Comparative analysis of the *Burkholderia cenocepacia* K56-2 essential genome reveals cell envelope functions that are uniquely required for survival in species of the genus *Burkholderia*

April S. Gislason, Keith Turner, Mike Domaratzki and Silvia T. Cardona

**Abstract**

*Burkholderia cenocepacia* K56-2 belongs to the *Burkholderia cepacia* complex, a group of Gram-negative opportunistic pathogens that have large and dynamic genomes. In this work, we identified the essential genome of *B. cenocepacia* K56-2 using high-density transposon mutagenesis and insertion site sequencing (Tn-seq circle). We constructed a library of one million transposon mutants and identified the transposon insertions at an average of one insertion per 27 bp. The probability of gene essentiality was determined by comparing the insertion density per gene with the variance of neutral datasets generated by Monte Carlo simulations. Five hundred and eight genes were not significantly disrupted, suggesting that these genes are essential for survival in rich, undefined medium. Comparison of the *B. cenocepacia* K56-2 essential genome with that of the closely related *B. cenocepacia* J2315 revealed partial overlapping, suggesting that some essential genes are strain-specific. Furthermore, 158 essential genes were conserved in *B. cenocepacia* and two species belonging to the *Burkholderia pseudomallei* complex, *B. pseudomallei* K96243 and *Burkholderia thailandensis* E264. Porins, including OpcC, a lysophospholipid transporter, LplT, and a protein involved in the modification of lipid A with aminoarabinose were found to be essential in *Burkholderia* genomes but not in other bacterial essential genomes identified so far. Our results highlight the existence of cell envelope processes that are uniquely essential in species of the genus *Burkholderia* for which the essential genomes have been identified by Tn-seq.

**DATA SUMMARY**

1. The Illumina sequencing reads generated and analysed during the current study are available in the NCBI Sequence Read Archive (SRA) repository, accession number SRP112587, https://www.ncbi.nlm.nih.gov/sra/?term=SRP112587.

2. The script used to identify and trim reads containing the transposon sequence and map the genomic regions can be accessed at https://github.com/khturner/Tn-seq/blob/master/TnSeq2.sh.

3. Custom scripts used to determine gene essentiality can be accessed at https://github.com/khturner/Tn-seq/tree/dockerize.

**INTRODUCTION**

Determining the essential genomes of bacteria has furthered our understanding about the fundamental processes required for survival [1–6] and provided a first step in identifying putative targets for developing antibacterial therapies [7–9]. The identification of essential genes is challenging due to the lethal phenotype that mutagenesis of essential genes cause. Yet, identification of essential genes in laboratory conditions can be achieved by recovering mutants growing in rich, undefined medium after concerted disruption of non-essential genes [10–13]. Next generation sequencing has facilitated the use of saturated transposon mutagenesis to identify the essential genomes of many bacteria. To identify essential genes from a high-density transposon mutant (HDTM) library, the...
The 

The large and dynamic genomes of *Burkholderia* are likely linked to their profound adaptability and pathogenicity. The differences found in the genomes of closely related *Burkholderia* strains suggest that their essential genomes may also vary. In this work, we used high-density transposon insertion sequencing to predict that the essential genome of *Burkholderia cenocepacia* K56-2 consists of more than five hundred genes. Despite *B. cenocepacia* K56-2 and J2315 belonging to the same clonal complex, comparison of their essential genomes revealed partial divergence. Further comparative analysis with the essential gene sets of two *Burkholderia pseudomallei* group pathogens identified 158 essential genes that are conserved among the four *Burkholderia* strains. Three genes within this common essential gene set encode proteins that contribute to the integrity of the cell envelope and were identified as uniquely essential in the three species of the genus *Burkholderia*. We propose that the normally non-essential functions of intrinsic antimicrobial resistance and cell envelope impermeability that characterize the genus *Burkholderia* are mechanisms tied to the survival of the species of the genus *Burkholderia* that were analysed.

**METHODS**

**Bacterial strains and growth conditions**

The bacterial strains and plasmids used in this study are listed in Table S1 (available in the online Supplementary Material). *B. cenocepacia* K56-2, a closely related strain of...
B. cenocepacia J2315 and isolated from a CF patient [49], was the strain used in this study. All bacteria were grown in Luria-Bertani (LB) media (BD) at 37°C with shaking at 220 r.p.m. in a New Brunswick Scientific E24 shaking incubator for broth cultures unless otherwise indicated. Escherichia coli strains were grown in 40 µg ml⁻¹ kanamycin (Kan40) or 50 µg ml⁻¹ trimethoprim (Tp50) when appropriate. B. cenocepacia transposon mutants were selected for in 100 µg ml⁻¹ trimethoprim (Tp100), 50 µg ml⁻¹ gentamicin (Gm50), and 0.2 % rhamnose (0.2 % rha). Additionally, the transposon mutants were grown in Tp100 with or without 0.2 % rha when appropriate. Growth was estimated by measuring the OD₆₀₀ using a Biotek Synergy 2 plate reader. All chemicals were ordered from Sigma Chemical Co. unless otherwise indicated.

Production of the one million high-density transposon mutant (HDTM) library

Triparental matings were carried out to conjugate pBRha-Boutgfp into wild-type B. cenocepacia K56-2. This was done using an E. coli SY327 donor strain [50] and an E. coli MM290 helper strain carrying the pRK2013 plasmid [51]. E. coli SY327 and MM290 were plated on LB agar supplemented with Tp50 or Kan40, respectively. B. cenocepacia K56-2 was grown in LB broth. All three strains were grown for 16–18 h at 37°C. Afterwards, B. cenocepacia K56-2 was subcultured in 5 ml LB broth at 37°C until an OD₆₀₀ of 0.3–0.6 was reached. Colonies of E. coli SY327 and MM290 were collected and resuspended in 5 ml LB in separate snap-cap tubes. From each of the E. coli SY327 and MM290 suspensions, volumes equivalent to 1.5 ml with an OD₆₀₀ of 0.3 were mixed with B. cenocepacia suspension equivalent to 0.5 ml with an OD₆₀₀ of 1.0. The mixture was vortexed then centrifuged for 1 min at 6000 r.p.m. The pellet was resuspended, then approximately seventy 20–25 µl aliquots of the triparental mating were spotted on LB agar plates containing 0.2 % rha, allowed to dry and incubated at 37°C for 2 h. Bacteria from the triparental mating spots were combined in 6–7 ml LB broth. Aliquots of 500 µl of bacteria were plated onto 8–10 Q-trays containing LB agar with 0.2 % rha, Tp100 and Gm50, and incubated for 48 h. After incubation, pools of approximately 20 000 bacterial cells were collected into 5 ml LB with 20 % glycerol (v/v) and stored at −80°C in 400 µl aliquots. Nine triparental mating experiments were performed to produce an HDTM library of approximately 1 000 000 mutants. To create the sequencing-ready HDTM library, glycerol stocks from each triparental mating were pooled together such that there was equal representation of each transposon mutant.

Molecular biology techniques

The DNA concentration of the samples was measured using the Qubit dsDNA BR Assay kit (Invitrogen). DNA (1–2 µl) was added to Qubit dsDNA BR working solution such that the total volume was 200 µl. The mixture was vortexed for 2–3 s, ensuring that no bubbles formed, and then incubated for 2–5 min at room temperature prior to the measurement of DNA concentration with a Qubit 2.0 fluorometer (Invitrogen).

DNA purification of amplicons was carried out using the Agencourt AMPure XP system (Beckman Coulter). Room temperature AMPure XP beads were mixed with DNA samples in low DNA-binding microcentrifuge tubes (Eppendorf), after which the mixtures were incubated for 5 min at room temperature. The tubes were then placed next to the magnets on the magnet rack to sit until all the beads were pulled towards the magnet. The supernatant was then removed, after which the beads were washed twice with 1 ml of 80 % ethanol while the tubes were still on the rack. After the supernatant was removed, the beads were allowed to air dry for 5 min on the magnet to allow evaporation of excess ethanol. The tubes were removed from the magnets, and DNA was eluted with EB buffer (Qiagen) The DNA samples were quantified after purification.

Genomic DNA isolation

Cell suspensions (0.5–1 ml) were washed twice, first in 500 µl 0.85 % saline followed by 500 µl TES (10 mM Tris, 30 mM EDTA, 150 mM NaCl in milliQ H₂O). Afterwards, the pellet was resuspended in 250 µl T10E25 (25 mM Tris, 62.5 mM EDTA in milliQ H₂O). Fifty microlitres of 2 mg ml⁻¹ lysozyme solution and 1 µl of 100 µg ml⁻¹ RNase (Qiagen) were added and then incubated for 15 min at 37°C. Next, 60 µl sarcosyl-protease solution (19.5 mM Tris, 0.975 mM EDTA, 100 mg ml⁻¹ sarcosyl, 5 g ml⁻¹ protease K in milliQ H₂O) was added followed by incubation for 16–18 h at 37°C. After incubation, the suspension was mixed with 361 µl PCI (50 % buffer-saturated phenol, 48 % chloroform, 2 % isoamyl alcohol; Invitrogen), then gently inactivated for 10 min until the mixture was homogenous. The suspension was transferred to a 1.5 ml phase-lock gel tube (Five Prime) and centrifuged for 5 min at 13 000 r.p.m. Afterwards, the upper aqueous phase was transferred to a low DNA-binding microcentrifuge tube. To precipitate the DNA, 0.1 volumes of 10 mM sodium acetate was added, followed by 0.54 volumes of 2-propanol. The suspension was allowed to sit at room temperature for at least 30 min then centrifuged for 30 min in a table-top centrifuge at 13 000 r. p. m. and 4°C. After removal of the supernatant, 1 ml ice-cold 70 % ethanol was added and the suspension was centrifuged for 10 min at 13 000 r.p.m. and 4°C. After removal of the ethanol, the DNA pellet was air-dried for 5 min to allow evaporation of excess ethanol. The DNA was resuspended in TE (Tris, EDTA) buffer and stored at 4°C. More TE buffer was added as necessary until the DNA was fully solubilized. Agarose gel electrophoresis was carried out to assess the quality of the DNA before quantifying by Qubit.

Tn-seq circle

The Tn-seq circle method [21] was used to enrich for the transposon–genome junction from the genomic DNA of transposon mutant libraries as described previously [21] with the modifications outlined here. Oligonucleotides used in this work are listed in Table S2. A sample of wild-type
Data analysis to determine gene essentiality

Gene essentiality was determined using modified versions of the custom scripts developed by Turner et al. [16] (see supporting data: https://github.com/khturner/Tn-seq/tree/dockerize). From the *B. cenocepacia* K56-2 gene feature format (gff) file, 10% from the 3’ end and 10% from the 5’ end was trimmed off of each gene so that insertions that are not likely to disrupt function are not included in the analysis. TnSeq2.sh (see supporting data: https://github.com/khturner/Tn-seq/tree/master/TnSeq2.sh) was used to identify reads containing the transposon sequence, remove the transposon tag and then map the remaining genomic sequences of the reads to the 17 contigs of the *B. cenocepacia* K56-2 draft genome (GenBank accession LAUA00000000) using Bowtie2 [52]. To avoid considering insertion sites that are over represented due to amplification bias, we discarded reads from the 100 most abundant insertion sites. To account for differential mutant abundance due to both genomic positional bias and PCR amplification bias we normalized the read data across the position of each contig, then based on G+C content of the region of insertion using the ‘im’ function in R. We totalled the number of reads per truncated gene to calculate the normalized number of reads per gene as well as the number of insertion sites per gene. A Monte Carlo simulation then was run where insertion sites were moved randomly across the genome to generate an ‘expected’ pseudo data set. This simulation was repeated 2000 times. The Monte Carlo simulations provided an estimate of the variance that would be expected if there were no essential genes. Since it is possible to have a gene with no insertions when comparing the real and pseudo data sets, to avoid dividing by zero, one was added to the read count for each gene.

The R package EdgeR [53] was used to compare the variance of the pseudo data set with the mean read count per gene of the real data. To determine the differences of mutant abundance between the real and pseudo data sets, EdgeR uses a negative binomial test. Since multiple comparisons increase the rate of identifying false positives [54], to control the false discovery rate (FDR), the Benjamini–Hochberg method was used, which computes an upper bound for the expected FDR and adjusts the *P* value accordingly to correct for multiple testing [55]. The R MCLUST package [56] was used to perform bimodal clustering of genes to either a ‘reduced’ or ‘unchanged’ mode, by fitting a parameterized bimodal Gaussian mixture model to the log2-transformed fold change mutant abundance. A gene was classified as essential if it was significantly depleted in the real data compared to the pseudo data (adjusted *P*<0.05, negative binomial test), and clustered in the ‘reduced’ mode (*P*<0.05, maximum-likelihood estimation).

Bioinformatics

The protein sequences from the 508 predicted essential genes of *B. cenocepacia* K56-2 were searched against all the bacterial essential genes in the database of essential genes (DEG 10, version 14.7, updated 24 October 2016) [57] using BLASTP with the default parameters provided in DEG and an E-value cut-off of 10⁻¹⁰. The essential homologues with the lowest E-value are listed in the Appendix. *B. cenocepacia* K56-2 homologues in *B. thailandensis* E264 [48], and *B. pseudomallei* K96243 [22] were identified as the best hits from performing BLASTP using Gencode software with an E-value cut-off of 10⁻¹⁰, minimum 30% sequence identity and 45% coverage. Operon predictions and clusters of orthologous categories (COG) [58] were assigned to genes of *B. cenocepacia* K56-2 from the corresponding homologous gene in *B. cenocepacia* J2315
listed in the Database of prOkaryotic OpeRons (DOOR) [59].

RESULTS
Production and sequencing of the HDTM libraries
With the goal of identifying the essential genome of B. cenocepacia K56-2, we used HDTM followed by Illumina sequencing of the transposon insertion sites. We generated a HDTM library of one million mutants by introducing the suicide plasmid pRBrahaBoutgfp [60] into B. cenocepacia K56-2 by triparental mating. To avoid insertions in non-suicide plasmid pRBrahaBoutgfp into a HDTM library of one million mutants by introducing the sequencing of the transposon insertion sites. We generated cepacia K56-2, we used HDTM followed by Illumina sequencing of the transposon insertion points.The first step in Tn-seq data analysis is to map the reads against the genome of the micro-organism under investigation (Fig. S5). As a complete genome of B. cenocepacia K56-2 is not available, we used the contigs from the draft genome of B. cenocepacia K56-2 [41] to map the Tn-seq reads and removed the 100 highest-read sites. After removing these reads, it was evident that there were positional effects on the insertion density and read counts (Fig. S6). These effects were not due to transposon insertion or Tn-seq method biases as the distribution of reads from sequencing the whole genome of B. cenocepacia K56-2 on the Illumina MiSeq platform exhibited the same trend. Notably, 17 of the 20 insertion sites with high read counts were in regions predicted to be genomic islands by Islandviewer3 [63] (Fig. 1). Overall, the reads were not evenly scaled across the contigs and the insertion density was strongly correlated with the G+C content (Fig. 1). To account for these biases, we created a model of read depth as a function of the position along each contig and G+C content. We then corrected the read count based on the model prediction to normalize the reads prior to the essentiality analysis, which minimized the effects of position and G+C content on read density (Fig. S7).

As species of the genus Burkhoderia have large multirepli-con genomes with approximately 67% G+C content [29], we first considered using a Hi-fidelity KAPA polymerase (KAPA bioscience) to amplify the transposon insertions by PCR. The KAPA DNA polymerase has been successfully used to increase the proportion of reads from transposon insertions in GC-rich regions in B. thailandensis [61] and has minimal amplification bias [62]. However, PCR-amplification of the transposon junctions may not be favoured by the KAPA DNA polymerase as the G+C content of the transposon sequence inserted into the genome is much lower (Fig. S2). To test the ability of the KAPA DNA polymerase to PCR-amplify the transposon junctions, the sequences of the HDTM library produced after PCR amplification with the KAPA DNA polymerase were compared with those obtained with the iTaq DNA polymerases (Bio-Rad). Sequencing the HDTM library after PCR-amplification with the KAPA polymerase resulted in 89,983 unique insertion sites with an average of 1 insert every 87 bases and a read G+C content of 61.0% (Table 1). However, PCR-amplification with the iTaq DNA polymerase revealed 293,568 unique insertion sites, with an average of 1 insert every 27 bases, a read G+C content of 59.7% (Table 1) and a lower proportion of insertion sites in GC-rich regions (Fig. S3). The total reads from PCR-amplification of the HDTM library with the iTaq DNA polymerase were more evenly distributed over the insertion sites, whereas the use of the KAPA DNA polymerase resulted in many insertions with a low read count and a large proportion of reads mapping to a small number of insertion sites (Fig. S4). For these reasons, the identification of essential genes was performed with the data produced using the iTaq DNA polymerase.

Table 1. Summary of results from sequencing the HDTM library

| HDTM library preparation | Total reads | No. of reads containing the transposon sequence and mapping to the genome | G+C content of reads | No. of unique insertion sites | Frequency of Tn insertion |
|--------------------------|-------------|-------------------------------------------------------------------------|---------------------|--------------------------------|--------------------------|
| KAPA                     | 20,459,975  | 4,370,457 (21 %)                                                         | 61.0 %              | 89,983                         | 1/87 bp                  |
| iTaq                     | 15,132,067  | 6,936,891 (46 %)                                                         | 59.7 %              | 293,568                        | 1/27 bp                  |
To identify which functional categories are represented in the essential genome of *B. cenocepacia* K56-2, we classified the essential and non-essential genes according to the cluster of orthologous group (COG) categories [58] previously identified for the K56-2 homologues in J2315 and listed in the Database of prOkaryotic OpeRons (DOOR) [59]. We assessed the enrichment of functions encoded by the predicted essential gene set for *B. cenocepacia* K56-2 with respect to non-essential genes in the genome. The essential gene set of *B. cenocepacia* K56-2 was enriched for genes.

**Fig. 1.** Distribution of reads in the HDTM library mapped to the *B. cenocepacia* K56-2 genome. From the outermost ring inwards: Ring 1: Contigs from the *B. cenocepacia* K56-2 assembly aligned to *B. cenocepacia* J2315 replicons using progressiveMauve [107]; RC, Reverse complemented. Ring 2: GC skew, positive (green), negative (purple). Ring 3: Alignment of reads from Tn-seq (Tn inserts, red). Ring 4: G+C content (black). Ring 5: Alignment of reads from the whole genome sequencing of *B. cenocepacia* K56-2 prepared with the Nextera XT kit (whole genome sequencing reads, turquoise). Rings 3 and 5: Genome regions with read coverage more than one standard deviation from the mean read coverage are indicated by blue regions. The innermost ring is the bp maker, with the genomic islands indicated by yellow arcs. Replicons are not to scale. Figure created using BLAST Ring Image Generator (BRIG 0.95) [108].
involved in cell division, cell-wall functions, protein synthesis, replication, and co-enzyme metabolism (Fig. 2). Genes related to translation, ribosomal structure and biogenesis were the most highly enriched category in *B. cenocepacia* K56-2 (Fig. 2). The enrichment pattern observed was similar to that of *E. coli* K-12 [10] with the exceptions of three categories: (i) energy production and conversion, (ii) replication, recombination and repair, and (iii) cell motility (Fig. 3), possibly representing species-specific differences.

**Comparison of the essential genomes of *B. cenocepacia* K56-2 and J2315**

As a recent study reported the essential genome of *B. cenocepacia* J2315 [23], we predicted that the set identified by our study would overlap substantially with the one reported by Wong et al. given that strains K56-2 and J2315 belong to the same clonal complex [49]. However, while our predicted essential gene set consists of 508 genes, Wong et al. predicted 383 essential genes and 439 genes required for additional growth in liquid LB medium [23]. We first compared the enrichment of essential gene functions between the two strains of *B. cenocepacia*. The functions enriched in the essential gene set of K56-2 were similar to those enriched in the previously identified essential genes of J2315 (Fig. 4a). In contrast to *E. coli*, the essential gene sets of both K56-2 and J2315 involved in the replication, recombination and repair category are significantly over-represented, while genes involved in cell motility are under-represented (Fig. 4a). Although essential genes in the energy production and conversion category are more significantly enriched in strain J2315 (1.74-fold, \( P < 0.05 \)) than in strain K56-2 (1.26-

---

**Fig. 2.** Functional categories of the predicted essential and non-essential genes of *B. cenocepacia* K56-2. Functional categories are based on the clusters of orthologous groups (COG) [58], which were assigned to genes of *B. cenocepacia* K56-2 from the corresponding homologous gene in *B. cenocepacia* J2315. Functional categories significantly enriched or under-represented in the essential gene set are indicated with an asterisk (\( P < 0.05 \), Fisher exact test).
fold, $P=0.103$) (Fig. 4a), these values are very different from those for the same category in *E. coli* (0.33-fold, $P=0.103$) (Fig. 3). Overall, essential genes are distributed across the same core biological functions in both strains of *B. cenocepacia*.

We next investigated the reasons for the discrepancy between the number of essential genes found between the two strains. From the *B. cenocepacia* K56-2 essential gene set, 294 genes have homologues to essential genes identified in *B. cenocepacia* J2315. Eighty-nine and 212 genes were uniquely essential to J2315 and K56-2, respectively. However, differences in the methodology of essential gene calling could cause this discrepancy. Our essentiality analysis used the read count per gene, which reflects the abundance of each transposon mutant in the library, whereas the analysis used to determine the essential genome of *B. cenocepacia* J2315 only considered the insertion sites per gene [15, 23]. We hypothesized that the 212 genes identified as essential in strain K56-2 but not in strain J2315 may correspond to genes with a strong effect on fitness. Therefore, these genes should be present in the set of genes required for growth after passage of the initial J2315 transposon mutant library through LB medium [23]. However, only 37 K56-2 essential genes out of 212 had homologues in J2315 that were identified as essential for survival in LB. In summary, from the 212 genes identified as essential in strain K56-2 but not in strain J2315, the essentiality of 175 genes appears to be specific to *B. cenocepacia* K56-2. In support of our findings, 107 out of these 175 genes have homologues that have been identified as essential in either *B. pseudomallei* or *B. thailandensis*, or have high similarity with essential genes found in the Database of Essential Genes (DEG, BLASTP, expect value cut-off $10^{-10}$) [57]. Ninety percent of the remaining 68 genes are annotated as hypothetical proteins (Table S3).

Of the 89 genes that are essential in strain J2315 but not in strain K56-2, eight do not have homologues in K56-2. For the remaining 81 J2315-specific essential genes that have homologues identified as non-essential in strain K56-2, we
Fig. 4. Comparison of the essential gene sets of *B. cenocepacia* K56-2 and *B. cenocepacia* J2315. (a) Fold enrichment of COG functional categories for the *B. cenocepacia* K56-2 and J2315 essential genes. Functional categories significantly enriched (purple) or under-

| COG Functional Category                                      | K56-2   | J2315  |
|-------------------------------------------------------------|---------|--------|
| CELLULAR PROCESSES AND SIGNALLING                           |         |        |
| Cell cycle control, cell division, chromosome partitioning  | 4.06    | 3.51   |
| Cell wall/membrane/envelope biogenesis                      | 2.10    | 1.58   |
| Cell motility                                               | 0.27    | 0.00   |
| Posttranslational modification, protein turnover, chaperones| 1.00    | 1.28   |
| Signal transduction mechanisms                              | 0.41    | 0.27   |
| Intracellular trafficking, secretion, and vesicular transport| 1.61    | 1.30   |
| Defense mechanisms                                          | 0.57    | 1.03   |
| Extracellular structures                                    | 0.00    | 0.00   |
| INFORMATION STORAGE AND PROCESSING                          |         |        |
| RNA processing and modification                             | 13.52   | 19.05  |
| Chromatin structure and dynamics                            | 0.00    | 0.00   |
| Translation, ribosomal structure and biogenesis             | 5.03    | 7.39   |
| Transcription                                               | 0.61    | 0.67   |
| Replication, recombination and repair                       | 1.87    | 2.59   |
| METABOLISM                                                  |         |        |
| Energy production and conversion                            | 1.26    | 1.74   |
| Amino acid transport and metabolism                         | 0.52    | 0.22   |
| Nucleotide transport and metabolism                         | 2.00    | 2.80   |
| Carbohydrate transport and metabolism                       | 0.59    | 0.59   |
| Coenzyme transport and metabolism                           | 2.53    | 3.20   |
| Lipid transport and metabolism                              | 1.32    | 1.08   |
| Inorganic ion transport and metabolism                      | 0.48    | 0.25   |
| Secondary metabolites biosynthesis, transport and catabolism | 0.65    | 0.24   |
| POORLY CHARACTERIZED                                        |         |        |
| General function prediction only                            | 0.46    | 0.39   |
| Function unknown                                            | 0.54    | 0.38   |

Enriched

Under-represented
confirmed that 49 of the K56-2 genes contain high numbers of reads corresponding to multiple insertions throughout the gene [an example is shown in Fig. 4b(i)]. From the remaining 32 genes, 14 had significantly less reads (EdgeR adjusted $P$-value <0.05, negative binomial test) but could not be classified as essential with a high level of certainty ($P$>0.05, maximum-likelihood estimation). Next, we considered the possibility that some transposon insertions in non-essential genes in the J2315 transposon library may cause polar effects on downstream essential genes causing the

confirmed that 49 of the K56-2 genes contain high numbers of reads corresponding to multiple insertions throughout the gene [an example is shown in Fig. 4b(i)]. From the remaining 32 genes, 14 had significantly less reads (EdgeR adjusted $P$-value <0.05, negative binomial test) but could not be classified as essential with a high level of certainty ($P$>0.05, maximum-likelihood estimation). Next, we considered the possibility that some transposon insertions in non-essential genes in the J2315 transposon library may cause polar effects on downstream essential genes causing the

confirmed that 49 of the K56-2 genes contain high numbers of reads corresponding to multiple insertions throughout the gene [an example is shown in Fig. 4b(i)]. From the remaining 32 genes, 14 had significantly less reads (EdgeR adjusted $P$-value <0.05, negative binomial test) but could not be classified as essential with a high level of certainty ($P$>0.05, maximum-likelihood estimation). Next, we considered the possibility that some transposon insertions in non-essential genes in the J2315 transposon library may cause polar effects on downstream essential genes causing the

confirmed that 49 of the K56-2 genes contain high numbers of reads corresponding to multiple insertions throughout the gene [an example is shown in Fig. 4b(i)]. From the remaining 32 genes, 14 had significantly less reads (EdgeR adjusted $P$-value <0.05, negative binomial test) but could not be classified as essential with a high level of certainty ($P$>0.05, maximum-likelihood estimation). Next, we considered the possibility that some transposon insertions in non-essential genes in the J2315 transposon library may cause polar effects on downstream essential genes causing the

confirmed that 49 of the K56-2 genes contain high numbers of reads corresponding to multiple insertions throughout the gene [an example is shown in Fig. 4b(i)]. From the remaining 32 genes, 14 had significantly less reads (EdgeR adjusted $P$-value <0.05, negative binomial test) but could not be classified as essential with a high level of certainty ($P$>0.05, maximum-likelihood estimation). Next, we considered the possibility that some transposon insertions in non-essential genes in the J2315 transposon library may cause polar effects on downstream essential genes causing the

confirmed that 49 of the K56-2 genes contain high numbers of reads corresponding to multiple insertions throughout the gene [an example is shown in Fig. 4b(i)]. From the remaining 32 genes, 14 had significantly less reads (EdgeR adjusted $P$-value <0.05, negative binomial test) but could not be classified as essential with a high level of certainty ($P$>0.05, maximum-likelihood estimation). Next, we considered the possibility that some transposon insertions in non-essential genes in the J2315 transposon library may cause polar effects on downstream essential genes causing the

confirmed that 49 of the K56-2 genes contain high numbers of reads corresponding to multiple insertions throughout the gene [an example is shown in Fig. 4b(i)]. From the remaining 32 genes, 14 had significantly less reads (EdgeR adjusted $P$-value <0.05, negative binomial test) but could not be classified as essential with a high level of certainty ($P$>0.05, maximum-likelihood estimation). Next, we considered the possibility that some transposon insertions in non-essential genes in the J2315 transposon library may cause polar effects on downstream essential genes causing the

confirmed that 49 of the K56-2 genes contain high numbers of reads corresponding to multiple insertions throughout the gene [an example is shown in Fig. 4b(i)]. From the remaining 32 genes, 14 had significantly less reads (EdgeR adjusted $P$-value <0.05, negative binomial test) but could not be classified as essential with a high level of certainty ($P$>0.05, maximum-likelihood estimation). Next, we considered the possibility that some transposon insertions in non-essential genes in the J2315 transposon library may cause polar effects on downstream essential genes causing the

confirmed that 49 of the K56-2 genes contain high numbers of reads corresponding to multiple insertions throughout the gene [an example is shown in Fig. 4b(i)]. From the remaining 32 genes, 14 had significantly less reads (EdgeR adjusted $P$-value <0.05, negative binomial test) but could not be classified as essential with a high level of certainty ($P$>0.05, maximum-likelihood estimation). Next, we considered the possibility that some transposon insertions in non-essential genes in the J2315 transposon library may cause polar effects on downstream essential genes causing the

confirmed that 49 of the K56-2 genes contain high numbers of reads corresponding to multiple insertions throughout the gene [an example is shown in Fig. 4b(i)]. From the remaining 32 genes, 14 had significantly less reads (EdgeR adjusted $P$-value <0.05, negative binomial test) but could not be classified as essential with a high level of certainty ($P$>0.05, maximum-likelihood estimation). Next, we considered the possibility that some transposon insertions in non-essential genes in the J2315 transposon library may cause polar effects on downstream essential genes causing the

confirmed that 49 of the K56-2 genes contain high numbers of reads corresponding to multiple insertions throughout the gene [an example is shown in Fig. 4b(i)]. From the remaining 32 genes, 14 had significantly less reads (EdgeR adjusted $P$-value <0.05, negative binomial test) but could not be classified as essential with a high level of certainty ($P$>0.05, maximum-likelihood estimation). Next, we considered the possibility that some transposon insertions in non-essential genes in the J2315 transposon library may cause polar effects on downstream essential genes causing the

Fig. 5. Comparison of the essential genes identified in B. cenocepacia K56-2, B. cenocepacia J2315 [23], B. thailandensis E264 [48] and B. pseudomallei K96243 [22] showing common and unique essential genes. Homologues were identified as the best hit from performing BLASTp using Geneious software with an E-value cut-off of $10^{-10}$, minimum 30 % sequence identity and 45 % coverage.
transposon-disrupted gene to be identified as essential. In our study, 10 of these 32 genes had reads mapping to insertions within the internal 10–90% of a gene and the rhamnose-inducible promoter (PrhaB) oriented to express the gene downstream of the insertion, three of which have a gene identified as essential in K56-2 directly downstream of the gene with PrhaB insertion [see example in Fig. 4b(ii)]. One gene, BCAL2054 had a similar insertion pattern, but also had an insertion where PrhaB is oriented in the opposite direction of the ORF [Fig. 4b(iii)]. BCAL2054 encodes a putative HEAT-like repeat protein involved in energy production and conversion. The insertions within BCAL2054 may not be disruptive, since it contains multiple phycobilisome (PBS) lyase HEAT-like repeat domains (IPR004155) [70]. For 21 genes identified as essential in strain J2315 but not in strain K56-2, our library contains mutants with insertions that only disrupt within the last 35% of the 3′ end of the ORF [Fig. 4b(iv)], suggesting that only insertions into the 5′ end are disruptive. It is possible that this trend is also true for BCAL2332, BCAL2334 and BCAL2337 [Fig. 4b(vi), red arrows]. BCAL2334 and BCAL2337 had inserts within the last 32% of the 3′end, while BCAL2337 had insertions disrupting the last 13% of the 3′end and 33% from the 5′ end. In addition, BCAL2334 and BCAL2337 had inserts with PrhaB in the same orientation as the ORF, allowing transcription of downstream coding regions [Fig. 4b(v), green arrows]. Taken together, from 89 genes identified as essential in B. cenocepacia J2315 but not in B. cenocepacia K56-2, 57 genes may correspond to actual strain-specific essential genes in B. cenocepacia J2315.

Comparison of four essential genomes of Burkholderia

The availability of essential genomes belonging to the two Burkholderia groups of human pathogens [22, 23, 48] allowed us to investigate conservation of essential genes with the goal of identifying common putative targets for antibacterial development. We found 158 genes that were commonly essential to the four Burkholderia strains (Fig. 5 and Table S4). The most significantly enriched essential functions of the 158 common essential gene set were translation, ribosomal structure and biogenesis, nucleotide transport and metabolism and cell wall/membrane/envelope biogenesis (P<0.001, Fisher’s exact test) (Fig. 6). In order to identify pathways enriched in essential genes common to the four Burkholderia strains, the 158 essential genes were mapped to B. cenocepacia J2315 pathways and a perturbation score (PPS) was computed using BioCyc [71]. The pathways of B. cenocepacia J2315 were scored to measure the extent of the enrichment of the 158 essential genes in a pathway by combining the essentiality of all reactions in the pathway. From essential genes identified as common to the four Burkholderia strains, we found that three pathways involved in the maintenance of the cell envelope, were highly enriched in essential genes: lipid IVₐ biosynthesis, peptidoglycan biosynthesis and 4-amino-4-deoxy-arabinose (Ara4N) biosynthesis (Fig. 7).

Very recently, the essential genome of a non-ET12 clone of B. cenocepacia (H111) was published [72]. We compared the 158 Burkholderia common essential genes with the essential genome of B. cenocepacia H111 and found that all but 17 were conserved in this strain as well [72]. We then compared the unique insertion density scores (UIDs) of these 17 genes with the cut-off value (UID=0.01) Higgins et al. used to differentiate essential and non-essential genes. While all 17 genes had UID scores above the cut-off of 0.01, 16 had UID scores ranging from 0.0103 to 0.0290. While these genes may not be strictly essential in B. cenocepacia H111, fewer insertions in these genes are tolerated compared to the majority (94–89%) of the genes in the genome, suggesting their importance for fitness. It is notable that the essential gene set of H111 also contains the genes involved in the pathways enriched in essential genes (Fig. 7).

While peptidoglycan and lipid A biosynthesis are essential in many Gram-negative bacteria [73, 74], Ara4N biosynthesis is not. Production of Ara4N is usually initiated in response to the presence of outer membrane stressors, such as the antimicrobial peptide polymyxin [75–78]. Ara4N replaces the phosphate groups of lipopolysaccharide (LPS), reducing the negative charge of the outer membrane and decreasing the affinity of cationic antimicrobial peptides [76]. Interestingly, polymyxin resistance is constitutive in many pathogenic species of the genus Burkholderia [79]. This finding led us to investigate whether there might be more essential functions in Burkholderia that are related to their particular biology. The amino acid sequences of the 158 common essential genes in B. cenocepacia, B. thailandensis and B. pseudomallei were then searched against all the strains in the DEG [57], excluding Burkholderia. This analysis identified three genes that are uniquely essential in the Burkholderia strains (BLASTP, expect value cut-off 10⁻⁵). These genes are annotated as encoding a polysaccharide deacetylase, a lysophospholipid transporter, LpIT, and a general porin, Omp38 (OpcP). Homologues of these three Burkholderia-specific essential genes were also recently identified as essential in B. cenocepacia H111 by Higgins et al. [72]. Remarkably, the three genes encode cell-envelope-related proteins.

The putative polysaccharide deacetylase encoded by WQ49_RS30025 is an orthologue of that encoded by BCAL1935 in B. cenocepacia J2313 (reciprocal best hit, BLASTP), BTH_I2189 in B. thailandensis E264 and BPSL1468 in B. pseudomallei K96243 [70]. This protein is a member of the Ara4N biosynthesis gene cluster in B. cenocepacia J2315 and had been previously identified as essential in B. cenocepacia [67]. LpIT functions to transfer membrane-disrupting lysophospholipids [80–82] across the inner membrane into the cell where they can be re-acylated by an acyltransferase/acyl-acyl carrier protein synthetase (Aas) [83]. B. cenocepacia K56-2 LpIT (WQ49_RS29130) is an orthologue of the protein encoded by BCAL2111 in B. cenocepacia J2313, BTH_I2006 in B. thailandensis and BPSL2180 in B pseudomallei [70] (reciprocal best hit, BLASTP). Omp38 (OpcP) is a homologue of a major
general porin, OmpF, in *E. coli* [84, 85] that is not essential for growth [10]. Omp38 (OpcP) is essential in *B. cenocepacia* K56-2 (WQ49_RS05300), *B. cenocepacia* J2315 (BCAM1931) [23] and *B. thailandensis* E264 (BTH_II1520) [70]. While the orthologue of Omp38 (OpcP) in *B. pseudomallei*, BPSS0879, (reciprocal best hit, BLASTP), was not identified as essential by Moule et al. [22], four putative porins were that have sequence similarity to Omp38 (OpcP): BPSL1655 (41% identity, 95% coverage), BPSL1674 (40% identity, 94% coverage), BPSL1728 (43% identity, 95% coverage), and BPSS0252 (34% identity, 100% coverage), and could possibly perform an analogous function to Omp (OpcP). Taken together these results suggest that protection of the outer membrane by surface charge modification and permeability modulation are uniquely essential functions in the three species of the genus *Burkholderia*.

**DISCUSSION**

In this work, one million mutants were collected and analysed to determine the essential genome of *B. cenocepacia* K56-2. The essential gene analysis was modified from the method of Turner et al., which identifies essential genes based on the relative abundance of reads mapping to each gene in the genome [16]. From our analysis, the most apparent bias was the high recovery reads and insertion sites corresponding to A/T-rich regions. Due to the low G+C content of the transposon sequence and subsequent reduced efficiency of PCR amplification, the use of KAPA polymerase did not improve recovery of insertion sites in GC-rich regions (Figs S3 and S4). Since the highest read counts were correlated with the G+C content and the read density mapping to each contig varied, we were able to minimize these biases by normalizing the reads to each contig (opposed to

| CELLULAR PROCESSES AND SIGNALLING | Fold enrichment |
|-----------------------------------|-----------------|
| Cell cycle control, cell division, chromosome partitioning | 4.35 | 4.86 | 4.42 | 4.25 |
| Cell wall/membrane/envelope biogenesis | 3.49 | 3.23 | 2.34 | 2.31 |
| Cell motility | 0.00 | 0.00 | 0.00 | 0.00 |
| Posttranslational modification, protein turnover, chaperones | 0.74 | 0.77 | 0.67 | 0.63 |
| Signal transduction mechanisms | 0.29 | 0.32 | 0.31 | 0.32 |
| Intracellular trafficking, secretion, and vesicular transport | 1.73 | 1.35 | 1.25 | 1.21 |
| Defense mechanisms | 0.61 | 0.62 | 0.61 | 0.55 |
| Extracellular structures | 0.00 | 0.00 | 0.00 | 0.00 |

| INFORMATION STORAGE AND PROCESSING | Fold enrichment |
|-----------------------------------|-----------------|
| RNA processing and modification | 0.00 | 0.00 | 0.00 | 0.00 |
| Chromatin structure and dynamics | 0.00 | 0.00 | 0.00 | 0.00 |
| Translation, ribosomal structure and biogenesis | 0.33 | 0.38 | 0.28 | 0.33 |
| Transcription | 0.44 | 0.64 | 0.60 | 0.59 |
| Replication, recombination and repair | 2.13 | 3.14 | 2.68 | 2.35 |

| METABOLISM | Fold enrichment |
|------------|-----------------|
| Energy production and conversion | 1.97 | 2.05 | 2.15 | 2.07 |
| Amino acid transport and metabolism | 0.29 | 0.28 | 0.32 | 0.32 |
| Nucleotide transport and metabolism | 0.94 | 0.79 | 0.99 | 0.89 |
| Carbohydrate transport and metabolism | 0.53 | 0.59 | 0.62 | 0.61 |
| Coenzyme transport and metabolism | 2.64 | 2.73 | 2.46 | 2.26 |
| Lipid transport and metabolism | 1.32 | 1.39 | 1.39 | 1.41 |
| Inorganic ion transport and metabolism | 0.16 | 0.15 | 0.15 | 0.15 |
| Secondary metabolites biosynthesis, transport and catabolism | 0.00 | 0.00 | 0.00 | 0.00 |

| POORLY CHARACTERIZED | Fold enrichment |
|---------------------|----------------|
| General function prediction only | 0.37 | 0.38 | 0.42 | 0.43 |
| Function unknown | 0.27 | 0.23 | 0.24 | 0.24 |

**Fig. 6.** Functional categories enriched in the *Burkholderia* common essential genes. Fold enrichment of COG functional categories for the 158 essential genes common to *B. cenocepacia* K56-2, *B. cenocepacia* J2315, *B. pseudomallei* K96243 and *B. thailandensis* E264 compared to all the genes in each respective genome. Functional categories enriched (green) or under-represented (blue) in the essential gene sets are colour coded by *p*-value (Fisher’s exact test, legend).
From sequencing a library of one million transposon mutants amplified using iTaq, we identified 293,568 unique insertion sites at an average of one insertion site for every 27 bp and 508 putative essential genes. Based on the number of essential genes in 14 bacteria with diverse genome sizes, we previously estimated the number of essential genes in *B. cenocepacia* to be 300–700 [60]. Putative essential genomes of three species of the genus *Burkholderia* have since been identified by HDTM [22, 23, 48, 72]. The identification of 508 essential genes is in agreement with the essential genes predicted for other species of the genus *Burkholderia*, including 406 in *B. thailandensis* E264 [48], 505 in *B. pseudomallei* K96243 [22] and recently, 383 essential genes were identified in a member of the epidemic ET12 lineage, *B. cenocepacia* J2315, which is closely related to *B. cenocepacia* K56-2, *B. cenocepacia* J2315, *B. thailandensis* E264 and *B. pseudomallei* K96243. Pathways enriched in essential genes were determined using the pathway perturbation score (PPS) computed by BioCyc [71]. Red, essential genes; black, non-essential genes.
B. cenocepacia K56-2 [23]. Within the B. cenocepacia K56-2 essential gene set, 291 and 211 have homologues to essential genes found in B. thailandensis E264 and B. pseudomallei K96243, respectively, while 296 have homologues to 294 essential genes found in B. cenocepacia J2315. The high correlation with the functions represented by the B. cenocepacia J2315 essential gene set when grown in similar conditions as B. cenocepacia K56-2, as well as the identification of experimentally determined essential genes, give support to the validity of the essentiality analysis in B. cenocepacia K56-2.

We identified 212 and 89 genes that were uniquely essential in B. cenocepacia K56-2 and B. cenocepacia J2315, respectively. However, identification of these strain-specific genes could be due to differences in the methodologies used to generate the HDTM libraries and assign genes as essential in the two studies. B. cenocepacia strains K56-2 and J2315 were both mutagenized with a Tn5 transposon, however, in contrast to the triparental mating used in this work, the B. cenocepacia J2315 HDTM libraries were constructed by transforming electrocompeent B. cenocepacia J2315 with an in vitro transposition system [23]. While both methods involve a 2 h incubation before selecting for transposon mutants, competition among replicating cells is more likely to occur during the triparental mating than during electroporation, as the electroporated cells are fragile and are not expected to replicate. Depletion of transposon mutants with poor fitness would cause the disrupted genes in those mutants to be identified as essential, consistent with the identification of 37 genes as essential in B. cenocepacia K56-2 that were only required for survival in LB in B. cenocepacia J2315 (Table S3). The presence of essential genes downstream and in the same operon of a non-essential gene interrupted by a transposon insertion could also erroneously identify the interrupted gene as essential due to polar effects of the insertion. We considered that this could be responsible for the B. cenocepacia J2315 genes classified as essential by Wong et al. that we did not identify in B. cenocepacia K56-2, as their transposon did not contain an outward-facing promoter. However, although we did identify essential operons with different essential genes between the two strains of B. cenocepacia (J2315 and K56-2), these differences were not correlated with the gene order (Fig. 4b), indicating that polar effects of the transposons used in this work and by Wong et al. are not responsible for the differences with respect to genes within operons. Since non-lethal insertions in essential genes are possible when the insertion site is located near the 3’ or 5’ ends of a gene, we identified a gene as essential if there are not a significant number of reads mapped to the inner 80 % of the gene. However, insertions into the central portion of essential genes encoding multidomain proteins could be non-lethal if the insertion site is in the region outside of the essential domain [11, 86], as could be the case in BCAL2054 (Fig. 4b). Tolerance of insertions within the essential region of a gene could also occur due to a transient merodiploid state during replication, which could result in recovery of reads that map to an essential gene.

The genomic differences in the closely related isolates of B. cenocepacia could account for some of the discrepancies in the essential genes identified in B. cenocepacia K56-2 and B. cenocepacia J2315. For example, eight of the genes identified as essential in B. cenocepacia J2315 have no homologues in B. cenocepacia K56-2, and a 57-kb region encoding 57 genes is duplicated in B. cenocepacia J2315 (BCAL0969 to BCAL1026 and BCAL2901 to BCAL2846) but not B. cenocepacia K56-2 [40]. Fourteen genes in this region are essential in B. cenocepacia K56-2, most of which have homologues to essential genes identified in B. thailandensis and B. pseudomallei (Table S3). While 23.2 and 41.7 % of essential genes were identified as uniquely essential in B. cenocepacia J2315 and K56-2, a previous study showed that approximately 17 and 26 % of Pseudomonas aeruginosa PAO1 and PA14 essential genes, respectively, were uniquely essential for growth in CF sputum media [16]. The genomes of Pseudomonas aeruginosa PAO1 (6.26 Mb) and PA14 (6.54 Mb) are highly similar with strain-specific regions comprising 4.2 and 8.3 % of each genome, respectively [87]. Whether the essential genes that are distinct in B. cenocepacia K56-2 and J2315 are actually uniquely essential in each strain remains to be confirmed by further experimental evidence.

By comparing the essential gene set of B. cenocepacia K56-2 with B. pseudomallei, B. thailandensis and B. cenocepacia J2315, we identified 158 essential genes that are shared amongst the four Burkholderia strains. B. thailandensis and B. pseudomallei are genetically similar species that have different lifestyles and diverged from a common ancestor over 40 million years ago [88]. While B. cenocepacia belongs to the Bcc, B. thailandensis and B. pseudomallei belong to the B. pseudomallei group. Both B. cenocepacia and B. pseudomallei cause infection in immunocompromised individuals. B. pseudomallei is classified as a category B biological threat agent [89] and is fatal in 10 to 40 % of melioidosis cases [90]. B. thailandensis rarely causes disease in humans [91], but can replicate in cultured mammalian cells and resists predation by amoeba [92, 93]. The finding of 158 shared genes between divergent strains highlights the importance of some essential processes in these pathogenic Burkholderia strains. One hundred and forty-one of these common essential genes were also recently identified as essential in the non-ET12 strain of B. cenocepacia H111 [72]. It is notable that 16 of the remaining 17 genes not identified as essential had fewer insertions per gene than approximately 90 % of genes in the genome of B. cenocepacia H111 [72], suggesting that these genes are important for growth in B. cenocepacia H111 as well. From the conserved essential gene set, we found that pathways involved in maintaining the integrity of the cell envelope are enriched in the common essential genes. Furthermore, we identified three genes that are uniquely essential in these four Burkholderia strains that are also involved in maintaining the integrity of the envelope.
BCAL1935 is transcribed with genes involved in L-Ara4N modification of lipid A, which has been previously identified as an essential process in B. cenocepacia [67]. BCAL1935 is included in a transcriptional unit that encodes the polysaccharide deacetylase and BCAL1936, which are predicted to synthesize UDP-Ara4N [67], however, the function of BCAL1935 is not known. While L-Ara4N-modified lipid A is required for resistance to antimicrobial peptides, it is not essential in most Gram-negative bacteria [76]. Recently it was discovered that in addition to L-Ara4N providing resistance to antimicrobial peptides, the presence of the L-Ara4N modification is required in order for LPS to be exported in Burkholderia strains [94], indicating how this modification is essential in B. cenocepacia.

The second gene identified as uniquely essential in the four Burkholderia strains encodes a lysophospholipid transporter (LpIT) that transfers membrane-disrupting lysophospholipids across the inner membrane where they can be re-acylated by Aas [83]. LpIT is not essential for growth in E. coli [10, 95] and the reason LpIT is essential for growth in Burkholderia is unknown. However, conservation of the gene lplT would provide an advantage for surviving within a microbial community as well as within the host, which both employ phospholipases as a defensive mechanism.

We also identified a porin, Omp38 (OpcP), that is uniquely essential in three of the four Burkholderia strains. While OpcP is not essential in B. pseudomallei, four other essential porins, BPSL1655, BPSL1674, BPSL1728 and BPSS0252, were identified as essential [22]. Although none of the orthologues of the porins that are essential in B. pseudomallei were predicted to be essential in B. cenocepacia J2315 [23] or K56-2, it is possible that these porins provide a similar function to Omp38 (OpcP). The reason for the essentiality of the Omp38 (OpcP) porin is unknown, however trimeric porins like Omp38 (OpcP) contain binding sites that stabilize the LPS in the outer membrane in Gram-negative bacteria [96]. The outer membrane of species of the genus Burkholderia has been shown to be 89% less permeable than E. coli [97] and the impermeability of the membrane in Gram-negative bacteria has been attributed to the functional characteristics of porins and LPS [76, 98].

Tn-seq of HDTM libraries provides a valuable resource for the identification of antibiotic resistance mechanisms [21, 99–101]. For example, clinical isolates of B. pseudomallei have been identified where loss of penicillin-binding protein 3 results in a high level of resistance to cefazidime [102]. Future work passaging HDTM libraries in the presence of antibiotics has the potential uncover other resistance mechanisms. In addition, the presence of uniquely essential processes identified in this work suggest a role in survival specific the genus Burkholderia. However, further study of the essential genomes of other species of the genus Burkholderia is needed to confirm this hypothesis. Efforts are ongoing to find alternatives to conventional treatments for pathogenic species of the genus Burkholderia [103–106]. The envelope proteins identified in this work may represent ideal candidates for immunogenic targeting, as conserved essential proteins are less prone to variability.

Data bibliography
STC was supported by research grants from Cystic Fibrosis Canada (CFC) and the Natural Sciences and Engineering Research Council of Canada (NSERC). ASG was supported by a Faculty of Science Graduate Studentship.

Conflicts of interest
The authors declare that there are no conflicts of interest.

References
1. Jordan IK, Rogozin IB, Wolf YI, Koonin EV. Essential genes are more evolutionarily conserved than are nonessential genes in bacteria. Genome Res 2002;12:962–968.
2. Rocha EP, Danchin A. Essentiality, not expressiveness, drives gene-strand bias in bacteria. Nat Genet 2003;34:377–378.
3. Rocha EP, Danchin A. Gene essentiality determines chromosome organisation in bacteria. Nucleic Acids Res 2003;31:6570–6577.
4. Rocha EP, Danchin A. An analysis of determinants of amino acids substitution rates in bacterial proteins. Mol Biol Evol 2004; 21:108–116.
5. Zhang R, Ou HY, Zhang CT. DEG: a database of essential genes. Nucleic Acids Res 2004;32:2710–2720.
6. Gao F, Zhang RR. Enzymes are enriched in bacterial essential genes. PLoS One 2011;6:e21683.
7. Fields FR, Lee SW, McConnell MJ. Using bacterial genomes and essential genes for the development of new antibiotics. Biochem Pharmacol 2017;134:74–86.
8. Mobegi FM, van Hijum SA, Burghout P, Bootsma HJ, de Vries SP et al. From microbial gene essentiality to novel antimicrobial drug targets. BMC Genomics 2014;15:958.
9. Murima P, McKinney JD, Petro K. Bacterial antibiotic resistance: globalisation and solution. Biochem Pharmacol 2014;124:1423–1432.
10. Baba T, Ara T, Hasegawa M, Takai Y, Okumura Y et al. Construction of Escherichia coli K-12 in-frame, single-gene knockout mutants: the Keio collection. Mol Syst Biol 2006;2:2006.0008.
11. Gerdes SY, Scholle MD, Campbell JW, Balázs G, Ravasz E et al. Experimental determination and system level analysis of essential genes in Escherichia coli MG1655. J Bacteriol 2003;185:5673–5684.
12. Kobayashi K, Ehrlich SD, Albertini A, Amati G, Andersen KK et al. Essential Bacillus subtilis genes. Proc Natl Acad Sci USA 2003;100:4678–4683.
13. Mori H, Isono K, Horiuchi T, Miki T. Functional genomics of Escherichia coli in Japan. Res Microbiol 2000;151:121–128.
14. Christen B, Abeliuk E, Collier JM, Kalogeraki VS, Passarelli B et al. The essential genome of a bacterium. Mol Syst Biol 2011;7:528–535.
15. Langridge GC, Phan M-D, Turner DJ, Perkins TT, Parts L et al. Simultaneous assay of every EScherichia coli–TnSeq transposon mutant. Genome Res 2009;19:2308–2316.
16. Turner KH, Wessel AK, Palmer GC, Murray JL, Whiteley M. Essential genome of Pseudomonas aeruginosa in cystic fibrosis sputum. Proc Natl Acad Sci USA 2015;112:4110–4115.
17. Weerdenburg EM, Abdallah AM, Rangkuti F, Abd El Ghany M, Otto TD et al. Genome-wide transposon mutagenesis indicates that Mycobacterium marinum customizes its virulence mecha-
nisms for survival and replication in different hosts. Infect Immun 2015;83:1778–1788.
18. Yang H, Krumholz EW, Brutinel ED, Palani NP, Sadowsky MJ et al. Genome-scale metabolic network validation of Shewanella oneidensis using transposon insertion frequency analysis. PLoS Comput Biol 2014;10:e1003848.
19. Gawronski JD, Wong SMS, Giannoukos G, Ward DV, Akerley BJ. Tracking insertion mutants within libraries by deep sequencing and a genome-wide screen for Haemophilus genes required in the lung. Proc Natl Acad Sci USA 2009;106:16422–16427.
20. Goodman AL, McNulty NP, Zhao Y, Leip D, Mitra RD et al. Identifying genetic determinants needed to establish a human gut symbiont in its habitat. Cell Host Microbe 2009;6:279–289.
21. Gallagher LA, Shendure J, Manoil C. Genome-scale identification of resistance functions in Pseudomonas aeruginosa using Tn-seq. MBio 2011;2:e00315-10.
22. Moule MG, Hemsley CM, Seet Q, Guerra-Assunção JA, Lim J et al. Genome-wide saturation mutagenesis of Burkholderia pseudomallei K96243 predicts essential genes and novel targets for antimicrobial development. MBio 2014;5:e00926-13.
23. Wong Y-C, Abd El Ghany M, Naeem R, Lee KW, Tan YC et al. Candidate essential genes in Burkholderia cenocepacia J2315 identified by genome-wide TraDIS. Front Microbiol 2016;7:1288.
24. Eberl L, Vandamme P. Members of the genus Burkholderia: good and bad guys. F1000Res 2016;5:1007.
25. Sahi JW, Vazquez AJ, Hall CM, Busch JD, Tuanyok A et al. The effects of signal erosion and core genome reduction on the identification of diagnostic markers. MBio 2016;7:e00846-16.
26. Glass MB, Steigerwalt AG, Jordan JG, Wikins PP, Gee JE. Burkholderia oklahomensis sp. nov., a Burkholderia pseudomallei-like species formerly known as the Oklahoma strain of Pseudomo-
nas pseudomallei. Int J Syst Evol Microbiol 2006;56:2171–2176.
27. Tuanyok A, Mayo M, Scholz H, Hall CM, Allender CJ et al. Burkholderia humphrydooensis sp. nov., a new species related to Burkholderia thailandensis and the fifth member of the Burkholderia pseudomallei complex. Appl Environ Microbiol 2017;83:e02802-16.
28. Wiersinga WJ, van der Pol T, White NJ, Day NP, Peacock SJ. Melioidosis: insights into the pathogenicity of Burkholderia pseudomallei. Nat Rev Microbiol 2006;4:272–282.
29. Mahenthiralingam E, Vandamme P. Taxonomy and pathogene-
sis of the Burkholderia cenocepacia complex. Chron Respir Dis 2005; 2:209–217.
30. Vanlaere E, Lipum a JJ, Baldwin A, Henry D, de Brandt E et al. Burkholderia latens sp. nov., Burkholderia diffusa sp. nov., Bur-
kholderia arboris sp. nov., Burkholderia seminalis sp. nov. and Burkholderia metallica sp. nov., novel species within the Burkh
holderia cenocepacia complex. Int J Syst Evol Microbiol 2008;58:1580–1590.
31. Vanlaere E, Baldwin A, Gevers D, Henry D, de Brandt E et al. Taxon K, a complex within the Burkholderia cenocepacia complex, comprises at least two novel species, Burkholderia contaminans sp. nov. and Burkholderia lata sp. nov. Int J Syst Evol Microbiol 2009;59:102–111.
32. Peeters C, Zlosnik JEA, Spilker T, Hird TJ, LiPuma JJ et al. Burkholderia pseudomultivirans sp. nov., a novel Burkholderia cepa-
cia complex species from human respiratory samples and the rhizosphere. Syst Appl Microbiol 2013;36:483–489.
33. de Smet B, Spilker T, Ginther JL, Currie BJ, Zlosnik JEA et al. Burkholderia stagnalis sp. nov. and Burkholderia territiori sp. nov., two novel Burkholderia cenocepacia complex species from environment-
al and human sources. Int J Syst Evol Microbiol 2015;65:2265–2271.
34. Mahenthiralingam E, Baldwin A, Dowson CG. Burkholderia cepacia complex bacteria: opportunistic pathogens with important natural history. J Appl Microbiol 2008;104:1539–1551.
35. Darling P, Chan M, Cox AD, Sokol PA. Siderophage production by cystic fibrosis isolates of Burkholderia cenocepacia. Infect Immun 1998;66:874–877.
36. Lipuma JJ, Spilker T, Gill LH, Campbell PW, Liu L et al. Dispro-
portionate distribution of Burkholderia cenocepacia complex species and transmissibility markers in cystic fibrosis. Am J Respir Crit Care Med 2001;164:92–96.
37. Cardona ST, Wopperer J, Eberl L, Valvano MA. Diverse patho-
genicity of Burkholderia cenocepacia complex strains in the Caenorhabditis elegans host model. FEMS Microbiol Lett 2005;250:97–104.
38. Vergunst AC, Meijer AH, Renshaw SA, O’Callaghan D. Burkh-
holderia cenocepacia creates an intramacrophage replication niche in zebrafish embryos, followed by bacterial dissemination and establishment of systemic infection. Infect Immun 2010;78: 1495–1508.
39. Mahenthiralingam E, Coenye T, Chung JW, Speert DP, Govan JR et al. Diagnostically and experimentally useful panel of strains from the Burkholderia cenocepacia complex. J Clin Microbiol 2000;38: 910–913.
40. Holden MT, Seth-Smith HM, Crossman LC, Sebaihia M, Bentley SD et al. The genome of Burkholderia cenocepacia J2315, an epi-
demic pathogen of cystic fibrosis patients. J Bacteriol 2009;191: 261–277.
41. Bloodworth R. Essential genes and genomes of the Burkholderia cenocepacia complex. 2013. http://mspace.lib.umanitoba.ca/xmlui/handle/1993/30912. Accessed 31 March 2017.
42. Baldwin A, Sokol PA, Parkhill J, Mahenthiralingam E. The Bur-
kholderia cenocepacia epidemic strain marker is a novel genomic island encoding both virulence and metabolism-associ-
ated genes in Burkholderia cenocepacia. Infect Immun 2004;72: 1537–1547.
43. Graindorge A, Menard A, Monnez C, Cournoyer B. Insertion sequence evolutionary patterns highlight congenic genetic
inactivations and recent genomic island acquisitions among epi-
demic Burkholderia cenocepacia. J Med Microbiol 2012;61:394–409.
44. Lessie TG, Hendrickson W, Manning BD, Devereux R. Genomic complexity and plasticity of Burkholderia cenocepacia. FEMS Microbiol Lett 1996;144:117–128.
45. Seo YS, Lim JY, Park J, Kim S, Lee HH et al. Comparative genome analysis of rice-pathogenic Burkholderia virulence
provides insight into capacity to adapt to different environments and hosts. BMC Genomics 2015;16:349.
46. Horvath P, Barranguet R. CRISPR/Cas, the immune system of bacteria and archaea. Science 2010;327:167–170.
47. Dennis JJ. Burkholderia cenocepacia virulence microevolution in the CF lung: variations on a theme. Virulence 2016;1:3–
48. Baugh L, Gallagher LA, Patrapuvich R, Clifton MC, Gardberg AS et al. Combining functional and structural genomics to sample the essential Burkholderia structure. PLoS One 2013;8:e53851.
49. Baldwin A, Mahenthiralingam E, Thickett KM, Honeybourne D, Maiden MC et al. Multilocus sequence typing scheme that pro-
vides both species and strain differentiation for the Burkholderia cenocepacia complex. J Clin Microbiol 2005;43:4665–4673.
50. Miller VL, Mekalanos JJ. A novel suicide vector and its use in construction of insertion mutations: osmoregulation of outer membrane proteins and virulence determinants in Vibrio chlo-
erae requires toXR. J Bacteriol 1998;170:2575–2583.
51. Figurski DH, Helinski DR. Replication of an origin-containing derivative of plasmid RK2 dependent on a plasmid function pro-
vided in trans. Proc Natl Acad Sci USA 1979;76:1648–1652.
52. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. Nat Methods 2012;9:357–359.
53. Robinson MD, McCarthy DJ, Smyth GK. edger: A Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 2010;26:139–140.

54. Noble WS. How does multiple testing correction work? Nat Biotechnol 2009;27:1135–1137.

55. Benjamin Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. J R Stat Soc Ser B Methodol 1995;57:289–300.

56. Fraley C, Raftery AE. MCLUST: software for model-based cluster analysis. J Classif 1999;16:297–306.

57. Luo H, Lin Y, Gao F, Zhang CT, Zhang R. DEG 10, an update of the database of essential genes that includes both protein-coding genes and noncoding genomic elements. Nucleic Acids Res 2014;42:D574–D580.

58. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. Nucleic Acids Res 2000;28:33–36.

59. Mao X, Ma Q, Zhou C, Chen X, Zhang H et al. DOOR 2.0: presenting operons and their functions through dynamic and integrated views. Nucleic Acids Res 2014;42:D654–D659.

60. Bloodworth RAM, Gislason AS, Cardona ST. Burkholderia cenocepacia conditional growth mutant library created by random promoter replacement of essential genes. Microbiogenopen 2013;2:243–258.

61. Gallagher LA, Ramage E, Patrapuvich R, Weiss E, Brittnacher M et al. Sequence-defined transposon mutant library of Burkholderia thailandensis. MBio 2013;4:e00604-13.

62. Quail MA, Otto TD, Gu Y, Harris SR, Skelly TF et al. Optimal enzymes for amplifying sequencing libraries. Nat Methods 2011; 9:10–11.

63. Dhillon BK, Laird MR, Shay JA, Winsor GL, Lo R et al. IslandViewer 3: more flexible, interactive genomic island discovery, visualization and analysis. Nucleic Acids Res 2015;43:W104–W108.

64. Cardona ST, Mueller C, Valvano MA. Identification of essential operons in Burkholderia cenocepacia with a rhamnose inducible promoter. Appl Environ Microbiol 2006;72:2547–2555.

65. Juhas M, Stark M, von Mering C, Lumijakstase P, Crook DW et al. High confidence prediction of essential genes in Burkholderia cenocepacia. PLoS One 2012;7:e40064.

66. Mohamed YF, Valvano MA. A Burkholderia cenocepacia MurJ (MvNh) homolog is essential for cell wall peptidoglycan synthesis and bacterial viability. Glycobiology 2014;24:546–576.

67. Ortega XP, Cardona ST, Brown AR, Loutet SA, Flannagan RS et al. A putative gene cluster for aminoarabinose biosynthesis is essential for Burkholderia cenocepacia viability. J Bacteriol 2007;189:3639–3644.

68. Bloodworth RAM, Ziltni S, Brown ED, Cardona ST. An electron flavoprotein is essential for viability and its depletion causes a rod-to-sphere change in Burkholderia cenocepacia. Microbiology 2015;161:1909–1920.

69. Gislason AS, Choy M, Bloodworth RA, Qu W, Stietz MS et al. Competitive growth enhances conditional growth mutant sensitivity to antibiotics and exposes a two-component system as an emerging antibacterial target in Burkholderia cenocepacia. Antimicrob Agents Chemother 2017;2016.

70. Winsor GL, Khaira B, Van Rossum T, Lo R, Whiteside MD et al. The Burkholderia genome database: facilitating flexible queries and comparative analyses. Bioinformatics 2008;24:2803–2804.

71. Caspi R, Billington R, Ferrer L, Foerster H, Fulcher CA et al. The MetaCyc database of metabolic pathways and enzymes and the BioCyc collection of pathway/genome databases. Nucleic Acids Res 2016;44:D471–D480.

72. Higgins S, Sanchez-Contreras M, Gualdi S, Pinto-Carbó M, Carlier A et al. The essential genome of Burkholderia cenocepacia H111. J Bacteriol 2017;199:e00260-17.

73. Onishi HR, Pelak BA, Gerckens LS, Silver LL, Kahan FM et al. Antibacterial agents that inhibit lipid A biosynthesis. Science 1996;274:980–982.

74. Tomasz A. The mechanism of the irreversible antimicrobial effects of penicillins: how the beta-lactam antibiotics kill and lyse bacteria. Annu Rev Microbiol 1979;33:113–137.

75. Moskowitz SM, Ernst RK, Miller SI. PmRAb, a two-component regulatory system of Pseudomonas aeruginosa that modulates resistance to cationic antimicrobial peptides and addition of aminolaevulinic acid to lipid A. J Bacteriol 2004;186:575–579.

76. Raetz CR, Reynolds CM, Trent MS, Bishop RE. Lipid A modification systems in gram-negative bacteria. Annu Rev Biochem 2007;76:295–329.

77. Wang X, Quinn PJ. Endotoxins: lipopolysaccharides of gram-negative bacteria. Subcell Biochem 2010;53:3–25.

78. Zhou Z, Ribeiro AA, Lin S, Cotter RJ, Miller SI et al. Lipid A modifications in polymyxin-resistant Salmonella typhimurium: PMRA-dependent 4-amino-4-deoxy-L-arabinose, and phosphoethanolamine incorporation. J Biol Chem 2001;276:43111–43121.

79. Loutet S, Valvano MA. Extreme antimicrobial peptide and polymyxin B resistance in the genus Burkholderia. Front Cell Infect Microbiol 2011;1:6.

80. Elsbach P, Weiss J. The bacterialcidal/permeability-increasing protein (BPI), a potent element in host-defense against Gram-negative bacteria and lipopolysaccharide. Immunobiology 1993;187:417–429.

81. Russell AB, Leroux M, Hathazi K, Agnello DM, Ishikawa T et al. Diverse type VI secretion phospholipases are functionally plastic antibiotic effectors. Nature 2013;496:508–512.

82. Wright GC, Weiss J, Kim KS, Verheij H, Elsbach P. Bacterial phospholipid hydrolysis enhances the destruction of Escherichia coli ingested by rabbit neutrophils. Role of cellular and extracellular phospholipases. J Clin Invest 1990;85:1925–1935.

83. Harvat EM, Zhang YM, Tran CV, Zhang Z, Frank MW et al. Lyso-phospholipid flipping across the Escherichia coli inner membrane catalyzed by a transporter (LptI) belonging to the major facilitator superfamily. J Biol Chem 2005;280:12028–12034.

84. Siripatawee J, Prinz H, Krittanan C, Suginta W. Expression and refolding of Omp38 from Burkholderia pseudomallei and Burkholderia thailandensis, and its function as a diffusion porin. Biochim J 2004;384:609–617.

85. Siripatawee J, Prinz H, Samosornsubk W, Ashley RH, Suginta W. Functional reconstitution, gene isolation and topology modeling of porins from Burkholderia pseudomallei and Burkholderia thailandensis. Biochem J 2004;377:579–587.

86. Zhang YJ, Loerger TR, Huttenhower C, Long JE, Sassetti CM et al. Global assessment of genomic regions required for growth in Mycobacterium tuberculosis. PLoS Pathog 2012;8:e1002946.

87. Lee DG, Urbach JM, Wu G, Liberati NT, Feinbaum RL et al. Genomic analysis reveals that Pseudomonas aeruginosa virulence is combinatorial. Genome Biol 2006;7:R90.

88. Yu Y, Kim HS, Chua H, Lin C, Sim S et al. Genomic patterns of pathogen evolution revealed by comparison of Burkholderia pseudomallei, the causative agent of melioidosis, to avirulent Burkholderia thailandensis. BMC Microbiol 2006;6:46.

89. Centers for Disease Control and Prevention (CDC). Department of Health and Human Services (HHS). Possession, use, and transfer of select agents and toxins; biennial review of the list of select agents and toxins and enhanced biosafety requirements. Final rule. Fed Regist 2017;82:6278–6294.

90. Chawrapreecha C, Holden MTG, Vehkala M, Välimäki N, Yang Z et al. Global and regional dissemination and evolution of Burkholderia pseudomallei. Nat Microbiol 2017;2:16263.

91. Glass MB, Gee JE, Steigerwalt AG, Cavuoti D, Barton T et al. Pneumonia and septicemia caused by Burkholderia thailandensis in the United States. J Clin Microbiol 2006;44:4601–4604.
92. Harley VS, Dance DA, Drasar BS, Tovey G. Effects of *Burkholderia pseudomallei* and other *Burkholderia* species on eukaryotic cells in tissue culture. *Microbios* 1998;96:71–93.

93. Hasselbring BM, Patel MK, Schell MA. Dictyostelium discoideum as a model system for identification of *Burkholderia pseudomallei* virulence factors. *Infect Immun* 2011;79:2079–2088.

94. Hamad MA, di Lorenzo F, Molinaro A, Valvano MA. Aminoarabinose is essential for lipopolysaccharide export and intrinsic antimicrobial peptide resistance in *Burkholderia cenocepacia*. *Mol Microbiol* 2012;85:962–974.

95. Lin Y, Bogdanov M, Tong S, Guan Z, Zheng L. Substrate selectivity of lysophospholipid transporter LptT involved in membrane phospholipid remodeling in *Escherichia coli*. *J Biol Chem* 2016;291:2136–2149.

96. Arunmanee W, Pathania M, Solovyova AS, Le Brun AP, Ridley H et al. Gram-negative trimeric porins have specific LPS binding sites that are essential for porin biogenesis. *Proc Natl Acad Sci USA* 2016;113:E5034–E5043.

97. Hancock RE. Resistance mechanisms in *Pseudomonas aeruginosa* and other nonfermentative gram-negative bacteria. *Clin Infect Dis* 1998;27 Suppl 1:593–599.

98. Pagès J-M, James CE, Winterhalter M. The porin and the permeating antibiotic: a selective diffusion barrier in Gram-negative bacteria. *Nat Rev Microbiol* 2008;6:893–903.

99. Blake KL, O’Neill AJ. Transposon library screening for identification of genetic loci participating in intrinsic susceptibility and acquired resistance to antistaphylococcal agents. *J Antimicrob Chemother* 2013;68:12–16.

100. Santa Maria JP, Sadaka A, Moussa SH, Brown S, Zhang YJ et al. Compound-gene interaction mapping reveals distinct roles for *Staphylococcus aureus* teichoic acids. *Proc Natl Acad Sci USA* 2014;111:12510–12515.

101. Shan Y, Lazinski D, Rowe S, Camilli A, Lewis K. Genetic basis of persister tolerance to aminoglycosides in *Escherichia coli*. *MBio* 2015;6:e00078-15.

102. Chantratita N, Rholl DA, Sim B, Wuthiekanun V, Limmathurotsakul D et al. Antimicrobial resistance to ceftazidime involving loss of penicillin-binding protein 3 in *Burkholderia pseudomallei*. *Proc Natl Acad Sci USA* 2011;108:17165–17170.

103. Pradenas GA, Myers JN, Torres AG. Characterization of the *Burkholderia cenocepacia* TonB mutant as a potential live attenuated vaccine. *Vaccines* 2017;5:33.

104. Titball RW, Burtnick MN, Bancroft GJ, Brett P. *Burkholderia pseudomallei* and *Burkholderia mallei* vaccines: are we close to clinical trials? *Vaccine* 2017;35:5981–5989.

105. Semler DD, Goudie AD, Finlay WH, Dennis JJ. Aerosol phage therapy efficacy in *Burkholderia cepacia* complex respiratory infections. *Antimicrob Agents Chemother* 2014;58:4005–4013.

106. Guang-Han O, Leang-Chung C, Vellasamy KM, Mariappan V, Li-Yen C et al. Experimental phage therapy for *Burkholderia pseudomallei* infection. *PLoS One* 2016;11:e0158213.

107. Darling AE, Mau B, Perna NT. progressiveMauve: multiple genome alignment with gene gain, loss and rearrangement. *PLoS One* 2010;5:e11147.

108. Alikhan NF, Petty NK, Ben Zakour NL, Beatson SA. BLAST Ring Image Generator (BRIG): simple prokaryote genome comparisons. *BM Genomics* 2011;12:402.

Five reasons to publish your next article with a Microbiology Society journal

1. The Microbiology Society is a not-for-profit organization.
2. We offer fast and rigorous peer review – average time to first decision is 4–6 weeks.
3. Our journals have a global readership with subscriptions held in research institutions around the world.
4. 80% of our authors rate our submission process as ‘excellent’ or ‘very good’.
5. Your article will be published on an interactive journal platform with advanced metrics.

Find out more and submit your article at microbiologyresearch.org.