



Comparative Evolutionary Patterns of *Burkholderia cenocepacia* and *B. multivorans* During Chronic Co-infection of a Cystic Fibrosis Patient Lung

OPEN ACCESS

Edited by:

Jianying Gu,
College of Staten Island, CUNY,
United States

Reviewed by:

Joanna Goldberg,
School of Medicine, Emory University,
United States
Jonathan J. Dennis,
University of Alberta, Canada

*Correspondence:

Professor Isabel Sá-Correia
isacorreia@tecnico.ulisboa.pt

† Present address:

A. Amir Hassan,
Laboratory of Microbiology – LM,
Department of Biochemistry and
Microbiology, Ghent University, Ghent,
Belgium

Specialty section:

This article was submitted to
Evolutionary and Genomic
Microbiology,
a section of the journal
Frontiers in Microbiology

Received: 20 June 2020

Accepted: 31 August 2020

Published: 25 September 2020

Citation:

Hassan AA, dos Santos SC,
Cooper VS and Sá-Correia I (2020)
Comparative Evolutionary Patterns of
Burkholderia cenocepacia and
B. multivorans During Chronic
Co-infection of a Cystic Fibrosis
Patient Lung.
Front. Microbiol. 11:574626.
doi: 10.3389/fmicb.2020.574626

A. Amir Hassan^{1,2†}, Sandra C. dos Santos², Vaughn S. Cooper³ and Isabel Sá-Correia^{1,2*}

¹ iBB - Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal,

² Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal, ³ Department of Microbiology and Molecular Genetics, University of Pittsburgh School of Medicine, Pittsburgh, PA, United States

During chronic respiratory infections of cystic fibrosis (CF) patients, bacteria adaptively evolve in response to the nutritional and immune environment as well as influence other infecting microbes. The present study was designed to gain insights into the genetic mechanisms underlying adaptation and diversification by the two most prevalent pathogenic species of the *Burkholderia cepacia* complex (Bcc), *B. cenocepacia* and *B. multivorans*. Herein, we study the evolution of both of these species during coinfection of a CF patient for 4.4 years using genome sequences of 9 *B. multivorans* and 11 *B. cenocepacia*. This co-infection spanned at least 3 years following initial infection by *B. multivorans* and ultimately ended in the patient's death by cepacia syndrome. Both species acquired several mutations with accumulation rates of 2.08 (*B. cenocepacia*) and 2.27 (*B. multivorans*) SNPs/year. Many of the mutated genes are associated with oxidative stress response, transition metal metabolism, defense mechanisms against antibiotics, and other metabolic alterations consistent with the idea that positive selection might be driven by the action of the host immune system, antibiotic therapy and low oxygen and iron concentrations. Two orthologous genes shared by *B. cenocepacia* and *B. multivorans* were found to be under strong selection and accumulated mutations associated with lineage diversification. One gene encodes a nucleotide sugar dehydratase involved in lipopolysaccharide O-antigen (OAg) biosynthesis (*wbil*). The other gene encodes a putative two-component regulatory sensor kinase protein required to sense and adapt to oxidative- and heavy metal-inducing stresses. This study contributes to understanding of shared and species-specific evolutionary patterns of *B. cenocepacia* and *B. multivorans* evolving in the same CF lung environment.

Keywords: *Burkholderia cepacia* complex, cystic fibrosis, chronic pulmonary infections, within-host evolution, comparative genomic analysis, *B. cenocepacia*, *B. multivorans*

INTRODUCTION

Chronic pulmonary infections are considered the leading cause for morbidity and premature death of patients at risk, as cystic fibrosis (CF) patients and immunocompromised individuals (Lipuma, 2010). Several opportunistic pathogens may populate the CF airways and the heterogeneous microbial populations usually present are difficult to eradicate (Lipuma, 2010). Among the bacterial pathogens, *Burkholderia cepacia* complex (Bcc) species are particularly feared by CF patients (Mahenthalingam et al., 2005). Currently, the Bcc includes 24 closely related bacterial species (De Smet et al., 2015; Depoorter et al., 2016; Ong et al., 2016; Bach et al., 2017; Weber and King, 2017; Martina et al., 2018). Pulmonary infection with Bcc bacteria can involve a single species or co-infection with more than one Bcc species and, besides interpatient transmission, these bacteria can cause chronic infections and a marked decline of lung functions and decreased life expectancy. In certain cases, some Bcc species can lead to a lethal uncontrolled clinical deterioration with septicemia and necrotizing pneumonia (the “cepacia syndrome”) (Mahenthalingam et al., 2005).

During chronic infection by Bcc bacteria, the original infecting strain(s) establishes a population that evolves and diversifies into genetically and phenotypically distinct lineages. These lineages are thought to acquire multiple adaptations to the CF lung environment (Coutinho et al., 2011; Lieberman et al., 2011, 2014; Madeira et al., 2011, 2013; Mira et al., 2011; Silva et al., 2011, 2016; Moreira et al., 2014, 2017; Maldonado et al., 2016; Cabral et al., 2017; Hassan et al., 2017, 2019a; Lee et al., 2017; Nunvar et al., 2017). Over the last decade, several virulence factors and adaptive traits have been identified in Bcc species (Mahenthalingam et al., 2005; Drevinek and Mahenthalingam, 2010; Loutet and Valvano, 2010; Zlosnik and Speert, 2010; Madeira et al., 2011, 2013; Mira et al., 2011; Zlosnik et al., 2011, 2014; Sousa et al., 2017; Hassan et al., 2019a). Several studies have used genome-wide methods to identify genetic changes during long term infection by *B. dolosa* (Lieberman et al., 2011, 2014), *B. multivorans* (Silva et al., 2016; Diaz Caballero et al., 2018) and *B. cenocepacia* (Lee et al., 2017; Nunvar et al., 2017) within individual hosts. For example, a retrospective study involving 112 sequential *B. dolosa* isolates from 14 CF patients, identified several mutated genes involved in genetic variation within individual patients, in particular genes required for expression of surface polysaccharides, lipopolysaccharide O-antigen (OAg) biosynthesis, outer membrane components, iron scavenging, and antibiotic resistance (Lieberman et al., 2011). A subsequent study using the same sequential *B. dolosa* isolates suggested parallel bacterial evolution where different lineages may coexist for many years within a patient and identified candidate pathogenicity genes (Lieberman et al., 2014). Another comparative genomic study including 22 isolates of *B. multivorans* recovered over 20 years from CF patients also found parallel adaptive variations resulting from host related selection pressures in additional genes involved in the above referred functional categories (Silva et al., 2016). A more recent comparative genomic analysis including 32 clonal variants of *B. cenocepacia* obtained from 8 CF patients (Nunvar et al., 2017) revealed that, in addition to the

aforementioned parallel mutations in gene functions previously described for *B. dolosa* and *B. multivorans* (Lieberman et al., 2011, 2014; Silva et al., 2016), genes related to transition metal metabolism are hotspots for nucleotide polymorphism (Nunvar et al., 2017). Another comparative study involving 215 genomes from serial *B. cenocepacia* isolates obtained from 16 CF patients during a 20 year-period also supports the above mentioned evolutionary trajectories during chronic Bcc infections and the well-established diverse-community model (Lee et al., 2017). This publication also reported the complete loss of chromosome III resulting in genome-size reduction as an adaptive trait of *B. cenocepacia* (Lee et al., 2017), consistent with a previously reported rare *in vivo* loss of the same mega-plasmid among *B. cenocepacia* clonal variants emerging during chronic infection (Moreira et al., 2017). Recently, a population genomic study focused on 111 *B. multivorans* isolates obtained from a Canadian CF patient showed potential parallel pathoadaptation involving genes associated with resistance toward multiple classes of antibiotics (Diaz Caballero et al., 2018).

Although the above genomic analyses have established the field of *Burkholderia* evolution during chronic infection of the CF airways, the comparison of the evolutionary patterns of strains of different species co-infecting the same CF patient is lacking. This was the objective of the present study focused on the most prevalent Bcc species among the CF community worldwide: *B. cenocepacia* and *B. multivorans* (Mahenthalingam et al., 2005; Drevinek and Mahenthalingam, 2010; Lipuma, 2010). Although *B. cenocepacia* has been identified as the most frequent species among the Bcc bacteria causing infections, with high potential for inter-patient transmission (Mahenthalingam et al., 2005; Drevinek and Mahenthalingam, 2010), in several countries *B. multivorans* has replaced *B. cenocepacia* as most frequent (Lipuma, 2010). Given that the CF environment is characterized by high pro-inflammatory cytokine levels, high antibiotic concentrations, high levels of oxidative stress and low oxygen concentration (Moriarty et al., 2007; Palmer et al., 2007; Reid et al., 2007; Williams et al., 2007), we hypothesize that the adaptive evolution of each Bcc species may vary in the same CF patient' lung-environment. This work was designed to compare the genome sequences of *B. cenocepacia* and *B. multivorans* clonal variants co-inhabiting the same host-selective environment to identify species-specific and shared evolutionary patterns. A retrospective study on twenty clonal variants derived from two ancestor strains (9 *B. multivorans* isolates and 11 *B. cenocepacia* isolates) was performed. They were sequentially retrieved from the same CF patient over a period of 4.4 years, from the onset of infection with *B. multivorans* followed by co-infection with *B. cenocepacia* until the patient's death from cepacia syndrome (Correia et al., 2008; Coutinho et al., 2011; Hassan et al., 2019a). These isolates were obtained in the major Portuguese CF treatment Center at Hospital de Santa Maria during consultation routines and pulmonary exacerbations that compelled the patient to hospitalization and intravenous therapy with gentamicin and ceftazidime (Correia et al., 2008; Coutinho et al., 2011). The design and successful execution of the present study provided relevant and useful information to contribute to a better understanding of the common and the

TABLE 1 | *Burkholderia cepacia* complex isolates examined, ordered based on the isolation date.

Bcc isolate	Isolation date	Bcc species
IST419	26 Feb 1998	<i>B. multivorans</i>
IST424	4 Jun 1998	<i>B. multivorans</i>
IST439	30 Jan 1999	<i>B. cenocepacia recA</i> lineage IIIA
IST453	19 Jul 1999	<i>B. multivorans</i>
IST455A/IST455B	1 Feb 2000	<i>B. multivorans</i>
IST461	4 Apr 2000	<i>B. multivorans</i>
IST495A/IST495B	29 May 2001	<i>B. multivorans</i>
IST4103	24 Jul 2001	<i>B. cenocepacia recA</i> lineage IIIA
IST4110	25 Sep 2001	<i>B. cenocepacia recA</i> lineage IIIA
IST4112	11 Oct 2001	<i>B. cenocepacia recA</i> lineage IIIA
IST4113	6 Nov 2001	<i>B. cenocepacia recA</i> lineage IIIA
IST4119	22 Jan 2002	<i>B. multivorans</i>
IST4116A/IST4116B	11 Feb 2002	<i>B. cenocepacia recA</i> lineage IIIA
IST4131	26 Feb 2002	<i>B. cenocepacia recA</i> lineage IIIA
IST4129	26 Mar 2002	<i>B. cenocepacia recA</i> lineage IIIA
IST4130	14 May 2002	<i>B. cenocepacia recA</i> lineage IIIA
IST4134	2 Jul 2002	<i>B. cenocepacia recA</i> lineage IIIA

They were obtained from a cystic fibrosis patient and their genomes were sequenced during this study.

specific evolutionary patterns occurring in *B. cenocepacia* and *B. multivorans* under selection in the same CF lung environment and host immune system during a co-infection that ultimately led to the cepacia syndrome.

MATERIALS AND METHODS

Bcc Bacterial Isolates and Growth Conditions

Eleven *B. cenocepacia* (*recA* lineage IIIA) sequential isolates and nine *B. multivorans* sequential isolates obtained from the same cystic fibrosis patient (patient J) (Cunha et al., 2003; Coutinho et al., 2011; Hassan et al., 2019a) were examined in this study. These isolates were collected during routine monitoring at Hospital de Santa Maria (HSM), Centro Hospitalar Lisboa Norte (CHLN) EPE, Lisbon, Portugal, from the sputum of a chronically infected CF patient who was under surveillance from February 1998 to July 2002 (Cunha et al., 2003; Correia et al., 2008; Coutinho et al., 2011; Hassan et al., 2019a; **Table 1**). These isolates were obtained from the onset of the Bcc infection until the patient's death with cepacia syndrome after 4.4 years of Bcc infection and were selected at random among the colonies isolated in selective *Burkholderia cepacia* Selectatab medium at the Hospital, at a specific date of isolation. Bacterial cultures are stored at -80°C in 1:1 (v/v) glycerol. Bacteria were grown in Lysogeny broth (LB; Conda, Pronadisa) at 37°C with shaking at 250 rpm or in LB agar plates.

Genomic DNA Sequencing, Assembly and Annotation

For genomic DNA extraction, bacterial cultures were prepared by suspending isolated colonies from LB agar plates in 3 mL LB broth, followed by overnight growth at 37°C with shaking

at 250 rpm. Genomic DNA was extracted and purified using a DNeasy Blood and Tissue kit (Qiagen, Germany) according to manufacturer instructions. DNA concentration and purity were assessed using a Nanodrop ND-1000 spectrophotometer.

Genomic DNA samples of all the studied clonal variants were processed according to Illumina's instructions for generating paired-end libraries. In brief, *Burkholderia cenocepacia* IST439, IST4113, IST4129 and IST4134 were sequenced using a whole-genome shotgun sequencing strategy and Illumina Genome Analyser sequencing technology at CD Genomics (New York, NY, United States), generating short 100-bp paired-end reads with high coverage ($\sim 300\times$), details of this protocol attached as **Supplementary Materials**. All of the *B. multivorans* clonal variants were sequenced on Illumina NextSeq 500 platform at the University of Pittsburgh, Pittsburgh, PA, United States; isolates of *B. cenocepacia* IST4103, IST4110, IST4112, IST4116A, IST4116B, IST4131 and IST4130 were sequenced using the 151-bp paired-end Illumina HiSeq platform at the University of New Hampshire Hubbard Center for Genomic Studies, after library preparation using a modified Illumina Nextera protocol designed for inexpensive library preparation of microbial genomes (Baym et al., 2015). Raw fastQ reads were analyzed using fastQC, which revealed that all isolates were sequenced at sufficient depth to accurately identify single nucleotide polymorphisms and indel mutations. Raw fastq paired-end files were processed for removal of Illumina adapters, trimming, and quality-based filtering using Trimmomatic v0.32 (Bolger et al., 2014).

Burkholderia multivorans IST419 and *B. cenocepacia* IST439, used as ancestor reference strains, were sequenced for a second time to generate a high-quality contiguous assembly. *B. multivorans* IST419 was likewise processed using Illumina NextSeq 500 platform producing a second sets of paired-ended reads, while *B. cenocepacia* IST439 was re-sequenced to generate a complete assembly by using a combination of single molecule, real-time (SMRT) Pacific Biosciences – PacBio reads and Illumina 100-bp paired-end reads (protocol described in **Supplementary Materials**). We used the hierarchical genome-assembly process workflow (HGAP3) to generate a completed assembly of *B. cenocepacia* IST439 and polished our assembly using the Quiver algorithm (Chin et al., 2013).

The trimmed reads of all the studied clonal variants (except *B. cenocepacia* IST439) were *de novo* assembled using VelvetOptimiser v2.2.4 (Zerbino and Birney, 2008; Zerbino, 2010) and SPAdes v3.11.1-linux (Bankevich et al., 2012) with automated optimization of the assembly parameters. The quality of the assemblies obtained was estimated by Quast v2.3 (Gurevich et al., 2013) and the best assembled outputs were automatically improved using Pilon v1.22 (Walker et al., 2014). The three trimmed data sets (two paired-end libraries and one mate-pair library) were used for scaffolding by SSPACE-standard v3.0 (Boetzer et al., 2011) followed by an automated improvement using Pilon v1.22 to gain the final draft genome sequences.

With draft genome sequences, contigs/scaffolds were reordered using MAUVE Contig Mover (Darling et al., 2010). All assembled-drafted genomes were reordered versus *B. cenocepacia* J2315 and *B. multivorans* ATCC_17616. We then annotated the completed assembly of *B. cenocepacia*

IST439 and the remaining draft genome sequences with the prokaryotic genome annotation tool prokka (v1.11) with a local, customized *Burkholderia* database of closely related Bcc strains (*B. cenocepacia* H111, *B. cenocepacia* J2315, *B. cenocepacia* ST32, *Burkholderia cenocepacia* VC1254, *B. dolosa* AU0158, *B. multivorans* ATCC_BAA-247, and *B. multivorans* ATCC 17616) (Seemann, 2014). BLAST Ring Image Generator (BRIG) was used to generate visual genome comparisons with the reference genomes of *B. cenocepacia* H111, *B. cenocepacia* J2315, *B. cenocepacia* ST32, *B. multivorans* ATCC_BAA-247, and *B. multivorans* ATCC 17616 (Alikhan et al., 2011). For enriched-functional annotation, eggNOG-mapper v4.5.1 was additionally applied for both strains (*B. cenocepacia* IST439 and *B. multivorans* IST419) that were used as references for the variant calling (Huerta-Cepas et al., 2017). This COG mapping was used to count the CDS per COG category for each species.

Variant Calling and SNP/INDEL Detections

Trimmed paired-end reads were mapped against the ancestor reference complete genome of *B. cenocepacia* IST439 (for all *B. cenocepacia* clonal variants) and/or against the ancestor reference draft genome sequence of *B. multivorans* IST419 for the corresponding clonal variants using BWA-MEM packages of Burrows-Wheeler Aligner (BWA v.0.7.10) (Li and Durbin, 2010) and NovoAlign¹. Variants [Single nucleotide polymorphisms (SNPs) and insertion-deletion mutations (INDELs)] were called as described previously by using two independent standard variant calling pipelines; GATK and SAMtools/BCFtools toolbox (Li et al., 2009; Van der Auwera et al., 2013; Dillon et al., 2015, 2017). A brief description of this approach is included in **Supplementary Materials**.

To perform a quality control of the results obtained by the above reported method, a parallel comparison was performed using Breseq (Barrick et al., 2014), and similar predictions were found in the case of base-substitutions, with some discrepancies in indels. All putative SNPs/INDELs were then manually inspected and evaluated using the Integrative Genomics Viewer – IGV (Robinson et al., 2011; Thorvaldsdottir et al., 2013), and discarded if the BWA and Novoalign-produced alignments did not provide enough confidence (poor coverage, as described in the **Supplementary Materials**). Thus, we are confident that nearly all base-substitution and indels identified in this study were genuine events that arose during the *in vivo* evolutionary process. Finally, Functional annotation of the called variants was performed by using SnpEff v3.1 (Cingolani et al., 2012) with manual BLAST verification against the NCBI Microbes genome database.

Assessment of Population Structure and Phylogeny

In silico whole genome multi-locus sequence typing (wgMLST) (Jolley and Maiden, 2010) was performed to confirm the isolate sequence type (ST) and the clonality of all the isolates examined.

To obtain a comprehensive phylogeny of the studied clonal variants, the SNP calls and their matrices were automatically processed by using a set of algorithms previously described (Kaas et al., 2014) via the online web server CSI phylogeny v1.4² with default settings; min depth at SNP position – 10x, relative depth at SNP position – 10%, minimum distance between SNPs – 10 bp, min SNP quality – 30, and min read mapping quality – 25 (Kaas et al., 2014). Phylogenetic trees were then constructed using 70 and 110 variant positions among *B. cenocepacia* and *B. multivorans* clonal variants, respectively, by using FastTree (Price et al., 2010).

Seven *B. cenocepacia* *recA* lineage IIIA, two *B. multivorans*, and one *B. dolosa* deposited-reference genomes in *Burkholderia* genome database³ were used to describe the phylogenetic relationship of the studied clinical isolates among the Bcc bacteria (**Supplementary Table 1**). Whole genome average nucleotide identity (ANI) was assessed by using fastANI script (Jain et al., 2018).

Genomic Features and Structural Analysis of the Ancestor *B. cenocepacia* and *B. multivorans* Strains Genomes

Burkholderia cenocepacia IST439 and *B. multivorans* IST419 were considered the ancestor reference strains and used to identify subsequent evolved mutations. Reported genomic islands (GIs: BcenGIs and GiST32-s) (Holden et al., 2009; Graindorge et al., 2012; Nunvar et al., 2017) were searched for to identify their presence in *B. cenocepacia* IST439 and *B. multivorans* IST419 genomes using basic local alignment search tool – BLAST (Altschul et al., 1990; Zhang et al., 2000).

The presence of Putative GIs was also predicted using IslandViewer 4 (Bertelli et al., 2017), unique GIs were manually confirmed and curated in the genome sequences of *B. cenocepacia* IST439 and *B. multivorans* IST419. For this confirmation, it was considered that they are unique continuous DNA regions absent in the available genome sequences of the published epidemic clones and longer than 10 kbp, flanked, at both sides, by homologous regions (especially those inserted immediately downstream of tRNA), and have lower (or higher) GC content compared with the rest of the genome (Guo et al., 2017; Nunvar et al., 2017). The absence of DNA regions homologous to these unique GIs in the genomes of epidemic clones – *B. cenocepacia* ET12 and ST32 (Holden et al., 2009; Nunvar et al., 2017) and *B. multivorans* ATCC_17616 and ATCC_BAA-247 (Komatsu et al., 2003; Johnson et al., 2015; Johnson et al., 2016) – was confirmed and visualized by Progressive MAUVE whole genome pairwise alignment and ACT/Artemis (Carver et al., 2005; Darling et al., 2010). Large DNA regions that respected those criteria but were not predicted by IslandViewer, were also manually identified and considered genuine-unique GIs.

The assembled genomes of IST419 and IST439 were also examined to identify acquired antimicrobial resistance –AMR – genes through different databases ResFinder 3.2 (Zankari, 2014)

¹www.novocraft.com

²<https://cge.cbs.dtu.dk/services/CSIphylogeny/>

³<https://www.burkholderia.com/>

and CARD (Jia et al., 2017) and virulence factors via VFAnalyzer from VF-database (Liu et al., 2019). All the pathoadaptive associated genes obtained *in silico* were confirmed by using ABRicate version 0.5⁴.

Ethics Statement

Studies involving the clinical isolates compared in this study were approved by Centro Hospitalar Lisboa Norte (CHLN) ethics committee and the anonymity of the patient is preserved. Informed consent for the use of these isolates in research was obtained from the patient and/or the legal guardians. The patient's isolates and clinical data was collected as part of the epidemiological survey of Bcc bacteria involved in pulmonary infections among the CF patients receiving care followed at Hospital de Santa Maria (Cunha et al., 2003; Correia et al., 2008).

Data and Nucleotide Sequence

Accession Numbers

DNA sequence reads for all isolates obtained by Illumina sequencing are available at the EMBL's European Nucleotide Archive (ENA) under accession number PRJEB20052 and PRJEB35836. The fully assembled and annotated genome of *B. cenocepacia* IST439 obtained by PacBio Sequencing is available at ERZ1345975, **Supplementary Table 1**. The other drafted and annotated genomes of *B. cenocepacia* and *B. multivorans* clonal variants obtained by Illumina Sequencing in the present study are available at **Supplementary Table 1**.

RESULTS

Genomic Analysis of *B. cenocepacia* and *B. multivorans* Ancestor Strains

The first isolates obtained from patient J, *B. cenocepacia* IST439 and *B. multivorans* IST419, were used as the ancestor reference strains for comparative genomic analysis. The genome of *B. cenocepacia* IST439 was sequenced using a combination of PacBio and Illumina sequences, yielding an assembly of three scaffolds corresponding to the three expected replicons (Holden et al., 2009) with a total genome size of 7.63 Mb (**Supplementary Table 2**). Plasmid pBCJ2315, present in the genome of *B. cenocepacia* J2315 strain (Holden et al., 2009), was not found in *B. cenocepacia* IST439 genome. The combination of two Illumina sequencing rounds of *B. multivorans* IST419 allowed the assembly of 76 scaffolds with a predicted genome size of ~ 6.48 Mb (**Supplementary Table 2**) corresponding to the three expected replicons (Komatsu et al., 2003; Holden et al., 2009).

The genome sequences of the ancestor reference strains, *B. cenocepacia* IST439 and *B. multivorans* IST419, were screened for the presence of previously reported Genomic Islands (GIs), BcenGIs and GiST32-s (Holden et al., 2009;

Graindorge et al., 2012; Nunvar et al., 2017). Most of these metabolic GIs were absent from both genomes but resistance GIs (in chromosome II) and pathogenic GI (Prophage BcepMu in chromosome III) were found (**Supplementary Table 3**). These shared GIs contain genes associated with arsenic resistance, antibiotic resistance, inorganic-ion and sulfate transporter, and stress response (Baldwin et al., 2004; Holden et al., 2009) and the *B. cenocepacia* low-oxygen-activated *lxa* locus (Sass et al., 2013). Another shared pathogenic GI, BcepMu prophage, contains genes involved in replication, regulation and pathogenesis (Summer et al., 2004) reported to be overexpressed in response to high doses of exogenous reactive oxygen species (ROS) (Peeters et al., 2010). Fourteen unique GIs were predicted in *B. cenocepacia* IST439 genome (MLST profile ST218) and termed BcenST218_GI1 to BcenST218_GI14 (**Supplementary Figure 1** – panel A and **Supplementary Table 4**) while 17 unique GIs were detected in *B. multivorans* IST419 (MLST profile ST836) and termed BmST836_GI1 to BmST836_GI17 (**Supplementary Figure 1** – panel B and **Supplementary Table 5**). All these unique GIs that were not shared by the *B. cenocepacia* and *B. multivorans* strains include multiple phage-related genes and/or genes associated with metabolism, replication, regulation and pathogenesis, and hypothetical membrane proteins (**Supplementary Tables 4, 5**).

Burkholderia cenocepacia IST439 (ST-218) was found to be closely related to the German *B. cenocepacia* H111 isolate [ST1506, 99.2% average nucleotide identity (ANI)] obtained from sputum of a CF patient (Carrier et al., 2014) and more distantly to the worldwide ET-12 epidemic clones J2315 (99.0% ANI), BC7 (98.9% ANI), and K56-2 (99.0% ANI) (Holden et al., 2009; Varga et al., 2013) (**Supplementary Figure 2** and **Supplementary Table 1**). *B. multivorans* IST419 (ST-836) is more related to the Belgian clone *B. multivorans* ATCC_BAA-247 (ST650, 99.0% ANI) retrieved from sputum of a CF patient (Johnson et al., 2015; Johnson et al., 2016) than to the U.S. soil reference isolate *B. multivorans* ATCC_17616 (ST21, 96.9% ANI) (Komatsu et al., 2003) (**Supplementary Figure 2** and **Supplementary Table 1**).

Comparative Analysis of the Genome Sequences of *B. cenocepacia* and *B. multivorans* Sequential Clonal Isolates

The genomes of the eleven *B. cenocepacia* isolates and of the nine *B. multivorans* isolates were examined to identify mutations arising in the same CF patient during 4.4 years of chronic airway infection (**Figure 1** – panel A). These isolates were sequentially obtained from the onset of the Bcc infection with the *B. multivorans* followed by a co-infection with *B. cenocepacia* for at least three years until the patient's death (Correia et al., 2008; Coutinho et al., 2011; Hassan et al., 2019a). The genome of two colonies of different morphotypes (assigned as A and B) that were isolated at the same isolation date for each species [**Figure 1**, (Coutinho et al., 2011; Hassan et al., 2019a)] were also compared.

All isolates from the same species shared the same MLST type, confirming their clonal relationship (**Supplementary**

⁴<https://github.com/tseemann/abricate>

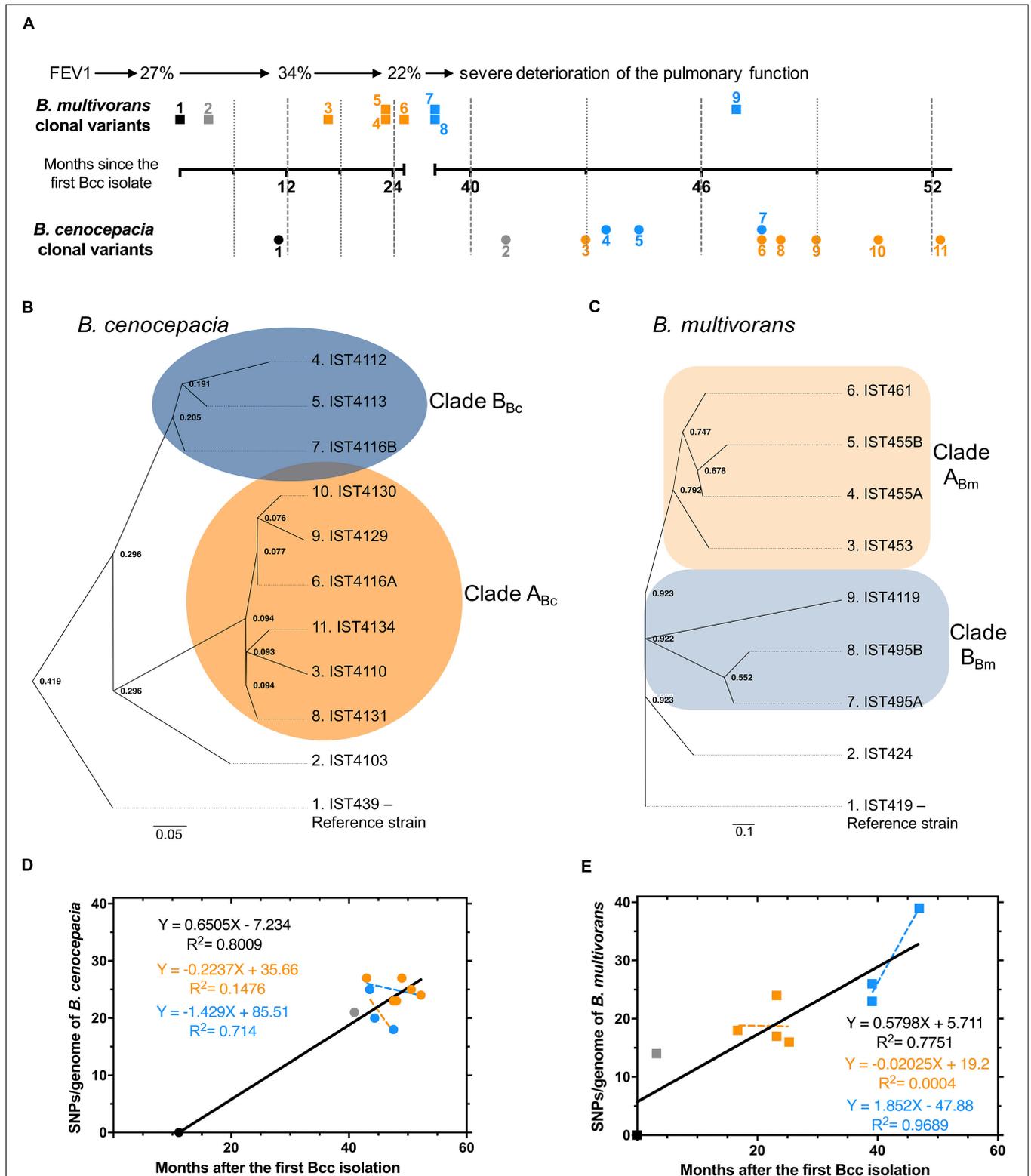


FIGURE 1 | (A) Schematic representation of the *B. cenocepacia* (●) and *B. multivorans* (■) isolates examined with information on lung function (Forced Expiratory Volume in the first second, FEV1%). Time points of isolation, in months after the first Bcc isolate was retrieved, are shown; the numerical order [also used in the phylogenetic trees of *B. cenocepacia* (B) and *B. multivorans* (C)] and the two colors used correlate with the two clades described for *B. cenocepacia* and *B. multivorans* clonal variants. (D and E) represent SNP accumulation during chronic infection and the relation between the time elapsed from the first Bcc isolation to the point of isolation of each isolate. A linear fit with a slope was plotted. Bc and Bm indicates *B. cenocepacia* and *B. multivorans*, respectively.

Table 2). Putative mutations were identified relative to the genomes of *B. cenocepacia* IST439 or *B. multivorans* IST419 (**Figure 1** – panels B and C), revealing 70 or 110 variant positions among *B. cenocepacia* and *B. multivorans* clonal variants, respectively. These positions were used to evaluate their phylogenetic relationships (**Figure 1**), which suggest that the two Bcc populations originated from a single colonization event followed by subsequent diversification leading to two different sub-populations (clades A_{Bc} and A_{Bm} ; B_{Bc} and B_{Bm}) for each Bcc species (**Figure 1** – panels B and C – orange and blue for clade A_{Bc} , A_{Bm} and B_{Bc} , B_{Bm} – respectively). The two *B. cenocepacia* clades (clades A_{Bc} and B_{Bc}) were detected emerging during a common infection period while *B. multivorans* clade B_{Bm} was detected subsequently to clade A_{Bm} (**Figure 1** – panels A and B). The genetic differentiation among isolates did not always follow the chronology of isolation. For instance, isolates IST4110 and IST4134 are closely related although they were isolated years apart, indicating long standing coexistence of the two *B. cenocepacia* subpopulations (**Figure 1** – panels A and B). Moreover, the sole pair of *B. cenocepacia* isolates IST4116A/B obtained at the same isolation date, three years after the isolation of the first *B. cenocepacia* isolate IST439, belongs to distinct lineages (**Figure 1** – panels A and B).

Evolutionary Dynamics of *B. cenocepacia* and *B. multivorans* Co-infecting Populations Inside the CF Lung

A total of 63 single-nucleotide polymorphisms (SNPs) and 9 insertions-deletions (INDELs) were identified among *B. cenocepacia* isolates, and 97 SNPs and 16 INDELs among *B. multivorans* clonal variants (**Supplementary Tables 6, 7**). SNP accumulation rates (calculated as linear regression slope, **Figure 1** – panels D and E) indicate that the mutation rates for *B. cenocepacia* and *B. multivorans* populations were 2.08 and 2.27 SNPs/year, respectively. Most of the mutations, especially those observed along with the clades' emergence and shared among multiple isolates, are non-synonymous (**Tables 2, 3, Supplementary Tables 6, 7**) and affect coding sequences (CDSs). For instance, 22 non-synonymous and 1 synonymous mutations were associated with two clades A_{Bc} and B_{Bc} that emerged within the *B. cenocepacia* population and 35 non-synonymous and 7 synonymous were related to the two clades A_{Bm} and B_{Bm} that emerged in the *B. multivorans* population. Such enrichment of non-synonymous mutations is greater than the reported substitution under neutral evolution given the codon usage and GC content of *B. cenocepacia* bacteria (Dillon et al., 2015). These results suggest that positive selection dominated during the evolution of Bcc in the CF patient lung (**Tables 2, 3 and Supplementary Figure 3**). Similar positive selection has been reported for other Bcc strains evolving in the CF airway, including *B. cenocepacia*, *B. multivorans* and *B. dolosa* species (Lieberman et al., 2011, 2014; Silva et al., 2016; Lee et al., 2017; Nunvar et al., 2017; Diaz Caballero et al., 2018).

Consistent with selection acting on *B. cenocepacia* clonal variants of clade B_{Bc} , only a few genes accumulated mutations

that were not also found in the other clade A_{Bc} (**Table 2 and Supplementary Table 6**). For instance, within clade B_{Bc} , the gene homologous to the IST439_04198 locus, a *penA* homolog from *B. cenocepacia* J2315 that encodes a putative beta-lactamase, was found to acquire a D262G mutation in clones IST4112, IST4113 and IST4116B. This mutation might be a result of the drastic intravenous therapy with gentamicin and ceftazidime and in agreement with the previously reported high resistance to the beta-lactam antibiotic ceftazidime (Coutinho et al., 2011). Also, among clade B_{Bc} variants, one indel was observed (P599_A601deletion) in the IST439_04244 gene, which is homologous to BCAM2211 in *B. cenocepacia* J2315 that is an ortholog of *acoR* from *Bacillus subtilis*. This *acoR* gene encodes a sigma-54 dependent transcriptional regulator that is involved in the regulation of the acetoin catabolic pathway (Huang et al., 1999; Ali et al., 2001). The acetoin catabolic pathway is believed to be a defense mechanism to avoid the lethal acidification produced by CF pathogens during fermentation and anaerobic respiration as a result of the decreased oxygen availability inside the CF lung (Whiteson et al., 2014). Moreover, other two mutations (S382L and V449L) were observed in genes homologous to IST439_00883 and IST439_00884 loci, respectively. These two loci are homologous to *orbA* and *pvdA*, respectively, that encode ornibactin siderophore biosynthesis and transport proteins (Agnoli et al., 2006). A *pvdA* deletion mutant was found to be less virulent than the parent strain in chronic and acute models of respiratory infection (Sokol et al., 1999) and a *B. cenocepacia* K56-2 *orbA* mutant (K56*orbA*:tp, *orbA* allelic exchange mutant) has also an attenuated virulence in the rat chronic respiratory infection model (Sokol et al., 2000). The reported functions of the four genes that were mutated within clade B_{Bc} suggest that positive selection might be driven by well-known environmental factors of the CF lung, in particular high antibiotic concentrations, low oxygen concentrations and low iron concentrations. On the other hand, *B. cenocepacia* clonal variants that form clade A_{Bc} accumulated mutations found in regulatory and metabolism related genes (**Table 2**). This suggests that *B. cenocepacia* tend to adapt within the CF airway by remodeling their metabolism during long term infection with periodic diversification as a function of variation within host environments. In *B. multivorans*, most of the mutated genes also affected metabolic functions (energy production and conservation and regulation of the carbohydrate metabolism) (**Table 3**).

Burkholderia cenocepacia and *B. multivorans* Genes Under Convergent Evolution in the CF Lung

The mutated genes during chronic infection belong to different clusters of orthologous groups – COG (**Supplementary Figure 3 and Tables 2, 3, Supplementary Tables 6, 7**). Genes with well-characterized functions that accumulated non-synonymous mutations (41 and 48 *B. cenocepacia* and *B. multivorans* CDSs, respectively) are schematized in **Supplementary Figure 3** and selected mutations are listed in **Tables 2, 3**.

TABLE 2 | Summary of selected non-synonymous mutations found among *B. cenocepacia* clonal variants associated with phylogenetic clades (clades A_{BC} and B_{BC}, **Figure 1** – panel B, depicted in lighter and darker colors, respectively).

Gene locus	Gene prediction, Annotation	COG category	Mutation effect	Predicted functional effect and impact	presence and type of mutation in each isolate of <i>B. cenocepacia</i>														
					IST4103	IST4110	IST4112	IST4113	IST4116A	IST4116B	IST4131	IST4129	IST4130	IST4134					
Cellular Processes and Signaling																			
IST439_01755	wbil, UDP-N-acetyl-alpha-D-glucosamine C6 dehydratase	Cell wall/membrane/envelope biogenesis	L493P	CDS: missense variant, moderate			SNP	SNP		SNP									
			E487K	CDS: missense variant, moderate		SNP			SNP		SNP	SNP	SNP	SNP					
IST439_06597	<i>fgrL/yycG_3</i> , Sensor histidine kinase, Two-component regulatory system	Signal transduction mechanisms	V176M	CDS: missense variant, moderate		SNP			SNP		SNP			SNP				SNP	
IST439_01475	Putative (ABC transporter) multidrug export ATP-binding membrane protein	Defense mechanisms	Q347*	CDS: stop gained, High	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	
IST439_04198	<i>penA</i> , Beta-lactamase precursor		D262G	CDS: missense variant, moderate			SNP	SNP		SNP									
Information storage and processing																			
IST439_00910	<i>bpeR/envR/tetR</i> , TetR family regulatory protein/HTH-type transcriptional regulator TtgR		C63R	CDS: missense variant, moderate			SNP	SNP		SNP									
			G196D	CDS: missense variant, moderate		SNP			SNP		SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	
IST439_00921	<i>rhsR</i> , Ribose operon repressor	Transcription	V287D	CDS: missense variant, moderate	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	
IST439_01925	<i>oxyR</i> , Hydrogen peroxide-inducible genes activator/oxidative stress regulatory protein		S53R	CDS: missense variant, moderate			SNP	SNP		SNP									
			T669C	CDS: missense variant, moderate	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	
IST439_04244	<i>acoR_2</i> , Acetoin catabolism regulatory protein		P599_A601del	CDS: frameshift variant, high			INDEL	INDEL		INDEL									

(Continued)

TABLE 2 | Continued

Gene locus	Gene prediction, Annotation	COG category	Mutation effect	Predicted functional effect and impact	presence and type of mutation in each isolate of <i>B. cenocepacia</i>										
					IST4103	IST4110	IST4112	IST4113	IST4116A	IST4116B	IST4131	IST4129	IST4130	IST4134	
IST439_05266	<i>cynR-13</i> , LysR family regulatory protein/HTH-type transcriptional regulator CynR		K259E	CDS: missense variant, moderate		SNP				SNP		SNP	SNP	SNP	SNP
IST439_05755	<i>rpoD/sigE</i> , RNA polymerase sigma factor RpoD		R167T	CDS: missense variant, moderate	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP
IST439_06667	<i>yhaO/NOPBF</i> , putative metallophosphoesterase YhaO/putative calcineurin-like phosphoesterase family protein	Replication, recombination and repair	G153W	CDS: missense variant, moderate		SNP			SNP		SNP		SNP	SNP	SNP
Metabolism															
IST439_01911	<i>hmpA</i> , Flavohemoprotein	Energy production and conversion	D337E	CDS: missense variant, moderate		SNP			SNP		SNP	SNP	SNP	SNP	SNP
IST439_02838	2-oxoacid ferredoxin oxidoreductase		V1062L	CDS: missense variant, moderate		SNP			SNP		SNP	SNP	SNP	SNP	SNP
IST439_00975	<i>flhY_1</i> , Cystine-binding periplasmic protein	Amino acid transport and metabolism	D13G	CDS: missense variant, moderate	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP
IST439_01981	<i>artP/glnQ</i> , Arginine transporter ATP-binding subunit/Glutamine transport ATP-binding protein		Q136*	CDS: stop gained, High	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP
IST439_05703	<i>BceL/ynfM_4</i> , major facilitator superfamily protein/Inner membrane transport protein YnfM	Carbohydrate transport and metabolism	T117A	CDS: missense variant, moderate		SNP			SNP		SNP	SNP	SNP	SNP	SNP
IST439_00216	<i>sulP/dauA/cynT</i> , putative sulfate transporter/C4-dicarboxylic acid transporter DauA	Inorganic ion transport and metabolism	L61P	CDS: missense variant, moderate		SNP			SNP		SNP	SNP	SNP	SNP	SNP
IST439_00883	<i>orbA</i> , Ornibactin receptor/Ferrichrome-iron receptor precursor		S382L	CDS: missense variant, moderate			SNP	SNP		SNP					
IST439_00884	<i>pvdA</i> , L-ornithine 5-monooxygenase	Secondary metabolites biosynthesis, transport, and catabolism	V449L	CDS: missense variant, moderate			SNP	SNP		SNP					
IST439_01356	<i>mdcE</i> , Malonate decarboxylase gamma subunit/Malonyl-S-ACP:biotin-protein carboxyltransferase MADD		D67E	CDS: missense variant, moderate		SNP			SNP		SNP	SNP	SNP	SNP	SNP
IST439_04949	<i>pcaC</i> , 4-Carboxymuconolactone decarboxylase family protein		S52R	CDS: missense variant, moderate	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP	SNP

Gene loci are ordered based on COG annotation and its categories. More details in **Supplementary Table 6**.

TABLE 3 | Summary of selected non-synonymous mutations found among *B. multivorans* clonal variants associated with phylogenetic clades (clades A_{Bm} and B_{Bm} , **Figure 1** – panel C, depicted in lighter and darker colors, respectively).

Gene locus	Gene prediction, Annotation	COG category	Mutation effect	Predicted functional effect and impact	presence and type of mutation in each isolate of <i>B. multivorans</i>									
					IST424	IST453	IST455A	IST455B	IST461	IST495A	IST495B	IST4119		
Cellular Processes and Signaling														
IST419_02260	<i>wbil</i> , UDP-N-acetyl-alpha-D-glucosamine C6 dehydratase	Cell wall/membrane/envelope biogenesis	M69T (+64bp)	CDS: frameshift variant, high		INDEL	INDEL	INDEL	INDEL					
IST419_02356	<i>rmlB/rflB</i> , dTDP-glucose 4,6-dehydratase		T192H	CDS: missense variant, moderate						SNP	SNP		SNP	
IST419_03735	patatin/ Putative patatin-like phospholipase		Q196*	CDS: stop gained, High						SNP	SNP			
IST419_00471	<i>flgD</i> , flagellar basal body rod modification protein	Cell motility	S219A	CDS: missense variant, moderate						SNP	SNP			
IST419_00702	<i>hprK</i> , HPr kinase/phosphorylase	Signal transduction mechanisms	S168N	CDS: missense variant, moderate		SNP	SNP	SNP	SNP					
IST419_05375	<i>fgrL/ yycG_3</i> , Sensor histidine kinase, Two-component regulatory system		L181F							SNP	SNP			
IST419_05223	<i>fhaB/ palA/ hylA</i> , filamentous hemagglutinin outer membrane protein/hemolysin	Intracellular trafficking, secretion, and vesicular transport	P115L	CDS: missense variant, moderate			SNP	SNP	SNP					
Information storage and processing														
IST419_01084	<i>rsuA/ rluF</i> , pseudouridine synthase, Rsu/ RNA-binding S4 domain-containing protein	Translation, ribosomal structure and biogenesis	V285A	CDS: missense variant, moderate	SNP	SNP	SNP	SNP	SNP					
Metabolism														
IST419_02576	molybdopterin oxidoreductase	Energy production and conversion	K360Q	CDS: missense variant, moderate		SNP	SNP	SNP	SNP					
IST419_05773	<i>tfdB</i> , monooxygenase FAD-binding protein		L190V	CDS: missense variant, moderate		SNP	SNP	SNP	SNP					
IST419_00066	<i>livF</i> , ABC transporter-like protein	Amino acid transport and metabolism	V185L	CDS: missense variant, moderate						SNP	SNP			

(Continued)

TABLE 3 | Continued

Gene locus	Gene prediction, Annotation	COG category	Mutation effect	Predicted functional effect and impact	presence and type of mutation in each isolate of <i>B. multivorans</i>							
					IST424	IST453	IST455A	IST455B	IST461	IST495A	IST495B	IST4119
IST419_00071	<i>livK</i> , extracellular ligand-binding receptor		Q139R	CDS: missense variant, moderate						SNP	SNP	
IST419_01537	<i>tcyB/ yecS</i> , polar amino acid ABC transporter inner membrane subunit		S165_1167dup	CDS: frameshift variant, high						INDEL	INDEL	
IST419_02455	<i>gltL/ aatP</i> , ABC transporter-like protein		D168N	CDS: missense variant, moderate		SNP	SNP	SNP	SNP			
IST419_03581	<i>soxA</i> , sarcosine oxidase subunit alpha		W385R	CDS: missense variant, moderate						SNP	SNP	
IST419_02163	<i>relA/ spoT</i> , (p)ppGpp synthetase I SpoT/ RelA	Nucleotide transport and metabolism	I493M	CDS: missense variant, moderate						SNP	SNP	
IST419_02029	<i>fabF2</i> , 3-oxoacyl-(acyl carrier protein) synthase II	Lipid transport and metabolism	G145A	CDS: missense variant, moderate						SNP	SNP	SNP
			F74T	CDS: missense variant, moderate		SNP	SNP	SNP	SNP			
IST419_05562	acetyl-CoA synthetase		A257T	CDS: missense variant, moderate	SNP		SNP	SNP				
			E484Q	CDS: missense variant, moderate		SNP			SNP	SNP	SNP	
IST419_04557	<i>pepM</i> , phosphoenolpyruvate phosphomutase	Carbohydrate transport and metabolism	D448E	CDS: missense variant, moderate						SNP	SNP	
IST419_04936	malto-oligosyltrehalose trehalohydrolase		D274N	CDS: missense variant, moderate						SNP	SNP	
IST419_03373	hypothetical methyltransferase protein	Secondary metabolites biosynthesis, transport, and catabolism	H225T	CDS: missense variant, moderate		SNP	SNP	SNP	SNP			
IST419_05081	PHB depolymerase family esterase		A272S	CDS: missense variant, moderate			SNP	SNP	SNP			

Gene loci are ordered based on COG annotation and its categories. More details in **Supplementary Table 7**.

The genes most likely under selection in the CF airway are those with non-synonymous mutations that become fixed within all subsequent samples. One of these is the *B. cenocepacia* gene linked to defense mechanisms against antibiotic stress (IST439_01475, encoding a putative ABC transporter ATP-binding membrane protein; homologous to BCAL1039 from *B. cenocepacia* J2315) (Table 4). Three other mutations affected genes involved in transcription regulation, encoding the ribose operon repressor RbsR and two sigma factor cytoplasmic proteins homologous to *B. cenocepacia* J2315 RpoC and RpoD proteins. Three mutations were also detected in genes associated with metabolism (BCAL1610 and *glnQ* homologs – encoding periplasmic cysteine-binding protein and glutamine ABC transporter ATP-binding protein – and *pcaC* homolog encoding 4-carboxymuconolactone decarboxylase, involved in the metabolism of amino acids and secondary metabolites, respectively). On the other hand, no single non-synonymous mutation became fixed among *B. multivorans* isolates during the chronic co-infection (Tables 3 and Supplementary Table 7). Although the simultaneous, apparent fixation of seven new mutations over 18 months may seem high, manual inspection of the evidence supporting these mutations indicates they are genuine. It is possible that these mutations may have accumulated in a lineage evolving over a longer period of time prior to the first sampling, when clones like IST439 and more evolved lineages co-occurred.

Among the mutated genes, two genes in *B. cenocepacia* and three genes in *B. multivorans* acquired multiple independent mutations, which provide strong evidence of selection (Figure 2). Although multiple mutations in a single gene can result from mismatched reads and/or hidden paralogy, the evidence for these mutations is robust following manual inspection. Notably, in *B. cenocepacia*, the gene IST439_00910 that is homologous to *B. cenocepacia* J2315 *bpeR* acquired three different mutations during co-infection; one mutation (G196D) became fixed in clade *A_{Bc}* variants while mutation C63R emerged in clade *B_{Bc}* and one INDEL (K113* [-57bp]) was only found in the second sequential clonal variant isolated, IST4103, the isolate that does not belong to any of the two defined clades (Figure 1 – panel B). The *bpeR* gene encodes a TetR family regulatory protein which is a conserved transcription regulator across the *Burkholderia* genus (Ramos et al., 2005). This gene was involved in the transcriptional control of multidrug efflux pumps, antibiotic biosynthesis, response to osmotic stress and toxic chemicals, and in pathogenicity in other Gram-negative pathogens (Ramos et al., 2005). The IST439_00910/*bpeR* is also an ortholog to *B. dolosa* AU0158 BDAG_00732 that was found before to acquire several mutations in the promotor region of 112 CF-clinical isolates (Lieberman et al., 2014). Among the three polymorphic genes in *B. multivorans* clonal variants, one polymorphic gene (IST419_01084) accumulated six different mutations and one of these variant positions (V285A) became fixed in the first 6 clonal variants (Figure 2 – panel B). This IST419_01084 gene encodes a putative RNA pseudouridine synthase whose structure contains a small RNA-binding protein domain (S4 domain) that delivers nucleotide-modifying enzymes to RNA, regulating translation

through structure specific RNA binding (Aravind and Koonin, 1999; Del Campo et al., 2001).

Orthologous Genes Experienced Positive Selection in Co-evolved *B. cenocepacia* and *B. multivorans* Populations

Two orthologous genes accumulated mutations related to lineage diversification in both *B. cenocepacia* and *B. multivorans* (Figure 3). These genes are homologs of *B. cenocepacia* J2315 *wbiI* and BCAS0619 and are predicted to encode a nucleotide sugar dehydratase involved in lipopolysaccharide O-antigen (OAg) biosynthesis (Ortega et al., 2005) and a putative two-component regulatory sensor kinase protein, respectively.

Regarding *wbiI*, E487K and L493P mutations were observed in IST439_01755 gene in all *B. cenocepacia* clonal variants forming clades *A_{Bc}* and *B_{Bc}*, respectively, while the M69T (+ 64bp) insertion was observed in IST419_02260 gene in all *B. multivorans* clonal variants forming clade *A_{Bm}*. Regardless of the Bcc species, all the isolates with a mutated *wbiI* gene expressed a lipopolysaccharide lacking the OAg (Hassan et al., 2019a). Remarkably, this *wbiI* gene was reported before to be under strong selection in *B. cenocepacia*, *B. multivorans* and *B. dolosa* during chronic lung infection (Lieberman et al., 2011, 2014; Silva et al., 2016; Hassan et al., 2017). Concerning BCAS0619, located at the third chromosome, mutation V176M was found within *B. cenocepacia* variants that form clade *A_{Bc}*, except for the IST4129 isolate that lacks chromosome III, while IST419_05375 gene has a L181F mutation in two *B. multivorans* variants (IST495A and IST495B).

DISCUSSION

This study was designed to gain insights into the molecular mechanisms underlying genomic diversification occurring in *B. cenocepacia* and *B. multivorans* strains that co-infected, for a period of at least 3 years, the same cystic fibrosis (CF) patient and to identify shared and relevant pathoadaptive mechanisms. Mutations with accumulation rates of 2.08 (*B. cenocepacia*) and 2.27 (*B. multivorans*) SNPs/year were found during the course of co-infection comparable with those reported for *B. dolosa* [2.1 SNPs/year] (Lieberman et al., 2011, 2014) and *B. multivorans* [2.4 SNPs/year] (Silva et al., 2016) and slightly higher than the mutation rate reported for the Czech epidemic *B. cenocepacia* ST32 strain [average rate 1.66 SNPs/year] (Nunvar et al., 2017). Such mutation rates are also comparable to those occurring in the major CF pathogen *P. aeruginosa* [2.7 SNPs/year] during CF chronic infection (Markussen et al., 2014; Marvig et al., 2015) but below the mutation rates reported for the non-Bcc member and etiological agent of melioidosis, *B. pseudomallei*, obtained from chronically infected CF patients [3.6 SNPs/year] (Viberg et al., 2017). The aforementioned mutation rates are believed to be associated with fast diversification of the infecting bacterial strains (Didelot et al., 2016) diverging into sublineages with their own functional and genomic signatures and rates of adaptation to different CF-lung environments (Markussen et al., 2014; Didelot et al., 2016).

TABLE 4 | List of non-synonymous mutations (remarked with ✓) that became fixed among the clonal variants of *B. cenocepacia* examined.

Gene locus/ homologous gene in J2315	Annotation	Mutational effect	COG	IST439	IST4103	IST4110	IST4112	IST4113	IST4116A	IST4116B	IST4131	IST4129	IST4130	IST4134
IST439_01475/ BCAL1039 homolog	Putative (ABC transporter) multidrug export ATP-binding/permease membrane protein	Q347*	V		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
IST439_00921/ <i>rbsR</i> homolog	Ribose operon repressor	V287D	K		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
IST439_02295/ <i>rpoC</i> homolog	DNA-directed RNA polymerase subunit beta	T669C	K		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
IST439_05755/ <i>rpoD/sigE</i> homolog	RpoD/SigE, RNA polymerase sigma factor RpoD	R167T	K		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
IST439_00975/ BCAL1610 homolog	FliY_1, Cystine-binding periplasmic protein	D13G	E		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
IST439_01981/ <i>glnQ</i> homolog	Glutamine transport ATP-binding protein GlnQ/ABC transporter-like protein	Q136*	E		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
IST439_4949/ <i>pcaC</i> homolog	PcaC, 4-Carboxymuconolactone decarboxylase family protein	S52R	Q		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓

The only mutations presented are those acquired by CDS associated with well-characterized functions as shown in **Supplementary Table 6**.

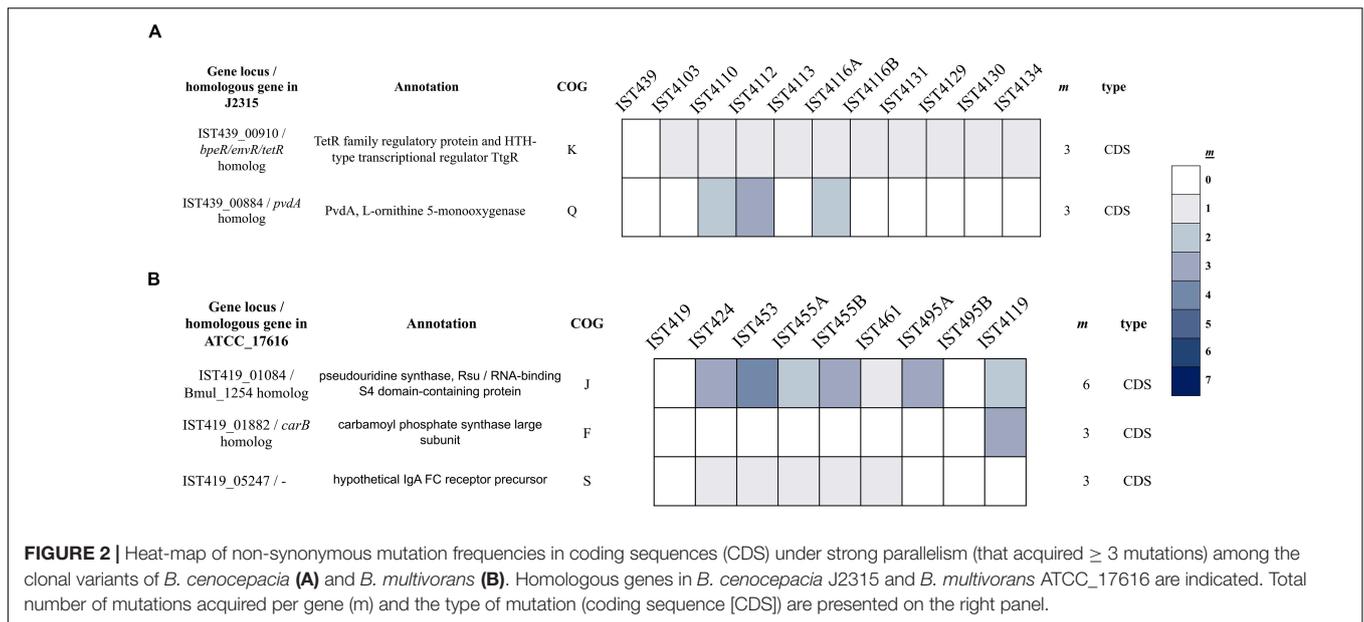


FIGURE 2 | Heat-map of non-synonymous mutation frequencies in coding sequences (CDS) under strong parallelism (that acquired ≥ 3 mutations) among the clonal variants of *B. cenocepacia* (A) and *B. multivorans* (B). Homologous genes in *B. cenocepacia* J2315 and *B. multivorans* ATCC_17616 are indicated. Total number of mutations acquired per gene (m) and the type of mutation (coding sequence [CDS]) are presented on the right panel.

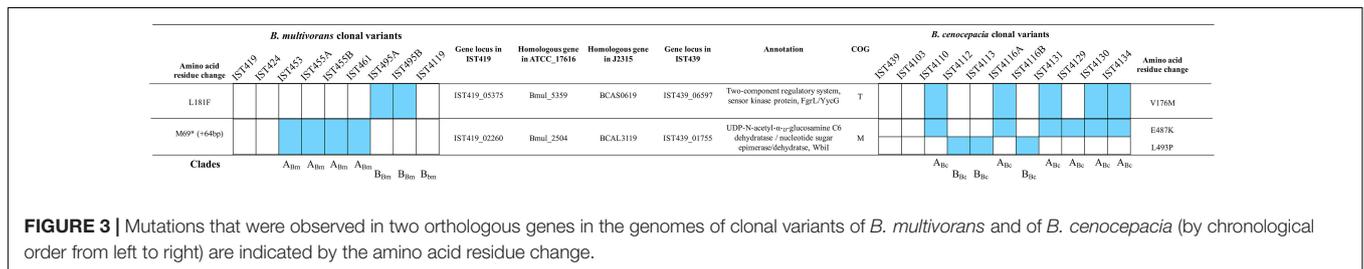


FIGURE 3 | Mutations that were observed in two orthologous genes in the genomes of clonal variants of *B. multivorans* and of *B. cenocepacia* (by chronological order from left to right) are indicated by the amino acid residue change.

In the co-infecting *B. cenocepacia* and *B. multivorans* populations examined, several genes were found to acquire different mutations either related or unrelated to lineage diversification. Remarkably, most of these genes have not been previously reported as being involved in parallel evolution in *B. dolosa*, *B. cenocepacia*, and *B. multivorans* (Lieberman et al., 2011, 2014; Silva et al., 2016; Nunvar et al., 2017). Specifically, most of the mutated genes found in this work are involved in cell envelope/wall/membrane biogenesis and in regulatory and metabolic processes, belonging to the functional categories reported before for Bcc strains as being under selective pressure in the CF lungs (Lieberman et al., 2011, 2014; Silva et al., 2016; Nunvar et al., 2017). Mutated genes with fixed non-synonymous variant positions over the entire period of chronic infection were found in the present work among the *B. cenocepacia* clonal variants examined but no single fixed non-synonymous variant position was found for *B. multivorans*. Results suggest that *B. cenocepacia* and *B. multivorans* infection involved periods of diversification dominated by positive selection, as previously suggested by others (Silva et al., 2016; Lee et al., 2017; Nunvar et al., 2017). Although we cannot discard the possibility that the spatiotemporal characteristics of the isolated clonal variants examined in this work are not representative of the population heterogeneity present at each isolation time in the CF airways, the distinct evolutionary patterns here attributed to *B. cenocepacia*

and *B. multivorans* when co-infecting the same CF environment for several years, appear to be consistent and in line with the literature. Population heterogeneity in the CF lungs was observed for both species and different colony morphotypes were isolated at the same isolation dates (Figure 1 – panels A, B and C). However, the very limited number of isolates obtained at the same isolation date and the fact that they were not always isolated randomly but to obtain different colony morphotypes limits our inferences of population dynamics. Nevertheless, heterogeneity within the *B. cenocepacia* population can be seen in the two *B. cenocepacia* isolates (IST4116A/B) obtained on the same day, which belong to different lineages (clades A_{Bc} and B_{Bc}).

Among the mutated genes in *B. cenocepacia* and *B. multivorans* during long-term co-infection of the CF patient are genes encoding oxidative stress regulatory proteins and heavy metal-sensing proteins (Tables 2, 3 and Supplementary Tables 6, 7). Such regulatory proteins (e.g., OxyR and FixK) are involved in bacterial protection against reactive oxygen species (ROS) produced by phagocytic leukocytes (Reyrat et al., 1993; Saito et al., 2003; Bobik et al., 2006; Panmanee et al., 2008). Consistent with this conclusion, the host immune system was proposed to play a fundamental role in driving *B. cenocepacia* and *B. multivorans* evolution during chronic CF infection and genes linked to the oxidative stress responses (homologous to *B. cenocepacia* J2315 *katG*, *yedY*, *moeA1*, *fixL*, *osmC* and *ompR*)

reported to suffer mutation (Silva et al., 2016; Nunvar et al., 2017). This selection driven by the host immune system of genes linked to the oxidative stress and heavy metal-sensing was not however proposed for within-patient evolution of either *B. dolosa* or *P. aeruginosa* (Lieberman et al., 2011, 2014; Winstanley et al., 2016).

Increased resistance to antibiotics is also a well-described evolutionary trait associated with many CF pathogens and genes related with antibiotic resistance have been among the genes mostly reported to be under selective pressure (Jeannot et al., 2008; Lieberman et al., 2011, 2014; Silva et al., 2016; Winstanley et al., 2016; Nunvar et al., 2017). Consistent with those reports, some of the mutated genes found in the present study in *B. cenocepacia* and *B. multivorans* are involved in defense mechanisms and their regulation, as it is the case of genes encoding PenA, EmrA, ABC-2 type transporter, and DNA gyrase proteins (Tables 2, 3 and Supplementary Tables 6, 7).

Another commonly reported evolved trait during chronic infection is mutation altering genes of the OAg biosynthetic cluster (Lieberman et al., 2011, 2014; Maldonado et al., 2016; Silva et al., 2016; Hassan et al., 2017; Nunvar et al., 2017). Longitudinal comparative genomic studies of Bcc isolates from CF patients with chronic lung infection revealed the conversion from smooth (in the early isolates) to rough LPS with no OAg at the late-stages of infections by *B. cenocepacia* and *B. multivorans* (Silva et al., 2016; Hassan et al., 2017). Remarkably, an extensive study focused on *B. dolosa* during chronic lung infections has shown that late isolates produce an LPS exhibiting the OAg that was absent from the LPS of the initial infecting isolate (Lieberman et al., 2011). The phenotypic switch of OAg presence/absence was recently examined among CF isolates of several Bcc species. In the most prevalent and feared species, *B. cenocepacia* and *B. multivorans*, a marked tendency to lose the OAg along chronic infection was seen, whereas in *B. dolosa* the OAg-chain was absent from the beginning of the 5.5-year infection until the patient death (Hassan et al., 2019a). It remains important to understand why LPS is under selection in *B. cenocepacia* or *B. multivorans*. While the most common explanation invokes selection for immune evasion, it is also possible that LPS modifications are adaptations to biofilm growth (Traverse et al., 2013) or antibiotic resistance (Scribner et al., 2020). Another explanation for loss of OAg is that it increases Bcc survival in phagocytic eukaryotic cells such as amoebae, epithelial cells and human macrophages (Saldias et al., 2009; Saldias and Valvano, 2009; Kotrange et al., 2011; Maldonado et al., 2016). It is also known that the OAg absence leads to increased internalization of *B. cenocepacia* into macrophages upon phagocytosis (Saldias et al., 2009; Kotrange et al., 2011) and to facilitate *B. multivorans* growth inside macrophages (Schmerk and Valvano, 2013). Moreover, the OAg presence was found to interfere with the adhesion of *B. cenocepacia* to bronchial epithelial cells (Saldias et al., 2009) although it was reported to increase the ability of *B. cenocepacia* to adhere to atomic force microscope Si₃N₄ tip (Hassan et al., 2019b). Collectively, the loss of the OAg seems advantageous for disease development associated with *B. cenocepacia* or *B. multivorans* infections. Different mutations were accumulated in relation to

the lineage emergence in the two co-infecting populations in genes IST439_01755 and IST419_02260 (homolog and ortholog to *wbil*, respectively – Figure 3) that encode a nucleotide sugar dehydratase, previously found to be under parallel evolution in *B. cenocepacia* or in *B. multivorans* (Silva et al., 2016; Hassan et al., 2017). Remarkably, the comparison of the ability of three of the *B. cenocepacia* isolates examined in the present work to subvert the host's immune function, assessed by internalization assays using human dendritic cells, showed that the late variants, IST4113 and IST4134, were significantly more internalized exhibiting increased survival within dendritic cells than the early isolate IST439 (Cabral et al., 2017) the sole with an LPS with OAg (Hassan et al., 2017). Altogether, mutated genes' functions involved in the observed adaptive evolution of the two species co-inhabiting the same host-selective environment are consistent with the idea that positive selection in *B. cenocepacia* and *B. multivorans* might be driven by the action of the host immune system.

Among the few mutated regulatory proteins identified in this study is the histidine kinase protein from *B. cenocepacia* IST439_06759 (an *hisS* homolog) and from *B. multivorans* IST419_05375 that enable bacterial cells to sense and adapt to oxidative- and heavy metal- inducing stresses (Wolanin et al., 2002; Skerker et al., 2005). The orthologous genes encoding the histidine kinase protein acquired two mutations, V176M and L181F, in the two co-infecting *B. cenocepacia* and *B. multivorans* populations, respectively (Figure 3). This is the first time that this transcription regulator is reported to undergo adaptive evolution during CF patient infection. Signal transduction histidine kinase-related proteins are very important regulatory proteins involved in signaling cascades contributing to several simultaneous coordinated responses (Wolanin et al., 2002; Skerker et al., 2005) to, for example, osmoregulation (Tomomori et al., 1999) and chemotaxis (Bilwes et al., 1999) in bacterial pathogens.

Burkholderia cenocepacia genes encoding oxidative stress- and heavy metal-sensing- proteins were hypothesized as being the targets of convergent evolution underlying higher affinity to intra-macrophage persistence and the development of the fatal cepacia syndrome (Lee et al., 2017; Nunvar et al., 2017). Although the mechanisms underlying the fatal cepacia syndrome still remains largely unknown, the results obtained in this study are in line with those described by Nunvar et al. (2017) who hypothesized that macrophages, whose bactericidal activity relies on hydrogen peroxide and copper, are putative key players behind the development of the cepacia syndrome.

In summary, the comparative genomic analysis performed in this work identified genes which are under strong positive selection in the most prevalent and most feared Bcc species that co-infected the lung of a CF patient who died from cepacia syndrome. Among them, polymorphic genes and others involved mainly in regulatory and metabolic functions were sorted out. Since cystic fibrosis is a genetic disorder associated with inflammation, sub-optimal antioxidant protection and the continuous use of antimicrobial therapy, all resulting in marked oxidative stress (Galli et al., 2012), it is likely that in the evolved strains several physiological processes might

be activated via the global stress responses which in turn can modulate the course of infection. Global stress responses were found to be under parallel evolution in experimentally evolved *P. aeruginosa* biofilms (Winstanley et al., 2016) and were recently observed during chronic lung infections, irrespective to the presence/absence of other members of the CF microbiome (Vandeplassche et al., 2019). Given that both co-infecting Bcc species are known for their ability of intracellular survival inside macrophages (Saldias et al., 2009; Saldias and Valvano, 2009; Kotrange et al., 2011; Schmerk and Valvano, 2013), the mutated genes associated with cell envelope biogenesis, including those responsible for lipopolysaccharide OAg biosynthesis, could promote Bcc persistence through intracellular survival.

This study provides useful information to better understand the shared and species-specific evolutionary patterns of *B. cenocepacia* and *B. multivorans* evolving in the same CF lung environment during co-infection. The two orthologous genes shared by *B. cenocepacia* and *B. multivorans* that accumulate mutations associated with lineage diversification strongly suggest that two traits, lipopolysaccharide O-antigen (OAg) biosynthesis and the sensing and response to oxidative- and heavy metal-induced stresses, are biological processes under strong selection during evolution in the CF lung environment.

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

SS and VC carried out the first comparative analysis of *B. cenocepacia* genome sequences and AH completed this analysis and performed the genome analysis of *B. multivorans* isolates in collaboration with VC. IS-C designed and coordinated the study and wrote the manuscript with contributions from AH and VC. All authors contributed to the article and approved the submitted version.

FUNDING

Funding received by iBB—Institute for Bioengineering and Biosciences from the Portuguese Foundation for Science and Technology (FCT) (UIDB/04565/2020) is acknowledged. This work was also supported by FCT through a Ph.D. fellowship to AH (FCT-Ph.D. Programme BIOTECnico—PD/BD/113629/2015) and by NIH R01GM110444 to VSC.

ACKNOWLEDGMENTS

The authors would like to thank Carla Coutinho, Robert Sebra and Marcus Dillon for their assistance with

genomic DNA extraction and PacBio sequencing and assembly.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2020.574626/full#supplementary-material>

Supplementary Figure 1 | Genomic islands (A) BcenST281_GIs and (B) BmST836_GIs in the reference strains, *B. cenocepacia* IST439 and *B. multivorans* IST419, respectively. The inner circles denote genomes of the reference isolates and the corresponding clonal variants, ordered as indicated in the legend; Bc – *B. cenocepacia* isolates and Bm – *B. multivorans* strains. Blue and orange colors denote different clades as indicated in **Figure 1**. All clonal variants were chronologically ordered based on the isolation date. Visualizations were carried out by BRIG (Alikhan et al., 2011). For more details, see **Supplementary Tables 4, 5**.

Supplementary Figure 2 | Phylogenetic tree based on the comparison of the whole-genome of *B. cenocepacia* and *B. multivorans* clonal variants and other sequenced strains available at *Burkholderia* genome database (**Supplementary Table 1**). ANI percentages are also presented relative to each studied reference strain, *B. cenocepacia* IST439 and *B. multivorans* IST419.

Supplementary Figure S3 | 3D plot of non-synonymous mutations observed in *B. cenocepacia* (A) and *B. multivorans* (B) clonal variants during the chronic co-infection. All mutated genes of well-characterized functions are depicted as green oval shapes (with SNPs) and/or blue shapes (with INDELS). The number of mutated genes per each COG category is illustrated above each corresponding shape as a ratio (number of genes per isolate/number of all mutated genes in the entire population for this particular COG). Accumulation of small green shapes indicates the number of SNP mutations that was observed per isolate. For a detailed list of mutations, see **Supplementary Tables 6, 7** for *B. cenocepacia* and *B. multivorans*, respectively. Abbreviation letters representing the COGs are detailed in the bottom panel.

Supplementary Table S1 | List of the whole genome sequences of Bcc isolates used in the present study for comparative genomic analyses. The isolation year and country are indicated as well as the multi-locus sequence typing (MLST). CS – cenocepacia syndrome.

Supplementary Table S2 | Metadata information of the studied clonal variants sequentially ordered based on the isolation date; details on genomic assembly and functional annotation are shown as well as the *in silico* MLST profiles.

Supplementary Table S3 | Published genomic islands (Holden et al., 2009; Graindorge et al., 2012; Nunvar et al., 2017) present in *B. cenocepacia* J2315 and ST32 and their eventual presence in *B. cenocepacia* IST439 and *B. multivorans* IST419.

Supplementary Table S4 | BcenST218_GIs as Genomic islands of *B. cenocepacia* IST439 and their presence in *B. multivorans* IST419, *B. cenocepacia* J2315 and ST32, *B. multivorans* ATCC_17616 and ATCC_BAA-247.

Supplementary Table S5 | BmST836_GIs as Genomic islands of *B. multivorans* IST419 and their presence in *B. cenocepacia* IST439, *B. cenocepacia* J2315 and ST32, *B. multivorans* ATCC_17616 and ATCC_BAA-247.

Supplementary Table S6 | List of mutations found among *B. cenocepacia* clonal variants. All mutated genes were grouped to predicted gene functions (COG). Isolates are chronologically ordered as indicated in **Figure 1**. A – GATK pipeline and B – SAMtools/BCFtools toolbox.

Supplementary Table S7 | List of mutations found among *B. multivorans* clonal variants. All mutated genes were grouped to predicted gene functions (COG). Isolates are chronologically ordered as indicated in **Figure 1**. A – GATK pipeline and B – SAMtools/BCFtools toolbox.

REFERENCES

- Agnoli, K., Lowe, C. A., Farmer, K. L., Husnain, S. I., and Thomas, M. S. (2006). The ornibactin biosynthesis and transport genes of *Burkholderia cenocepacia* are regulated by an extracytoplasmic function sigma factor which is a part of the Fur regulon. *J. Bacteriol.* 188, 3631–3644. doi: 10.1128/jb.188.10.3631-3644.2006
- Ali, N. O., Bignon, J., Rapoport, G., and Debarbouille, M. (2001). Regulation of the acetoin catabolic pathway is controlled by sigma L in *Bacillus subtilis*. *J. Bacteriol.* 183, 2497–2504. doi: 10.1128/JB.183.8.2497-2504.2001
- Alikhan, N. F., Petty, N. K., Ben Zakour, N. L., and Beatson, S. A. (2011). BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BMC Genomics* 12:402. doi: 10.1186/1471-2164-12-402
- Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1016/S0022-2836(05)80360-2
- Aravind, L., and Koonin, E. V. (1999). Novel predicted RNA-binding domains associated with the translation machinery. *J. Mol. Evol.* 48, 291–302.
- Bach, E., Sant'Anna, F. H., Magrich Dos Passos, J. F., Balsanelli, E., de Baura, V. A., Pedrosa, F. O., et al. (2017). Detection of misidentifications of species from the *Burkholderia cepacia* complex and description of a new member, the soil bacterium *Burkholderia catarinensis* sp. nov. *Pathogens Dis.* 75:ftx076. doi: 10.1093/femspd/ftx076
- Baldwin, A., Sokol, P. A., Parkhill, J., and Mahenthalingam, E. (2004). The *Burkholderia cepacia* epidemic strain marker is part of a novel genomic island encoding both virulence and metabolism-associated genes in *Burkholderia cenocepacia*. *Infect. Immun.* 72, 1537–1547. doi: 10.1128/iai.72.3.1537-1547.2004
- Bankevich, A., Nurk, S., Antipov, D., Gurevich, A. A., Dvorkin, M., Kulikov, A. S., et al. (2012). SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J. Comp. Biol.* 19, 455–477. doi: 10.1089/cmb.2012.0021
- Barrick, J. E., Colburn, G., Deatherage, D. E., Traverse, C. C., Strand, M. D., Borges, J. J., et al. (2014). Identifying structural variation in haploid microbial genomes from short-read resequencing data using breseq. *BMC Genomics* 15:1039. doi: 10.1186/1471-2164-15-1039
- Baym, M., Kryazhimskiy, S., Lieberman, T. D., Chung, H., Desai, M. M., and Kishony, R. (2015). Inexpensive multiplexed library preparation for megabase-sized genomes. *PLoS One* 10:e0128036. doi: 10.1371/journal.pone.0128036
- Bertelli, C., Laird, M. R., Williams, K. P., Simon Fraser University Research, Computing, Group, Lau, B. Y., et al. (2017). IslandViewer 4: expanded prediction of genomic islands for larger-scale datasets. *Nucleic Acids Res.* 45, W30–W35. doi: 10.1093/nar/gkx343
- Bilwes, A. M., Alex, L. A., Crane, B. R., and Simon, M. I. (1999). Structure of CheA, a signal-transducing histidine kinase. *Cell* 96, 131–141. doi: 10.1016/S0092-8674(00)80966-6
- Bobik, C., Meilhoc, E., and Batut, J. (2006). FixJ: a major regulator of the oxygen limitation response and late symbiotic functions of *Sinorhizobium meliloti*. *J. Bacteriol.* 188, 4890–4902. doi: 10.1128/JB.00251-06
- Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D., and Pirovano, W. (2011). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27, 578–579. doi: 10.1093/bioinformatics/btq683
- Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170
- Cabral, M., Pereira, M., Silva, Z., Iria, I., Coutinho, C., Lopes, A., et al. (2017). Using dendritic cells to evaluate how *Burkholderia cenocepacia* clonal isolates from a chronically infected cystic fibrosis patient subvert immune functions. *Med. Microbiol. Immunol.* 206, 111–123. doi: 10.1007/s00430-016-0488-4
- Carlier, A., Agnoli, K., Pessi, G., Suppiger, A., Jenul, C., Schmid, N., et al. (2014). Genome sequence of *Burkholderia cenocepacia* H111, a cystic fibrosis airway isolate. *Genome Announc.* 2:e00298-14. doi: 10.1128/genomeA.00298-14
- Carver, T. J., Rutherford, K. M., Berriman, M., Rajandream, M. A., Barrell, B. G., and Parkhill, J. (2005). ACT: the artemis comparison tool. *Bioinformatics* 21, 3422–3423. doi: 10.1093/bioinformatics/bti553
- Chin, C. S., Alexander, D. H., Marks, P., Klammer, A. A., Drake, J., Heiner, C., et al. (2013). Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nature Methods* 10, 563–569. doi: 10.1038/nmeth.2474
- Cingolani, P., Platts, A., Wang le, L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly (Austin)* 6, 80–92. doi: 10.4161/fly.19695
- Correia, S., Nascimento, C., Pereira, L., Cunha, M. V., Sá-Correia, I., and Barreto, C. (2008). Infecção respiratória por bactérias do complexo *cepacia*: evolução clínica em doentes com fibrose quística. *Rev. Portuguesa Pneumol.* 14, 5–26.
- Coutinho, C. P., de Carvalho, C. C., Madeira, A., Pinto-de-Oliveira, A., and Sa-Correia, I. (2011). *Burkholderia cenocepacia* phenotypic clonal variation during a 3.5-year colonization in the lungs of a cystic fibrosis patient. *Infect. Immun.* 79, 2950–2960. doi: 10.1128/IAI.01366-10
- Cunha, M. V., Leitão, J. H., Mahenthalingam, E., Vandamme, P., Lito, L., Barreto, C., et al. (2003). Molecular analysis of *Burkholderia cepacia* complex isolates from a Portuguese cystic fibrosis center: a 7-year study. *J. Clin. Microbiol.* 41, 4113–4120. doi: 10.1128/JCM.41.9.4113-4120.2003
- Darling, A. E., Mau, B., and Perna, N. T. (2010). progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 5:e11147. doi: 10.1371/journal.pone.0011147
- De Smet, B., Mayo, M., Peeters, C., Zlosnik, J. E., Spilker, T., Hird, T. J., et al. (2015). *Burkholderia stagnalis* sp. nov. and *Burkholderia territorii* sp. nov., two novel *Burkholderia cepacia* complex species from environmental and human sources. *Int. J. Syst. Evol. Microbiol.* 65, 2265–2271. doi: 10.1099/ij.s.0.000251
- Del Campo, M., Kaya, Y., and Ofengand, J. (2001). Identification and site of action of the remaining four putative pseudouridine synthases in *Escherichia coli*. *Rna* 7, 1603–1615.
- Depoorter, E., Bull, M. J., Peeters, C., Coenye, T., Vandamme, P., and Mahenthalingam, E. (2016). *Burkholderia*: an update on taxonomy and biotechnological potential as antibiotic producers. *Appl. Microbiol. Biotechnol.* 100, 5215–5229. doi: 10.1007/s00253-016-7520-x
- Diaz Caballero, J., Clark, S. T., Wang, P. W., Donaldson, S. L., Coburn, B., Tullis, D. E., et al. (2018). A genome-wide association analysis reveals a potential role for recombination in the evolution of antimicrobial resistance in *Burkholderia multivorans*. *PLoS Pathog.* 14:e1007453. doi: 10.1371/journal.ppat.1007453
- Didelot, X., Walker, A. S., Peto, T. E., Crook, D. W., and Wilson, D. J. (2016). Within-host evolution of bacterial pathogens. *Nat. Rev. Microbiol.* 14, 150–162. doi: 10.1038/nrmicro.2015.13
- Dillon, M. M., Sung, W., Lynch, M., and Cooper, V. S. (2015). The rate and molecular spectrum of spontaneous mutations in the GC-rich multichromosome genome of *Burkholderia cenocepacia*. *Genetics* 200, 935–946. doi: 10.1534/genetics.115.176834
- Dillon, M. M., Sung, W., Sebra, R., Lynch, M., and Cooper, V. S. (2017). Genome-wide biases in the rate and molecular spectrum of spontaneous mutations in *Vibrio cholerae* and *Vibrio fischeri*. *Mol. Biol. Evol.* 34, 93–109. doi: 10.1093/molbev/msw224
- Drevinek, P., and Mahenthalingam, E. (2010). *Burkholderia cenocepacia* in cystic fibrosis: epidemiology and molecular mechanisms of virulence. *Clin. Microbiol. Infect.* 16, 821–830. doi: 10.1111/j.1469-0691.2010.03237.x
- Galli, F., Battistoni, A., Gambari, R., Pompella, A., Bragonzi, A., Pilolli, F., et al. (2012). Oxidative stress and antioxidant therapy in cystic fibrosis. *Biochim. Biophys. Acta (BBA)* 1822, 690–713. doi: 10.1016/j.bbadis.2011.12.012
- Graindorge, A., Menard, A., Monnez, C., and Cournoyer, B. (2012). Insertion sequence evolutionary patterns highlight convergent genetic inactivations and recent genomic island acquisitions among epidemic *Burkholderia cenocepacia*. *J. Med. Microbiol.* 61(Pt 3), 394–409. doi: 10.1099/jmm.0.036822-0
- Guo, F. B., Xiong, L., Zhang, K. Y., Dong, C., Zhang, F. Z., and Woo, P. C. (2017). Identification and analysis of genomic islands in *Burkholderia cenocepacia* AU 1054 with emphasis on pathogenicity islands. *BMC Microbiol.* 17:73. doi: 10.1186/s12866-017-0986-6
- Gurevich, A., Saveliev, V., Vyahhi, N., and Tesler, G. (2013). QUASt: quality assessment tool for genome assemblies. *Bioinformatics* 29, 1072–1075. doi: 10.1093/bioinformatics/btt086
- Hassan, A. A., Coutinho, C. P., and Sá-Correia, I. (2019a). *Burkholderia cepacia* complex species differ in the frequency of variation of the lipopolysaccharide O-antigen expression during cystic fibrosis chronic respiratory infection. *Front. Cell. Infect. Microbiol.* 9:273. doi: 10.3389/fcimb.2019.00273
- Hassan, A. A., Maldonado, R. F., Dos Santos, S. C., Di Lorenzo, F., Silipo, A., Coutinho, C. P., et al. (2017). Structure of O-antigen and hybrid biosynthetic locus in *Burkholderia cenocepacia* clonal variants recovered from a cystic fibrosis patient. *Front. Microbiol.* 8:1027. doi: 10.3389/fmicb.2017.01027

- Hassan, A. A., Vitorino, M. V., Robalo, T., Rodrigues, M. S., and Sa-Correia, I. (2019b). Variation of *Burkholderia cenocepacia* cell wall morphology and mechanical properties during cystic fibrosis lung infection, assessed by atomic force microscopy. *Sci. Rep.* 9:16118. doi: 10.1038/s41598-019-52604-9
- Holden, M. T., Seth-Smith, H. M., Crossman, L. C., Sebaihia, M., Bentley, S. D., Cerdeno-Tarraga, A. M., et al. (2009). The genome of *Burkholderia cenocepacia* J2315, an epidemic pathogen of cystic fibrosis patients. *J. Bacteriol.* 191, 261–277. doi: 10.1128/JB.01230-08
- Huang, M., Oppermann-Sanio, F. B., and Steinbuchel, A. (1999). Biochemical and molecular characterization of the *Bacillus subtilis* acetoin catabolic pathway. *J. Bacteriol.* 181, 3837–3841.
- Huerta-Cepas, J., Forslund, K., Coelho, L. P., Szklarczyk, D., Jensen, L. J., von Mering, C., et al. (2017). Fast genome-wide functional annotation through orthology assignment by eggNOG-mapper. *Mol. Biol. Evol.* 34, 2115–2122. doi: 10.1093/molbev/msx148
- Jain, C., Rodriguez-R, L. M., Phillippy, A. M., Konstantinidis, K. T., and Aluru, S. (2018). High throughput ANI analysis of 90K prokaryotic genomes reveals clear species boundaries. *Nat. Commun.* 9:5114. doi: 10.1038/s41467-018-07641-9
- Jeannot, K., Elsen, S., Kohler, T., Attree, I., van Delden, C., and Plesiat, P. (2008). Resistance and virulence of *Pseudomonas aeruginosa* clinical strains overproducing the MexCD-OprJ efflux pump. *Antimicrob. Agents Chemother.* 52, 2455–2462. doi: 10.1128/AAC.01107-07
- Jia, B., Raphenya, A. R., Alcock, B., Waglechner, N., Guo, P., Tsang, K. K., et al. (2017). CARD 2017: expansion and model-centric curation of the comprehensive antibiotic resistance database. *Nucleic Acids Res.* 45, D566–D573. doi: 10.1093/nar/gkw1004
- Johnson, S. L., Bishop-Lilly, K. A., Ladner, J. T., Daligault, H. E., Davenport, K. W., Jaissle, J., et al. (2016). Correction for Johnson et al., complete genome sequences for 59 *Burkholderia* isolates, both pathogenic and near neighbor. *Genome Announc.* 4:e00313-16. doi: 10.1128/genomeA.00313-16
- Johnson, S. L., Bishop-Lilly, K. A., Ladner, J. T., Daligault, H. E., Davenport, K. W., Jaissle, J., et al. (2015). Complete genome sequences for 59 *Burkholderia* isolates, both pathogenic and near neighbor. *Genome Announc.* 3:e00159-15. doi: 10.1128/genomeA.00159-15
- Jolley, K. A., and Maiden, M. C. J. (2010). BIGSdb: scalable analysis of bacterial genome variation at the population level. *BMC Bioinformatics* 11:595. doi: 10.1186/1471-2105-11-595
- Kaas, R. S., Leekitcharoenphon, P., Aarestrup, F. M., and Lund, O. (2014). Solving the problem of comparing whole bacterial genomes across different sequencing platforms. *PLoS One* 9:e104984. doi: 10.1371/journal.pone.0104984
- Komatsu, H., Imura, Y., Otori, A., Nagata, Y., and Tsuda, M. (2003). Distribution and organization of auxotrophic genes on the multichromosomal genome of *Burkholderia multivorans* ATCC 17616. *J. Bacteriol.* 185, 3333–3343. doi: 10.1128/JB.185.11.3333-3343.2003
- Kotrange, S., Kopp, B., Akhter, A., Abdelaziz, D., Abu Khweek, A., Caution, K., et al. (2011). *Burkholderia cenocepacia* O polysaccharide chain contributes to caspase-1-dependent IL-1 β production in macrophages. *J. Leukoc. Biol.* 89, 481–488. doi: 10.1189/jlb.0910513
- Lee, A. H., Flibotte, S., Sinha, S., Paiero, A., Ehrlich, R. L., Balashov, S., et al. (2017). Phenotypic diversity and genotypic flexibility of *Burkholderia cenocepacia* during long-term chronic infection of cystic fibrosis lungs. *Genome Res.* 27, 650–662. doi: 10.1101/gr.213363.116
- Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26, 589–595. doi: 10.1093/bioinformatics/btp698
- Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352
- Lieberman, T. D., Flett, K. B., Yelin, I., Martin, T. R., McAdam, A. J., Priebe, G. P., et al. (2014). Genetic variation of a bacterial pathogen within individuals with cystic fibrosis provides a record of selective pressures. *Nature Genet.* 46, 82–87. doi: 10.1038/ng.2848
- Lieberman, T. D., Michel, J. B., Aingaran, M., Potter-Bynoe, G., Roux, D., Davis, M. R. Jr., et al. (2011). Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nature Genet.* 43, 1275–1280. doi: 10.1038/ng.997
- Lipuma, J. J. (2010). The changing microbial epidemiology in cystic fibrosis. *Clin. Microbiol. Rev.* 23, 299–323. doi: 10.1128/CMR.00068-09
- Liu, B., Zheng, D., Jin, Q., Chen, L., and Yang, J. (2019). VFDB 2019: a comparative pathogenomic platform with an interactive web interface. *Nucleic Acids Res.* 47, D687–D692. doi: 10.1093/nar/gky1080
- Loutet, S. A., and Valvano, M. A. (2010). A decade of *Burkholderia cenocepacia* virulence determinant research. *Infect. Immun.* 78, 4088–4100. doi: 10.1128/IAI.00212-10
- Madeira, A., dos Santos, S. C., Santos, P. M., Coutinho, C. P., Tyrrell, J., McClean, S., et al. (2013). Proteomic profiling of *Burkholderia cenocepacia* clonal isolates with different virulence potential retrieved from a cystic fibrosis patient during chronic lung infection. *PLoS One* 8:e83065. doi: 10.1371/journal.pone.0083065
- Madeira, A., Santos, P. M., Coutinho, C. P., Pinto-de-Oliveira, A., and Sá-Correia, I. (2011). Quantitative proteomics (2-D DIGE) reveals molecular strategies employed by *Burkholderia cenocepacia* to adapt to the airways of cystic fibrosis patients under antimicrobial therapy. *Proteomics* 11, 1313–1328. doi: 10.1002/pmic.201000457
- Mahenthalingam, E., Urban, T. A., and Goldberg, J. B. (2005). The multifarious, multireplicon *Burkholderia cepacia* complex. *Nature Rev. Microbiol.* 3, 144–156. doi: 10.1038/nrmicro1085
- Maldonado, R. F., Sá-Correia, I., and Valvano, M. A. (2016). Lipopolysaccharide modification in Gram-negative bacteria during chronic infection. *FEMS Microbiol. Rev.* 40, 480–493. doi: 10.1093/femsre/fuw007
- Markussen, T., Marvig, R. L., Gomez-Lozano, M., Aanaes, K., Burleigh, A. E., Hoiby, N., et al. (2014). Environmental heterogeneity drives within-host diversification and evolution of *Pseudomonas aeruginosa*. *mBio* 5:e01592-14. doi: 10.1128/mBio.01592-14
- Martina, P., Leguizamón, M., Prieto, C. I., Sousa, S. A., Montanaro, P., Draghi, W. O., et al. (2018). *Burkholderia puraquae* sp. nov., a novel species of the *Burkholderia cepacia* complex isolated from hospital settings and agricultural soils. *Int. J. Syst. Evol. Microbiol.* 68, 14–20. doi: 10.1099/ijsem.0.002293
- Marvig, R. L., Dolce, D., Sommer, L. M., Petersen, B., Ciofu, O., Campana, S., et al. (2015). Within-host microevolution of *Pseudomonas aeruginosa* in Italian cystic fibrosis patients. *BMC Microbiol.* 15:218. doi: 10.1186/s12866-015-0563-9
- Mira, N. P., Madeira, A., Moreira, A. S., Coutinho, C. P., and Sá-Correia, I. (2011). Genomic expression analysis reveals strategies of *Burkholderia cenocepacia* to adapt to cystic fibrosis patients' airways and antimicrobial therapy. *PLoS One* 6:e28831. doi: 10.1371/journal.pone.0028831
- Moreira, A. S., Coutinho, C. P., Azevedo, P., Lito, L., Melo-Cristino, J., and Sá-Correia, I. (2014). *Burkholderia dolosa* phenotypic variation during the decline in lung function of a cystic fibrosis patient during 5.5 years of chronic colonization. *J. Med. Microbiol.* 63(Pt 4), 594–601. doi: 10.1099/jmm.0.069849-0
- Moreira, A. S., Mil-Homens, D., Sousa, S. A., Coutinho, C. P., Pinto-de-Oliveira, A., Ramos, C. G., et al. (2017). Variation of *Burkholderia cenocepacia* virulence potential during cystic fibrosis chronic lung infection. *Virulence* 8, 782–796. doi: 10.1080/21505594.2016.1237334
- Moriarty, T. F., McElroy, J. C., Elborn, J. S., and Tunney, M. M. (2007). Sputum antibiotic concentrations: implications for treatment of cystic fibrosis lung infection. *Pediatric Pulmonol.* 42, 1008–1017. doi: 10.1002/ppul.20671
- Nunvar, J., Capek, V., Fiser, K., Fila, L., and Drevinec, P. (2017). What matters in chronic *Burkholderia cenocepacia* infection in cystic fibrosis: insights from comparative genomics. *PLoS Pathog.* 13:e1006762. doi: 10.1371/journal.ppat.1006762
- Ong, K. S., Aw, Y. K., Lee, L. H., Yule, C. M., Cheow, Y. L., and Lee, S. M. (2016). *Burkholderia paludis* sp. nov., an antibiotic-siderophore producing novel *Burkholderia cepacia* complex species, isolated from Malaysian tropical peat swamp soil. *Front. Microbiol.* 7:2046. doi: 10.3389/fmicb.2016.02046
- Ortega, X., Hunt, T. A., Loutet, S., Vinion-Dubiel, A. D., Datta, A., Choudhury, B., et al. (2005). Reconstitution of O-specific lipopolysaccharide expression in *Burkholderia cenocepacia* strain J2315, which is associated with transmissible infections in patients with cystic fibrosis. *J. Bacteriol.* 187, 1324–1333. doi: 10.1128/JB.187.4.1324-1333.2005
- Palmer, K. L., Aye, L. M., and Whiteley, M. (2007). Nutritional cues control *Pseudomonas aeruginosa* multicellular behavior in cystic fibrosis sputum. *J. Bacteriol.* 189, 8079–8087. doi: 10.1128/JB.01138-07
- Panmanee, W., Gomez, F., Witte, D., Pancholi, V., Britigan, B. E., and Hassett, D. J. (2008). The peptidoglycan-associated lipoprotein OprL helps protect a *Pseudomonas aeruginosa* mutant devoid of the transactivator OxyR from

- hydrogen peroxide-mediated killing during planktonic and biofilm culture. *J. Bacteriol.* 190, 3658–3669. doi: 10.1128/JB.00022-08
- Peeters, E., Sass, A., Mahenthalingam, E., Nelis, H., and Coenye, T. (2010). Transcriptional response of *Burkholderia cenocepacia* J2315 sessile cells to treatments with high doses of hydrogen peroxide and sodium hypochlorite. *BMC Genomics* 11:90. doi: 10.1186/1471-2164-11-90
- Price, M. N., Dehal, P. S., and Arkin, A. P. (2010). FastTree 2—approximately maximum-likelihood trees for large alignments. *PLoS One* 5:e9490. doi: 10.1371/journal.pone.0009490
- Ramos, J. L., Martínez-Bueno, M., Molina-Henares, A. J., Terán, W., Watanabe, K., Zhang, X., et al. (2005). The TetR family of transcriptional repressors. *Microbiol. Mol. Biol. Rev. MMBR* 69, 326–356. doi: 10.1128/MMBR.69.2.326-356.2005
- Reid, D. W., Misso, N., Aggarwal, S., Thompson, P. J., and Walters, E. H. (2007). Oxidative stress and lipid-derived inflammatory mediators during acute exacerbations of cystic fibrosis. *Respirology* 12, 63–69. doi: 10.1111/j.1440-1843.2006.00962.x
- Reyrat, J. M., David, M., Blonski, C., Boistard, P., and Batut, J. (1993). Oxygen-regulated in vitro transcription of *Rhizobium meliloti* nifA and fixK genes. *J. Bacteriol.* 175, 6867–6872. doi: 10.1128/jb.175.21.6867-6872.1993
- Robinson, J. T., Thorvaldsdottir, H., Winckler, W., Guttman, M., Lander, E. S., Getz, G., et al. (2011). Integrative genomics viewer. *Nature Biotechnol.* 29, 24–26. doi: 10.1038/nbt.1754
- Saito, K., Ito, E., Hosono, K., Nakamura, K., Imai, K., Iizuka, T., et al. (2003). The uncoupling of oxygen sensing, phosphorylation signalling and transcriptional activation in oxygen sensor FixL and FixJ mutants. *Mol. Microbiol.* 48, 373–383.
- Saldias, M. S., Ortega, X., and Valvano, M. A. (2009). *Burkholderia cenocepacia* O antigen lipopolysaccharide prevents phagocytosis by macrophages and adhesion to epithelial cells. *J. Med. Microbiol.* 58(Pt 12), 1542–1548. doi: 10.1099/jmm.0.013235-0
- Saldias, M. S., and Valvano, M. A. (2009). Interactions of *Burkholderia cenocepacia* and other *Burkholderia cepacia* complex bacteria with epithelial and phagocytic cells. *Microbiology* 155(Pt 9), 2809–2817. doi: 10.1099/mic.0.031344-0
- Sass, A. M., Schmerk, C., Agnoli, K., Norville, P. J., Eberl, L., Valvano, M. A., et al. (2013). The unexpected discovery of a novel low-oxygen-activated locus for the anoxic persistence of *Burkholderia cenocepacia*. *ISME J.* 7, 1568–1581. doi: 10.1038/ismej.2013.36
- Schmerk, C. L., and Valvano, M. A. (2013). *Burkholderia multivorans* survival and trafficking within macrophages. *J. Med. Microbiol.* 62(Pt 2), 173–184. doi: 10.1099/jmm.0.051243-0
- Scribner, M. R., Santos-Lopez, A., Marshall, C. W., Deitrick, C., and Cooper, V. S. (2020). Parallel evolution of tobramycin resistance across species and environments. *mBio* 11:e00932-20. doi: 10.1128/mBio.00932-20
- Seemann, T. (2014). Prokka: rapid prokaryotic genome annotation. *Bioinformatics* 30, 2068–2069. doi: 10.1093/bioinformatics/btu153
- Silva, I. N., Ferreira, A. S., Becker, J. D., Zlosnik, J. E., Speert, D. P., He, J., et al. (2011). Mucooid morphotype variation of *Burkholderia multivorans* during chronic cystic fibrosis lung infection is correlated with changes in metabolism, motility, biofilm formation and virulence. *Microbiology* 157(Pt 11), 3124–3137. doi: 10.1099/mic.0.050989-0
- Silva, I. N., Santos, P. M., Santos, M. R., Zlosnik, J. E., Speert, D. P., Buskirk, S. W., et al. (2016). Long-term evolution of *Burkholderia multivorans* during a chronic cystic fibrosis infection reveals shifting forces of selection. *mSystems* 11:e00029-16. doi: 10.1128/mSystems.00029-16
- Skerker, J. M., Prasol, M. S., Perchuk, B. S., Biondi, E. G., and Laub, M. T. (2005). Two-component signal transduction pathways regulating growth and cell cycle progression in a bacterium: a system-level analysis. *PLoS Biol.* 3:e334. doi: 10.1371/journal.pbio.0030334
- Sokol, P. A., Darling, P., Lewenza, S., Corbett, C. R., and Kooi, C. D. (2000). Identification of a siderophore receptor required for ferric ornibactin uptake in *Burkholderia cepacia*. *Infect. Immun.* 68, 6554–6560. doi: 10.1128/iai.68.12.6554-6560.2000
- Sokol, P. A., Darling, P., Woods, D. E., Mahenthalingam, E., and Kooi, C. (1999). Role of ornibactin biosynthesis in the virulence of *Burkholderia cepacia*: characterization of pvdA, the gene encoding L-ornithine N(5)-oxygenase. *Infect. Immun.* 67, 4443–4455.
- Sousa, S. A., Feliciano, J. R., Pita, T., Guerreiro, S. I., and Leitao, J. H. (2017). *Burkholderia cepacia* complex regulation of virulence gene expression: a review. *Genes (Basel)* 8:43. doi: 10.3390/genes8010043
- Summer, E. J., Gonzalez, C. F., Carlisle, T., Mebane, L. M., Cass, A. M., Savva, C. G., et al. (2004). *Burkholderia cenocepacia* phage BcepMu and a family of Mu-like phages encoding potential pathogenesis factors. *J. Mol. Biol.* 340, 49–65. doi: 10.1016/j.jmb.2004.04.053
- Thorvaldsdottir, H., Robinson, J. T., and Mesirov, J. P. (2013). Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration. *Brief. Bioinform.* 14, 178–192. doi: 10.1093/bib/bbs017
- Tomomori, C., Tanaka, T., Dutta, R., Park, H., Saha, S. K., Zhu, Y., et al. (1999). Solution structure of the homodimeric core domain of *Escherichia coli* histidine kinase EnvZ. *Nature Struct. Mol. Biol.* 6, 729–734. doi: 10.1038/11495
- Traverse, C. C., Mayo-Smith, L. M., Poltak, S. R., and Cooper, V. S. (2013). Tangled bank of experimentally evolved *Burkholderia* biofilms reflects selection during chronic infections. *Proc. Natl. Acad. Sci. U.S.A.* 110, E250–E259. doi: 10.1073/pnas.1207025110
- Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del Angel, G., Levy-Moonshine, A., et al. (2013). From FastQ data to high confidence variant calls: the genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinform.* 43, 11.10.1–11.10.33. doi: 10.1002/0471250953.bi1110s43
- Vandeplassche, E., Sass, A., Lemarcq, A., Dandekar, A. A., Coenye, T., and Crabbe, A. (2019). In vitro evolution of *Pseudomonas aeruginosa* AA2 biofilms in the presence of cystic fibrosis lung microbiome members. *Sci. Rep.* 9:12859. doi: 10.1038/s41598-019-49371-y
- Varga, J. J., Losada, L., Zelazny, A. M., Kim, M., McCarrison, J., Brinkac, L., et al. (2013). Draft genome sequences of *Burkholderia cenocepacia* ET12 lineage strains K56-2 and BC7. *Genome Announc.* 1:e00841-13. doi: 10.1128/genomeA.00841-13
- Viberg, L. T., Sarovich, D. S., Kidd, T. J., Geake, J. B., Bell, S. C., Currie, B. J., et al. (2017). Within-host evolution of *Burkholderia pseudomallei* during chronic infection of seven Australasian cystic fibrosis patients. *mBio* 8:e00356-17. doi: 10.1128/mBio.00356-17
- Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. doi: 10.1371/journal.pone.0112963
- Weber, C. F., and King, G. M. (2017). Volcanic Soils as Sources of Novel CO-Oxidizing *Paraburkholderia* and *Burkholderia*: *Paraburkholderia hiiakeae* sp. nov., *Paraburkholderia metrosideri* sp. nov., *Paraburkholderia paradisi* sp. nov., *Paraburkholderia peleaee* sp. nov., and *Burkholderia alpina* sp. nov. a member of the *Burkholderia cepacia* complex. *Front. Microbiol.* 8:207. doi: 10.3389/fmicb.2017.00207
- Whiteson, K. L., Meinardi, S., Lim, Y. W., Schmieder, R., Maughan, H., Quinn, R., et al. (2014). Breath gas metabolites and bacterial metagenomes from cystic fibrosis airways indicate active pH neutral 2,3-butanedione fermentation. *ISME J.* 8, 1247–1258. doi: 10.1038/ismej.2013.229
- Williams, H. D., Zlosnik, J. E., and Ryall, B. (2007). Oxygen, cyanide and energy generation in the cystic fibrosis pathogen *Pseudomonas aeruginosa*. *Adv. Microb. Physiol.* 52, 1–71. doi: 10.1016/S0065-2911(06)52001-6
- Winstanley, C., O'Brien, S., and Brockhurst, M. A. (2016). *Pseudomonas aeruginosa* evolutionary adaptation and diversification in cystic fibrosis chronic lung infections. *Trends Microbiol.* 24, 327–337. doi: 10.1016/j.tim.2016.01.008
- Wolanin, P. M., Thomason, P. A., and Stock, J. B. (2002). Histidine protein kinases: key signal transducers outside the animal kingdom. *Genome Biol.* 3:REVIEWS3013. doi: 10.1186/gb-2002-3-10-reviews3013
- Zankari, E. (2014). Comparison of the web tools ARG-ANNOT and ResFinder for detection of resistance genes in bacteria. *Antimicrob. Agents Chemother.* 58:4986. doi: 10.1128/AAC.02620-14
- Zerbino, D. R. (2010). Using the Velvet de novo assembler for short-read sequencing technologies. *Curr. Protoc. Bioinform.* Chapter11:Unit11.5. doi: 10.1002/0471250953.bi1105s31
- Zerbino, D. R., and Birney, E. (2008). Velvet: algorithms for de novo short read assembly using de Bruijn graphs. *Genome Res.* 18, 821–829. doi: 10.1101/gr.074492.107
- Zhang, Z., Schwartz, S., Wagner, L., and Miller, W. (2000). A greedy algorithm for aligning DNA sequences. *J. Comp. Biol.* 7, 203–214. doi: 10.1089/10665270050081478
- Zlosnik, J. E., Costa, P. S., Brant, R., Mori, P. Y., Hird, T. J., Fraenkel, M. C., et al. (2011). Mucooid and nonmucooid *Burkholderia cepacia* complex bacteria

- in cystic fibrosis infections. *Am. J. Respir. Crit. Care Med.* 183, 67–72. doi: 10.1164/rccm.201002-0203OC
- Zlosnik, J. E., Mori, P. Y., To, D., Leung, J., Hird, T. J., and Speert, D. P. (2014). Swimming motility in a longitudinal collection of clinical isolates of *Burkholderia cepacia* complex bacteria from people with cystic fibrosis. *PLoS One* 9:e106428. doi: 10.1371/journal.pone.0106428
- Zlosnik, J. E., and Speert, D. P. (2010). The role of mucoidy in virulence of bacteria from the *Burkholderia cepacia* complex: a systematic proteomic and transcriptomic analysis. *J. Infect. Dis.* 202, 770–781. doi: 10.1086/655663

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.

Copyright © 2020 Hassan, dos Santos, Cooper and Sá-Correia. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.