household waste from the deceased patient (mixed with vomit), and one sample of vomit from the patient who passed away in the hospital.

### Ethics approval and consent to participate

This study was conducted in compliance with the Declaration of Helsinki (http://www.wma.net/en/30publications/10policies/b3/index.html) and relevant national ethical guidelines. Informed consent was obtained from the patients or their legal representatives prior to sample collection. The study protocol was approved by the Institutional Review Board (IRB) of the Ethics Committee of the Kunming Center for Disease Control and Prevention (Approval No. NO2020012). All personal identifiers were anonymized to protect patient privacy.

### Isolation and identification of pathogenic strains

The three samples were enriched using cooked meat medium as follows, 25 g of each sample were placed in sterile homogenization bags, and 225 mL of physiological saline was added to each bag. The samples were homogenized for 2 min. A 2 mL aliquot of the homogenized supernatant was inoculated into 10 mL of cooked meat medium and incubated anaerobically at 35 °C for 5 days. The phenotypic characteristics of the isolated strains were observed microscopically. To further characterize the strains and their toxin types, nucleic acids were extracted from the enriched supernatant, followed by qualitative detection using a *C. botulinum* nucleic acid detection kit (fluorescent PCR method) (#SKY-8145, Shenzhen Bioscience Co., Ltd., Shenzhen, China). The kit used the following primers: upstream primer F: GAATAYACWATAATWAATTGYAT, and downstream primers R1: TATCCTGYAAAGTCCAGATTATY, R2: TCARTTAARGTCCATATTAT, and R3:

TCTTGYAAWGTCCAAATTAT. To determine the toxin types of the strains, qPCR was performed using detection kits for botulinum toxin genes BonT/A, BonT/B, BonT/E, and BonT/F. The *C. botulinum* type A nucleic acid detection kit (fluorescent PCR method) (#SKY-8146, Shenzhen Bioscience Co., Ltd., Shenzhen, China) used primers F: GTGATACAACCAGATGGTAGTTAT and R: AAAAAA CAAGTCCCAATTATTAACTTT. The type B kit (#SKY-8147) used primers F: GAGATGTTTGTGAATATTATG ATCCAG and R: GTTCATGCATTAATATCAAGGCTG G. The type E kit (#SKY-8148) used primers F: TATCCA AAATGATGCTTATATACCAAA and R: GGCACTTTC TGTGCATCTAAATA. The type F kit (#SKY-8149) used primers F: GCTTCATTAAAGAACGGAAGCAGTGC T and R: GTGGCGCCTTTGTACCTTTTCTAGG. The qPCR procedure included 10 min of pre-denaturation at 95 °C, followed by 5 s of denaturation at 95 °C, 60 s of annealing at 60 °C, and 8 s of extension at 72 °C, repeated for 45 cycles.

#### Whole-genome sequencing and sequence analysis

The whole genome DNA of the *C. botulinum* strain was sequenced by using the MinION and Illumina MiSeq platforms. The MinION sequence reads were assembled and corrected by Illumina MiSeq with unicycler. The prediction of ORFs and their annotations were performed using Prokka v1.14.6. The BLASTn was served for further sequence comparisons. The Illumina short reads and MinION long reads were mapped to the sequence of KM001 by Burrows-Wheeler Aligner(BWA v0.7.17-r1188), and the corresponding depths were calculated by Mosdepth v0.3.3.

**Table 1** *C. botulinum* genomic information used to construct a phylogenetic tree

| Type | Name | Taxonomy ID |
|---|---|---|
| A3 | *C. botulinum* A2 str Kyoto | 536,232 |
| | *C. botulinum* A str Hall | 441,771 |
| | *C. botulinum* A3 str Loch Maree | 498,214 |
| B3 | *C. botulinum* B str Eklund 17B (NRP) | 935,198 |
| | *C. botulinum* Ba4 str 657 | 515,621 |
| | *C. botulinum* B1 str Okra | 498,213 |
| C1 | *C. botulinum* C str Stockholm | 929,505 |
| D1 | *C. botulinum* D str 16,868 | 1,443,117 |
| E12 | *C. botulinum* E1 str 'BoNT E Beluga' | 536,233 |
| | *C. botulinum* E10 str GA1101E1BB | 1,358,960 |
| | *C. botulinum* E3 str Alaska E43 | 508,767 |
| | *C. botulinum* E1 str E-RUSS | 1,358,950 |
| | *C. botulinum* E1 str E1 Dolman | 1,358,951 |
| | *C. botulinum* E10str FE9709EBB2 | 1,358,954 |
| | *C. botulinum* E10 str FWKR11E1 | 1,358,956 |
| | *C. botulinum* E11 str SO329E1 | 1,358,995 |
| | *C. botulinum* E11 str SWKR24E1 | 1,359,002 |
| | *C. botulinum* E3 str211 VH Dolman | 1,359,004 |
| | *C. botulinum* E3 str FE9708E1PI | 1,359,019 |
| | *C. botulinum* E3 str ME0702E1CS | 1,359,063 |
| *F3* | *C. botulinum* 202 F | 1,415,774 |
| | *C. botulinum* F str 230,613 | 758,678 |
| | *C. botulinum* F str 610 F | 1,359,131 |

#### Functional annotation of genome and phylogenetic trees

The gene sequences predicted were compared using BLAST analysis with various functional databases such as COG, KEGG, and GO to annotate gene functions. After summarizing the function of genes from three aspects: cell components, molecular functions and biological processes, the secondary classification of the top 20 GOslims with the most annotations under each classification is selected for drawing. We utilized the Virulence Searcher tool to predict virulence factors (VF) within the genome. The whole genome sequences of *C. botulinum* of different genotypes in 23 different countries or regions were selected from the GenBank module of NCBI website (Table 1). The 23 bacterial genomes were compared by HomBlocks, and the sequence was pruned by Gblock method to obtain the collinear block for the construction of evolutionary tree. Finally, execute 1000 bootstrap replicates using RAxML (v8.2.9) to infer the evolutionary maximum likelihood (ML) tree of bacteria.

#### Comparative genomic analysis

The whole genomes of four strains(as shown in Table 2) were compared with KM001 by using ORTHOFINDER v2.5.5 (https://github.com/davidemms/orthofinder). The genome comparative analysis circle diagram was by CGView Comparison Tool (CCT) [11]. Covariance analysis of genomes of similar strains was using mauve (http://darlingLab.org/mauve/).

In the analysis of common and unique genetic traits, a gene family is defined as a group of genes derived from a common ancestor. Single-copy and multi-copy gene families are identified through the clustering of homologous genes. These families are generally conserved across species and serve to elucidate interspecies relationships. Additionally, species-specific genes and families can be identified, potentially linked to unique phenotypic traits, allowing for the examination of evolutionary relationships at the genomic level. Using ORTHOFINDER (parameters: -M msa -A mafft -T raxml), we performed genetic family cluster analysis of the generated genomes and related species. The results of the gene family analysis were classified, and the number of genes in each category was tallied.

## Results

### *C. botulinum* identified as the causative agent of food poisoning

Laboratory analysis of the three samples was conducted using anaerobic culturing on meat broth at 35 °C for five

**Table 2** Genomic information for comparative genomic analysis

| Organism Name | Organism Groups | Strain | BioSample | Assembly | link |
|---|---|---|---|---|---|
| E3 str Alaska E43 GCF_000020285.1 | Bacteria; Firmicutes; Clostridia; Eubacteriales; Clostridiaceae; Clostridium; *C. botulinum*; *C. botulinum* E3 str Alaska E43 | E3 str | **SAMN02603537** | **GCA_000020285.1** | https://www.ncbi.nlm.nih.gov/bioproject/?term=E3+str+Alaska+E43+GCF_000020285.1 |
| *C. botulinum* B str Eklund 17B (NRP) | Bacteria; Firmicutes; Clostridia; Eubacteriales; Clostridiaceae; Clostridium; *C. botulinum*; *C. botulinum* B str Eklund 17B (NRP) | **B str** | **SAMEA2272780** | **GCA_000307125.1** | https://www.ncbi.nlm.nih.gov/bioproject/60761 |
| *C. botulinum* B str Eklund 17B (NRP) [Taxonomy ID: 935198] | Bacteria; Firmicutes; Clostridia; Eubacteriales; Clostridiaceae; Clostridium; *C. botulinum*; *C. botulinum* B str Eklund 17B (NRP) | B str | **SAMN02603538** | **GCA_000020165.1** | https://www.ncbi.nlm.nih.gov/bioproject/28857 |
| **C. botulinum 202 F** | Bacteria; Terrabacteria group; Firmicutes | 202 F | SAMN03222822 | GCA_000789355.1 | https://www.ncbi.nlm.nih.gov/genome/726?genome_assembly_id=214045 |

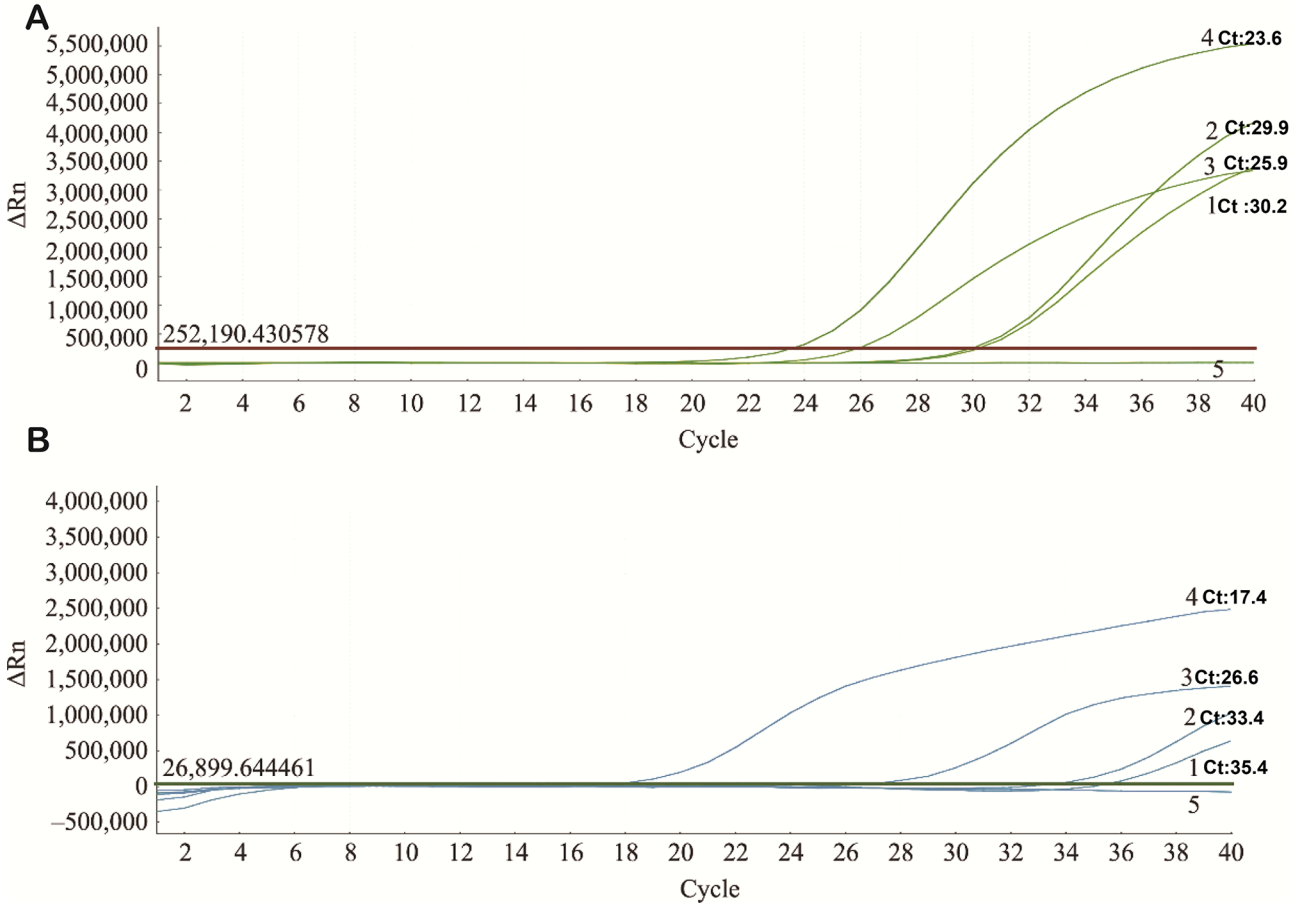


**Fig. 1** Detection of *Clostridium botulinum* gene by quantitative Real-Time PCR (qPCR). **A**. qPCRfor the detection of C. botulinum. **B**. Detection of the neurotoxin gene encoding Clostridium botulinum type E (BoNT/E). Lane 1: Remaining marinated pig liver; Lane 2: Household garbage from the patient (mixed with vomit); Lane 3: Vomit from the patient; Lane 4: Positive control; Lane 5: Negative control

days. All 3 samples were culture positive. All samples demonstrated black coloration, gas production, and a distinct foul odor. Microscopic examination revealed the presence of Gram-positive bacilli with oval spores, larger than the vegetative cells and located at the terminal end, displaying a characteristic "tennis racket" shape consistent with *C. botulinum*. Real-time PCR results confirmed the presence of *C. botulinum* and type E botulinum toxin genes in all three samples. The genes for type A, B, and F botulinum toxins were negative, with Ct values greater than 37.5. The amplification curves for the *C. botulinum* nucleic acid detection samples were shown in Fig. 1A, and the amplification curves for the BonT/E toxin gene detection were shown in Fig. 1B.

According to guidelines for food poisoning diagnosis and management, combined with epidemiological investigations, clinical presentations, and laboratory findings, this incident was classified as bacterial food poisoning due to the consumption of home-prepared pickled pig liver contaminated with type E botulinum toxin.

*C. botulinum* is a Gram-positive anaerobic bacterium that is relatively rare and difficult to survive in normal environments. Its pathogenicity is mainly due to botulinum toxin, a potent neurotoxin that inhibits the release of acetylcholine at the neuromuscular junction, causing paralysis of major muscle groups, including ocular and respiratory muscles [12]. Botulinum toxin is estimated to be 10,000 times more toxic than potassium cyanide, with a lethal dose for humans around 0.1 micrograms [13]. These results confirm that the incident was caused by the same strain of *C. botulinum* producing type E toxin, isolated from these samples. The strain from sample 3, named KM001, underwent next-generation sequencing for whole genome analysis, while no further sequencing was performed on the other two samples.

## Genomic characteristics

Whole genome sequencing of the isolated strain revealed a genome size of 3,713,470 bp, with an average sequencing depth of 428.6x for the second-generation sequencing and 87x for the third-generation sequencing, achieving a genome coverage of 99.17%. The analysis predicted 7,239 genes, including 3,509 coding sequences (CDs), 79 tRNA genes, and 31 rRNA genes. The detailed genome characteristics are illustrated in Fig. 2, which presents the GC skew, GC content, and annotations for rRNA and tRNA.

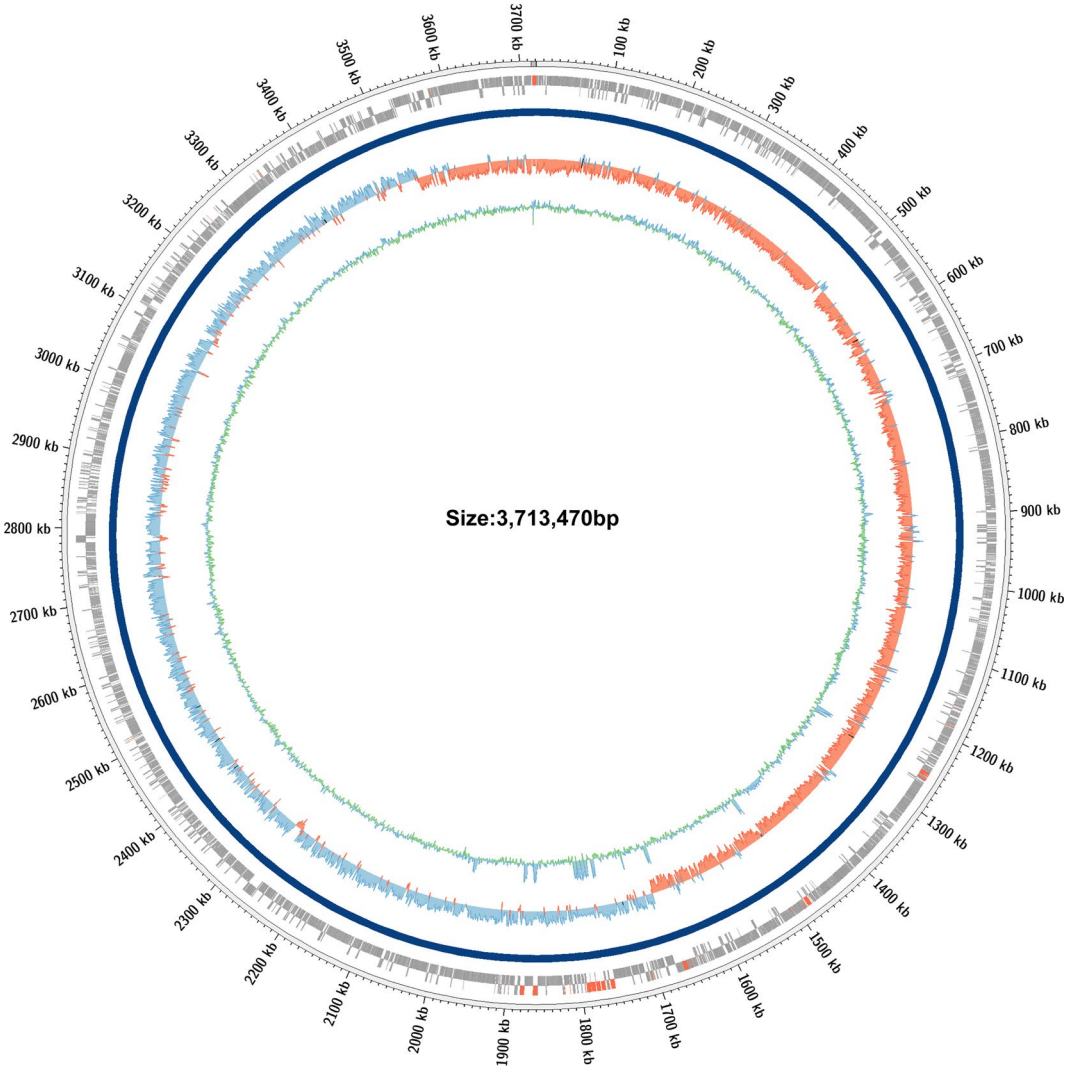


**Fig. 2** Circos genome circle map of KM001, the bacterial chromosome is 3.71 Mb in size. The distribution of the circle from the outside indicates the genome size, forward CDS, reverse CDS, GC ratio (blue and orange indicate regions where the GC ratio is higher than average and lower than average, respectively), and CG skew positive (blue) and negative (green)

### Whole genome functional annotation

To explore the complex biological functions associated with these genes, KEGG annotation was conducted, resulting in the classification of 1,840 genes into 44 metabolic pathways as shown in Fig. 3A. The most frequently annotated pathways included signaling and cellular processes, metabolism, and genetic information processing. Furthermore, the GO annotation analysis, which categorized genes into cell composition, molecular function, and biological processes, indicated that the primary activities involve oxidation-reduction processes and metabolic functions, as depicted in Fig. 3B.

The prediction results of genomic islands (GIs) reveal that there are 131 segments in the strain's genome predicted as putative alien genes (PUTAL), with the maximum length reaching 3734 bp. Most of these segments are short fragments below 3000 bp. In the COG analysis, the majority of contigs were enriched in the S category (Signal Transduction) (Fig. 3.C). In the analysis of the virulence factors of *C. botulinum*, a total of several virulence-related elements were identified (Fig. 3.D). Among these, the *bont* was found to be the most prominent, accounting for 27% of all predicted virulence factors. Specifically, when comparing the *bont* within different strains, it constituted 8% of the virulence factors detected

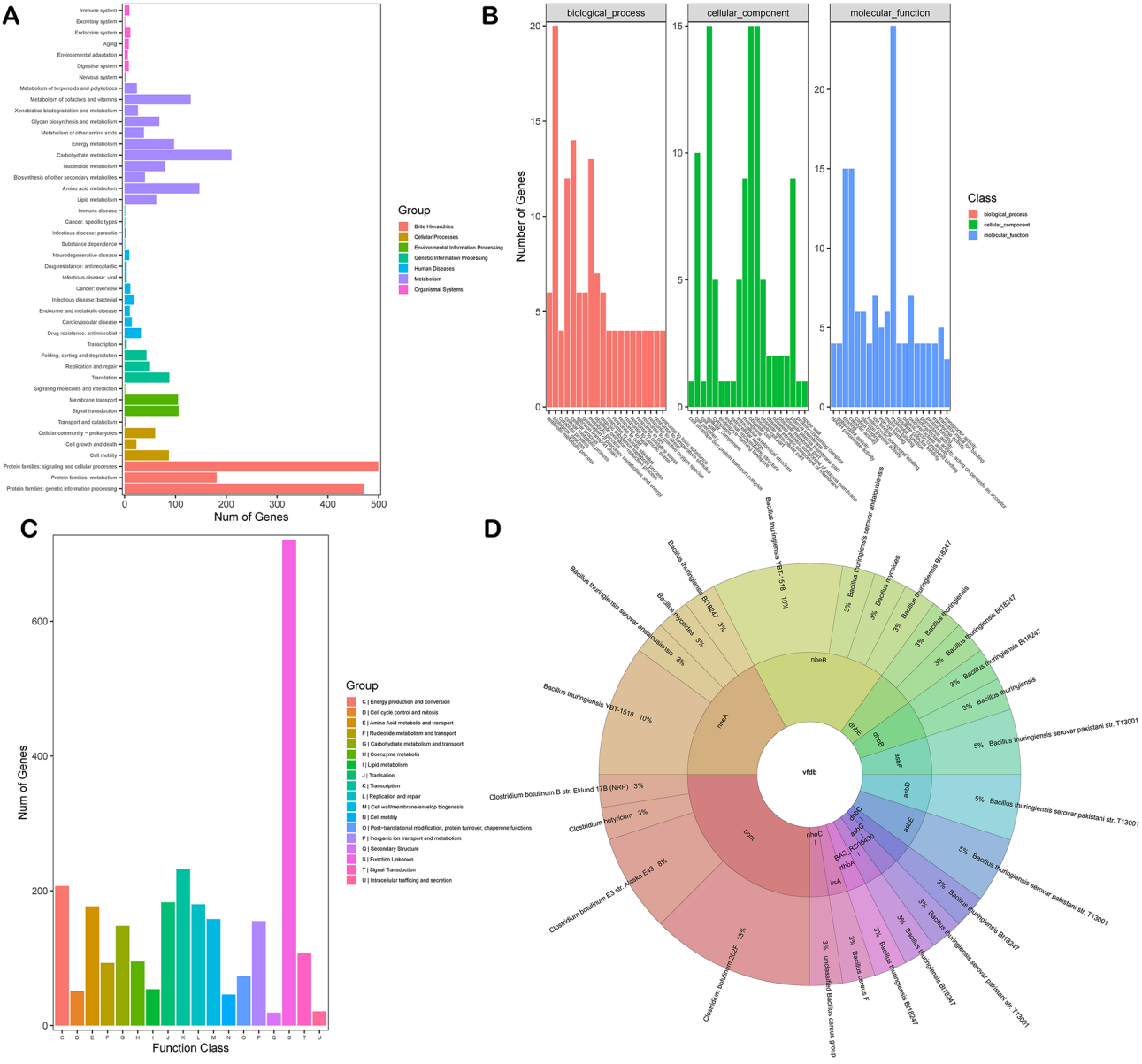


**Fig. 3** Genome analysis results of KM001. **A**. Functional annotation of KEGG databases of KM001. **B**. Functional annotation of GO databases of KM001. **C**. Functional annotation of COG databases of KM001. **D**. Distribution of Predicted Virulence Factors in the Complete Genome of KM001

in *C. botulinum* E3 strain Alaska E43, with a coverage of 98.67% and a consistency of 92.83%. In contrast, in *C. botulinum* 202 F, the *bont* represented 13% of the virulence factors, showing a coverage of 98.64% and a consistency of 92.37%. Additionally, the *bont* aligned with other strains, including *C. botulinum* B strain Eklund 17B (NRP) and Clostridium butyricum. Furthermore, other virulence factors such as *nheA*, *nheB*, *asbD*, *asbE*, *asbF* and *dhbB* were also identified, all corresponding to various strains of Bacillus thuringiensis.

### Phylogenetic analysis of *C. botulinum*

The amino acid difference between serotypes A to G ranges from 37 to 70% [14]. We obtained 23 complete genome sequences of *C. botulinum* from the NCBI database, and together with KM001, constructed a phylogenetic tree based on whole-genome sequences. The KM001 strain detected in this study showed similarity to published E-type strains available on GenBank. As shown in Fig. 4, this strain is most closely related to two isolates from Canada, namely E1 str E1 Dolman (2022) and E1 str BoNT E Beluga (2009). The strains isolated from food samples in Canada (salmon roe, seal meat in oil, and pickled herring) all belong to E-type. However, *C. botulinum* samples isolated from food in Japan, the United States, and Canada exhibited varying serotypes: A, B, and E, respectively. These serotypes are the most common ones affecting humans, with types A and B frequently found in soil samples from North America and other continents, while type E is predominantly associated with marine or estuarine sediments and the intestines of fish and marine mammals. In Alaska, Canada's Arctic regions, and northern Russia, Inuit populations traditionally store 'fermented' seal or whale blubber outdoors. Inadequate storage conditions can facilitate the germination and growth of spores in partially anaerobic environments [15]. The distribution of botulinum types across different countries may be influenced by local food characteristics and processing practices.

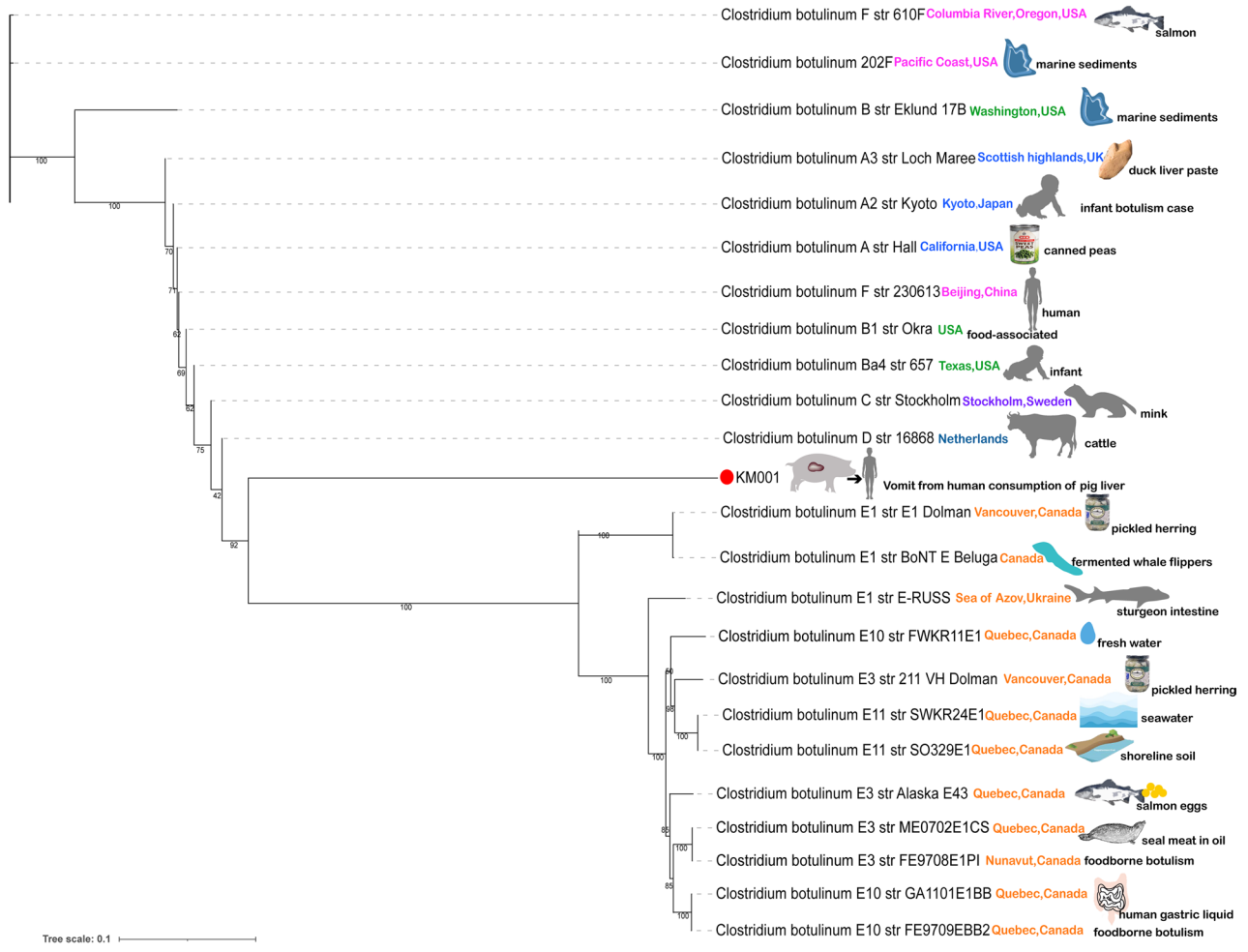


**Fig. 4** Dendrogram constructed by average nucleotide identification of published whole-genome assembly sequences and neighbor-joining tree methods

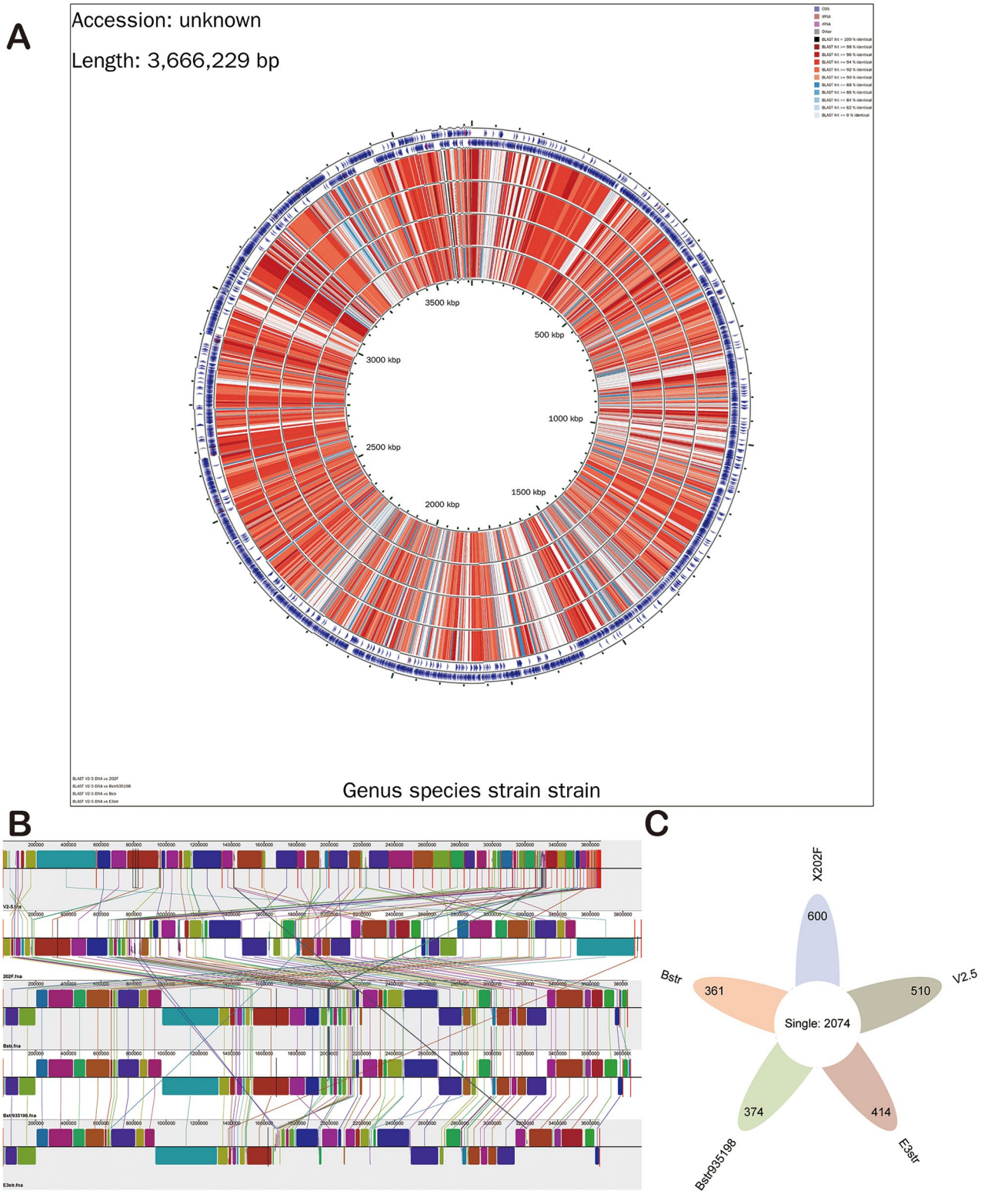


**Fig. 5** (See legend on next page.)

(See figure on previous page.)

**Fig. 5** Comparative Genomic Analysis of *Clostridium botulinum* Strains. **A**. CCT map comparing the *C. botulinum* strain to other four type of *C. botulinum* genomes. This map starting from the outermost ring the feature rings depict: (1) COG functional categories for forward strand coding sequences; (2) Forward strand sequence features; (3) Reverse strand sequence features; (4) COG functional categories for reverse strand coding sequences. The next four rings show regions of sequence similarity detected by BLAST comparisons conducted between CDS translations from the reference genome and five *C. botulinum* comparison genomes. The 5 circles from inside to outside represent the genomes of 5 *C. botulinum* strains. The color of each of the 5 circles corresponds to the homology score of BLAST HIT. Red indicates the sites with homology score greater than 90%, and some blue and white sites have low homology scores even within the same species. Colored triangular arrows refer to the forward and reverse chains of coding regions, and the functional classification of cogs in the same region has cross and unique functions. **B**. A Mauve alignment of five *C. botulinum* strains. It shows collinear set of matched colored regions, Each sequence of identically colored blocks represents a collinear set of matching regions. One connecting line is drawn per collinear block. Most regions of *C. botulinum* species have common conserved segments. **C**. Petals of single copy genes and species-specific genes in each sample

A total of five genomes were compared to analyze differences among related bacterial genomes (Fig. 5.A). In comparative genomic analyses (Table 3), each strain contains a substantial number of genes, ranging from 3,160 to 3,523. Notably, the 202 F strain has the highest gene count (3,523), while the E3str strain exhibits the lowest (3,160). The number of single genes varies across the strains, with the highest being in the v2.5 strain (469), and unique genes are most prominent in the 202 F strain (344). Interestingly, Bstr935198 and E3str strains lack unique genes entirely. The percentage of unclustered genes also varies, with the highest number found in the v2.5 strain (510), indicating a diverse genomic landscape. The gene families are relatively consistent across strains(Fig. 5.B). The number of genes within families ranges from 2,997 in the E3str strain to 3,406 in the Bstr strain, highlighting differences in gene family composition. The average number of genes per family is relatively similar across the strains(Fig. 5.C), ranging from 1.04 to 1.1, suggesting a stable family structure despite variations in total gene count.

## Discussion

In this study, we enriched and isolated *C. botulinum* from a foodborne illness outbreak, followed by real-time quantitative PCR detection and toxin typing for strain identification and whole-genome sequencing. A phylogenetic tree was constructed using 23 *C. botulinum* strains from the NCBI database. The results indicated that three individuals developed neurological symptoms, including dizziness, weakness, tongue numbness, slurred speech, blurred vision, and vomiting, after consuming pickled pig liver stored in a refrigerator for a week. The presence of *C. botulinum* and type E botulinum toxin in the pickled liver was confirmed, with the potent neurotoxin inhibiting acetylcholine release at the neuromuscular junction and causing paralysis of major muscle groups, including ocular and respiratory muscles [16].Type E botulism cases in mainland China are primarily reported from the Tibetan Plateau, where raw meat from autumn slaughtering of cattle, sheep, pigs, and marmots is susceptible to spore contamination and toxin production during storage. The region's typical food preparation methods and lower boiling points at high altitudes further hinder toxin inactivation, along with the local practice of consuming raw meat, which increases the risk of botulism [10]. Therefore, understanding the contamination pathways of *C. botulinum* in food and preventive measures is crucial. Pickled pig liver stored in a refrigerator for a week can serve as an optimal growth medium for *C. botulinum* under suitable conditions, resulting in toxin production. Strict adherence to hygiene standards during meat slaughtering and handling is essential. In early environmental studies, samples from Qinhuangdao, Shanghai, and Hainan Island showed that 3 out of every 5 tested positive for type E botulinum neurotoxin. Additionally, among 302 soil samples from the Tibetan Plateau (Qinghai and Tibet), 28 were positive, while 5 out of 249 soil samples from the Yunnan-Guizhou Plateau (Guizhou, Guangxi, and Yunnan) also detected *C. botulinum* [17]. This indicates that type E *C. botulinum* is widespread in soil environments without clear regional characteristics, highlighting the importance of timely diagnosis and effective treatment for patients exhibiting typical symptoms. However, clinical diagnostic clues are often insufficient, as conditions like Guillain-Barré syndrome and myasthenia gravis may present similarly to botulism [18].Special tests are necessary to rule these out. Definitive diagnosis can be made by detecting botulinum toxin in food, gastric or intestinal contents, vomit, or feces. Most hospitals lack comprehensive commercial diagnostic kits such as PCR or ELISA and do not routinely use specialized culture media. In cases of severe symptoms or outbreaks, metagenomic next-generation sequencing (mNGS) can be employed for pathogen detection without predefined infection targets [19], thereby enhancing the evidence for diagnosis.

Subsequently, we analyzed the functional genes of *C. botulinum*. Through whole-genome sequencing and functional annotation, the primary genes of this strain were found to be involved in metabolism, signal transduction, and genetic information processing. The functions of these genes elucidate the bacterium's cellular behavior and survival strategies. KEGG annotation revealed that 1,840 genes were categorized into 44 metabolic pathways, with the highest enrichment observed in signaling

**Table 3** Genomic characterization of *Clostridium botulinum* strains: gene distribution and family analysis

| Strains | Single | Multi | Unique | Other | Unclustered | Genes Number | Genes Number In Families | Unclustered Genes | Family Number | Unique Families Number | Average Genes Number Per Family |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 202 F | 2,074 | 407 | 344 | 442 | 256 | 3,523 | 3,267 | 256 | 3,015 | 72 | 1.08 |
| Bstr | 2,074 | 362 | 46 | 614 | 315 | 3,411 | 3,406 | 5 | 3,286 | 0 | 1.04 |
| Bstr935198 | 2,074 | 366 | 0 | 619 | 374 | 3,433 | 3,428 | 5 | 3,302 | 0 | 1.04 |
| E3str | 2,074 | 418 | 0 | 254 | 414 | 3,160 | 2,997 | 163 | 2,776 | 20 | 1.08 |
| v2.5 | 2,074 | 469 | 0 | 283 | 510 | 3,336 | 3,045 | 291 | 2,764 | 50 | 1.1 |

Single: a family in which all species have only one gene

Mutil: None of the species has zero, but there is at least one family with more than one species

Unique: Except for its own species, the remaining species are zero;

Other: It is more than zero, and the remaining species are not 0, but at least one is 0 family;

Unclustered: Nothing to clustering with other species, it is not clustered

Genes Number: The total number of genes corresponding to each species;

Genes number in families: Total number of clusters;

Unclustered genes number: Total number of unclustered genes

and cellular processes as well as metabolism. This indicates that *C. botulinum* possesses a complex metabolic capacity, which is linked to its adaptability to various environments and pathogenic potential, enabling it to utilize a broader range of nutritional sources and enhancing its survival and toxin production in diverse ecological niches [19–21]. However, this may also be related to the growth phase of the laboratory-cultured strains. A proteomic study of C. botulinum indicated that the most significantly expressed proteins between the exponential and stationary phases are enriched in functions related to metabolic processes. Proteins predominant in the exponential phase include ribosomal proteins and those involved in acetate fermentation and other growth-related functions such as transcription and translation. In contrast, the stationary phase is characterized by butyrate fermentation, amino acid metabolism, and the expression of neurotoxins in pathogenic strains [23]. It is likely that our cultured strains are primarily in the exponential phase. GO annotation analysis further suggests that the main cellular activities involve redox processes and metabolic functions. In the analysis of C. botulinum virulence factors, bont accounts for 27% of the predicted virulence factors, with 3% of the bont sequences aligning with Clostridium butyricum, which is also capable of producing BoNT/E [14].Plasmids containing bont can be transferred through conjugation among different types of C. botulinum, even to non-toxic Clostridium strains [24]. Virulence genes reflect genomic similarities and potential evolutionary relationships among different serotypes and related species. Comparative genomics reveal variations in genome size and gene composition among closely related strains, which may relate to their adaptability to different environments or hosts. The number of genes within gene families remains relatively stable (between 1.04 and 1.1). Although genomic sizes vary among different strains, the fundamental structure of gene families is conserved [25], possibly reflecting the core biological requirements of *C. botulinum*.

This study has certain limitations. Firstly, only samples from fatal cases associated with the foodborne illness outbreak were isolated, which does not reflect the overall strains responsible for the poisoning. Secondly, the whole-genome analysis lacked functional validation of virulence factors and did not address evolutionary differences among *C. botulinum* strains from different regions. Additionally, the mechanisms by which virulence factors contribute to the pathogenicity of *C. botulinum* remain unclear.

## Conclusion

This study analyzed samples from a foodborne illness outbreak, identifying *C. botulinum* as the causative agent. The strain was isolated, identified, subjected to whole-genome sequencing, and functionally analyzed. The findings provide significant insights into the biological characteristics, pathogenicity, and evolutionary relationships of *C. botulinum*. These results hold important implications for future food safety regulations, disease diagnosis and prevention, and microbiological research. However, this study also has certain limitations that require further exploration in subsequent research.

**Acknowledgements**

The authors thank all the staff contributed to the study.

**Author contributions**

Project design: LYY, LGM, WHW and ZY. Methodology: LYY, HM and LGM. Specimens and epidemiologic data collection: WHW, YJ and DZW. Project administration: LYY. Sequencing data generation: LGM, PL and XXH. Data analyses: CQY, ZSM and HM. Manuscript writing—original draft preparation: LYY and HM. Writing—review and editing: LYY, LGM, WHW, ZY, YJ, XXH CQY HM and ZSW. All authors provided final approval of the version submitted for publication.

**Funding**
There is no funding in this article.

**Data availability**
The sequences data of C. botulinum genome in this study have been deposited in GenBank under project numbers PRJNA1079894 and PRJNA1079899. The other data are available from the corresponding author on reasonable request.

## Declarations

**Ethics approval and consent to participate**
The study has been approved by the Ethics Committee of the Kunming Center for Disease Control and Prevention. The protocol number is: NO2020012. The collection and processing of all relevant test materials was reasonable and informed consent was obtained from all subjects and their legal guardians.

**Consent for publication**
Not applicable.

**Competing interests**
The authors declare no competing interests.

Received: 17 October 2023 / Accepted: 20 February 2025
Published online: 06 March 2025

**References**

1. Arnon SS, Schechter R, Inglesby TV, Henderson DA, Bartlett JG, Ascher MS, Eitzen E, Fine AD, Hauer J, Layton M, et al. Botulinum toxin as a biological weapon: medical and public health management. JAMA. 2001;285(8):1059–70.
2. Horowitz BZ. Botulinum toxin. Crit Care Clin. 2005;21(4):825–39. viii.
3. Brunt J, van Vliet AHM, Stringer SC, Carter AT, Lindström M, Peck MW. Pan-Genomic analysis of Clostridium botulinum group II (Non-Proteolytic *C. botulinum*) associated with foodborne botulism and isolated from the environment. Toxins 2020, 12(5).
4. Yılmaz M, Yiğin A, Kılıç Altun S, Aydemir M. Assessing the potential public health risk of clostridium botulinum toxin genes in canned food: A laboratory experience in TÜrklye. Middle East J Sci 2024;10.
5. Arnaboldi S, Magagna G, Monastero P, Todeschi S, Bornati L, Marina Nadia L, Finazzi G, From genome inspection to clinical significance.: source attribution of a foodborne botulism outbreak. 2023.
6. Lin Y, Jiang Y, Gong Z, Wang Y, Jiang M, Chen Q, Li C, Hu Q, Shi X. Investigation and identification of food poisoning caused by Clostridium botulinum type B1 in Shenzhen, China. Foodborne Pathog Dis. 2022;19(3):226–31.
7. Raphael BH, Lautenschlager M, Kalb SR, de Jong LI, Frace M, Lúquez C, Barr JR, Fernández RA, Maslanka SE. Analysis of a unique Clostridium botulinum strain from the Southern hemisphere producing a novel type E botulinum neurotoxin subtype. BMC Microbiol. 2012;12:245.
8. Hielm S, Hyytiä E, Andersin AB, Korkeala H. A high prevalence of Clostridium botulinum type E in Finnish freshwater and Baltic sea sediment samples. J Appl Microbiol. 1998;84(1):133–7.
9. Horowitz BZ. Type E botulism. Clin Toxicol (Phila). 2010;48(9):880–95.
10. Fu SW, Wang CH. An overview of type E botulism in China. Biomed Environ Sci: BES. 2008;21(4):353–6.
11. Grant JR, Arantes AS, Stothard P. Comparing thousands of circular genomes using the CGView Comparison Tool. BMC genomics. 2012;13:202.
12. Nagalli. ATS: Clostridium botulinum Infection(Archived). StatPearls [Internet]. Updated 2024 Jul 2.
13. Bai L, Peng X, Liu Y, Sun Y, Wang X, Wang X, Lin G, Zhang P, Wan K, Qiu Z. Clinical analysis of 86 botulism cases caused by cosmetic injection of botulinum toxin (BoNT). Medicine. 2018;97(34):e10659.
14. Hill KK, Xie G, Foley BT, Smith TJ. Genetic diversity within the botulinum neurotoxin-producing bacteria and their neurotoxins. Toxicon: Official J Int Soc Toxinology. 2015;107(Pt A):2–8.
15. Todd ECD. Foodborne Diseases: Overview of Biological Hazards and Foodborne Diseases. In: *Encyclopedia of Food Safety.* edn. Edited by Motarjemi Y. Waltham: Academic Press 2014;221–242.
16. Nigam PK, Nigam A. Botulinum toxin. Indian J Dermatology. 2010;55(1):8–14.
17. Qingyi Gao HL, Yao J, Wu J, Chen X, Dai C, Liu X, et al. Investigation of Clostridium perfringens contamination in soil and seafood in coastal areas of China. Health Res. 1983;18(03):28–32.
18. Chaudhry R. Botulism: a diagnostic challenge. Indian J Med Res. 2011;134(1):10–2.
19. Han D, Li Z, Li R, Tan P, Zhang R, Li J. mNGS in clinical microbiology laboratories: on the road to maturity. Crit Rev Microbiol. 2019;45(5–6):668–85.
20. Granato ET, Meiller-Legrand TA, Foster KR. The evolution and ecology of bacterial warfare. Curr Biol. 2019;29(11):R521–37.
21. Passalacqua KD, Charbonneau ME, O'Riordan MXD. Bacterial metabolism shapes the Host-Pathogen interface. Microbiol Spectr. 2016;4(3).
22. Shi A, Fan F, Broach JR. Microbial adaptive evolution. J Ind Microbiol Biotechnol. 2022;49(2).
23. Deatherage Kaiser BL, Hill KK, Smith TJ, Williamson CHD, Keim P, Sahl JW, Wahl KL. Proteomic analysis of four Clostridium botulinum strains identifies proteins that link biological responses to proteomic signatures. PLoS ONE. 2018;13(10):e0205586.
24. Smith TJ, Hill KK, Raphael BH. Historical and current perspectives on Clostridium botulinum diversity. Res Microbiol. 2015;166(4):290–302.
25. Ma X, Li K, Li F, Su J, Meng W, Sun Y, Sun H, Sun J, Yuan Y, Lin Y, et al. Tracing foodborne botulism events caused by Clostridium botulinum in Xinjiang Province, China, using a core genome sequence typing scheme. Microbiol Spectr. 2022;10(6):e0116422.

## Publisher's note