Genomic Characterization of Multidrug-Resistant *Escherichia coli* BH100 Sub-strains

Rodrigo Carvalho¹, Flavia Aburjaile²³, Marcus Canario², Andréa M. A. Nascimento², Edmar Chartone-Souza², Luis de Jesus², Andrey A. Zamyatin Jr.¹⁴, Bertram Brenig⁵, Debmalya Barh⁶, Preetam Ghosh⁷, Aristoteles Goes-Neto⁸, Henrique C. P. Figueiredo⁵, Siomar Soares⁸, Rommel Ramos⁹, Anne Pinto² and Vasco Azevedo²*

¹Institute of Molecular Medicine, Sechenov First Moscow State Medical University, Moscow, Russia, ²Departamento de Genética, Ecologia e Evolução, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil, ³Departamento de Genética, Universidade Federal de Pernambuco, Recife, Brazil, ⁴Belozersky Institute of Physico-Chemical Biology, Lomonosov Moscow State University, Moscow, Russia, ⁵Institute of Veterinary Medicine, University of Göttingen, Göttingen, Germany, ⁶Institute of Integrative Omics and Applied Biotechnology, Purba Medinipur, India, ⁷Department of Computer Science, Virginia Commonwealth University, Richmond, VA, United States, ⁸Departmento de Microbiologia, Imunologia e Parasitologia, Universidade Federal do Triângulo Mineiro, Uberaba, Brazil, ⁹Universidade Federal do Pará, Belém, Brazil

The rapid emergence of multidrug-resistant (MDR) bacteria is a global health problem. Mobile genetic elements like conjugative plasmids, transposons, and integrons are the major players in spreading resistance genes in uropathogenic *Escherichia coli* (UPEC) pathotype. The *E. coli* BH100 strain was isolated from the urinary tract of a Brazilian woman in 1974. This strain presents two plasmids carrying MDR cassettes, pBH100, and pAp, with conjugative and mobilization properties, respectively. However, its transposable elements have not been characterized. In this study, we attempted to unravel the factors involved in the mobilization of virulence and drug-resistance genes by assessing genomic rearrangements in four BH100 sub-strains (BH100 MG2014, BH100 MG2017, BH100L MG2017, and BH100N MG2017). Therefore, the complete genomes of the BH100 sub-strains were achieved through Next Generation Sequencing and submitted to comparative genomic analyses. Our data shows recombination events between the two plasmids in the sub-strain BH100 MG2017 and between pBH100 and the chromosome in BH100L MG2017. In both cases, *IS3* and *IS21* elements were detected upstream of *Tn21* family transposons associated with MDR genes at the recombined region. These results integrated with Genomic island analysis suggest pBH100 might be involved in the spreading of drug resistance through the formation of resistance islands. Regarding pathogenicity, our results reveal that BH100 strain is closely related to UPEC strains and contains many *IS3* and *IS21*-transposase-enriched genomic islands associated with virulence. This study concludes that those IS elements are vital for the evolution and adaptation of BH100 strain.

**Keywords:** antibiotic resistance, genomic sequencing, mobile genetic elements, urinary tract infection, UPEC
INTRODUCTION

*Escherichia coli* are Gram-negative bacteria naturally found in the intestinal tract of several animal species and humans, where some strains can coexist with the host in a symbiotic relationship while others may cause disease (Moulin-Schouleur et al., 2007; Römer et al., 2012; Poolman and Wacker, 2016; Kittana et al., 2018). A subset of *E. coli* strains is capable of causing extra-intestinal disease, including urinary tract infection (UTI), in which the pathotype uropathogenic *E. coli* (UPEC) accounts for 90% of global cases (Marrs et al., 2005). In this context, multidrug-resistant (MDR) strains have been considered as a major bottleneck limiting the effective control of infections. Indeed, the abuse of antibiotics has contributed to the selection of strains with the ability to acquire drug-resistance genes via horizontal transfer mechanisms (Davies and Davies, 2010; Leungtongkam et al., 2018). Previous studies suggest that conjugative R plasmids and integrons are highly prevalent in UPEC isolates and are likely maintained in bacterial populations under the selective pressure of antimicrobials (Oliveira-Pinto et al., 2017).

In the last decade, Whole-genome sequencing (WGS) approaches have been considered as valuable tools to improve the understanding of *E. coli* populations and the relationship between strains and clinical significance (Klemm and Dougan, 2016; Nakano et al., 2019). *E. coli* BH100 was isolated in the city of Belo Horizonte in 1974 from urine samples from women with UTI (Chartone-Souza, 2017). This bacterium is a non-colicin producer, lactose fermenter (lac +) and harbors two plasmids, pBH100 and pAp, of approximately 60 and 10 MDa respectively. Previous studies have focused on the phenotype characterization of these strains based on antibiotic susceptibility and plasmid transference experiments. It has been shown that pBH100 presents conjugative properties and is responsible for the resistance to inorganic mercury, chloramphenicol, kanamycin, tetracycline, and potential resistance to streptomycin. Furthermore, the smaller plasmid, pAp, can be mobilized by pBH100 and seems to provide resistance to β-lactams (Nascimento et al., 1992; Nascimento and Chartone-Souza, 2003). These previous phenotyping and molecular approaches allowed the description of the individual and small clusters of resistance elements. However, a WGS approach of this strain is required to provide a comprehensive investigation of the diversity and evolutionary relationship of *E. coli* BH100. Interestingly, the long-term cultivation of BH100 in artificial conditions has promoted genetic variability, which resulted in the emergence of sub-strains (Chartone-Souza, 2017).

This single scenario represents a favorable condition to investigate the factors contributing to the genome plasticity of this strain. Although previous studies have suggested the activity of transposons in the plasmids, a large-scale molecular characterization of mobile genetic elements (MGE), involved in the mobilization of drug-resistance genes and virulence factors have not been investigated. Therefore, in this present study, we aimed to assess the full genetic diversity of BH100 sub-strains using WGS to elucidate the factors that might be involved in the mobilization of virulence and drug-resistance genes and their evolutionary relationship with UPEC strains.

MATERIALS AND METHODS

Bacterial Strains

BH100 wild type strain, harboring the pBH100 (conjugative) and pAp (mobilizable) plasmids with antibiotic resistance, was isolated from urine sample of a Brazilian woman from Belo Horizonte in 1974. After isolation, the strain was maintained in tubes containing Lignières medium at room temperature (Parker Hitchens, 1921) and cultures were transferred every 2 years to a new fresh medium. This strain was sequenced in 2014, and referred to in this study as MG2014, and was considered for the investigation of its evolutionary relationship to *E. coli* phylogroups and pathotypes. Two sub-strains were obtained in 1987 by elimination of the plasmids through overnight cultivation at 37°C in nutrient broth containing a subinhibitory concentration of ethidium bromide as previously described (Bouanchaud et al., 1968). The plasmid-cured sub-strains were hereafter referred to as BH100L and BH100N for containing only pBH100 and none of the plasmids, respectively. These sub-strains have been transferred every 6 years to a new fresh medium. All the sub- strains, including BH100 were sequenced lately in 2017 and referred here as MG2017. The sub-strains were selected for investigating the intra-strain activity of mobile elements. All the strains (Supplementary Table 1) were kindly provided by the Microorganism Genetics Laboratory (LGM) of the Federal University of Minas Gerais, Belo Horizonte, Brazil.

Growth Conditions for Genomic DNA Preparation

The MG2014 strain was grown in 30 mL of BHI medium (HIMEDIA) at 37°C for 18 h and centrifuged at 2,500 g for 15 min. MG2017 sub-strains were grown in 5.0 mL LB medium at 37°C for 18 h. Chloramphenicol (15.0 µg/mL) and 50.0 µg/mL of ampicillin (Sigma) were added in the growth media of the BH100 variants when needed. Genomic DNA was extracted following the protocol of Pacheco et al. (2007) for all bacterial cultures (Pacheco et al., 2007).

Sequencing of Complete Genomes

BH100 MG2014 was sequenced on the Ion Torrent PGM platform. The construction of a 200 bp fragment library was performed from 1.0 µg DNA using the Ion Xpress™ Plus Fragment Library Kit (Thermo Fisher) as recommended by the manufacturer. The BH100 MG2014 was also sequenced using 3 and 6 kb mate-pair libraries from 10.0 µg of DNA using 5500 SOLiD® Mate-Pair Library Kit, as per Thermo Fisher technical recommendations. All MG2017 sub-strains (BH100, BH100L, and BH100N) were sequenced on the Illumina Hi-Seq 2500 platform from 0.1 µg DNA using paired-end libraries (2 × 150 bp) and approximately 500 bp fragments, as recommended by the manufacturer.

Assembly of Genomes

The quality assessment of the reads was performed using FastQC software. All sub-strains contigs were assembled using Newbler
2.9 and the quality assessment of each assembly was evaluated in QUAST and the best K-mer values were estimated using KmerGenie 1.7023 (Chikhi and Medvedev, 2014). Scaffolding of contigs was performed on CONTIGUATOR software 2.7.3 using the complete genome from \textit{E. coli} 536 (RefSeq: NC_008253.1) as a reference and Gap closure was accomplished via CLC Genome Workbench 7.0 (Galandrini et al., 2011). The other bacterial variant reads were submitted to the same pipeline, though using the complete genome \textit{E. coli} BH100 MG2014 as a reference.

**Assembly of Plasmids**

For the five plasmids of \textit{E. coli} BH100 sub-strains (NZ_CP024651.1, NZ_CP024652.2, NZ_CP025252.1, NZ_CP025253.1, and CP025139.1) the reads were assembled via PlasmidSpades (Antipov et al., 2016) and the contigs generated were investigated using BLASTn for observation of possible homologous plasmid sequences. Scaffolding of contigs from the pBH100 plasmids was performed in the same manner described above, however using plasmid NRI (DQ364638.1) as a reference. No gaps were found in pAp contigs.

**Structural and Functional Annotation**

The sequences from \textit{E. coli} BH100 sub-strains were annotated using an \textit{in-house} script for annotation transfer from \textit{E. coli} 536. The sequences were also submitted to the RAST automatic annotation tool. Resistance genes were identified using the Comprehensive Antibiotic Resistance Database (CARD) (Alcock et al., 2020). Annotation of insertion sequences (IS) was performed using the ISFinder platform (Siguier et al., 2006). The identification of phage sequences was performed at PHASTER (Arndt et al., 2016). The serotype of BH100 strain was determined via SerotypeFinder 2.0 (Joensen et al., 2015). The Artemis software 16.0 (Carver et al., 2012) was used in the manual curation of annotations based on search results at Uniprot and Pfam databases. Finally, the prediction of the incompatibility groups of each plasmid occurred via PlasmidFinder 1.3 (Carattoli et al., 2014).

**Genomic Synteny Analysis**

In order to investigate possible genomic rearrangement and inversion, the synteny between the two plasmids in each strain was investigated through a multiple alignment using Mauve 2.4 (Darling et al., 2004).

**Genomic Islands Prediction**

The prediction of genomic islands (GEI) in \textit{E. coli} BH100 variants was performed using the software IslandViewer 4 (Bertelli and Brinkman, 2018) and Genomic Island Prediction Software (GIPSY) (Soares et al., 2016). Virulence factors were predicted through an alignment via BLAST against the virulence factor database (VFDB) (Chen et al., 2005). The \textit{E. coli} K12 MG1655 (RefSeq: NC_000913.3) and \textit{E. coli} 536 (RefSeq: NC_008253.1) strains were used as controls of non-pathogenic and pathogenic strains, respectively. Visualization of GEI was performed in BRIG v0.95 (Alikhan et al., 2011).

**Phylogenetic Analysis**

Two approaches were used for phylogenomic analysis, the Phylogenetic Tree Building Service available in the Pathosystems Resource Integration Center (PATRIC) (Wattam et al., 2017) and the PGADB-builder tool (Liu et al., 2016). The complete genome sequence of \textit{E. coli} BH100 MG2014 and other 14 genomes of this species from GenBank representing the phylogroups A, B1, B2, C, D, E, and F were considered in these analysis (Supplementary Table 2). The genome of \textit{E. fergusonii} ATCC 35469 (NC_011740) was used as an outgroup. For PATRIC’s Phylogenetic Tree Building Service, we used the Codon Tree method with 1,000 single-copy gene families. In this pipeline, protein families were identified using PGFams (Davis et al., 2016). The protein sequences are aligned using MUSCLE (Edgar, 2004) and using the codon_align function of BioPython (Cock et al., 2009). The alignments are used as input to estimate the samples phylogeny using the Maximum Likelihood method implemented in RaxML (Stamatakis, 2014), with using 100 rounds of rapid bootstrapping (Stamatakis et al., 2008). Concomitantly, a tree was generated from the whole genome multilocus sequence typing (wgMLST) using the PGADB-builder, which uses a database of alleles from the pan-genome of these strains. Both trees were visualized using iTOL (Letunic and Bork, 2019).

**RESULTS AND DISCUSSION**

**Sequencing and Assembly**

The \textit{E. coli} BH100 sub-strains were sequenced using different Next Generation Sequencing platforms (Ion Torrent and Illumina). A total of 4,269,641 fragment reads, 3,596,096 3 SE, single-end; PE, paired-end; MP, mate-pair.

**TABLE 1 | \textit{E. coli} BH100 variants sequencing and assembly statistics.**

| Sub-strain | Platform/library | Replicon | Size | Depth |
|------------|------------------|----------|------|-------|
| BH100      | Ion torrent/M pAp | Chromosome | 5,072,848 | 157× |
| MG2014     | and SE           |           |       |       |
|            |                  | Plasmid pBH100-1 | 107,274 | 157× |
|            |                  | Plasmid pAp     | 14,241  | 956×  |
| BH100      | Illumina         | Chromosome | 5,131,212 | 439× |
| MG2017     | Hi-seq/PE        |           |       |       |
|            |                  | Plasmid pBH100-1 | 105,801 | 1,464× |
|            |                  | Plasmid pAp     | 33,924  | 1,050× |
| BH100L     | Illumina         | Chromosome | 5,033,217 | 368× |
| MG2017     | Hi-seq/PE        |           |       |       |
|            |                  | Plasmid pBH100alpha | 103,103 | 1,408× |
| BH100N     | Illumina         | Chromosome | 5,116,036 | 368× |
| MG2017     | Hi-seq/PE        |           |       |       |

**Comparative Genomics of \textit{Escherichia coli} BH100**

Frontiers in Microbiology | www.frontiersin.org 3 January 2021 | Volume 11 | Article 549254
generated. These genomes have been assembled, annotated, and deposited as complete sequences in GenBank/NCBI (Table 1).

**Genomic Synteny of E. coli BH100 Sub-strains**

The synteny analysis of the pBH100 variants performed in Mauve revealed a conserved structure among strains BH100 MG2014, BH100 MG2017, and BH100L MG2017. However, the pApR variant from strain BH100 MG2017, presented a dramatic difference in terms of nucleotide similarity when compared to pAp from BH100 MG2014 (Figure 1). In fact, a portion of ~15 kb from pApR was shown to be identical to a sequence located in pBH100 variants, suggesting a recombination event between the two plasmids in the BH100 MG2017. Other small portions of pAp also seem to be inserted in pBH100 in the MG2014 and BH100 MG2017.

**Differences in Structural and Functional Properties of E. coli BH100 Sub-strains**

A summary of the structural annotations from E. coli BH100 variants chromosomes and plasmids can be seen in Supplementary Tables 3, 4. Distribution of Insertion Sequences (IS) along the chromosomes is illustrated in Figure 2 and Supplementary Figure 1. In general, our results reveal considerable differences in the number of coding sequences (CDS) and pseudogenes between BH100 variants chromosomes. It has been demonstrated by other studies that IS elements play an essential role in pathogen evolution by providing the sites for genome rearrangements, duplications and deletions (Jackson et al., 2011; Proença et al., 2017). Previous studies suggest IS3 can function as a mobile promoter in E. coli due to the presence of an outward promoter (Charlier et al., 1982). A gene expression approach was not carried out in the present study. In this context, we suggest that a further study comparing the transcriptome profile should be conducted to investigate the promoter activity of IS3 identified in this work. In summary, these results suggest that the activity of the IS3 and IS21 are important factors in the evolution of this strain once they provide genetic variability.

**Drug Resistance Genes and Associated Transposons**

The genetic map of transposons, and resistance genes, representing the original location and synteny in pBH100 and pAp from BH100 MG2014 are illustrated in Figures 3, 4. Additional information for pBH100 is available in the Supplementary Table 5. Among the MGE observed, we highlight Tn21, which was detected in all variants of pBH100. This transposon contains...
the class 1 integron, In2, and it is possibly responsible for the mobilization of the mer operon and the streptomycin resistance gene, aadA1. Tn21 was first described in the R100 plasmid, also known as NR1 from Shigella flexneri, which also determines resistance to inorganic mercury. Other studies have reported a conserved structure of this transposon in

![FIGURE 2](image)

**FIGURE 2** | Distribution of IS elements and GEI on E. coli BH100L MG2017. An alignment between the genomic sequences of E. coli BH100 MG2017 and BH100 MG2014 is shown as indicated by the rings in light blue (70–100% identity). GEI identified via IslandViewer4 in E. coli MG2017 are indicated by the gray bars. The IS elements and other transposases, predicted by the ISfinder tool, are indicated by arrows in green (IS21), black (IS3), blue (IS4), orange (IS110), purple (Tn3), and red (Tn21). GC content is represented by the inner ring in black.
many other R plasmids from different Gram-negative bacteria, suggesting it is widely distributed (Clennel et al., 1995; Kiyono et al., 2009). The presence of \textit{In2} has been considered one of the most important factors contributing to MDR widespread in Gram-negative pathogens as they have the ability to capture and eventually accumulate gene cassettes that confer adaptive advantages (Chang et al., 2009; Ahangarzadeh Rezaee et al., 2012; Firoozeh et al., 2019). In this study, we show that a Tn21 was detected in the recombined region (locus tag BH100Bp05113) of pApR (NZ_CP025252.1) from BH100 MG2017 (Figure 5 and Supplementary Figure 2). Furthermore, IS21 and IS3 transposases were found located upstream of this transposon, suggesting they could play a role in the mobilization of MDR genes in \textit{E. coli} BH100 strain (Figure 5). Intriguingly, the presence of Tn21 has also been detected in the \textit{E. coli} BH100L MG2017 chromosome but not in the remaining replicons, indicating that it was probably transferred from pBH100 (Figure 2).

All pAp variants presented Tn3, which carries the gene \textit{bla}, encoding the β-lactamase TEM-1 type (Gerlach and Wiedemann, 1985). Surprisingly, Tn3 was also found in pBH100-1 from both BH100 MG2014 and BH100 MG2017 but not in the pBH100alpha from the pAp-cured strain BH100L MG2017. Thus, this finding suggests Tn3 was disseminated from pAp to other replicons.

The description of other transposons associated with MDR genes which were not present in rearrangements can be found in the Supplementary Material. In addition to the resistance genes located in the plasmids, analysis performed by CARD and UniProtKB/Swiss-Prot databases in this study.
predicted up to 18 drug transmembrane transport genes and 29 transcriptional regulators associated with multidrug resistance, including fluoroquinolone class in the chromosome of all BH100 sub-strains (80–100% identity) (Supplementary Table 6).

Therefore, our findings show that several resistance genes from *E. coli* BH100 are contained in functional transposable elements as we confirm their activity as being involved in genomic rearrangements, thus representing the potential ability of this strain to mobilize antibiotic resistance genes.

**Genomic Islands Prediction**
The genomes from *E. coli* BH100 sub-strains presenting rearrangements were analyzed for the prediction of GEI (Figure 2 and Supplementary Figure 1). In total, around 80 islands were predicted in BH100 variants. Moreover, we predicted 72 transposase genes in BH100 MG2014 GEI, in which the majority, represented by 16 and 12 genes, were classified into IS3 and IS21 families, respectively. The GEI distribution was shown to be conserved among the strains, though BH100L MG2017 presented two integrated sequences of Tn21, associated with Mer operon and streptomycin resistance genes, in 2 different islands. This result suggests that the genetic variability promoted by IS21 and IS3 could have an important role in the formation of resistance islands. Interestingly, a recent study has reported UPEC strains in Mexico containing Tn21 with similar MDR cassette, but integrated in resistance islands in the chromosome (Paniagua-Contreras et al., 2019). In this context, our work supports the hypothesis that pBH100 and other related R100 plasmids might be involved for MDR spreading through the development of Resistance Islands alternatively to conjugation mechanisms.

In order to identify genes coding for uropathogenic virulence factors, the results of BH100 MG2014 were compared to the

---

**FIGURE 4** | Schematic representation of pAp from *E. coli* BH100 MG2014. The position and orientation of all CDS of pAp annotated in RAST Server are represented in purple arrows.
non-pathogenic *E. coli* K-12 MG1655 and the UPEC 536, as shown in Figure 6 and Supplementary Figures 3, 4). This analysis shows the majority of GEI are shared between *E. coli* BH100 and 536 but are not present in the non-pathogenic K12 strain, suggesting these GEI are probably pathogenicity islands (PAI). The gene content related to virulence of those possible PAI are discussed below.

An important uropathogenic virulence factor is the hemolysin toxin, which causes epithelial cell exfoliation and promotes the dissemination of the pathogen (Justice and Hunstad, 2012). Our data reveal the presence of eight genes (two *hlyABCD* operons) involved in transport, activation, and secretion of hemolysin both in the BH100 strain and in the *E. coli* 536 while, as expected, no hemolysin related gene was found in K-12. Recent studies have reported iron uptake genes as being vital in the pathogenesis of UPEC strains to overcome the scarcity of iron (Robinson et al., 2018; Bauckman et al., 2019). Interestingly, three siderophore genes were found in the *E. coli* strains 536 and BH100 MG2014 but none in the K-12.

In summary, the presence of IS3 and IS21 in the predicted genomic islands containing genes associated with virulence in *E.
coli BH100 suggest their potential role in providing mobility of virulence factors.

**Evolutionary Relationships of *E. coli* BH100 Replicons With UPEC and Other MDR Bacteria**

Both phylogenomic approaches revealed *E. coli* BH100 is clustered together with the 536 and S88 strains representing the B2 phylogroup thus showing the robustness of the methods used in this study (Figure 7 and Supplementary Figure 5). Furthermore, this result supports that BH100 might be considered as UPEC pathotype once the *E. coli* 536 genome is a model for the genetic basis of uropathogenic virulence (Dobrindt et al., 2001). Though some differences between the two strains can be highlighted as indicated by the serotype analysis revealing that BH100 (O6: H31) does not contain the capsular antigens, which are present in the 536 strain (O6:

![Distribution of IS elements and GEI on E. coli 536](image-url)
K15: H31) (Wiles et al., 2008). Moreover, when contemplating the IS3 family, it has been rarely reported in E. coli 536 (Terlizzi et al., 2017). Therefore, these differences might suggest the acquisition of genetic features, including MGE from different sources.

To determine the incompatibility groups of E. coli BH100 plasmids, the nucleotide sequence was analyzed on PlasmidFinder. These results show that pBH100 and pAp belong to IncFII and ColRNAI groups (100% identity), respectively. Interestingly, plasmids from these incompatibility groups have not been reported in 536 strain, suggesting BH100 may have acquired them from another strain. The alignment via BLASTn against Non-redundant sequences of the NCBI database indicates that the NR1 sequence (GenBank: DQ364638.1), a representative plasmid of IncFII, is very similar to pBH100alpha from E. coli BH100 MG2014 (99.74% identity, 92% coverage, E value: 0.0). The nucleotide sequence of Tn21 present in pBH100 proved to be identical to that described in plasmid NR1, which carries the In2 associated with the mer operon. Previous studies have reported the wide distribution of Tn21/In2 in enterobacteria, might be explained by the exposure of toxic metals in agricultural and industrial practice at the same time when the clinical use of antibiotic drugs was rapidly increasing (Trieu-Cuot et al., 1993; Nascimento and Chartone-Souza, 2003).

Regarding pAp from E. coli BH100 MG2014, it was possible to observe a high similarity to many Klebsiella sp. plasmids through the alignment via BLASTn, especially pNJST258N4 from Klebsiella pneumoniae 30660/NJST258_1 (GenBank DQ298019.1) (99.93% identity, 92% coverage, E value: 0.0). The closest E. coli plasmid sequence revealed to be pECAZ161_KPC (GenBank: CP019010.1). However, the significant difference is the presence of the Klebsiella pneumoniae carbapenemase enzyme gene located in pECAZ161_KPC. This result is very interesting, as carbapenemases have appeared over the last 18 years as the most clinically relevant antimicrobial resistance in Enterobacteriaceae (Stoesser et al., 2017). Although the present study does not support the direction of genetic flow between these strains, we suppose that BH100 might have either the ability to acquire this plasmid from Klebsiella sp. or vice versa.

![Phylogenomic tree of E. coli BH100 and GenBank sequences inferred by PATRIC.](image-url)
CONCLUSION
This study has described E. coli BH100 genetic diversity, especially concerning functional mobile elements associated with MDR and virulence genes potentially related to uropathogenesis. To our knowledge, it is the first time that the IS3 family has been associated with UPEC pathotype. Furthermore, our results have shown that conserved regions of E. coli BH100 are shared not only with UPEC strains but also with other MDR Enterobacteriaceae. In this context, we highlight the importance of future studies further to investigate the evolutionary role of IS3 in UPEC strains and other enterobacteria.

DATA AVAILABILITY STATEMENT
The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/Supplementary Material.

AUTHOR CONTRIBUTIONS
RC performed structural and functional bioinformatic analysis of the E. coli strains, interpreted the data and was the major contributor in writing the manuscript. FA was responsible for bioinformatic analysis, interpretation of data. MC performed the phylogenomic analysis and contributed to manuscript writing. EC-S and AN was responsible for ceding the E. coli BH100 strains and contributed to data interpretation. LJ performed DNA extraction. AZ contributed to manuscript writing and data interpretation. BB was responsible for samples sequencing and manuscript writing. DB, PG, and AG-N contributed to data interpretation and writing the manuscript. HF performed sample sequencing and manuscript writing. SS was responsible for genomic island prediction and explanation of the analysis. RR and AP contributed to the analysis and interpretation of genomic data. VA provided interpretation of genomic data and was a major contributor to the revision of the manuscript.

FUNDING
This work was funded by CAPES, CNPq, Fapemig, and PRPq-UFGM. Grant number:312045/2020-4.

ACKNOWLEDGMENTS
The authors would like to thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Fundação de Amparo a Pesquisa do Estado de Minas Gerais (Fapemig), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Pró-Reitoria de Pesquisa da UFMG for their financial support and fellowships. We would also like to show our gratitude to G de Oliveira, who provided insight and expertise that greatly assisted the research.

SUPPLEMENTARY MATERIAL
The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmicb.2020.549254/full#supplementary-material

REFERENCES
Ahangarzadeh Rezaee, M., Langarizadeh, N., and Aghazadeh, M. (2012). First report of class 1 and class 2 integrons in multidrug-resistant Klebsiella pneumoniae isolates from northwest Iran. Immu. J. Infect. Dis. 65, 256–259. doi: 10.7883/venken.65.256
Alcock, B. P., Raphenya, A. R., Lau, T. T. Y., Tsang, K. K., Bouchard, M., Edalatmand, A., et al. (2020). CARD 2020: antibiotic resistome surveillance with the comprehensive antibiotic resistance database. Nucleic Acids Res. 48, D517–D525. doi: 10.1093/nar/gkz935
Alikhan, N. F., Petty, N. K., Ben Zakour, N. L., and Beaton, S. A. (2011). BLAST Ring Image Generator (BRIG): Simple prokaryote genome comparisons. BMC Genomics 12:402. doi: 10.1186/1471-2164-12-402
Antipov, D., Hartwick, N., Shen, M., Raiko, A., Mabrouk, N., and Pezarov, P. A. (2016). plasmidSPAdes: assembling plasmids from whole genome sequencing data. Bioinformaticas 32, 3380–3387. doi: 10.1093/bioinformatics/btw493
Arndt, D., Grant, J. R., Marcu, A., Sajed, T., Pon, A., Liang, Y., et al. (2016). PHASTER: a better, faster version of the PHAST phage search tool. Nucleic Acids Res. 44, W16–21. doi: 10.1093/nar/gkw387
Bauckman, K. A., Matsuda, R., Higgins, C. B., DeBosch, B. J., Wang, C., and Mysorekar, I. U. (2019). Dietary restriction of iron availability attenuates UPEC pathogenesis in a mouse model of urinary tract infection. Am. J. Physiol. Renal Physiol. 316, F814-F822. doi: 10.1152/ajprenal.0133.2018
Bertelli, C., and Brinkman, F. S. L. (2018). Improved genomic island predictions with IslandPath-DIMOB. Bioinformaticas 34, 2161–2167. doi: 10.1093/bioinformatics/bty095
Bouanchaud, D. H., Scavizzi, M. R., and Chabbert, Y. A. (1968). Elimination by ethidium bromide of antibiotic resistance in enterobacteria and staphylococci. J. General Microbiol. 54, 417–425. doi: 10.1099/00221287-54-3-417
Carattoli, A., Zankari, E., García-Fernández, A., Voldby Larsen, M., Lund, O., Villa, L., et al. (2014). In silico detection and typing of plasmids using PlasmidFinder and plasmid multilocus sequence typing. Antimicrob. Agents Chemother. 58, 3895–3903. doi: 10.1128/AAC.02412-14
Carver, T., Harris, S. R., Berriman, M., Parkhill, J., and McQuillan, J. A. (2012). Artemis: an integrated platform for visualization and analysis of high-throughput sequence-based experimental data. Bioinformatics 28, 464–469. doi: 10.1093/bioinformatics/btt703
Chang, Y., Fang, T., Tsai, M., Chang, L., and Yu, L. (2009). Characterization of class 1 integrons and gene cassettes in clinical isolates of Klebsiella pneumoniae from Taiwan. Diagn. Microbiol. Infect. Dis. 65, 214–216. doi: 10.1016/j.diagmicrobio.2009.06.005
Charlier, D., Piette, J., and Glansdorff, N. (1982). IS3 can function as a mobile promoter in E. coli. Nucleic Acids Res. 10, 5935–5948. doi: 10.1093/nar/10.19.5935
Charlone-Souza, E. (2017). Personnal communication.
Chen, L., Yang, J., Yu, Z., Sun, L., Shen, Y., et al. (2005). VFDB: a reference database for bacterial virulence factors. Nucleic Acids Res. 33, D325–D328. doi: 10.1093/nar/gkj008
Chikhi, R., and Medvedev, P. (2014). Informed and automated k-mer size selection for genome assembly. Bioinformatics 30, 31–37. doi: 10.1093/bioinformatics/btt310
Clennel, A. M., Johnston, B., and Rawlings, D. E. (1995). Structure and function of Tn5467, a Tn21-like transposon located on the Thiodacillus ferrooxidans.
broad-host-range plasmid PT-F2C2. Appl. Environ. Microbiol. 61, 4223–4229. doi: 10.1128/AEM.61.12.4223-4229.1995
Cock, P. J., Antao, T., Chang, J. T., Chapman, B. A., Cox, C. J., Dalke, A., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics 25, 1422-1423. doi:10.1093/bioinformatics/btp163
Darling, A. C. E., Mau, B., Blattner, F. R., and Perna, N. T. (2004). Mauve: multiple alignment of conserved genomic sequence with rearrangements. Genome Res, 14, 1394–1403. doi:10.1101/gr.2289704
Davies, J., and Davies, D. (2010). Origins and evolution of antibiotic resistance. Microbiol. Mol. Biol. Rev. 74, 417–433. doi:10.1128/MMBR.00016-10
Davies, J. I., Gerdes, S., Olsen, G. J., Olson, R., Pusch, G. D., Shukla, M., et al. (2016). PATyFams: protein families for the microbial genomes in the PATRIC database. Front. Microbiol,7, 1–12. doi: 10.3389/fmicb.2016.00118
Dobrindt, U., Blum-Oehler, G., Hartsch, T., Gottschalk, G., Ron, E. Z., Funfstück, R., et al. (2001). S-Fimbria-encoding determinant sia(f) is located on pathogenicity island III (536) of uropathogenic Escherichia coli strain 536. Infect. Immun. 69, 4248–4256. doi:10.1128/IAI.69.7.4248-4256.2001
Edgar, R. C. (2004). MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. 32, 1792-1797. doi:10.1099/nar.gkb340
Firnrozh, F., Mahluji, Z., Khoshadi, A., and Zibaee, M. (2019). Molecular characterization of class 1, 2 and 3 integrons in clinical multi-drug resistant Klebsiella pneumoniae isolates. Antimicrob. Resist. Infect. Control 8:59. doi:10.1186/s13756-019-0509-3
Galardini, M., Biondi, E. G., Bazzicalupo, M., and Mengoni, A. (2011). CONTIGuator: a bacterial genomes finishing tool for structural insights on draft genomes. Source Code Biol. Med. 6:11. doi:10.1186/1751-0473-6-11
Gerlach, B. A., and Wiedemann, B. (1985), Tn3 as the molecular basis for antibiotic resistance. FEMS Microbiol. Lett. 252, 183–190. doi:10.1016/j.femsle.2005.08.028
Moulin-Schouleur, M., Réperant, M., Laurent, S., Brée, A., Mignón-Grasteau, S., Germon, P., et al. (2007). Extraintestinal pathogenic Escherichia coli strains of avian and human origin: link between phylogenetic relationships and common virulence factors. J. Clin. Microbiol. 45, 3366–3376. doi:10.1128/JCM.00037-07
Nakano, S., Fujisawa, T., Ito, Y., Chang, B., Matsumura, Y., Yamamoto, M., et al. (2019). Whole-genome sequencing analysis of multidrug-resistant serotype 15A Streptococcus pneumoniae in Japan and the emergence of a highly resistant serotype 15A-STW084 clone. Antimicrob. Agents Chemother. 63:e02579-18. doi:10.1128/AAC.02579-18
Nascimento, A. M. A., Azevedo, M. O., Astolfi-Filho, S., and Chartone-Souza, E. (1992). Cloning of a mercuric ion-resistance operon into Escherichia coli SK using the mini-plasmid technique. Biotechnol. Tech, 6, 139–142. doi: 10.1080/095228600086468
Pacheco, L. G. C., Pena, R. R., Castro, T. L. P., Dorella, F. A., Bahia, R. C., Carminati, R., et al. (2007). Multiplex PCR assay for identification of Corynebacterium pseudotuberculosis from pure cultures and for rapid detection of this pathogen in clinical samples. J. Med. Microbiol. 56, 480–486. doi:10.1099/jmm.0.46997-0
Paniagua-Contreras, G. L., Monroy-Pérez, E., Díaz-Velásquez, C. E., Uribe-García, A., Labastida, A., Peñaloza-Figueroa, F., et al. (2019). Whole-genome sequence analysis of multidrug-resistant uropathogenic strains of Escherichia coli from Mexico. IDR Vol, 12, 2363–2377. doi:10.2147/IDR.S203661
Parker Hitchens, A. (1921). Advantages of culture mediums containing small percentages of agar. J. Infect. Dis. 29, 390–407. doi: 10.1093/infdis/29.5.390
Poolman, J. T., and Wacker, M. (2016). Extraintestinal pathogenic Escherichia coli, a common human pathogen: challenges for vaccine development and progress in the field. J. Infect. Dis. 213, 6–13. doi: 10.1093/infdis/jiv429
Pöoena, J. T., Barral, D. C., and Gordo, I. (2017). Commensal-to-pathogen transition: one-single transposon insertion results in two pathoadaptive traits in Escherichia coli -macrophage interaction. Sci. Rep. 7, 1–12. doi:10.1038/s41598-017-04808-1
Robinson, A. E., Heffernan, J. R., and Henderson, J. P. (2018). The iron hand of uropathogenic Escherichia coli: the role of transition metal control in virulence. Future Microbiol. 13, 745–756. doi:10.2217/fmb-2017-0295
Römer, A., Wieler, L. H., and Schierack, P. (2012). Analyses of intestinal commensal Escherichia coli strains from wild boars suggest adaptation to conventional pig production conditions. Vet. Microbiol. 161, 122–129. doi:10.1016/j.vetmic.2012.07.009
Siguier, P., Perochon, J., Lestrade, L., Mahillon, J., and Chandler, M. (2006). ISfinder: the reference centre for bacterial insertion sequences. Nucleic Acids Res. 34, D32–D36. doi: 10.1093/nar/gkj014
Soares, S. C., Geyik, H., Ramos, R. T. J., de Sá, P. H. C. G., Barbosa, E. G. V., Baumbach, J., et al. (2016). GIPsy: genomic island prediction software. J. Biotechnol. 232, 2–11. doi:10.1016/j.jbiotec.2015.09.008
Stamatakis, A. (2014). RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30, 1312–1313. doi:10.1093/bioinformatics/btu033
Liu, Y. Y., Chiou, C. S., and Chen, C. C. (2016). PGAdb-builder: A web service tool for creating pan-genome allele database for molecular fine typing. Sci. Rep. 6:36213. doi:10.1038/srep36213
Marrs, C. F., Zhang, L., and Foxman, B. (2005). Escherichia coli mediated urinary tract infections: are there distinct uropathogenic E. coli (UPEC) pathotypes? FEMS Microbiol. Lett. 252, 183–190. doi:10.1016/j.femsle.2005.08.028
Frontiers in Microbiology | www.frontiersin.org 12
January 2021 | Volume 11 | Article 549254
Stamatakis, A., Sheppard, A. E., Pfeifer, G., Anson, L. W., Pankhurst, L., Sebra, R., et al. (2017). Genomic epidemiology of global Klebsiella pneumoniae carbapenemase (KPC)-producing Escherichia coli. Sci. Rep. 7:5917. doi: 10.1038/s41598-017-06256-2

Terlizzi, M. E., Gribaudo, G., and Maffei, M. E. (2017). Uropathogenic Escherichia coli (UPEC) Infections: virulence factors, bladder responses, antibiotic, and non-antibiotic antimicrobial strategies. Front. Microbiol. 8:1566. doi: 10.3389/fmicb.2017.01566

Trieu-Cuot, P., de Cespédés, G., Bentorcha, F., Delbos, F., Gaspar, E., and Horaud, T. (1993). Study of heterogeneity of chloramphenicol acetyltransferase (CAT) genes in streptococci and enterococci by polymerase chain reaction: characterization of a new CAT determinant. Antimicrob. Agents Chemother. 37, 2593–2598. doi: 10.1128/AAC.37.12.2593

Wattam, A. R., Davis, J. J., Assaf, R., Boisvert, S., Brettin, T., Bun, C., et al. (2017). Improvements to PATRIC, the all-bacterial bioinformatics database and analysis resource center. Nucleic Acids Res. 45, D535–D542. doi: 10.1093/nar/gkw1017

Wiles, T. J., Kulesus, R. R., and Mulvey, M. A. (2008). Origins and virulence mechanisms of uropathogenic Escherichia coli. Exp. Mol. Pathol. 85, 11–19. doi: 10.1016/j.yexmp.2008.03.007

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.