Molecular identification, Prevalence, and Phylogeny of *Burkholderia cepacia* Complex (Bcc) Species in the Respiratory Tract of Hospitalized Patients

Vance Franso Esho\(^1\), Hemin Esmael Othman*\(^1\)

\(^1\)Department Biology, College of Science, University of Duhok, Duhok, Kurdistan Region, Iraq.

Received 15 October 2023; revised 24 February 2024; accepted 24 February 2023; available online 14 March 2024

DOI: 10.24271/PSR.2024.420817.1408

**ABSTRACT**

*Burkholderia cepacia* complex (Bcc) is a group of at least 25 phenotypically similar but genetically distinct and closely related bacterial species, causing severe pulmonary inflammation with high mortality rates among hospitalized patients. There was a gap in knowledge regarding the prevalence and accurate species status of Bcc strains in the Kurdistan region of Iraq. Therefore, the current study aimed to directly molecular identification and characterization of Bcc species in respiratory tract infections of hospitalized patients in Duhok province based on PCR amplification and sequencing of the Bcc recA gene. The investigation also targeted the detection of the epidemic marker BCESM (esmR gene) and the transmissibility marker cblA gene. Results revealed that seven Bcc isolates were identified out of 309 sputum and oropharyngeal swabs. The overall prevalence rate of Bcc was 2.26%, with a higher percentage of isolation in sputum specimens (8.51%, 4/47) compared with oropharyngeal samples (1.14%, 3/262). The results of direct antibiotic susceptibility testing revealed that Imipenem was the most effective antibiotic against Bcc isolates (100% sensitive); contrastingly, Cefixime showed no effect on the isolates. Based on the recA gene sequencing, six Bcc species have been identified, including *B. cepacia*, *B. multivorans*, *B. cenocepacia*, *B. anthina*, *B. contaminans*, and *B. lata*, whose sequence identities ranged from 99.71–100%. Each of *B. anthina* (OR662134), *B. lata* (OR662135), *B. cenocepacia* (OR662137), and *B. contaminans* (OR662140) were recorded as the first identified and reported Bcc species in Iraq. The transmissibility marker was found in three isolates of the identified Bcc complex species (42.85%, 3/7). Also, *B. cepacia* was found to be a transmissible epidemic strain (cblA+/esmR+). Identifying transmissible and epidemic strains of the Bcc complex in this region puts hospitalized patients at high risk of cross-infection, which may require active surveillance and further investigation.

**Keywords:** *Burkholderia Cepacia Complex*, Bcc, BCESM, cblA, PCR, RecA Gene, Phylogeny.

1. **Introduction**

The Gram-negative opportunistic human pathogen, *Burkholderia cepacia* complex (Bcc), is a group of at least 25 phenotypically very similar but genetically distinct and closely related species that inhabit diverse environmental niches and cause severe and life-threatening infections, including pulmonary inflammation in both cystic fibrosis and non-cystic fibrosis patients with high mortality rates among immunocompromised patients and hospitalized neonates\(^1\).

This group of opportunistic pathogens has been increasingly isolated as human pathogens due to their capability to cause serious infections and spread between patients. They have a potential role in declining lung function with necrotizing pneumonia and frequently fatal septicemia known as *B. cepacia* syndrome \(^2\). The clinical manifestations of cepacia syndrome include fever, pulmonary infiltration, respiratory distress, bronchopneumonia, leukocytosis, and confluent septicemia\(^3\).

Although relatively low prevalence rates of infections (<5 %) by Bcc strains have been reported in cystic fibrosis and non-cystic fibrosis populations worldwide, they are of significant, with great concern, and highly problematic clinical consequences. This is mainly due to their inherent resistance to most antimicrobial drugs and their capability to form biofilm structures, which lead to persistent infections and make them hard to eradicate \(^4\). Furthermore, person-to-person transmissibility, cross-infection between patients, device-related infections, and epidemic outbreaks have been documented within and outside hospitals\(^5\,\,6\).

Various markers have been reported to be associated with transmissible strains and the epidemic spread of Bcc strains. These markers include extracellular appendages called cable pili (encoded by the *cblA* gene), which mediates adherence to respiratory mucins. And the *B. cepacia* epidemic strain marker
(BCESM, encoded by the esmR gene), which is a conserved 1.4-kb open reading frame that encodes both virulence and metabolism-associated genes, is present in epidemic strains and is absent in sporadic strains[5,7].

Initially, members of the Bcc complex were thought to be a single bacterial species known as B. cepaci, as they share similar phenotypic, cultural, and biochemical characteristics. However, to date, 25 properly named species in B. cepaci complex have been identified based on molecular analysis, which are phenotypically similar but genotypically distinct and closely related bacterial species[8].

Accurately and precisely identifying the Bcc complex species with conventional phenotypic and biochemical methods is problematic and questionable, as persistent colonization of patients by Bcc often leads to isolates losing their phenotypic characteristics or growth-required conditions that ultimately lead to inaccurate or misidentification of the bacterial species. Furthermore, even the most commercially available identification systems cannot dependably discriminate them from each other due to their high rates of similarities. In addition, the discrimination of this complex from other related taxonomic bacteria, including Ralstonia, Pandoraea, Cupriavidus, Brevundimonas, Achromobacter, Delftia, and Comamonas species, is challenging[9].

Therefore, with the advances in molecular biology technologies and DNA sequencing strategies in recent decades, many reliable molecular identification strategies have been proposed to distinguish between species of the Bcc complex. These included molecular targeting and sequencing of 16S rRNA, rpsU, recA, hisA genes, and WGS of Bcc species. However, among these, nucleotide sequence variation analysis of Bcc-specific recA gene showed higher levels of accurate identification, discrimination power, and reliability for species-level identification and discrimination of Bcc species within this complex[9,10,11].

To date, and to our knowledge, the literature contains no reports addressing molecular identification and species status (genomovar prevalence) of Bcc in the Kurdistan region of Iraq and there was a gap in knowledge in this regard. Therefore, the current study was designed to (i) accurately identify and characterize Burkholderia cepacia complex species, targeting the direct PCR amplification and sequencing of the Bcc recA gene, (ii) evaluate the prevalence of Bcc infection in hospitalized patients in Duhok province, (iii) assess the frequency of transmissibility and epidemic markers (cblaA and BCESM markers), and phylogenetic analysis of the Bcc complex strains.

2. Materials and Methods

2.1. Sample Collection

Following the Manual of Clinical Microbiology [12,13], a total of 309 specimens (262 oropharyngeal swabs and 47 sputum samples) were collected from patients (169 males and 140 females). Oropharyngeal swabs were taken from patients suffering from respiratory infections who attended Khabat Health Center (209) in Duhok city and Mangesh Health Center (53) in Duhok province. Sputum samples were from patients in the intensive care unit (ICU) in the Emergency Teaching Hospital (10) and the Specialized Respiratory Center (37) in Duhok during the period extended from June 2022 to February 2023. Both males and females aged between (1-86) years enrolled in this study.

2.2. Processing of Specimens

Sputum samples were collected directly by expectoration into a sterilized plastic receptacle. Oropharyngeal swabs were collected by direct friction of a sterile cotton swab in the posterior pharynx with the help of a tongue depressor. The respiratory specimens (sputum or oropharyngeal swabs) were cultured into 5ml nutrient broth incubated at 37°C for up to 24 hours. From the 24 hours of fresh cultured samples, 2ml were preserved in 50% glycerol, and the rest were used for direct antimicrobial susceptibility testing (DST) and DNA extraction for downstream procedures.

2.3. Antibiotic Susceptibility Testing

In this study, the direct antimicrobial susceptibility testing (DST) method was used to predict the antibiotic properties of the identified Bcc isolates using disk diffusion on Difco™ Mueller Hinton Agar plates[14,15,16]. The susceptibility toward ten different antibiotics (Bioanalyse/Turkey) was investigated by Kirby-Bauer disc diffusion method including (Ciprofloxacin 5µg, Levofloxacin 5µg, Nitrofurantoin 300µg, Gentamicin 10µg, Azithromycin 15µg, Ceftriazone 30µg, Cefixime 5µg, Trimethoprim 5µg, Imipenem 10µg, and Norfloxacin 10µg). DST plates were incubated overnight at 37°C. Inhibition zones were measured in millimeters and interpreted as resistant (R), susceptible (S), or intermediate susceptible (I), according to the Clinical and Laboratory Standards Institute[17].

2.4. DNA Extraction, quantification, and qualification

All the 309 fresh 24-hour-old specimen cultures were subjected to DNA extraction using the High Pure PCR Template Preparation Kit (#11796828001 Roche, Germany), following the manufacturer’s instruction manual. The purity and concentration of the extracted DNA samples were measured and calculated using NanoDrop™ 2000/2000c Spectrophotometer (Thermo Fisher Scientific™/Catalog number: ND2000CLAPTOP). The extracted DNA samples were stored in the Freezer at -20°C until use.

2.5. PCR Methods for detection of recA, cbla, and esmR genes:

The direct molecular identification of the Bcc complex was performed by subjecting all the 309 extracted DNA samples to a conventional PCR assay using the recA gene of the Bcc complex in clinical samples implementing BCR1 and BCR2 primers (Table 1). The PCR detection of the transmissibility marker (cbla gene) was performed using a primer pair CBLA1 and CBLA2 (Table 1). The PCR detection of the Bcc Epidemic strain marker BCESM (esmR gene) was carried out using a primer pair BCESM1 and BCESM2 (Table 1). The PCR assays for the above reactions were carried out in a total volume of 25 µl reaction system containing 12.5 µl Taq PCR MasterMix (add bio, Korea), 1.0 µl of each forward and reverse primers, 8.5 nuclelease-free water, and 2 µl of extracted DNA template. The PCR reaction program was set as follows: for the recA gene, one cycle of 94°C
for 2 minutes, followed by 30 cycles of 45 seconds at 94°C, 45 seconds at 60°C, and 90 seconds at 72°C. The final extension step was set on one cycle of 7 minutes at 72°C. For the cblA gene, one cycle of 94°C for 5 minutes, followed by 30 cycles of 30 seconds at 94°C, 45 seconds at 58°C, and 45 seconds at 72°C. The final extension step was set on one cycle of 5 minutes at 72°C. For the esmR gene, one cycle of 94°C for 5 minutes, followed by 30 cycles of 30 seconds at 94°C, 45 seconds at 63°C, and 90 seconds at 72°C. The final extension step was set on one cycle of 7 minutes at 72°C, using the ABI Geneamp 9700 PCR -Thermal Cycler (Applied Biosystems). The PCR products were run on a 1% agarose gel in 0.5% Tris-borate-EDTA buffer at 85 voltages for 30 minutes, stained with Safe Gel Stain (add bio, Korea), visualized by UV-transluminator, and photographed[5,18].

### Table 1: Oligonucleotide primers used in this study.

<table>
<thead>
<tr>
<th>Target</th>
<th>Primers</th>
<th>Sequences (5’-3’)</th>
<th>Product Size</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>B. cepacia</em> complex recA gene</td>
<td><strong>BCR1</strong></td>
<td>TGACCCCGCAGAAGAGCACCA</td>
<td>1043bp</td>
<td>[18]</td>
</tr>
<tr>
<td></td>
<td><strong>BCR2</strong></td>
<td>CTCTTCTCTGCTCCACGCCTC</td>
<td>1400bp</td>
<td>[5]</td>
</tr>
<tr>
<td>esmR (BCESM)</td>
<td><strong>BCESM1</strong></td>
<td>CCACGGACGTCGAATAACA</td>
<td>1026bp</td>
<td></td>
</tr>
<tr>
<td></td>
<td><strong>BCESM2</strong></td>
<td>CGTCCATCCGAACAGAT</td>
<td>1026bp</td>
<td></td>
</tr>
<tr>
<td>cblA gene</td>
<td><strong>CBLA1</strong></td>
<td>GTTCGGATCGCTGCTGCT</td>
<td>643bp</td>
<td>This study</td>
</tr>
<tr>
<td></td>
<td><strong>CBLA2</strong></td>
<td>ACGCGATGTTCCATCACCAT</td>
<td>643bp</td>
<td></td>
</tr>
</tbody>
</table>

2.6. Sequencing of the recA gene

The PCR products of the recA gene were purified using a MinElute PCR Purification Kit (50) (Qiagen) following the supplier's instructions. The purified PCR products were sequenced by Macrogen Inc. (Seoul, South Korea), implementing an automated ABI3730XL Genetic analyzer, and the primer pair of BCR1 and BCR2 were used individually in separate reactions.

3. Sequence data analyses

3.1. Species identification of Bcc complex

The nucleotide sequences of both BCR1 and BCR2 were trimmed, edited, verified, and analyzed to produce consensus sequences in Geneious, version R8.1, Biomatters, and BioEdit version 7.2 software programs and saved as FASTA format. The partial DNA sequences of the recA gene (1022-1026bps) were obtained and were searched for their specificity by comparing them with their counter-available sequences at NCBI GenBank (National Center for Biotechnology Information). Sequences that showed >99% identity were retrieved by using the Basic Local Alignment Search Tool (BLAST) against recA gene sequences (www.ncbi.nlm.nih.gov/BLAST).

3.2. Phylogenetic analysis

The phylogenetic relationship among the Bcc strains was constructed based on the nucleotide sequence alignment of the recA gene using the ClustalW algorithm and default parameters. The Neighbor-Joining method supported by the bootstrap values (1000 replicates) was implemented to build the Phylogenetic tree. The evolutionary distances were computed using the Jukes-Cantor method. Evolutionary analyses were conducted in MEGA11[19].

3.3. Accession numbers

The nucleotide sequences of the recA gene, determined in the present study for all the identified Bcc species, were submitted to the NCBI GenBank to assign accession numbers.

4. Results and Discussion

This study involved a total of 309 oropharyngeal swabs and sputum specimens from hospitalized patients who were infected with respiratory tract infections, from which seven isolates were identified and confirmed to be *B. cepacia* complex by direct PCR amplification and sequencing of Bcc-specific recA gene. The NCBI GenBank accession numbers for the identified six Bcc species based on recA gene sequencing are *Burkholderia anthina* strain BCC-1 (OR662134), *Burkholderia lata* strain BCC-2 (OR662135), *Burkholderia multivorans* strain BCC-5 (OR662136), *Burkholderia cenocepacia* strain BCC-7 (OR662137), *Burkholderia cepacia* strain BCC-58 (OR662138), *Burkholderia cepacia* strain BCC-68 (OR662139), *Burkholderia contaminans* strain BCC-234 (OR662140).

The overall prevalence rate of the Bcc complex was 2.26% (17/7309), with a higher prevalence rate in sputum samples (8.51%, 4/47). The overall prevalence rate of Bcc reported in this study was in line with the overall prevalence rate of Bcc infections (2-4%) worldwide[20]. Even though the prevalence rate of infection is relatively low, they are of great clinical concern and troublesome due to their high rate of mortality, inherent resistance to antibiotics, and their ability to produce biofilm structures, which make them hard to eradicate and lead to persistent infections[21,22]. A high prevalence rate of Bcc in sputum specimens may require active surveillance as part of infection control policies due to the ability of the Bcc members to produce cross-infection between patients and device-related infections[23,24].

The results of direct antibiotic susceptibility testing of the Bcc isolates revealed that Imipenem was the most effective antibiotic (100%), followed by Levofloxacin (85.7%). Further, Ciprofloxacin, Gentamicin, Norfloxacin, and Ceftriaxone showed (71.4%) sensitivity to Bcc strains. However, much lower sensitivity rates were documented for Trimethoprim (57%), Azithromycin and Nitrofurantoin (42.8%). In contrast, Cefixime showed no effect on the isolates (100% resistant). Although direct antibiotic susceptibility testing (DST) might be controversial since the inoculant is not properly standardized, when applied selectively to clinical samples and interpreted carefully, DST is considerably beneficial for critically ill patients and their management, as it leads to shortening the time considerably by approximately 24 hours[25]. In addition, the Bcc complex species...
has become increasingly reported as highly resistant strains to the majority of antibiotics and cause severe acute respiratory syndrome with a high mortality rate, which may lead to greater use of healthcare resources and a poor prognosis [1]. Therefore, early diagnosis and their rapid predictive antimicrobial properties are crucial to administering suitable and proper antibiotics in time. This may lead to appropriate antimicrobial administration, decreasing antibiotic misuse, and eventually, fewer antibiotic-related complications.

The nucleotide sequence variations in the recA gene were analyzed to determine B. cepacia complex species and genomovar status. Results revealed that six Bcc species have been identified, of which two isolates belonged to B. cepacia (genomovar I), representing the most frequent species. Others were identified as B. multivorans (genomovar II), B. cenocepacia (genomovar III), B. anthina (genomovar VIII), B. contaminans, and B. lata. It is noteworthy to mention that each of B. anthina (accession number: OR662134), B. lata (accession number: OR662135), B. cenocepacia (accession number: OR662137), and B. contaminans (accession number: OR662140) are considered the first identifications and records of these species in Iraq; as to our knowledge, they have not been identified and not been reported previously in this region.

The recA gene nucleotide sequences for all the identified species were submitted to the NCBI GenBank, and the accession numbers were assigned to each species individually. All seven Bcc isolates were correctly categorized to their respective genomovor or species level within the Bcc complex using the nucleotide sequence variations in the recA gene. The sequence identities of the identified species ranged from 99.71 to 100% when compared to their NCBI GenBank counter-reference sequences. Many previous studies have used this technique for the same purpose and with accurate and satisfactory results [21,22]. Further, it has been reported that the recA gene exhibits 94-95% sequence similarity among different Bcc species and around 98-99% sequence similarity for the same species [23]. Furthermore, this gene, which codes for the conserved protein RecA, is found in a single copy on the largest replicon found in Bcc genomes and has significant nucleotide variations that allow for species discrimination and speciation [24]. These make the recA gene an excellent candidate molecular-based marker for Bcc species identification and discrimination.

This study was further extended to investigate the presence of the epidemic strain marker (BCESM) encoded by the esmR gene and the transmissibility marker encoded by the cblA gene within the respiratory infected patients in the Kurdistan of Iraq. Results revealed that the cblA gene was found in three Bcc complex isolates out of seven, representing (42.85%, 3/7). The Bcc complex species that exhibited this transmissibility marker was B. anthina isolated from a sputum sample of a 60-year-old patient from Sharya/Duhok and two strains of B. cepacia, one isolated from oropharyngeal swabs taken from a 33-year-old male patient from Mangesh who attended Mangesh Health Center. The other was from a 26-year-old female patient from Khabat who attended Khabat Health Center in Duhok province. Further, the epidemic marker (BCESM) was identified in one Bcc isolate out of seven, representing (14.28%, 1/7), which belonged to B. cepacia isolated from a swab sample of the 33 years old male patient from Mangesh who attended Mangesh Health Center.

The genotypic profiles of the Bcc isolates in terms of cblA and esmR genes were cblA*/esmR* (one isolate of B. cepacia) known as transmissible epidemic strain, cblA*/esmR* (one isolate of B. cepacia and an isolate of B. anthina) known as transmissible strain. The rest belonged to the cblA+/esmR- genotype, which included B. multivorans, B. cenocepacia, B. contaminans, and B. lata, which were all non-transmissible sporadic strains of these bacteria. Besides the differences in antibiotic sensitivity profiles of the isolates, the species and genotypic diversities in the Bcc complex may explain the heterogeneity in outcomes and manifestations among respiratory-infected patients with Bcc complex strains. For instance, some patients may carry Bcc complex for a long period without any visible complications, while others may suffer from a dramatic decline in pulmonary functions. This contraindication in Bcc infection, and increased mortality due to Bcc infection make the accurate identification and characterization of Bcc complex greatly important and very helpful for epidemiological research, prognosis assessment, as well as clinical therapy because the increase in risk from species to species appears to vary significantly. Additionally, the identification of transmissible epidemic strains in pulmonary-infected patients is alarming. It may raise questions regarding optimal control measuring policies, particularly since the strain has the potential to cause cross-infections both in the hospital and the community.

The phylogenetic analysis based on the nucleotide sequences of the recA gene, was constructed to illustrate the taxonomic status of the identified Bcc species. Results showed that all the identified Bcc complex species have clustered into four major clusters supported by strong bootstrap values, including C1, C2, C3, and C4 (Figure 1). The cluster C1 was further sub-grouped into B. contaminans and B. lata. Cluster C2 comprised B. cenocepacia, C3 comprised B. cepacia, and C4 sub-divided further into B. multivorans and B. anthina. Additionally, the genetic distances ranged from 0.001 to 0.065 among identified Bcc species. These small values reflect that they are genetically closely related species and are in line with the previously reported taxonomic status of Bcc complex species [24]. However, the nucleotide sequence variations within the recA gene and phylogenetic analysis efficiently and reliably discriminated the Bcc complex members and allowed them to resolve at the species level and lineage. This might further confirm the reliability of recA sequencing as a valuable tool for accurate identification and epidemiological investigations of Bcc complex species.
enBank accession numbers, OR662134, OR662135, OR662137, and OR662140, respectively. Isolates assigned NCBI GenBank accession numbers, OR662134, OR662135, OR662137, and OR662140, respectively. Isolates showed the coexistence of both transmissible and epidemic strains identified in this study. Identification of transmissible epidemic strains makes patients infected with Bcc complex at high risk due to their ability to cause cross-infections. Therefore, rapid and accurate identification of Bcc complex species is essential in terms of clinical prognosis and epidemiology, and to evaluate specific risks, which may have positive consequences in terms of antimicrobial therapy, decreasing disease severity, and infection control policies. Finally, conducting a large-scale and systematic study using whole genome sequences may provide further insights into their taxonomic status, phylogeny, and adaptation of Bcc species to diverse environmental niches.

Conclusions

The overall prevalence rate of the Bcc complex was found to be higher in sputum samples compared with that of the oropharyngeal swabs. The nucleotide sequence variation of the Bcc recA gene reliably discriminated the Bcc complex members and assigned them to their species level. Diverse species of this opportunistic group of bacteria were identified, in which four species were identified as the first record in Iraq, including B. anthina, B. lata, B. cenocepacia, and B. contaminans, with their assigned NCBI GenBank accession numbers, OR662134, OR662135, OR662137, and OR662140, respectively. Isolates showed the coexistence of both transmissible and epidemic strains identified in this study. Identification of transmissible epidemic strains makes patients infected with Bcc complex at high risk due to their ability to cause cross-infections. Therefore, rapid and accurate identification of Bcc complex species is essential in terms of clinical prognosis and epidemiology, and to evaluate specific risks, which may have positive consequences in terms of antimicrobial therapy, decreasing disease severity, and infection control policies. Finally, conducting a large-scale and systematic study using whole genome sequences may provide further insights into their taxonomic status, phylogeny, and adaptation of Bcc species to diverse environmental niches.

Conflict of interests

The authors declare no conflict of interest.

Author Contribution and Funding Information

Both authors contributed fully to this research. The first author mainly contributed to the experimental part and writing. The second author contributed to the supervision, edition, and revision. The University of Duhok partially funded this research.

Ethical Statement

This study was approved by the Research Ethics Committee (Ministry of Higher Education, University of Duhok, and Ministry of Health, Duhok Directorate General of Health). The approval letter reference number (1011201-11-1), issued on November 10, 2021, was provided. Patients permitted and allowed the collection of the various clinical specimens used in this current study.

References

https://doi.org/10.1186/s12941-019-0306-0


