




Long-Term Evolution of *Burkholderia multivorans* during a Chronic Cystic Fibrosis Infection Reveals Shifting Forces of Selection

Inês N. Silva,^a  Pedro M. Santos,^b Mário R. Santos,^a James E. A. Zlosnik,^c David P. Speert,^c Sean W. Buskirk,^d Eric L. Bruger,^e Christopher M. Waters,^e  Vaughn S. Cooper,^f  Leonilde M. Moreira^{a,g}

Institute for Bioengineering and Biosciences, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal^a; Centre of Molecular and Environmental Biology, Department of Biology, Universidade do Minho, Braga, Portugal^b; Department of Pediatrics, Faculty of Medicine, Centre for Understanding and Preventing Infection in Children, University of British Columbia, Vancouver, Canada^c; Department of Molecular, Cellular and Biomedical Sciences, University of New Hampshire, Durham, New Hampshire, USA^d; Department of Microbiology and Molecular Genetics, Michigan State University, East Lansing, Michigan, USA^e; Department of Microbiology and Molecular Genetics, University of Pittsburgh, Pittsburgh, Pennsylvania, USA^f; Department of Bioengineering, Instituto Superior Técnico, Universidade de Lisboa, Lisbon, Portugal^g

ABSTRACT *Burkholderia multivorans* is an opportunistic pathogen capable of causing severe disease in patients with cystic fibrosis (CF). Patients may be chronically infected for years, during which the bacterial population evolves in response to unknown forces. Here we analyze the genomic and functional evolution of a *B. multivorans* infection that was sequentially sampled from a CF patient over 20 years. The population diversified into at least four primary, coexisting clades with distinct evolutionary dynamics. The average substitution rate was only 2.4 mutations/year, but notably, some lineages evolved more slowly, whereas one diversified more rapidly by mostly nonsynonymous mutations. Ten loci, mostly involved in gene expression regulation and lipid metabolism, acquired three or more independent mutations and define likely targets of selection. Further, a broad range of phenotypes changed in association with the evolved mutations; they included antimicrobial resistance, biofilm regulation, and the presentation of lipopolysaccharide O-antigen repeats, which was directly caused by evolved mutations. Additionally, early isolates acquired mutations in genes involved in cyclic di-GMP (c-di-GMP) metabolism that associated with increased c-di-GMP intracellular levels. Accordingly, these isolates showed lower motility and increased biofilm formation and adhesion to CFBE410⁻ epithelial cells than the initial isolate, and each of these phenotypes is an important trait for bacterial persistence. The timing of the emergence of this clade of more adherent genotypes correlated with the period of greatest decline in the patient's lung function. All together, our observations suggest that selection on *B. multivorans* populations during long-term colonization of CF patient lungs either directly or indirectly targets adherence, metabolism, and changes in the cell envelope related to adaptation to the biofilm lifestyle.

IMPORTANCE Bacteria may become genetically and phenotypically diverse during long-term colonization of cystic fibrosis (CF) patient lungs, yet our understanding of within-host evolutionary processes during these infections is lacking. Here we combined current genome sequencing technologies and detailed phenotypic profiling of the opportunistic pathogen *Burkholderia multivorans* using sequential isolates sampled from a CF patient over 20 years. The evolutionary history of these isolates highlighted bacterial genes and pathways that were likely subject to strong selection within the host and were associated with altered phenotypes, such as biofilm pro-

Received 9 March 2016 Accepted 25 April 2016 Published 24 May 2016

Citation Silva IN, Santos PM, Santos MR, Zlosnik JEA, Speert DP, Buskirk SW, Bruger EL, Waters CM, Cooper VS, Moreira LM. 2016. Long-term evolution of *Burkholderia multivorans* during a chronic cystic fibrosis infection reveals shifting forces of selection. *mSystems* 1(3):e00029-16. doi:10.1128/mSystems.00029-16.

Editor Jonathan Eisen, UC Davis Genome Center

Copyright © 2016 Silva et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Leonilde M. Moreira, lmoreira@tecnico.ulisboa.pt, or Vaughn S. Cooper, vaughn.cooper@pitt.edu.

duction, motility, and antimicrobial resistance. Importantly, multiple lineages coexisted for years or even decades within the infection, and the period of diversification within the dominant lineage was associated with deterioration of the patient's lung function. Identifying traits under strong selection during chronic infection not only sheds new light onto *Burkholderia* evolution but also sets the stage for tailored therapeutics targeting the prevailing lineages associated with disease progression.

KEYWORDS: *Burkholderia multivorans*, biofilm, c-di-GMP signaling, chronic infection, cystic fibrosis, evolution, lung function decline, respiratory infection, within-host adaptation

Bacteria have evolved many strategies to adapt within their hosts, such as changes in their surface antigens to evade the immune system, remodeling of the host epigenetic machinery, and disruption of host signaling (1–3). While some host-microbe interactions have been extensively studied, less is known about the evolutionary process of bacterial adaptation inside individual hosts. Bacterial infections of the respiratory airways of cystic fibrosis (CF) patients present a dynamic and highly relevant model for studying microbial adaptation. Feared pathogens include *Pseudomonas aeruginosa* and strains of the *Burkholderia cepacia* complex, as they are difficult to eradicate from the lungs of CF patients due, in part, to their intrinsic resistance to antimicrobial compounds (4). One of the first genome comparisons of longitudinally collected *P. aeruginosa* isolates recovered from infected CF patients led to the identification of mutations in genes important for host adaptation, such as *mexZ*, encoding a negative regulator of a multidrug efflux pump, and *lasR*, encoding the primary quorum-sensing regulator (5). More recently, by performing whole-genome sequencing of more than 400 *P. aeruginosa* isolates recovered from 34 CF patients, Marvig and colleagues identified 52 genes that convergently evolved, most of them involved in antibiotic resistance, motility, and biofilm formation (6).

Infection of CF patient airways by *B. cepacia* complex bacteria can be transient but in most cases leads to chronic colonization, causing a progressive deterioration of lung function. In a subset of patients, *B. cepacia* complex infection develops into the “cepacia syndrome,” characterized by fatal necrotizing pneumonia sometimes accompanied by bacteremia (7). Although many studies of putative virulence factors of the *B. cepacia* complex have been performed (reviewed in references 8 to 11), very few characterized the determinants of adaptation by these bacteria to CF patient lungs. A retrospective study of 112 *Burkholderia dolosa* isolates from 14 patients with CF who were infected in an epidemic outbreak over the course of 16 years (12) identified several bacterial genes that adaptively evolved by tracking recurrent patterns of mutations during infection of multiple individuals. These mutations affected important phenotypes, such as antibiotic resistance and expression of surface polysaccharides (12). In a subsequent study, sputum samples from five patients who had been infected with *B. dolosa* were analyzed for intraspecies diversity, which illustrated that several lineages may coexist for many years within a patient and that genes involved in outer membrane components, iron scavenging, and antibiotic resistance are under strong selection pressure (13).

The first comprehensive study covering more than 600 *B. cepacia* complex isolates from the environment and from chronic lung infection samples from 100 CF patients analyzed variation in colony mucoidy due to exopolysaccharide biosynthesis (14). Beyond demonstrating that all species of the *B. cepacia* complex can express the mucoid phenotype, it showed that sequential clonal isolates from 15 patients may undergo switches in this phenotype, the majority being mucoid-to-nonmucoid transitions in the species *Burkholderia multivorans* (9 patients) and *Burkholderia cenocepacia* (3 patients) (14). Further, patients colonized with nonmucoid isolates tended to exhibit more-rapid deterioration in lung function and poorer clinical outcome than those from whom mucoid isolates were recovered (15). *In vitro* studies have shown that mucoid-to-nonmucoid transitions in *B. cepacia* complex organisms may occur under stress

conditions imposed by a long stationary phase, oxidative and osmotic stresses, and subinhibitory concentrations of antibiotics (16). As the molecular mechanisms underlying within-host genotypic and phenotypic variation are unknown, our aim was to identify genome-wide adaptations that could explain the long-term persistence of *B. cepacia* complex bacteria in CF patient lungs. Thus, we sequenced the genomes of 22 *Burkholderia multivorans* isolates recovered over 20 years of chronic colonization in a patient from whom isolates with a mucoid-to-nonmucoid transition had been obtained (14). This represents the first retrospective study of the genomic and phenotypic evolution of longitudinally collected *B. multivorans* specimens, which is notable because *B. multivorans* has recently been shown to be the predominant *Burkholderia* species infecting CF patients in Canada, the United States, New Zealand, and Europe (17–20). Our primary objective was to identify pathoadaptive mechanisms to this particular host environment. Here, we provide evidence that several *B. multivorans* clades coexisted at any given time but evolved at different rates. Mutations that defined new clades and others that evolved independently in multiple isolates affected a limited number of genes encoding regulatory proteins, lipid and amino acid metabolic enzymes, and proteins involved in envelope biogenesis. Overall, the evolution of this *B. multivorans* population was marked by periods of strong diversifying selection followed by periods of relative stasis. This dynamic suggests that monitoring these evolutionary and molecular patterns might be used to design responsive therapies designed to limit population diversity and disease progression.

RESULTS

Evolution of *B. multivorans* within the airway of a CF patient involves diversification yet a slow and steady accumulation of selected mutations. The CF patient from whom *B. multivorans* isolates were recovered carried a complex genotype (*CFTR* Δ F508/*CFTR* I444S) and is one of the nine patients chronically colonized by *B. multivorans* in which a colony morphotype mucoid-to-nonmucoid transition was identified previously (14). This patient was hospitalized once due to CF-related respiratory illness but had been treated on multiple occasions with antibiotics, including cephalosporins, fluoroquinolones, β -lactams, azithromycin, and sulfamethoxazole-trimethoprim. During the course of this patient's regular semiannual examinations, *B. multivorans* was first isolated in 1993 and recovered periodically until 2013, yielding a total of 22 isolates for analysis (see Table S1 in the supplemental material). In 1996, 2006, 2011, and 2013, two colonies of different morphologies were analyzed, but at all other times, only one isolate was evaluated. Although CF patient airways harbor numerous microorganisms that evade detection by culture-based methods, the information available from these methods showed that this patient was initially infected by *Staphylococcus aureus*, *Haemophilus influenzae*, and *Pseudomonas aeruginosa* (years, 1989 to 1993), followed by a period of coinfection of these microorganisms with *B. multivorans* (1993 to 1996). After that period, except for *Candida* and *Staphylococcus aureus* isolation, only *B. multivorans* was reported (Table S2).

The analysis of the mucoid phenotype, assessed by quantification of the dry weight ethanol precipitate present in the culture supernatant, indicated that isolates BM5, BM12, and BM22 were the only ones unable to produce detectable amounts of exopolysaccharide (see Fig. S1 in the supplemental material). Despite some statistically nonsignificant variation in the total amounts, all the other isolates were considered highly mucoid.

Aiming to understand the evolutionary adaptation of *B. multivorans* in the CF host environment and to identify genetic mutations possibly leading to phenotypic variation, we sequenced and analyzed the genomes of the 22 isolates. Single-nucleotide polymorphisms (SNPs) and indels were identified by mapping sequence reads for each isolate against the genome sequence of the first isolate (BM1). A total of 399 SNPs and 88 indels were detected over 20 years of chronic colonization (see Table S3 in the supplemental material). The last isolate (BM22) alone accounted for 196 new mutations, likely due to a frameshift mutation in the mismatch repair gene *mutL* which conferred

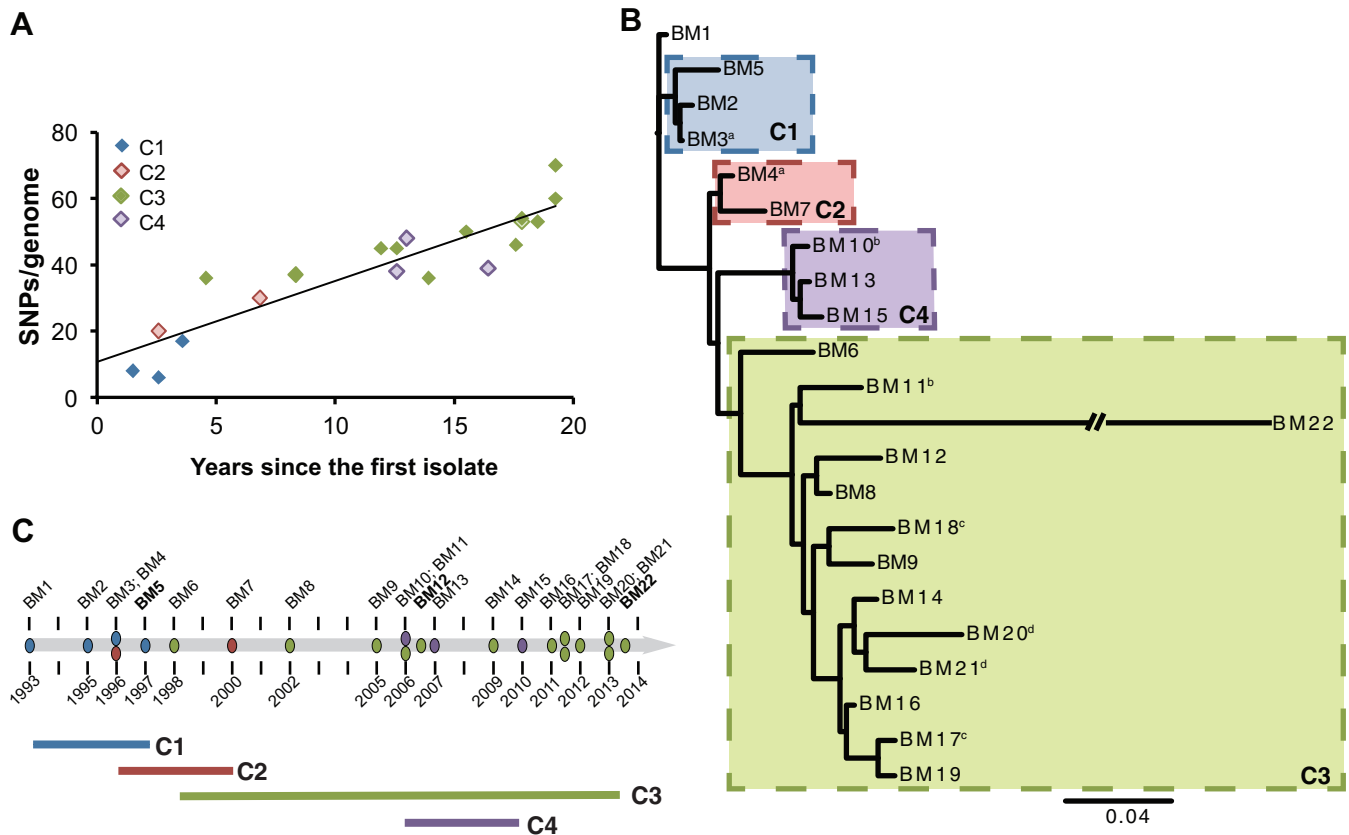


FIG 1 Genomic phylogeny reveals the long-term coexistence of diverse clades. Colors denote phylogenetic clades C1 to C4. (A) Number of SNPs distinguishing each isolate from the first (BM1) over time. The mutator isolate BM22 was excluded from this analysis. A linear fit with a slope of 2.4 mutations per year is shown. (B) Maximum *a posteriori* phylogenetic tree of the 22 isolates modeling both SNPs and indels as independent mutational processes. Shared superscripts in the strain labels indicate that isolates were obtained on the same date. (C) The temporal distribution of isolate samplings enables the inference that clades coexisted. Isolate labels in bold font are nonmucoid (see Fig. S1 in the supplemental material).

a hypermutator phenotype. The estimated mutation rate of all isolates was 4.0 SNPs/year ($r = 0.6$), but if the last isolate is excluded, this rate drops to 2.4 ($r = 0.9$) (Fig. 1A), which is within the range observed for other long-term infections with *P. aeruginosa* and *B. dolosa* in CF patients (5, 12). We used both SNPs and indels as independent mutational processes (21) to determine the evolutionary relationship among the 22 isolates (Fig. 1B). The Bayesian maximum *a posteriori* phylogenetic tree revealed the evolution of an early clade, C1, from the first isolate, followed shortly by a lineage that quickly diversified into clades C2, C3, and C4. Note that this analysis produces an unrooted tree that nevertheless points to the node between BM1 and C1 as the common ancestor, and it reveals that indels resolve the phylogeny more completely than an SNP-based tree, which places the roots of C2 to C4 in an unresolved polytomy. The distribution of these clades over time (Fig. 1C) suggests that at least 2 to 3 distinct lineages coexisted at any given time, although the last 7 isolates were all from clade C3. Moreover, isolates BM3/BM4 and BM10/BM11, each a pair that was recovered simultaneously, clustered in different clades in the phylogenetic tree, consistent with the coexistence of different lineages within the same infected CF patient. These results provide further compelling evidence that sampling single clones from CF patient airways over time significantly undersamples the diversity within the prevailing infection.

Early diversification results from strong selection on genes affecting regulatory functions, metabolism, and envelope biogenesis. We identified mutations associated with each node that defined the four clades in the phylogenetic tree (Fig. 1B and 2A, and see Table S4 in the supplemental material). Most mutations,

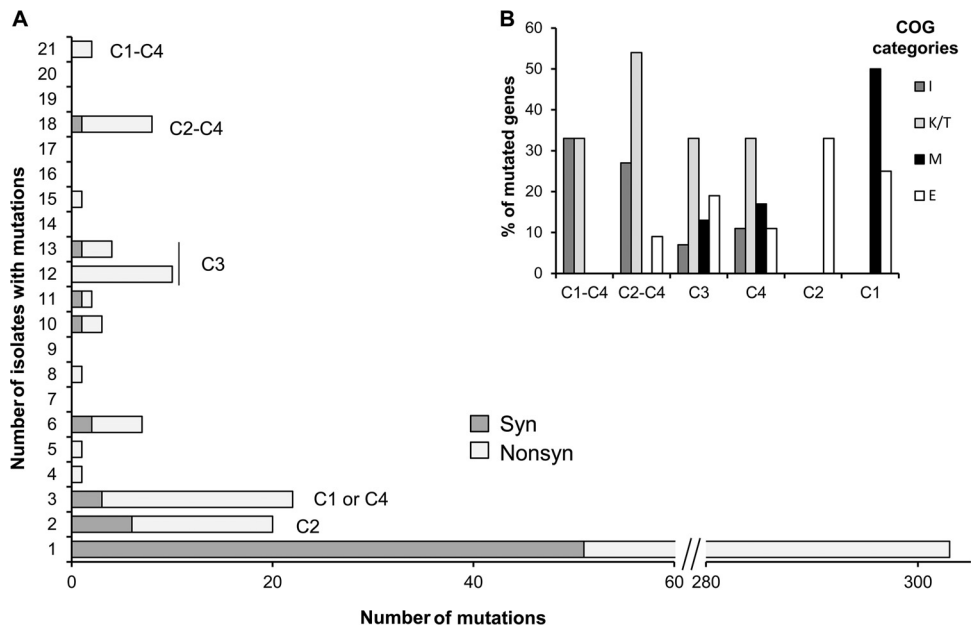


FIG 2 Most mutations were nonsynonymous, especially at the nodes of new clades, with affected genes encoding regulatory functions, lipid metabolism, and envelope biogenesis. (A) Numbers of synonymous and nonsynonymous mutations shared by each of the isolates and associated with nodes in the phylogeny (Fig. 1B); (B) distribution of the synonymous and nonsynonymous mutations at each node by clusters of the following orthologous groups: lipid transport and metabolism (COG I), transcription and signal transduction (COGs K and T), cell wall/membrane/envelope biogenesis (COG M), and amino acid transport and metabolism (COG E).

especially those defining clades and shared among multiple isolates, are nonsynonymous (Fig. 2A) (Fisher exact test, one tailed, $P = 0.0155$). Of the 64 mutations associated with lineage splitting, 41 were either nonsynonymous or indels, 5 were synonymous, and 18 were intergenic, most of which were proximal to coding sequences (CDSs) and hence near promoters. This enrichment of nonsynonymous coding mutations (41/46) is significantly greater than the expectation of 27.8% synonymous mutations under neutral evolution given the codon usage and percent G+C content of this species (22). This suggests that the infection involved periods of diversification dominated by positive selection followed by intervals of relatively neutral evolution.

The nodes at the base of the four clades were enriched in functional categories that shifted over time. The first mutations shared by all clades, or all except C1, were enriched in mutations (SNPs and indels) in the categories of transcription/signal transduction (clusters of orthologous groups [COGs] K and T) and lipid metabolism (COG I) (Fig. 2B; Table S4 in the supplemental material). Mutations producing clades C3 and C4 were also enriched in transcription/signal transduction mutations (COGs K and T) but also in amino acid transport and metabolism (COG E) and cell wall/membrane/envelope biogenesis (COG M) mutations. The mutations at the roots of clades C1 and C2 affected genes from COGs E and M, and COG E, respectively (Fig. 2B; Table S4). This analysis reveals that new lineages evolved mainly by mutations in genes with regulatory/signaling roles (31.3%) and in genes whose proteins are involved in lipid, amino acid, and carbohydrate metabolism (32.8%), possibly having an impact on envelope biogenesis.

The most obvious candidates for genes under selection during this chronic *B. multivorans* colonization are mutations that became fixed over time. Only three mutations were found in all isolates from clades C1 to C4, including a nonsynonymous mutation in the gene encoding a FabD homolog (a malonyl coenzyme A [CoA]:acyl carrier protein [ACP] transacylase) involved in fatty acid biosynthesis, an insertion of 76 amino acids into cell division protein FtsK, and a point mutation 181 bp upstream of the coding region of a DNA-binding regulatory protein (Table 1). Twelve mutations were found in all isolates from clades C2, C3, and C4. These include genes encoding the diguanylate

TABLE 1 List of mutations that fixed in the infecting population and their association with phylogenetic clades C1 to C4, as shown in Fig. 1^a

Gene locus	Annotation	Mutation category	Type of mutation	Effect in protein	Presence of mutation in clade:			
					C1	C2	C3	C4
BMD20_16520	ACP S-malonyltransferase FabD	CDS	SNP, nonsyn	Y89H	+	+	+	+
BMD20_24750	Cell division protein FtsK	CDS	Indel, 76-aa insertion		+	+	+	+
BMD20_24880	DNA-binding regulatory protein, YebC/PmpR family	INT (181-bp upstream CDS)	SNP		+	+	+	+
BMD20_02495	Diguanylate cyclase/phosphodiesterase RpfR	CDS	SNP, nonsyn	E185D	–	+	+	+
BMD20_04285	Acetyl-CoA carboxylase, carboxyl transferase subunit beta AccD	CDS	SNP, nonsyn	T240M	–	+	+	+
BMD20_05445	Transcriptional regulator/2-aminoadipate aminotransferase	CDS	SNP, syn	A362A	–	+	+	+
BMD20_06065	Two-component heavy-metal response transcriptional regulator	CDS	SNP, nonsyn	E229D	–	+	+	+
BMD20_08755	LysR transcriptional regulator	CDS	SNP, nonsyn	A24E	–	+	+	+
BMD20_10060	DNA-binding protein	CDS	SNP, nonsyn	I228M	–	+	+	+
BMD20_11660	OmpR-like response regulator	INT (16-bp upstream CDS)	SNP		–	+	+	+
BMD20_14360	ABC transporter permease	CDS	SNP, nonsyn	P12T	–	+	+	+
BMD20_14695	Multidrug resistance efflux pump membrane protein EmrA	CDS	SNP, nonsyn	Q97H	–	+	+	+
BMD20_16530	Glycerol-3-phosphate acyltransferase PlsX	INT (9-bp upstream CDS)	SNP		–	+	+	+
BMD20_16530	Glycerol-3-phosphate acyltransferase PlsX	INT (10-bp upstream CDS)	SNP		–	+	+	+
BMD20_17700	Nucleoside-diphosphate sugar epimerase Wbil	CDS	Indel, frameshift		–	+	+	–

^aCDS, coding sequence; INT, intergenic region; aa, amino acid; nonsyn, nonsynonymous; syn, synonymous.

cyclase/phosphodiesterase RpfR (also known as YciR), which is also the receptor of the *Burkholderia* diffusible signal factor (BDSF), the subunit beta of acetyl-CoA carboxylase, which is involved in fatty acid biosynthesis, four proteins involved in transcription/signal transduction, and two transporters, including a multidrug resistance efflux pump membrane protein homolog to EmrA. Three other mutations were in intergenic putative-promoter regions, two of them upstream of the coding sequence of protein PlsX, involved in the biosynthesis of phospholipids, and another one upstream of the coding sequence of a OmpR-like response regulator (Table 1). A last mutation common to all isolates from clades C2 and C3 was a frameshift within gene *wbil* (BMD20_17700), encoding a nucleoside diphosphate sugar epimerase involved in the synthesis of the O-antigen repeats of the lipopolysaccharide (LPS). Thus, two haplotypes, one derived from the other, rose to high frequency within this infection in its first 3 years.

However, the 41 genes in which we detected multiple independent mutations may provide even stronger and more informative evidence of selection, given that some mutations at nodes may have risen to high frequency by linkage to another selected mutation and not by any directly selected phenotype. Of the 218 independent mutations found in all isolates except the last hypermutator isolate, we expected none to occur in the same gene three or more times by chance ($P < 0.005$) among the 5,809 genes in this *B. multivorans* genome (Fig. 3A). Notably, eight genes were found to have three or more different mutations, five genes had four or more mutations, and two genes had five or more mutations. Eight of these polymorphic genes with three or more mutations within the coding sequence and two other genes with mutations upstream of the CDS are detailed in Fig. 3B (see also Table S4 in the supplemental material). Three of these polymorphic genes encode proteins involved in regulating transcription and signal transduction that may influence pathogenesis. One gene with six mutations (BMD20_13505, with 5 nonsynonymous mutations and 1 frameshift mutation) that evolved independently in clades C2, C3, and C4 encodes a Fis family transcriptional

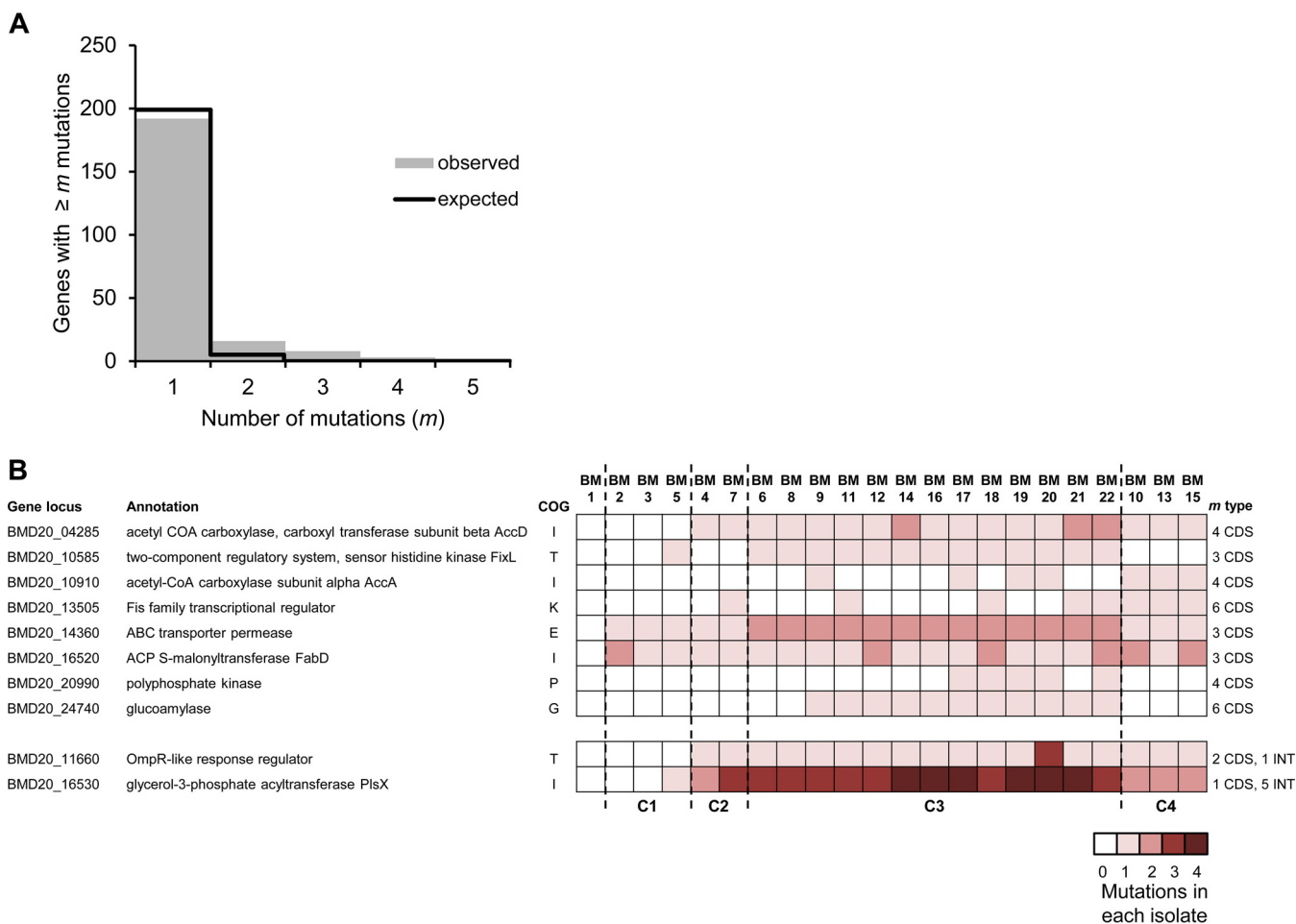


FIG 3 Strong parallelism in mutated genes defines adaptations to the environment of CF patient lungs. (A) Observed and expected numbers of genes with at least *m* mutations given the number of total mutations found. Genes with multiple mutations are in significant excess ($P < 0.005$). (B) Heatmap of mutation frequencies in the 10 loci that acquired ≥ 3 mutations, including point mutations upstream of the coding sequence that might be relevant for gene expression. Isolates are clustered by clade. (Right) Total numbers of mutations acquired per gene (*m*) and the type of mutation (coding sequence [CDS] or intergenic [INT]).

regulator conserved across the *Burkholderia* genus, but the genes that it regulates are currently unknown. Two genes with three mutations were the *fixL* homolog BMD20_10585 and the *ompR*-like response regulator of a two-component signal transduction system (BMD20_11660). BMD20_10585 is a homolog of BDAG_01161 from *B. dolosa*, the *fixL* homolog that accumulated the highest number of mutations in the 112 clinical isolates analyzed by Lieberman and colleagues (12). Homologs of this gene have been implicated in oxygen sensing in other microorganisms (23), and due to the low-oxygen tension of CF patient lungs, it is possible that alterations in this *Burkholderia* regulator are involved in adaptation to microaerophilic conditions. Another gene, BMD20_11660, encodes a response regulator homologous to the osmolarity response regulator *OmpR* and is located upstream of a region encoding a sensor signal histidine kinase. Multiple mutations in the ortholog of this locus in *B. cenocepacia* populations were associated with adaptation during long-term experimental evolution in a biofilm system (24). Although the genes regulated by this two-component signal transduction system are unknown, we demonstrate that the point mutation 16 nucleotides upstream of the start codon increased the transcription of the BMD20_11660 gene in two isolates tested from clade C3 and one from clade C4 relative to that of isolate BM1 or BM2, which lacked any mutation (Table S5). To summarize, mutations in genes associated with regulation of the cellular response to reduced oxygen levels and perhaps also to osmolarity were under selection in multiple independent isolates.

The second group of genes that acquired 3 to 6 different mutations (encoding acetyl-CoA carboxylase subunits alpha and beta, the FabD homolog, and the glycerol-3-phosphate acyltransferase *PlsX*) are all involved in the metabolism of fatty acids/phospholipids. Particularly notable are the six mutations that accumulated in the *plsX* gene (1 nonsynonymous mutation and 5 point mutations in the untranslated region upstream of the start codon). Three other genes that acquired multiple mutations encode an ABC transporter permease (3 mutations), a glucoamylase (4 nonsynonymous point mutations and 2 frameshift mutations, all in clade C3 isolates) that was also mutated in *B. dolosa* isolates (12), and polyphosphate kinase (all mutations are in clade C3 isolates), an enzyme important for virulence and response to stress in many bacterial pathogens (25). In summary, the coexisting subpopulations within this infection each faced similar selective pressures favoring independent mutations in several regulatory and metabolic genes.

The first-evolved phenotypes included reduced growth rates and altered O-antigen repeats. Given strong genetic evidence of positive selection during the infection, we sought to determine the associated mutant phenotypes that could have been selected. First, the *in vitro* doubling times and biomass yields of the isolates were estimated in synthetic cystic fibrosis medium. Compared to the doubling times of 6 h for isolate BM1 and of 6.2 to 6.3 h for clade C4 isolates, all isolates from clades C1 ($P < 0.01$), C2, and C3 ($P < 0.05$) needed between 7 and 15 h to duplicate (see Fig. S2A in the supplemental material). These differences are even more pronounced for biomass, in which all isolates after BM1 exhibited reduced output ($P < 0.001$) (Fig. S2B). Only four mutations distinguish isolate BM1 from all subsequent isolates, implying that at least one is responsible for this slow-growth phenotype according to the logic of parsimony. It remains possible that multiple causes of slow growth may have evolved, however. One such mutation likely to influence growth is a 76-amino-acid insertion at position 220 of protein FtsK, a septum-located DNA translocase that coordinates the late stages of cytokinesis and chromosome segregation during cell division. The other mutations altered a noncoding region upstream of an uncharacterized DNA-binding protein similar to YebC/PmpR, a malonyl CoA:acyl carrier protein (ACP) transacylase, and an ABC transporter permease possibly transporting amino acids. Multiple independent mutations involved this permease gene, which is evidence of strong selection for a putative transport function.

The finding of three different mutations accumulating in a locus homologous to the *wbi* gene cluster led us to predict that biosynthesis of the LPS O antigen had been altered (26). LPS plays an important role in *Burkholderia* virulence, and O-antigen repeats are known to interfere with adhesion to abiotic surfaces and bronchial epithelial cells (27, 28). One mutation in the *wbiG* gene (BMD20_17710), a nucleotide sugar epimerase/dehydratase, was found in all isolates from clade C1 (BM2, BM3, and BM5). This nonsynonymous mutation altered the conserved region ¹⁶⁸PPLV¹⁷³ to ¹⁶⁸PLLV¹⁷³. We assessed whether this mutation altered the LPS O-antigen repeats in the bacterial outer membrane and found that the two tested isolates from clade C1 do not produce the O-antigen repeats of LPS but that isolate BM1 presents a smooth LPS (Fig. 4). Complementation of BM2 and BM3 isolates with the BM1 *wbiG* allele expressed from pBBR1MCS under the native promoter restored the O-antigen repeats in those isolates (Fig. 4). Another mutated gene was *wbil* (BMD20_17700), encoding another nucleotide sugar epimerase/dehydratase. Deletion of one nucleotide in *wbil* occurred in all isolates from clades C2 and C3 and led to a frameshift mutation. As predicted, representative isolates from clades C2 and C3 also lacked the O-antigen portion of LPS (Fig. 4). However, the three isolates from clade C4 had no *wbi* mutations, potentially by reversion of the *wbil* frameshift mutation, and displayed the same LPS O-antigen pattern as isolate BM1 (Fig. 4). Further, isolate BM22 has a nonsynonymous mutation in *wbiF* (BMD20_17715), encoding a glycosyltransferase, in addition to the *wbil* mutation. We conclude that clones with mutations in the *wbi* locus were favored during selection and produced altered LPS O antigen.

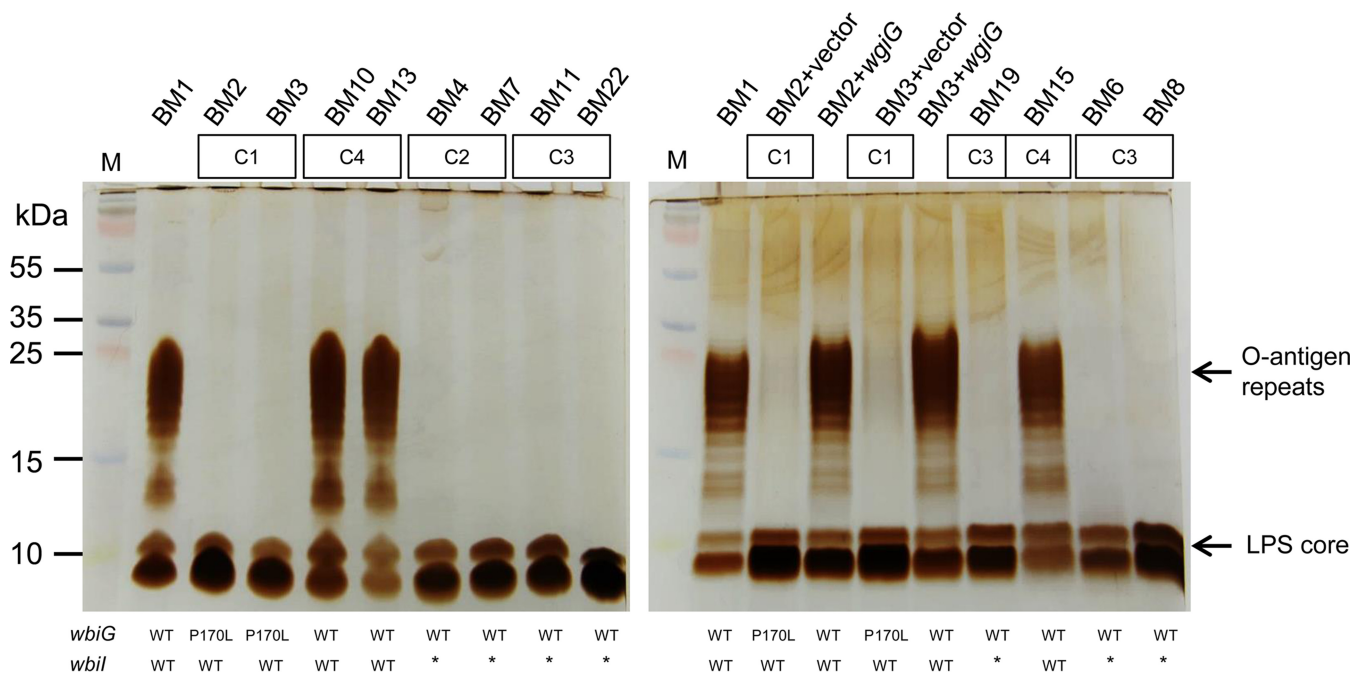


FIG 4 Evolved changes in the LPS pattern and the causative mutation. (Left) Electrophoretic profiles of LPSs extracted from *B. multivorans* isolates representative of all clades and *wbi* alleles, in which the *wgiG* mutation resulting in P169L and a frameshift mutation in *wbiI* are associated with the loss of the O antigen. (Right) BM2 and BM3 were complemented with the empty vector (pBBR1MCS) or a vector carrying a functional *wgiG* copy from isolate BM1, which restores the O antigen. Lanes M, protein markers. WT, wild-type allele; *, *wbiI* frameshift mutation.

Evolutionarily successful lineages originated from mutations affecting antibiotic resistance and cyclic-diguanylate signaling.

The phylogeny (Fig. 1B) shows that most of the evolved diversity and all successful lineages trace to the root of clades C2, C3, and C4, which dates to the first 2.5 years of the *B. multivorans* infection. There are 12 mutations at this node (7 nonsynonymous, 1 synonymous, and 4 non-coding), which makes identifying the driver(s) of this diversification challenging. One potential and obvious target of selection is increased antimicrobial resistance. Among the 12 mutations at this node, one (Q97H) occurred in BMD20_14695, a homolog of the *emrA/hlyD* gene encoding the periplasmic adapter protein (also known as the membrane fusion protein) of a bacterial tripartite efflux pump putatively involved in the extrusion of toxic compounds (29). To test this hypothesized function, the different isolates were assayed for resistance to piperacillin plus tazobactam, a combined antibiotic therapy frequently prescribed to individuals with cystic fibrosis. Although piperacillin and tazobactam are a β -lactam and a β -lactamase inhibitor, respectively, *Burkholderia* resistance against these antibiotics may involve efflux pumps rather than β -lactamases, as has been suggested for *P. aeruginosa* (30). Compared to that of BM1 and the early clade C1 isolates, resistance among the 18 isolates harboring the mutated allele of the *emrA* gene was increased ($P < 0.001$) (Fig. 5A). The Q97H mutation maps to the beginning of the α -helical hairpin domain, which in other EmrA homologs has a coiled-coil arrangement for interaction with TolC-like outer membrane proteins to form sealed periplasmic channels (31). Alignment of the amino acid sequence of EmrA with the Q-to-H mutation at position 97 (EmrA_{Q97H}) to sequences of other *Burkholderia* homologs showed the replacement of Q97 by the positively charged amino acids R or K, but never H. It remains to be determined whether this mutation is the basis for the increased resistance to piperacillin plus tazobactam of isolates from clades C2, C3, and C4. Consistently with this data, the antimicrobial susceptibilities to piperacillin plus tazobactam of the *B. multivorans* isolates under study and of the discarded ones carried out by the clinical laboratory showed that the first isolates were susceptible, followed by a period during which the majority were resistant (Fig. 5B). The arrows in Fig. 5B indicate the periods of antimicrobial therapy administration to this patient. Interest-

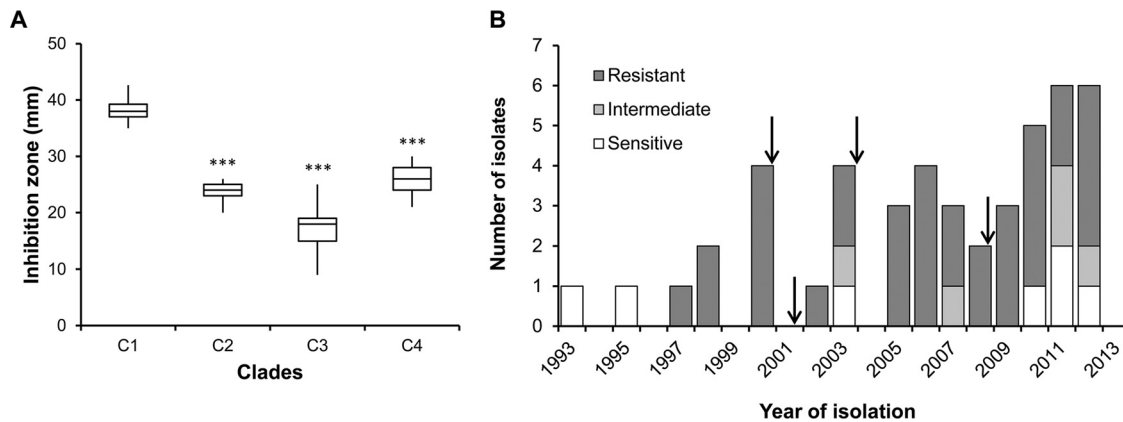


FIG 5 The evolution of antibiotic resistance associates with a mutation in an efflux pump. (A) Susceptibilities to piperacillin plus tazobactam determined at 37°C after 24 h of incubation by measuring the diameter of cell growth inhibition. Isolates from clades C2, C3, and C4 have the *emrA* mutation. Clade C1 isolates (including also isolate BM1) lacking the mutation differed significantly from those with the mutation (***, $P < 0.001$) by Tukey's HSD multiple-comparison test for unequal group sample sizes. (B) Temporal variations in levels of resistance to piperacillin plus tazobactam of all *B. multivorans* colonies recovered from the CF patient under study. Arrows indicate the date that the patient was subjected to an antibiotic therapy.

ingly, following antibiotic therapy, only resistant bacteria were isolated, but in the subsequent 1 to 3 years of no-antibiotic pressure, some susceptible isolates were again recovered. Similar trends were observed with other antibiotics (see Fig. S3 in the supplemental material).

Another mutation at the node preceding the diversification of clades C2, C3, and C4 and thus shared by all derived isolates arose in gene BMD20_02495, the *rpfr* homolog that encodes the cis-2-dodecenoic acid (BDSF) receptor (Table S4). This protein is involved in cyclic di-GMP (c-di-GMP) metabolism mainly as a diguanylate phosphodiesterase, which has been shown to promote motility and reduced biofilm formation in *Burkholderia cenocepacia* (24, 32, 33). This mutation is the only 1 of the 12 at this node with a clear connection to c-di-GMP metabolism. The E185D amino acid substitution occurred at a strictly conserved glutamate residue (see Fig. S4 in the supplemental material) that is part of the Per-Arnt-Sim (PAS) or BDSF-sensing domain. As BDSF molecules have been shown to bind the PAS domain of *B. cenocepacia* RpfR and stimulate its phosphodiesterase activity (34, 35), alteration of this domain can alter affinity to BDSF signaling and consequently disrupt the phosphodiesterase activity of RpfR. We therefore quantified levels of c-di-GMP in cells grown for 5 h in salt-mannitol (SM) medium (16). As predicted, all tested isolates with the RpfR_{E175D} mutation exhibited increased c-di-GMP levels ($P < 0.05$ for BM12 and BM20; $P < 0.01$ for BM21; $P < 0.001$ for BM11) (Fig. 6A). However, two tested isolates (BM12 and BM21) acquired subsequent mutations in genes with putative diguanylate cyclase (GGDEF) or phosphodiesterase (EAL) domains, which may also have altered c-di-GMP levels, but this compensatory phenotype has not been tested. Because deleting RpfR reduces *B. cenocepacia* motility (35), we also measured this trait for these isolates. Figure 6B confirmed the decrease in swimming motility of the isolates harboring the RpfR_{E175D} mutation ($P < 0.001$), suggesting that increased levels of c-di-GMP reduce swimming motility.

Increased biofilm formation and adhesion to epithelial cells play a role in bacterial adaptation within CF patient lungs. Since c-di-GMP is known to regulate several virulence determinants, including biofilm formation (36–41), we quantified biofilm production by all isolates (Fig. 7A). Isolates from clades C2 and C3 produce more biofilm than the first isolate as well as isolates from clades C1 and C4 ($P < 0.001$). Many factors beyond c-di-GMP levels, including adhesion mechanisms, the biosynthesis of various polymers, quorum-sensing signaling, and LPS (38), contribute to biofilm formation. The higher levels of c-di-GMP in clade C3 isolates and possibly also in C2 and C4 isolates (all harboring RpfR_{E185D} and mutations in lipid biosynthesis metabolic enzymes) correlated with the observed increase in biofilm biomass, except in isolates

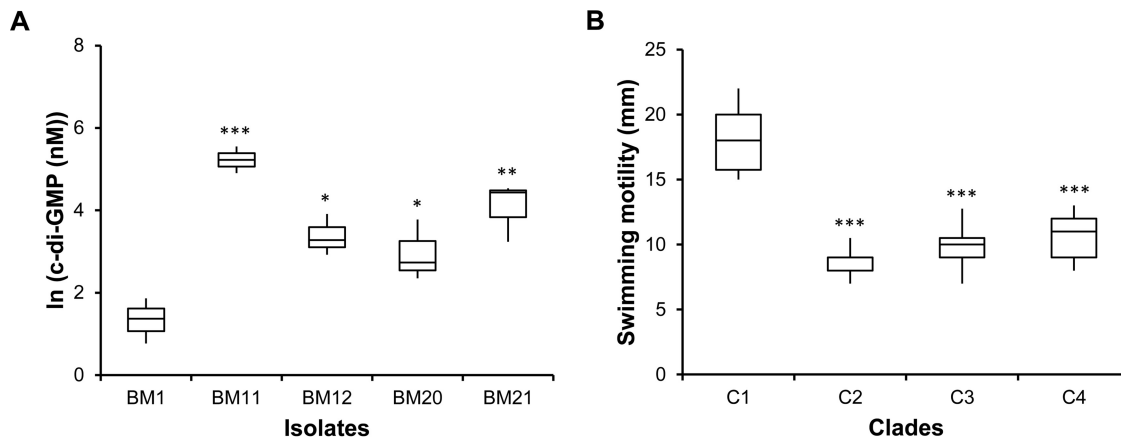


FIG 6 Later isolates from the dominant clade, C3, exhibit higher c-di-GMP levels and less motility than earlier isolates. (A) Intracellular c-di-GMP concentrations in BM1 and clade C3 (BM11, BM12, BM20, and BM21) isolates following 5 h of growth, as measured by LC-MS/MS. Significantly greater c-di-GMP levels were detected in evolved isolates than in the first isolate, BM1 (*, $P < 0.05$; **, $P < 0.01$; ***, $P < 0.001$), by Tukey's HSD multiple-comparison test. (B) Swimming motility (assayed in 0.3% of agar plates incubated at 37°C for 24 h) was significantly less for isolates from clades C2 to C4 than for isolates from C1 (which includes isolate BM1) (***, $P < 0.001$ by Tukey's HSD multiple-comparison test).

BM10, BM13, and BM15 from clade C4. Nevertheless, clade C4 isolates display a smooth LPS in which the O-antigen repeats can possibly mask surface structures affecting the adhesion of bacteria to surfaces, as was reported for *Burkholderia pseudomallei* (42).

In agreement with the biofilm assays on polystyrene, isolates from clades C2 and C3 also adhered better to cystic fibrosis patient bronchial epithelial CFBE410⁻ cells ($P < 0.001$ for C3 and $P < 0.05$ for C2) than did isolate BM1 and clade C1 isolates (Fig. 7B). Again, isolates from clade C4 were the least adherent, potentially owing to the presence of the LPS O-antigen repeats. In *B. cenocepacia*, the lipopolysaccharide O antigen prevents adhesion to epithelial cells (27). In summary, variation in biofilm production and adherence to epithelial cells were likely traits selected to increase during this chronic infection, but notably, variation in adherence was observed throughout the infection.

The highest rate of lung function deterioration parallels genotypic and phenotypic diversification of the isolates and alteration of the lung microbiota. The vital lung capacity of this patient, as determined by forced expiratory volume in 1 s (FEV₁) (percentage predicted) showed a 1.9% decline per year, indicative of a mild infection. Nevertheless, the plot of FEV₁ over time apparently followed three periods with different rates of decline (Fig. 8A). To test this interpretation, we fit multiple continuous models to the temporal FEV₁ measures for comparison to a linear model and evaluated for greatest quality using the Akaike information criterion (AIC) and JMP

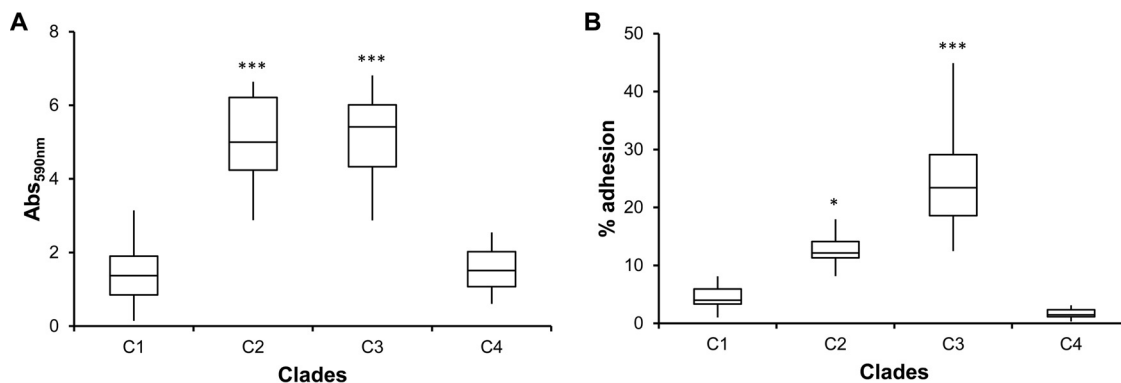


FIG 7 Biofilm formation and adhesion to cystic fibrosis bronchial epithelial cells increase during long-term colonization of CF patient lungs. Biofilm formation (A) and adhesion to CFBE410⁻ cystic fibrosis lung epithelial cells (B) increase for clade C2 and C3 isolates (*, $P < 0.05$; ***, $P < 0.001$ by Tukey's HSD multiple-comparison test). Clade C1 also includes isolate BM1.

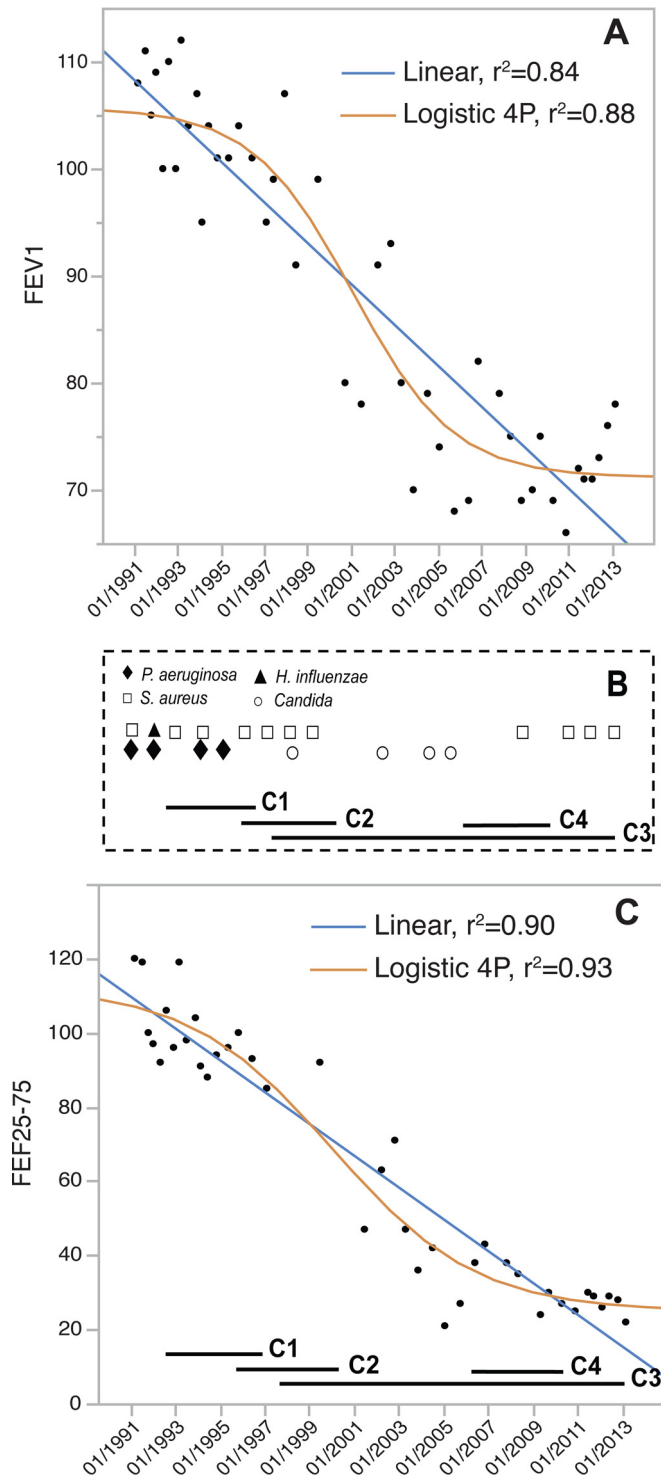


FIG 8 Lung function decline correlates with the evolution of clade C3 and mutations influencing LPS biosynthesis, motility, adhesion to surfaces, and cyclic diguanylate signaling. Forced expiratory volume in 1 s (FEV_1) (A) and forced expiratory flow over the middle half of the forced vital capacity (FEF_{25-75}) (C) as predicted percentages. Patient respirometry data and clinical microbiological reports from sputa were collected during a previous retrospective study (15) and were available 2 years prior to the detection of *B. multivorans* and in the subsequent 20 years of infection. Both linear and 4-parameter (4P) logistic functions (shown), as well as quadratic, cubic, and logistic 3-parameter and 5-parameter models (not shown), were fit to the data and evaluated for the greatest quality using the Akaike information criterion (AIC). The upward trend in FEV_1 in the last 2 years slightly favored a cubic fit over other models (FEV_1 , linear AIC = 296.5, quadratic AIC = 292.4, logistic 5 parameters = 288.5, logistic 4 parameters = 287.2, cubic = 285.1), whereas the logistic

(Continued)

v.11.0. We found that a cubic function producing a sigmoidal curve, followed closely by a four-parameter logistic function, provided a better fit to the data than a linear function or any other models (3- or 5-parameter logistic functions or quadratic functions) (see the legend of Fig. 8A). This supports a period of more rapid decline consistent with the rise of clade C3 isolates. In the first interval from 1989 to 1997, FEV₁ declined 1.2% per year, in which isolates from clade C1 prevailed in coinfections with *P. aeruginosa* and *Staphylococcus aureus* (Fig. 8B). Between 1998 and 2006 the decline was 3.2% per year and corresponds to the rise of clade C3 isolates. The cooccurring pathogens during this period also differed, with frequent detection of *Candida*. After 2006, the decline was 0.7% per year, corresponding to the sampling of isolates mainly from clade C3 but also the three isolates from clade C4. In this period, no *Candida* species were reported in the sputum, but *S. aureus* was isolated. When the forced expiratory flow over the middle half of the forced vital capacity (FEF_{25–75%}; thought to be indicative of the function of small/peripheral airways) was plotted against time, the same dynamics were observed but with greater magnitudes of functional loss, namely, 3.1%, 8.5%, and 2.3% declines per year for the three intervals (Fig. 8C). Here, a four-parameter logistic function fit to these FEF_{25–75%} data provided greater explanatory power than a linear function or any other models (Fig. 8C legend). Although the sigmoidal models only modestly increase the fit to the data over a linear function that explains these spirometry dynamics well (increased r^2 values of 0.05 for FEV₁ and 0.03 for FEF_{25–75%}), they also reduce the root mean square error (RMSE) by 15% and hence increase the predictive powers of both functions. This relative increase in the rate of decline in lung capacity coincides temporally with the rise of clade C3 as the dominant fraction of the *B. multivorans* infection. To summarize, both the phylogenetic composition of the *B. multivorans* isolates and the coinfecting pathogens correlated with the rate of decay in lung function.

DISCUSSION

B. multivorans is the most commonly isolated *Burkholderia* species from chronic infections of the airways of CF patients, with an overall prevalence in the United States of 0.68% (43). However, our understanding of the traits required for bacterial colonization and persistence, as well as the molecular mechanisms underlying this adaptation, is limited. In this comprehensive study, we combined current genome sequencing technologies and phenotypic profiling of 22 longitudinally collected *B. multivorans* isolates from a single CF patient to gain new insight into the cause and consequences of pathogen diversification in chronic CF infections. We also associated these changes in the bacterial population with changes in the patient's lung function. Not only did multiple evolutionary lineages coexist within this single patient for decades, multiple mutations whose phenotypes were almost certainly selected as adaptations evolved in relatively few loci and may have contributed to lung function decline.

Comparing the genomes of the 22 *B. multivorans* isolates revealed the steady accumulation of relatively subtle mutations over time, including 38 short insertions, 50 deletions, and 399 single-nucleotide changes. Similar evolutionary patterns were reported previously for *B. dolosa* (12, 13) and *P. aeruginosa* (5, 6, 44). The exception to this dynamic involved the final isolate from this study, BM22, which represents a lineage that acquired a frameshift mutation in *mutL* that led to an increased mutation rate and 196 new mutations. Mutators have often been described in *P. aeruginosa* isolates from CF patients (5, 45) and recently also for *B. cepacia* complex CF isolates (46). Isolate BM22, isolated in the year 2013, clusters with BM11 sampled in 2006, indicating that the

Figure Legend Continued

4-parameter model provided the best fit for FEF_{25–75%} (linear = 307.5, quadratic = 303.0, logistic 5 parameters = 300.5, cubic = 299.4, logistic 4 parameters = 298.1 [lower is better]). This supports a period of more rapid decline consistent with the rise of clade C3. (B) Temporal distribution of microorganisms (see Table S2 in the supplemental material) coinfecting the lungs of the CF patient under study, as well as the phylogenetic structure of the detected *B. multivorans* isolates.

mutL mutation might have occurred between 2006 and 2013. During the period of 2006 to 2013, the patient was treated with ceftazidime, tazobactam, and cloxacillin in 2008 and azithromycin, sulfamethoxazole, and trimethoprim in 2013, 8 months before the isolation of BM22, but neither these therapeutic interventions nor other coinfecting species correlated clearly with the *mutL* mutation. Mutator alleles likely encode no direct benefit but instead rise in frequency as traits linked to other adaptive mutations, but the large number of new mutations in this strain prevents inference of the conditions under which it evolved. Nevertheless, the side effect of mutator lineages is their accelerated evolutionary rate, making them inherently more difficult to control with antibiotics (47).

The phylogenetic analysis demonstrated that an initial infecting strain diversified into at least four major subpopulations (clades C1, C2, C3, and C4), each with its distinct genetic signature of mutations, with at least two of them coexisting at any time. Despite this diversity, clade C3 became the dominant lineage, not only because most isolates belong to this clade but also because they span a colonization period of 15 years. It is difficult to associate the success of the C3 population with any particular mutation, but the finding of multiple mutations in a sensor histidine kinase homologous to FixL in all isolates bears mentioning. A homolog of FixL has been shown to be involved in oxygen sensing and regulation of gene expression (23), and this gene was previously found to be the most polymorphic among 112 *B. dolosa* clinical isolates in a clonal outbreak of CF (12). Additional mutations found in clade C3 isolates, including those encoding a glucoamylase (also mutated in *B. dolosa* clinical isolates), a porin, and a type III secretion protein, and mutations in genes contributing to the metabolism of amino acids and lipids likely also influenced the success of this clade. These altered proteins may reflect selection on the basis of particular nutritional needs as well as on the basis of altered envelope biogenesis and host interactions. In total, these mutations associated with the diversification of clade C3, and shifts in other coinfecting species also coincided with the period of enhanced decline in lung function of the infected patient (Fig. 8). This evidence motivates the hypothesis that mutations in one or more of these genes were adaptive in this lung infection and may have accelerated the decline in patient lung function.

The mutations at the nodes of each of the clades provide overwhelming evidence of positive selection; only five synonymous mutations accompany 41 nonsynonymous mutations. In addition, most of the mutated genes at these nodes (some of which are also the most polymorphic) are involved mainly in transcription/signal transduction and lipid/amino acid metabolism. This role of regulatory genes in adaptation is echoed by the finding that common mutations in the *P. aeruginosa* *mutA* gene have extensive pleiotropic effects on gene expression and a general reduction of virulence factor expression (48). Here, the *B. multivorans* *rpfR*, *fixL*, and *ompR*-like genes encode proteins with regulatory functions and are likely candidates for governing global regulatory changes that influence host adaptation.

We also observed conspicuous changes in the physiology of these isolates, beginning with a reduced growth rate and biomass. Growth rate declines have also been observed for *P. aeruginosa* isolates obtained during chronic CF patient lung infections (48). Again, the temporal analysis of when these traits evolved suggests that coinciding mutations may explain them, including early mutations in genes encoding cell division protein FtsK and a DNA-binding regulatory protein from the YebC family. Which of these mutations is ultimately responsible and why a reduced growth rate may be adaptive are still unclear, though it might be a response to the new nutritional environment of CF patient sputum or a side effect of greater investment in adhesion and biofilm production.

Another trait under negative selection was the expression of the O-antigen repeats of LPS, a putative virulence factor in other *Burkholderia* organisms (26, 28). Except for BM1 and clade C4 isolates, all the other isolates displayed a rough phenotype from a lack of O-antigen repeats due to mutations in sugar epimerase enzymes. Conversely, *B. dolosa* exhibits a gain-of-function mutation that restores O-antigen presentation in

isolates from several patients (12). It has also been shown that early isolates of the highly transmissible ET12 lineage of *B. cenocepacia* (J2315, BC7, and C5424) lack the O antigen, but the later isolate K56-2 expresses it (26). On one hand, the O antigen is known to interfere with adhesion to abiotic surfaces and bronchial epithelial cells and to stimulate the host immune response (27, 28), and its loss may be beneficial for cells entering a biofilm lifestyle like the one in CF patient lungs. On the other hand, it may inhibit phagocytosis (27), and its loss may enhance bacterial elimination. Notably, *P. aeruginosa* isolates from CF patients with chronic lung infections often have rough LPS, possibly allowing these bacteria to evade the host's antibody-mediated immune response (49). Furthermore, a study of long-term adaptation in biofilm conditions *in vitro* by *B. cenocepacia* revealed mutations in LPS-associated loci, which implies that these mutations may enhance adherence or aggregation independently of host selection (24). Clearly, altered LPS biosynthesis can drive biofilm adaptation, and how this contributes to infections deserves further study.

Yet another metabolic trait under selection during this infection was lipid metabolism. Mutated genes included *fabD* (ACP *S*-malonyltransferase), acetyl-CoA carboxylase, and glycerol-3-phosphate acyltransferase. Remarkably, some strains acquired multiple nonsynonymous mutations in the same gene (*fabD*) (Table S3), suggesting that selection acted to continually refine this trait. The reason for this strong selection is unclear. Changes in lipid metabolism influence membrane composition and can alter the dynamics of membrane-bound signaling complexes, as described previously for the production of small colony variants (SCVs) (50). Further, one study showed a reduction of the degree of saturation of fatty acids produced by infecting *B. cenocepacia* organisms that coincided with the deterioration of pulmonary function in CF patients (51).

Still another evolved trait in some isolates was elevated c-di-GMP pools, which may be expected given its role in regulating adherence. As expected, increased c-di-GMP levels influenced other phenotypes known to be dependent on this signaling molecule, such as motility and biofilm formation (38). The decreased motility of isolates from clades C2, C3, and C4 correlates with increased biofilm formation and adhesion to CF lung epithelial cells, with the exception of clade C4 isolates that express O-antigen repeats in the outer membrane. These implicate loss of motility and increased biofilm production as two other important adaptations that occur either as a direct or an indirect response to selection in CF patient lungs.

Perhaps the best-studied trait associated with bacterial adaptation to CF patient lungs is increased resistance to antimicrobial agents. Several reports describe mutations in genes encoding efflux pumps and its regulators, DNA gyrase, among others, as under selection in *P. aeruginosa* (5, 52) and *B. dolosa* (12). We also observed a mutation in a gene encoding a component of an efflux pump and confirmed that isolates carrying that mutation had enhanced antibiotic resistance (Fig. 5A). All together, our results demonstrate that only a few genetic changes can fundamentally alter the physiology and fitness of bacteria evolving during chronic infections and that evolutionary adaptation can involve mutations of structural genes involved in specific traits, but also regulatory genes possibly affecting global gene expression. These mutations therefore shift the colonizing bacteria toward a "chronic infection phenotype" of slow growth, higher biofilm production, reduced motility, and greater antibiotic resistance (48).

Fortunately, the long-term *B. multivorans* infection of the CF patient under study has been mild, requiring only a few periods of antibiotic therapy and a single hospitalization. Nevertheless, a general decay of lung function occurred at various rates, as is commonly observed from a variety of contributing host and microbiological factors (53). Data available for this patient suggest a higher diversity of the bacterial community before the onset of *B. multivorans* infection, since 2 to 3 years following the first isolate, *P. aeruginosa* was eradicated and only *Candida* and *S. aureus* were identified as coinfecting organisms. Supporting this observation, a recent report shows that CF-associated pathogens such as *Burkholderia*, *Pseudomonas*, *Stenotrophomonas*, and *Achromobacter*, have a strong tendency to dominate the bacterial community and that lower community diversity correlates with worse lung function (54). Therefore, it is not

surprising that the period of accelerated lung function decline was associated with the presence of *B. multivorans* isolates displaying phenotypes usually attributed to chronic colonization, such as lower motility, increased biofilm formation, loss of O-antigen repeats, and antibiotic resistance, as well as diminished overall bacterial diversity. This was not the only change in the lung microbial community at this time, and it is worthy of mention that *Candida* was often isolated in this period, which is consistent with recent observations that yeasts/fungi such as *Candida albicans* contribute to lung function decline (55, 56). Whether the evolution of the *B. multivorans* population was shaped by interactions with other species and/or whether microbial interactions influenced the deterioration of lung function remain important open questions.

We have presented a detailed analysis of the temporal sequence of all of the genetic and many phenotypic changes that lead to the successful establishment of a *B. multivorans* chronic CF patient lung infection. The evolutionary dynamics revealed an initial 2- to 3-year period of adaptation to the CF patient airways, with extensive phenotypic changes and genetic diversification followed by a long period of more-limited phenotypic changes despite relatively constant genomic evolution (Fig. 1). Although *B. multivorans* adaptation to the CF lung seems to follow patterns of mutation similar to those described for *B. dolosa* and even *P. aeruginosa* (5, 12, 13), more isolates of *B. multivorans* from chronic infections should be analyzed to understand the success of this species among bacterial pathogens of CF patient airways. In addition, studying how *Burkholderia* species interact with the microbiomes of patients and correlate with disease status will provide further insights into the biology and decline in function of CF patient lungs.

In summary, we developed a detailed genome-wide model of evolution within a diverse population of *B. multivorans* isolates over more than 2 decades of chronic infection. This analysis clearly identified genes that are under strong positive selection, including mutations that produced long-lived lineages (e.g., *rpfR*) and multiple independent mutations in the same gene that recurrently altered key functions (e.g., *fixL*, *plsX* gene, *hfq*, an *ompR*-like gene, and a *fis*-like gene). Remarkably, 10 loci were found with three or more independent mutations irrespective of the genetic background of the isolates. As these genes are not especially prone to mutation (22), this provides strong evidence that their predicted phenotypes are exceptionally important for successful persistence in the CF patient lung environment. We also assayed a broad array of phenotypes that were either directly or strongly associated with the predicted functions of these driver mutations (e.g., elevated c-di-GMP levels, increased antibiotic resistance, and altered O antigen). Collectively, these strongly selected traits represent a panel that should enable genome-based diagnostics for more-personalized treatment of airway infections in patients with CF and perhaps other chronic infections. This knowledge also points to potential therapeutic targets and can direct efforts for the development of effective treatments against *B. cepacia* complex bacteria.

MATERIALS AND METHODS

Bacterial strains and growth conditions. The 22 investigated bacterial isolates (see Table S1 in the supplemental material) consist of a single clone of *Burkholderia multivorans* sampled from a patient attending a CF clinic in Vancouver, Canada, spanning the period between 1993 and 2013. This patient was chosen based on the fact that the initial isolates and most of the remaining ones were mucoid, but three isolates collected in the years 1997, 2006, and 2013 were nonmucoid and thus examples of the mucoid-to-nonmucoid transition (14). *Burkholderia* isolates were grown in LB or in extracellular polymeric substance (EPS)-producing salt-mannitol (SM) medium (16) at 37°C and were stored at -80°C with 30% glycerol. *E. coli* strains were grown in LB at 37°C. When it was necessary to maintain the selective pressure, growth media were supplemented with antibiotics at the following concentrations: for *B. multivorans*, chloramphenicol at 200 µg ml⁻¹, and for *E. coli*, chloramphenicol at 25 µg ml⁻¹ and kanamycin at 50 µg ml⁻¹.

Genome sequencing, assembly, and annotation. Genomic DNA from the *B. multivorans* clinical isolates was extracted and purified using the DNeasy blood and tissue kit (Qiagen). DNA samples were processed according to Illumina's instructions for generating paired-end libraries. Paired-end libraries of *B. multivorans* D2095 (other designations: VC13401; BM11) and D2214 (VC13534; BM12) were sequenced using the Illumina HiSeq2000 platform at Fasteris SA (Switzerland), the mate-pair library of *B. multivorans* D2095 and the paired-end library of *B. multivorans* VC5602 (BM1) were sequenced using an Illumina HiSeq2500 at BaseClear BV (Netherlands), and paired-end libraries of the remaining 19 *B. multivorans*

isolates were sequenced using an Illumina MiSeq system at the Instituto Gulbenkian de Ciência (Portugal).

B. multivorans D2095 (BM11) was used as the reference genome, and sequencing and *de novo* assembly of this genome were described previously (BioProject accession no. PRJNA239698) (57). In more detail, raw paired-end reads (two 100-bp reads with an insert size of around 300 bp and ~1,000-fold coverage), and an extra paired-end library of D2095 (BM11) obtained with an Illumina HiSeq2000 (~80-fold coverage) and two 50-bp mate-pair reads with an insert size of 5 kb (~70-fold coverage) were corrected using Quake v.0.3 (58). These Quake-corrected data sets were processed using the fastq-mcf tool (59) and filtered based on Phred quality scores. Adapter contamination and low-quality and ambiguous nucleotides were trimmed from the remaining reads. The overall quality assessment of each corrected data set was carried out using FastQC (60). Paired-end reads from the HiSeq2000 were *de novo* assembled using Edena v.3.131028 (61) and SPAdes v.3.0.0 (62). The quality of assemblies obtained by using different parametrizations was estimated by assessing the number of single-nucleotide variants (SNVs) and indels when the filtered HiSeq2000 paired-end reads were mapped against the *de novo* assemblies with Bowtie2 (63), SAMtools v.0.1.19 (64), and SNVer 0.5.2 (65) for format file conversion and single-nucleotide polymorphism (SNP) detection, respectively. The assemblies from Edena and Spades assemblers were merged using GAM-NGS (66), followed by automated improvement using Pilon v.1.6 (67). The three filtered data sets (two paired-end libraries and one mate-pair library) from the Illumina HiSeq2000 were used for scaffolding by SSPACE v.2.3 (68), gap closure by GapFiller v.1.5 (69), and GapCloser v.1.12 (70). The final draft genome sequence of *B. multivorans* D2095 was obtained after automatic improvement using Pilon and manual curation. Genome annotation was carried out using the National Center for Biotechnology Information Prokaryotic Genomes Automatic Annotation Pipeline (NCBI PGAAP), v.2.9.

Detection of SNP and indel mutations. Raw paired-end reads from 21 *B. multivorans* isolates generated with an Illumina MiSeq were filtered based on Phred quality scores. Adapter contamination and ambiguous nucleotides were trimmed off using the fastq-mcf tool (59). Only reads with Phred scores higher than 30 (99.9% accuracy) were considered for subsequent analysis, yielding a mean of 92% of the genomes with ≥ 20 -fold coverage (see Table S6 in the supplemental material). Filtered paired-end data sets were mapped against the reference draft genome sequence of *B. multivorans* D2095 (BM11) using BWA-MEM of BWA v.0.7.10 (71, 72) and NovoAlign v.3.02.07 (Novocraft; website last accessed 6 June 2015) to determine a first list of putative mutations. Each SNP was analyzed according to the Phred score of SNP calls, the SNP calls on forward and reverse reads, and the read depth of the inspected region. Only SNPs and indels detected in alignments using both BWA-MEM and NovoAlign and occurring in at least three forward and reverse reads considered high-quality SNPs were further analyzed. In parallel, SNVer 0.5.3 (65) was used to call variants and to produce .vcf files for both BWA-MEM- and NovoAlign-generated assemblies. A frequency of 90% was considered to call a SNP real, and reads with coverage of lower than 20-fold were carefully evaluated by visual inspection using Geneious v.6.1.8 (73), with which the existence of other variable nucleotides in the same region was considered a strong indicator of error. The SNPs and indels shown in Table S3 were normalized against the genome sequence of the first isolate, VC5602 (BM1).

Phylogenetic analysis. To infer the phylogeny of the 22 *B. multivorans* isolates, SNPs and indel mutations (337 positions) were concatenated separately for each isolate (Table S3) and analyzed as two distinct substitution models (substitutions: Tamura-Nei and indel RS07) with the Bali-Phy package (21). Following multiple runs with a Bayesian Markov chain Monte Carlo (MCMC) approach, a consensus maximum *a posteriori* tree was generated. This approach is explicitly aware of the distinct mutational processes and spectra of SNPs and indels.

Distribution of mutations expected for neutral evolution. To determine the distribution of mutations expected under a model of neutral evolution, we randomly generated 192 mutations in the 5,809 genes of the *B. multivorans* D2095 genome and counted the distribution of the number of mutations obtained per gene. This was repeated 1,000 times, and mutations occurring in the last mutator isolate, BM22, were excluded. A permutation test (R v.3.2.2, Biostrings package v.2.36.3) was then used to determine the probability that any given gene is mutated any number of times.

Molecular cloning and in trans complementation experiments. The region upstream of *BMD20_17715* and the coding sequences of *BMD20_17715* (*wbiF*) and *BMD20_17710* (*wbiG*) were amplified from *B. multivorans* BM1 genomic DNA by PCR using the primers P1 (CTAGGGTACCGTGCCGACCTATCCTGCC [italics indicate restriction sites]) and P2 (CGTCTAGATATCGGGAGATGCATGTCG). The amplification conditions were 2 min at 95°C, 30 cycles of 95°C for 45 s, 59°C for 30 s, and 72°C for 2 min, and a final extension of 7 min at 72°C. The resulting 1,951-bp product was digested with KpnI and XbaI and cloned in the same restriction sites of pBBR1MCS. After *E. coli* DH5 α transformation and clone selection, the plasmid obtained was named pLM015-12 (see Table S1 in the supplemental material). Plasmid pLM015-12 was mobilized into *B. multivorans* isolates by triparental mating using *E. coli* DH5 α (pRK2013) as a helper strain (74). Exconjugants from the transfer of pLM015-12 were selected on LB agar plates supplemented with 200 μ g of chloramphenicol ml⁻¹ and 20 μ g of gentamicin ml⁻¹.

Growth rate and doubling time determination. Isolates were grown overnight in liquid synthetic cystic fibrosis medium (SCFM) (75) at 37°C. A volume was then centrifuged (1 min, 10,000 \times g), and the pellet was washed with saline solution (0.9% NaCl) and used to inoculate a microplate with 200 μ l of SCFM in each well (5 wells for each isolate; 12 isolates per plate), generating an initial optical density at 640 nm (OD₆₄₀) of 0.05. Microplates were incubated in a microplate reader at 37°C with shaking for 5 s every 30 min, and growth rates were measured by monitoring the OD₆₄₀ for 22 h. Growth rates were calculated from the exponential phase of growth from at least three independent experiments. Statistical

analysis was carried out using Excel with QI Macros 2016. One-way analysis of variance (ANOVA) and Tukey's honestly significant difference (HSD) multiple-comparison test for unequal group sample sizes were performed to determine statistically significant differences. Differences were considered statistically significant if the *P* value was lower than 0.05.

Analysis of lipopolysaccharide and exopolysaccharide. LPS was extracted as described previously (76), with minor modifications. Briefly, cells from overnight plate cultures were suspended in a lysis buffer containing DNase and incubated for 30 min at 37°C, followed by proteinase K digestion for 4 h at 60°C. Then, LPS was purified by hot phenol extraction and a subsequent extraction of the aqueous phase with ether. LPS samples were resolved by electrophoresis in 16% polyacrylamide gels with a Tricine-SDS system (77) and silver stained.

The amount of exopolysaccharide being produced was evaluated based on the dry weight of the ethanol-precipitated polysaccharide recovered from 100-ml cultures grown in SM medium for 6 days at 37°C with orbital agitation (78). Data are based on results from duplicate flasks from three independent experiments.

Antimicrobial susceptibility. The agar disc diffusion method (79) was used to assess isolate susceptibility against paper discs containing piperacillin (75 µg) plus tazobactam (10 µg). The discs were applied onto the surfaces of Muller-Hinton (Difco Laboratories) agar plates previously inoculated with 100 µl of a suspension at an OD₆₄₀ of 0.1 prepared from exponential-phase cells growing on LB medium. The diameter of the growth inhibition zone was measured after 24 h of incubation at 37°C. Three experiments, each with three discs per isolate, were performed. Statistical analysis was conducted by ANOVA and Tukey's HSD multiple-comparison test.

Motility. The swimming motility of all isolates was tested on 0.3% (wt/vol) agarose supplemented with 1% tryptone and 0.5% NaCl. Plates were spot inoculated with a 3-µl drop of a culture at an OD₆₄₀ of 1.0 and were incubated for 24 h at 37°C. The motility zones of three replicates from three independent experiments were measured. Statistical analysis was conducted by ANOVA and Tukey's HSD multiple-comparison test.

Measurement of c-di-GMP levels. For c-di-GMP extraction, *B. multivorans* isolates were cultured in triplicate in SM medium for 5 h. Cells from a 40-ml culture were harvested by centrifugation (10,000 × *g*), and pelleted cells were resuspended in extraction buffer (40% methanol, 40% acetonitrile, 0.1 N formic acid). This mixture was incubated at -20°C for 30 min, after which insoluble cell debris were precipitated by centrifugation at 14,000 × *g* for 5 min (80). A total volume of 200 µl of the supernatants was neutralized with 8 µl of 15% (wt/vol) NH₄HCO₃. To measure the c-di-GMP of the lysates, 10 µl of each sample was analyzed using liquid chromatography-tandem mass spectrometry (LC-MS/MS) as described previously (80). Concentrations are reported as nanomolar concentrations of c-di-GMP, assuming cylinder-shaped bacteria with a volume of 8.3 × 10⁻¹⁶ liters and taking into account the number of CFU in each sample. Statistical analysis was conducted by ANOVA, followed by Tukey's HSD multiple-comparison test.

Biofilm formation. Biofilm formation assays were performed as described previously (81). Bacteria were grown in LB medium at 37°C to mid-exponential phase and diluted to an OD₆₄₀ of 0.05, and 200-µl samples of these cell suspensions were used to inoculate 96-well polystyrene microtiter plates. Plates were incubated at 37°C statically for 48 h, after which the wells were washed three times with 0.9% (wt/vol) NaCl. The biofilm was stained with 1% (wt/vol) crystal violet solution for 20 min, followed by dye solubilization with ethanol and measurement of the solution's A₅₉₀ in a microplate reader. Results are the means of results from six replicates of five independent experiments. ANOVA and Tukey's HSD multiple-comparison test were performed to determine statistically significant differences.

Adhesion to epithelial cells. The *B. multivorans* isolates under study were examined for adhesion to the bronchial epithelial cell line CFBE41o⁻, derived from a patient homozygous for the cystic fibrosis transmembrane conductance regulator ΔF508 mutation (82). Host cell attachment was performed as described previously (83). Briefly, bacterial strains grown in SM medium to an OD₆₄₀ of 0.6 (log phase) were used to infect epithelial cells at a multiplicity of infection (MOI) of 10 (10 bacterial cells to 1 epithelial cell). Bacteria were applied to a 24-well plate previously seeded with CFBE41o⁻ cells in minimal essential medium-supplemented medium, and the plates were centrifuged at 700 × *g* for 5 min. The plates were incubated for 30 min at 37°C in an atmosphere of 5% CO₂. Each well was then washed three times with phosphate-buffered saline (PBS) to remove unbound bacteria, and cells were lysed with lysis buffer (0.01 M PBS, 20 mM EDTA, 0.5% [vol/vol] Triton X-100; pH 7.4) for 30 min at 4°C. Serial dilutions were plated onto LB agar, and adhesion was quantified by determining CFU counts after 48 h of incubation at 37°C. Duplicates with each strain were performed per assay, and the results presented were obtained from three independent experiments. Results are expressed as the percentage of adhesion, which was calculated as the number of CFU recovered divided by the number CFU applied to the epithelial cells multiplied by 100. Results are the mean values from two replicates from three independent experiments. ANOVA followed by Tukey's HSD multiple-comparison test was performed.

Ethics statement. The bacterial samples used are from an anonymous patient, and none of the samples was taken specifically for this study. The patient's clinical data were collected as part of a previously published chart review study, which was approved by the University of British Columbia's Research Ethics Boards (approval H07-01396).

Nucleotide sequence accession numbers. The DNA sequence reads for assemblies of the several genomes are available in the EMBL's European Nucleotide Archive (ENA) under accession number [PRJEB13383](https://www.ebi.ac.uk/ena/entry/PRJEB13383). Fully assembled and annotated genomes of *B. multivorans* VC13401 (D2095) and VC13534

(D2214) are available in the NCBI database under accession numbers PRJNA239698 and PRJNA239704, respectively. Alignment and SNP call files are in figshare (<https://figshare.com/s/58adae59e3ebd57550da>).

SUPPLEMENTAL MATERIAL

Supplemental material for this article may be found at <http://dx.doi.org/10.1128/mSystems.00029-16>.

Table S1, DOCX file, 0.04 MB.
 Table S2, DOCX file, 0.04 MB.
 Table S3, XLSX file, 0.1 MB.
 Table S4, XLSX file, 0.05 MB.
 Table S5, DOCX file, 0.03 MB.
 Table S6, DOCX file, 0.04 MB.
 Figure S1, TIF file, 0.7 MB.
 Figure S2, TIF file, 0.6 MB.
 Figure S3, TIF file, 1.4 MB.
 Figure S4, TIF file, 1.7 MB.

ACKNOWLEDGMENTS

J.E.A.Z. and D.P.S. are grateful to the microbiology laboratories at British Columbia Children's Hospital and St. Paul's Hospital for their provision of isolates. V.S.C. acknowledges technical advice from Marcus Dillon (UNH).

FUNDING INFORMATION

This work, including the efforts of Leonilde M. Moreira and Inês N. Silva, was funded by Programa Operacional Regional de Lisboa 2020 (LISBOA-01-0145-FEDER- 007317). This work, including the efforts of Christopher M. Waters and Vaughn S. Cooper, was funded by HHS | National Institutes of Health (NIH) (GM109259 and R01-GM110444). This work, including the efforts of Christopher M. Waters, was funded by National Science Foundation (NSF) (DBI-0939454). This work, including the efforts of James E.A. Zlosnik and David P. Speert, was funded by Cystic Fibrosis Canada (Fibrose kystique Canada) (20R40868 and 20R32976). This work, including the efforts of Leonilde M. Moreira and Inês N. Silva, was funded by Ministry of Education and Science | Fundação para a Ciência e a Tecnologia (FCT) (PTDC/QUI-BIQ/118260/2010, UID/BIO/04565/2013, and SFRH/BPD/86475/2012).

The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

REFERENCES

1. Brodsky IE, Medzhitov R. 2009. Targeting of immune signalling networks by bacterial pathogens. *Nat Cell Biol* **11**:521–526. <http://dx.doi.org/10.1038/ncb0509-521>.
2. Bierre H, Cossart P. 2012. When bacteria target the nucleus: the emerging family of nucleomodulins. *Cell Microbiol* **14**:622–633. <http://dx.doi.org/10.1111/j.1462-5822.2012.01758.x>.
3. Deitsch KW, Lukehart SA, Stringer JR. 2009. Common strategies for antigenic variation by bacterial, fungal and protozoan pathogens. *Nat Rev Microbiol* **7**:493–503. <http://dx.doi.org/10.1038/nrmicro2145>.
4. Aaron SD, Ferris W, Henry DA, Speert DP, Macdonald NE. 2000. Multiple combination bactericidal antibiotic testing for patients with cystic fibrosis infected with *Burkholderia cepacia*. *Am J Respir Crit Care Med* **161**:1206–1212. <http://dx.doi.org/10.1164/ajrccm.161.4.9907147>.
5. Smith EE, Buckley DG, Wu Z, Saenphimmachak C, Hoffman LR, D'Argenio DA, Miller SI, Ramsey BW, Speert DP, Moskowitz SM, Burns JL, Kaul R, Olson MV. 2006. Genetic adaptation by *Pseudomonas aeruginosa* to the airways of cystic fibrosis patients. *Proc Natl Acad Sci U S A* **103**:8487–8492. <http://dx.doi.org/10.1073/pnas.0602138103>.
6. Marvig RL, Sommer LM, Molin S, Johansen HK. 2015. Convergent evolution and adaptation of *Pseudomonas aeruginosa* within patients with cystic fibrosis. *Nat Genet* **47**:57–64. <http://dx.doi.org/10.1038/ng.3148>.
7. Isles A, Macluskay I, Corey M, Gold R, Prober C, Fleming P, Levison H. 1984. *Pseudomonas cepacia* infection in cystic fibrosis: an emerging problem. *J Pediatr* **104**:206–210. [http://dx.doi.org/10.1016/S0022-3476\(84\)80993-2](http://dx.doi.org/10.1016/S0022-3476(84)80993-2).
8. Ferreira AS, Silva IN, Oliveira VH, Cunha R, Moreira LM. 2011. Insights into the role of extracellular polysaccharides in *Burkholderia* adaptation to different environments. *Front Cell Infect Microbiol* **1**:16. <http://dx.doi.org/10.3389/fcimb.2011.00016>.
9. Leitão JH, Sousa SA, Ferreira AS, Ramos CG, Silva IN, Moreira LM. 2010. Pathogenicity, virulence factors, and strategies to fight against *Burkholderia cepacia* complex pathogens and related species. *Appl Microbiol Biotechnol* **87**:31–40. <http://dx.doi.org/10.1007/s00253-010-2528-0>.
10. Mahenthiralingam E, Urban TA, Goldberg JB. 2005. The multifarious, multireplicon *Burkholderia cepacia* complex. *Nat Rev Microbiol* **3**:144–156. <http://dx.doi.org/10.1038/nrmicro1085>.
11. Saldias MS, Valvano MA. 2009. Interactions of *Burkholderia cenocepacia* and other *Burkholderia cepacia* complex bacteria with epithelial and phagocytic cells. *Microbiology* **155**:2809–2817. <http://dx.doi.org/10.1099/mic.0.031344-0>.
12. Lieberman TD, Michel J-B, Aingaran M, Potter-Bynoe G, Roux D, Davis MR, Skurnik D, Leiby N, LiPuma JJ, Goldberg JB, McAdam AJ, Priebe GP, Kishony R. 2011. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nat Genet* **43**:1275–1280. <http://dx.doi.org/10.1038/ng.997>.
13. Lieberman TD, Flett KB, Yelin I, Martin TR, McAdam AJ, Priebe GP,

- Kishony R.** 2014. Genetic variation of a bacterial pathogen within individuals with cystic fibrosis provides a record of selective pressures. *Nat Genet* **46**:82–87. <http://dx.doi.org/10.1038/ng.2848>.
14. **Zlosnik JE, Hird TJ, Fraenkel MC, Moreira LM, Henry DA, Speert DP.** 2008. Differential mucoid exopolysaccharide production by members of the *Burkholderia cepacia* complex. *J Clin Microbiol* **46**:1470–1473. <http://dx.doi.org/10.1128/JCM.02273-07>.
 15. **Zlosnik JE, Costa PS, Brant R, Mori PY, Hird TJ, Fraenkel MC, Wilcox PG, Davidson AG, Speert DP.** 2011. Mucoid and nonmucoid *Burkholderia cepacia* complex bacteria in cystic fibrosis infections. *Am J Respir Crit Care Med* **183**:67–72. <http://dx.doi.org/10.1164/rccm.201002-0203OC>.
 16. **Silva IN, Tavares AC, Ferreira AS, Moreira LM.** 2013. Stress conditions triggering mucoid morphotype variation in *Burkholderia* species and effect on virulence in *Galleria mellonella* and biofilm formation *in vitro*. *PLoS One* **8**:e82522. <http://dx.doi.org/10.1371/journal.pone.0082522>.
 17. **Zlosnik JEA, Zhou G, Brant R, Henry DA, Hird TJ, Mahenthalingam E, Chilvers MA, Wilcox P, Speert DP.** 2015. *Burkholderia* species infections in patients with cystic fibrosis in British Columbia, Canada. 30 years' experience. *Ann Am Thorac Soc* **12**:70–78. <http://dx.doi.org/10.1513/AnnalsATS.201408-395OC>.
 18. **Lipuma JJ.** 2010. The changing microbial epidemiology in cystic fibrosis. *Clin Microbiol Rev* **23**:299–323. <http://dx.doi.org/10.1128/CMR.00068-09>.
 19. **Govan JRW, Brown AR, Jones AM.** 2007. Evolving epidemiology of *Pseudomonas aeruginosa* and the *Burkholderia cepacia* complex in cystic fibrosis lung infection. *Future Microbiol* **2**:153–164. <http://dx.doi.org/10.2217/17460913.2.2.153>.
 20. **Pope CE, Short P, Carter PE.** 2010. Species distribution of *Burkholderia cepacia* complex isolates in cystic fibrosis and non-cystic fibrosis patients in New Zealand. *J Cyst Fibros* **9**:442–446. <http://dx.doi.org/10.1016/j.jcf.2010.08.011>.
 21. **Redelings BD, Suchard MA.** 2007. Incorporating indel information into phylogeny estimation for rapidly emerging pathogens. *BMC Evol Biol* **7**:40. <http://dx.doi.org/10.1186/1471-2148-7-40>.
 22. **Dillon MM, Sung W, Lynch M, Cooper VS.** 2015. The rate and molecular spectrum of spontaneous mutations in the GC-rich multichromosome genome of *Burkholderia cenocepacia*. *Genetics* **200**:935–946. <http://dx.doi.org/10.1534/genetics.115.176834>.
 23. **Marteyn B, West NP, Browning DF, Cole JA, Shaw JG, Palm F, Mounier J, Prévost M-C, Sansonetti P, Tang CM.** 2010. Modulation of *Shigella* virulence in response to available oxygen *in vivo*. *Nature* **465**:355–358. <http://dx.doi.org/10.1038/nature08970>.
 24. **Traverse CC, Mayo-Smith LM, Poltak SR, Cooper VS.** 2013. Tangled bank of experimentally evolved *Burkholderia* biofilms reflects selection during chronic infections. *Proc Natl Acad Sci U S A* **110**:E250–E259. <http://dx.doi.org/10.1073/pnas.1207025110>.
 25. **Rao NN, Gómez-García MR, Kornberg A.** 2009. Inorganic polyphosphate: essential for growth and survival. *Annu Rev Biochem* **78**:605–647. <http://dx.doi.org/10.1146/annurev.biochem.77.083007.093039>.
 26. **Ortega X, Hunt TA, Loutet S, Vinion-Dubiel AD, Datta A, Choudhury B, Goldberg JB, Carlson R, Valvano MA.** 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. <http://dx.doi.org/10.1128/JB.187.4.1324-1333.2005>.
 27. **Saldías MS, Ortega X, Valvano MA.** 2009. *Burkholderia cenocepacia* O antigen lipopolysaccharide prevents phagocytosis by macrophages and adhesion to epithelial cells. *J Med Microbiol* **58**:1542–1548. <http://dx.doi.org/10.1099/jmm.0.013235-0>.
 28. **Vinion-Dubiel AD, Goldberg JB.** 2003. Lipopolysaccharide of *Burkholderia cepacia* complex. *J Endotoxin Res* **9**:201–213. <http://dx.doi.org/10.1179/096805103225001404>.
 29. **Hinchliffe P, Symmons MF, Hughes C, Koronakis V.** 2013. Structure and operation of bacterial tripartite pumps. *Annu Rev Microbiol* **67**:221–242. <http://dx.doi.org/10.1146/annurev-micro-092412-155718>.
 30. **Bonfiglio G, Laksai Y, Franchino L, Amicosante G, Nicoletti G.** 1998. Mechanisms of beta-lactam resistance amongst *Pseudomonas aeruginosa* isolated in an Italian survey. *J Antimicrob Chemother* **42**:697–702. <http://dx.doi.org/10.1093/jac/42.6.697>.
 31. **Hinchliffe P, Greene NP, Paterson NG, Crow A, Hughes C, Koronakis V.** 2014. Structure of the periplasmic adaptor protein from a major facilitator superfamily (MFS) multidrug efflux pump. *FEBS Lett* **588**:3147–3153. <http://dx.doi.org/10.1016/j.febslet.2014.06.055>.
 32. **Deng Y, Boon C, Eberl L, Zhang LH.** 2009. Differential modulation of *Burkholderia cenocepacia* virulence and energy metabolism by the quorum-sensing signal BDSF and its synthase. *J Bacteriol* **191**:7270–7278. <http://dx.doi.org/10.1128/JB.00681-09>.
 33. **Ryan RP, McCarthy Y, Watt SA, Niehaus K, Dow JM.** 2009. Intraspecies signaling involving the diffusible signal factor BDSF (cis-2-dodecenoic acid) influences virulence in *Burkholderia cenocepacia*. *J Bacteriol* **191**:5013–5019. <http://dx.doi.org/10.1128/JB.00473-09>.
 34. **McCarthy Y, Yang L, Twomey KB, Sass A, Tolker-Nielsen T, Mahenthalingam E, Dow JM, Ryan RP.** 2010. A sensor kinase recognizing the cell-cell signal BDSF (cis-2-dodecenoic acid) regulates virulence in *Burkholderia cenocepacia*. *Mol Microbiol* **77**:1220–1236. <http://dx.doi.org/10.1111/j.1365-2958.2010.07285.x>.
 35. **Deng Y, Schmid N, Wang C, Wang J, Pessi G, Wu D, Lee J, Aguilar C, Ahrens CH, Chang C, Song H, Eberl L, Zhang L-H.** 2012. Cis-2-dodecenoic acid receptor PpfR links quorum-sensing signal perception with regulation of virulence through cyclic dimeric guanosine monophosphate turnover. *Proc Natl Acad Sci U S A* **109**:15479–15484. <http://dx.doi.org/10.1073/pnas.1205037109>.
 36. **Boyd CD, O'Toole GA.** 2012. Second messenger regulation of biofilm formation: breakthroughs in understanding c-di-GMP effector systems. *Annu Rev Cell Dev Biol* **28**:439–462. <http://dx.doi.org/10.1146/annurev-cellbio-101011-155705>.
 37. **Hengge R.** 2013. Novel tricks played by the second messenger c-di-GMP in bacterial biofilm formation. *EMBO J* **32**:322–323. <http://dx.doi.org/10.1038/emboj.2012.351>.
 38. **Fazli M, Almlblad H, Rybtke ML, Givskov M, Eberl L, Tolker-Nielsen T.** 2014. Regulation of biofilm formation in *Pseudomonas* and *Burkholderia* species. *Environ Microbiol* **16**:1961–1981. <http://dx.doi.org/10.1111/1462-2920.12448>.
 39. **Tischler AD, Camilli A.** 2004. Cyclic diguanylate (c-di-GMP) regulates *Vibrio cholerae* biofilm formation. *Mol Microbiol* **53**:857–869. <http://dx.doi.org/10.1111/j.1365-2958.2004.04155.x>.
 40. **Lindenberg S, Klauck G, Pesavento C, Klauck E, Hengge R.** 2013. The EAL domain protein YciR acts as a trigger enzyme in a c-di-GMP signaling cascade in *E. coli* biofilm control. *EMBO J* **32**:2001–2014. <http://dx.doi.org/10.1038/emboj.2013.120>.
 41. **Harmsen M, Yang L, Pamp SJ, Tolker-Nielsen T.** 2010. An update on *Pseudomonas aeruginosa* biofilm formation, tolerance, and dispersal. *FEMS Immunol Med Microbiol* **59**:253–268. <http://dx.doi.org/10.1111/j.1574-695X.2010.00690.x>.
 42. **Anuntaganon N, Wuthiekanun V, White NJ, Currie BJ, Sermswan RW, Wongratanchewin S, Taweekhaisupapong S, Chaiyaroj SC, Sirinaha S.** 2006. Lipopolysaccharide heterogeneity among *Burkholderia pseudomallei* from different geographic and clinical origins. *Am J Trop Med Hyg* **74**:348–352.
 43. **Salsgiver EL, Fink AK, Knapp EA, LiPuma JJ, Olivier KN, Marshall BC, Saiman L.** 2016. Changing epidemiology of the respiratory bacteriology of patients with cystic fibrosis. *Chest* **149**:390–400. <http://dx.doi.org/10.1378/chest.15-0676>.
 44. **Markussen T, Marvig RL, Gómez-Lozano M, Aanæs K, Burleigh AE, Høiby N, Johansen HK, Molin S, Jelsbak L.** 2014. Environmental heterogeneity drives within-host diversification and evolution of *Pseudomonas aeruginosa*. *mBio* **5**:e01592–14. <http://dx.doi.org/10.1128/mBio.01592-14>.
 45. **Oliver A, Cantón R, Campo P, Baquero F, Blázquez J.** 2000. High frequency of hypermutable *Pseudomonas aeruginosa* in cystic fibrosis lung infection. *Science* **288**:1251–1254. <http://dx.doi.org/10.1126/science.288.5469.1251>.
 46. **Martina P, Feliziani S, Juan C, Bettiol M, Gatti B, Yantorno O, Smania AM, Oliver A, Bosch A.** 2014. Hypermutation in *Burkholderia cepacia* complex is mediated by DNA mismatch repair inactivation and is highly prevalent in cystic fibrosis chronic respiratory infection. *Int J Med Microbiol* **304**:1182–1191. <http://dx.doi.org/10.1016/j.ijmm.2014.08.011>.
 47. **Oliver A, Mena A.** 2010. Bacterial hypermutation in cystic fibrosis, not only for antibiotic resistance. *Clin Microbiol Infect* **16**:798–808. <http://dx.doi.org/10.1111/j.1469-0691.2010.03250.x>.
 48. **Rau MH, Hansen SK, Johansen HK, Thomsen LE, Workman CT, Nielsen KF, Jelsbak L, Høiby N, Yang L, Molin S.** 2010. Early adaptive developments of *Pseudomonas aeruginosa* after the transition from life in the environment to persistent colonization in the airways of human cystic fibrosis hosts. *Environ Microbiol* **12**:1643–1658. <http://dx.doi.org/10.1111/j.1462-2920.2010.02211.x>.
 49. **Hancock RE, Mutharia LM, Chan L, Darveau RP, Speert DP, Pier GB.** 1983. *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis:

- a class of serum-sensitive, nontypeable strains deficient in lipopolysaccharide O side chains. *Infect Immun* **42**:170–177.
50. **Blanka A, Düvel J, Dötsch A, Klinkert B, Abraham W-R, Kaefer V, Ritter C, Narberhaus F, Häussler S.** 2015. Constitutive production of c-di-GMP is associated with mutations in a variant of *Pseudomonas aeruginosa* with altered membrane composition. *Sci Signal* **8**:ra36. <http://dx.doi.org/10.1126/scisignal.2005943>.
 51. **Coutinho CP, de Carvalho CCCR, Madeira A, Pinto-de-Oliveira A, Sá-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. <http://dx.doi.org/10.1128/IAI.01366-10>.
 52. **Jeannot K, Elsen S, Köhler T, Attree I, van Delden C, Plésiat P.** 2008. Resistance and virulence of *Pseudomonas aeruginosa* clinical strains overproducing the MexCD-OprJ efflux pump. *Antimicrob Agents Chemother* **52**:2455–2462. <http://dx.doi.org/10.1128/AAC.01107-07>.
 53. **Filkins LM, O'Toole GA.** 2015. Cystic fibrosis lung infections: polymicrobial, complex, and hard to treat. *PLoS Pathog* **11**:e1005258. <http://dx.doi.org/10.1371/journal.ppat.1005258>.
 54. **Coburn B, Wang PW, Diaz Caballero J, Clark ST, Brahma V, Donaldson S, Zhang Y, Surendra A, Gollig Y, Tullis DE, Yau YC, Waters VJ, Hwang DM, Guttman DS.** 2015. Lung microbiota across age and disease stage in cystic fibrosis. *Sci Rep* **5**:10241. <http://dx.doi.org/10.1038/srep10241>.
 55. **Chotirmall SH, O'Donoghue E, Bennett K, Gunaratnam C, O'Neill SJ, McElvaney NG.** 2010. Sputum *Candida albicans* presages FEV₁ decline and hospital-treated exacerbations in cystic fibrosis. *Chest* **138**:1186–1195. <http://dx.doi.org/10.1378/chest.09-2996>.
 56. **Paganin P, Fiscarelli EV, Tuccio V, Chianciani M, Bacci G, Morelli P, Dolce D, Dalmastrì C, De Alessandri A, Lucidi V, Taccetti G, Mengoni A, Bevivino A.** 2015. Changes in cystic fibrosis airway microbial community associated with a severe decline in lung function. *PLoS One* **10**:e0124348. <http://dx.doi.org/10.1371/journal.pone.0124348>.
 57. **Silva IN, Santos PM, Moreira LM.** 2015. Draft genome sequence of two *Burkholderia multivorans* sequential isolates from a chronic lung infection of a cystic fibrosis patient. *Genome Announc* **3**:e01531-1. <http://dx.doi.org/10.1128/genomeA.01531-14>.
 58. **Kelley DR, Schatz MC, Salzberg SL.** 2010. Quake: quality-aware detection and correction of sequencing errors. *Genome Biol* **11**:R116. <http://dx.doi.org/10.1186/gb-2010-11-11-r116>.
 59. **Aronesty E.** 2013. Comparison of sequencing utility programs. *Open Bioinforma J* **7**:1–8. <http://dx.doi.org/10.2174/1875036201307010001>.
 60. **Andrews S.** 2010. FastQC: a quality control tool for high throughput sequence data. Babraham Bioinformatics, Babraham Institute, Cambridge, United Kingdom. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc/>.
 61. **Hernandez D, François P, Farinelli L, Osterås M, Schrenzel J.** 2008. *De novo* bacterial genome sequencing: millions of very short reads assembled on a desktop computer. *Genome Res* **18**:802–809. <http://dx.doi.org/10.1101/gr.072033.107>.
 62. **Bankevich A, Nurk S, Antipov D, Gurevich AA, Dvorkin M, Kulikov AS, Lesin VM, Nikolenko SI, Pham S, Prjibelski AD, Pyshkin AV, Sirotkin AV, Vyahhi N, Tesler G, Alekseyev MA, Pevzner PA.** 2012. SPAdes: a new genome assembly algorithm and its applications to single-cell sequencing. *J Comput Biol* **19**:455–477. <http://dx.doi.org/10.1089/cmb.2012.0021>.
 63. **Langmead B, Salzberg SL.** 2012. Fast gapped-read alignment with bowtie 2. *Nat Methods* **9**:357–359. <http://dx.doi.org/10.1038/nmeth.1923>.
 64. **Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R, 1000 Genome Project Data Processing Subgroup.** 2009. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**:2078–2079. <http://dx.doi.org/10.1093/bioinformatics/btp352>.
 65. **Wei Z, Wang W, Hu P, Lyon GJ, Hakonarson H.** 2011. SNVer: a statistical tool for variant calling in analysis of pooled or individual next-generation sequencing data. *Nucleic Acids Res* **39**:1–13. <http://dx.doi.org/10.1093/nar/gkq742>.
 66. **Vicedomini R, Vezzi F, Scalabrini S, Arvestad L, Policriti A.** 2013. GAM-NGS: genomic assemblies merger for next generation sequencing. *BMC Bioinformatics* **14**(Suppl 7):S6. <http://dx.doi.org/10.1186/1471-2105-14-S7-S6>.
 67. **Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, Cuomo CA, Zeng Q, Wortman J, Young SK, Earl AM.** 2014. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement 9. *PLoS One* **9**:e112963. <http://dx.doi.org/10.1371/journal.pone.0112963>.
 68. **Boetzer M, Henkel CV, Jansen HJ, Butler D, Pirovano W.** 2011. Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* **27**:578–579. <http://dx.doi.org/10.1093/bioinformatics/btq683>.
 69. **Boetzer M, Pirovano W.** 2012. Toward almost closed genomes with GapFiller. *Genome Biol* **13**:R56. <http://dx.doi.org/10.1186/gb-2012-13-6-r56>.
 70. **Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, He G, Chen Y, Pan Q, Liu Y, Tang J, Wu G, Zhang H, Shi Y, Liu Y, Yu C, Wang B, Lu Y, Han C, Cheung DW, Yiu S-M, Peng S, Xiaoqian Z, Liu G, Liao X, Li Y, Yang H, Wang J, Lam T-W, Wang J.** 2012. SOAPdenovo2: an empirically improved memory-efficient short-read *de novo* assembler. *GigaScience* **1**:18. <http://dx.doi.org/10.1186/2047-217X-1-18>.
 71. **Li H, Durbin R.** 2010. Fast and accurate long-read alignment with burrows-Wheeler transform. *Bioinformatics* **26**:589–595. <http://dx.doi.org/10.1093/bioinformatics/btp698>.
 72. **Li H.** 26 May 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. arXiv:1303.3997v2. <http://arxiv.org/abs/1303.3997>. Accessed 6 June 2015.
 73. **Kearse M, Moir R, Wilson A, Stones-Havas S, Cheung M, Sturrock S, Buxton S, Cooper A, Markowitz S, Duran C, Thierer T, Ashton B, Meintjes P, Drummond A.** 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. *Bioinformatics* **28**:1647–1649. <http://dx.doi.org/10.1093/bioinformatics/bts199>.
 74. **Figurski DH, Helinski DR.** 1979. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function provided in trans. *Proc Natl Acad Sci U S A* **76**:1648–1652. <http://dx.doi.org/10.1073/pnas.76.4.1648>.
 75. **Palmer KL, Aye LM, Whiteley M.** 2007. Nutritional cues control *Pseudomonas aeruginosa* multicellular behavior in cystic fibrosis sputum. *J Bacteriol* **189**:8079–8087. <http://dx.doi.org/10.1128/JB.01138-07>.
 76. **Marolda CL, Welsh J, Dafoe L, Valvano MA.** 1990. Genetic analysis of the O7-polysaccharide biosynthesis region from the *Escherichia coli* O7:K1 strain VW187. *J Bacteriol* **172**:3590–3599.
 77. **Lesse AJ, Campagnari AA, Bittner WE, Apicella MA.** 1990. Increased resolution of lipopolysaccharides and lipooligosaccharides utilizing tricine-sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *J Immunol Methods* **126**:109–117. [http://dx.doi.org/10.1016/0022-1759\(90\)90018-Q](http://dx.doi.org/10.1016/0022-1759(90)90018-Q).
 78. **Silva IN, Ferreira AS, Becker JD, Zlosnik JEA, Speert DP, He J, Mil-Homens D, Moreira LM.** 2011. Mucoid morphotype variation of *Burkholderia multivorans* during chronic cystic fibrosis lung infection is correlated with changes in metabolism, motility, biofilm formation and virulence. *Microbiology* **157**:3124–3137. <http://dx.doi.org/10.1099/mic.0.050989-0>.
 79. **Bauer AW, Kirby WM, Sherris JC, Turck M.** 1966. Antibiotic susceptibility testing by a standardized single disk method. *Am J Clin Pathol* **45**:493–496.
 80. **Waters CM, Lu W, Rabinowitz JD, Bassler BL.** 2008. Quorum sensing controls biofilm formation in *Vibrio cholerae* through modulation of cyclic di-GMP levels and repression of VpsT. *J Bacteriol* **190**:2527–2536. <http://dx.doi.org/10.1128/JB.01756-07>.
 81. **Ferreira AS, Leitão JH, Sousa SA, Cosme AM, Sá-Correia I, Moreira LM.** 2007. Functional analysis of *Burkholderia cepacia* genes *bceD* and *bceF*, encoding a phosphotyrosine phosphatase and a tyrosine autokinase, respectively: role in exopolysaccharide biosynthesis and biofilm formation. *Appl Environ Microbiol* **73**:524–534. <http://dx.doi.org/10.1128/AEM.01450-06>.
 82. **Gonc KK, Feeney L, Gruenert DC.** 1999. Differential sensitivity of normal and cystic fibrosis airway epithelial cells to epinephrine. *Br J Pharmacol* **128**:227–233. <http://dx.doi.org/10.1038/sj.bjpp.0702772>.
 83. **Ferreira AS, Silva IN, Fernandes F, Pilkington R, Callaghan M, McClean S, Moreira LM.** 2015. The tyrosine kinase BceF and the phosphotyrosine phosphatase BceD of *Burkholderia contaminans* are required for efficient invasion and epithelial disruption of a cystic fibrosis lung epithelial cell line. *Infect Immun* **83**:812–821. <http://dx.doi.org/10.1128/IAI.02713-14>.