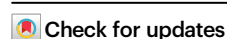














De novo mutations mediate phenotypic switching in an opportunistic human lung pathogen

Received: 13 January 2025

Accepted: 12 June 2025

Published online: 23 July 2025



Alexandra J. Poret ^{1,2,3}, Matthew Schaefer ^{4,5}, Christina Merakou ^{4,5}, Kathryn E. Mansour⁴, Connor D. Ahern⁴, Georgia K. Lagoudas ^{3,6}, Alyssa Haynes ^{1,2}, Ashley R. Cross⁷, Joanna B. Goldberg ⁷, Roy Kishony ⁸, Ahmet Z. Uluer^{9,10,11}, Alexander J. McAdam^{12,13}, Paul C. Blainey ^{3,6,14}, Sara O. Vargas¹³, Tami D. Lieberman ^{1,2,6,16}  & Gregory P. Priebe ^{4,5,6,15,16} 

Bacteria evolving within human hosts encounter selective tradeoffs that render mutations adaptive in one context and deleterious in another. Here, we report that the cystic fibrosis-associated pathogen *Burkholderia dolosa* overcomes in-human selective tradeoffs by acquiring successive point mutations that alternate phenotypes. We sequenced the whole genomes of 931 respiratory isolates from two recently infected cystic fibrosis patients and an epidemiologically-linked, chronically-infected patient. These isolates are contextualized using 112 historical genomes from the same outbreak strain. Within both newly infected patients, convergent mutations that disrupt O-antigen expression quickly arose, comprising 29% and 63% of their *B. dolosa* communities by 3 years. The selection for loss of O-antigen starkly contrasts with our previous observation of parallel O-antigen-restoring mutations after many years of chronic infection in the historical outbreak. Experimental characterization reveals that O-antigen loss increases uptake in immune cells while decreasing competitiveness in the mouse lung. We propose that the balance of these pressures, and thus whether O-antigen expression is advantageous, depends on tissue localization and infection duration. These results suggest that mutation-driven phenotypic alternation may be underestimated without dense temporal sampling, particularly for microbes with prolonged infection or colonization.

Bacteria colonizing humans are subjected to environmental variations across both space and time. The optimal survival strategy in one condition can be counterproductive in another, creating selective tradeoffs^{1–3}. Bacteria employ a range of strategies to navigate tradeoffs, including evolving environment-responsive regulatory elements⁴, acquiring and discarding genetic programs via horizontal gene transfer⁵, and developing sequence motifs prone to high contraction and expansion rates⁶.

The acquisition of de novo mutations to toggle a phenotype on and off has traditionally been considered inefficient due to the low

likelihood of restoring a pathway's original function by random mutation^{7,8}. However, reversion has occasionally been observed during in vitro⁹ and in vivo laboratory evolution experiments¹⁰, consistent with theoretical work suggesting that reversion is likely under conditions of high mutation rate and high population size^{11,12}. Yet, concrete, in-human, evidence of recurrent de novo mutation driving phenotypic switching has remained elusive.

Here, we report a bacterial pathogen acquiring successive de novo mutations in the same pathway, including reversion of a stop codon, to navigate environmental fluctuations. We compare early versus long-

A full list of affiliations appears at the end of the paper.  e-mail: tami@mit.edu; Gregory.Priebe@childrens.harvard.edu

term chronic infections with *Burkholderia dolosa*, which causes decades-long lung infections in people with cystic fibrosis (CF)¹³, using 931 whole genomes isolated from newly acquired and long-term chronic infections. Our phylogenetic analysis reveals that *B. dolosa* switches O-antigen expression on and off via de novo mutation, and that O-antigen loss is adaptive at early, but not late, stages of chronic infection. In vivo experiments in mice demonstrate that while O-antigen expression is beneficial for *B. dolosa* survival in the lung, it is detrimental in the spleen. Together, our results show how *B. dolosa* repeatedly mutates the same pathway to navigate a selective landscape that changes over the course of chronic infection.

Results

Selection for O-antigen expression during long-term infection

In the 1990s, an outbreak of *B. dolosa* spread among CF patients in the Boston area¹⁴. We previously studied the genomic evolution of *B. dolosa* during this outbreak using a longitudinal collection of 112 isolates¹⁵ and culture-based metagenomics¹⁶ from sputum samples collected after 7–9 years of chronic infection. These studies revealed genes under parallel evolution and pointed to key survival strategies of bacteria in the CF lung^{15,16}.

One of the most notable signatures for adaptation during this outbreak was seen in the pathway encoding for *B. dolosa*'s O-antigen, the sugar-chain decorating the lipopolysaccharide (LPS)¹⁵. Surprisingly, phylogenetic analyses indicated that this outbreak was initiated by a strain carrying a stop codon in *wbaD* (AK34_RS24375, previously annotated as BDAG_02317), a glycosyltransferase gene essential for the expression of O-antigen. This stop codon underwent at least 10 independent reversions, reinstating O-antigen expression. As these chronic infections persisted, the proportion of isolates expressing O-antigen increased steadily. Variation in O-antigen expression and length has also been widely documented in other *Burkholderia* species that chronically infect the CF lung^{17–20}.

The ability of a single mutation to restore function suggests this gene was pseudogenized only in the recent past. Moreover, phylogenetic data suggest that initial infections during this outbreak were established with O-antigen-absent strains¹⁵. These observations raise the intriguing possibility that O-antigen presence is disadvantageous during transmission or early infection, despite its advantage during later stages of disease.

New infection cluster was established by *B. dolosa* genotypes expressing O-antigen

With the guiding hypothesis that O-antigen expression may be a liability during transmission or early stages of infection, our lens focused on a new *B. dolosa* transmission cluster in the Boston vicinity after 9 years of no new cases. This cluster comprised 3 adults with CF: Patient J (the suspected index patient), Patient Q (co-worker of Patient J), and Patient R (sibling of Patient Q) (Fig. 1). Patient J was known to be infected with *B. dolosa* as part of the historical outbreak. Siblings Q and R, the new cases, lived together and were not followed at the hospital of the historical outbreak.

To test the hypothesis that O-antigen expression is disadvantageous during transmission or early infection, we characterized the *B. dolosa* populations in this new Boston-area transmission cluster. We collected three sputum samples from Patients Q and R between 4 and 38 months after first diagnosis and sequenced the genomes of 5–24 single-colony *B. dolosa* isolates from each sample (122 total; Fig. S1). The death of Patient J due to influenza and cepacia syndrome in the same year enabled us to study Patient J's *B. dolosa* population in great detail: we sequenced 650 isolates collected during autopsy and 159 isolates from multiple blood cultures drawn in the week prior to death (Supplementary Dataset S1). We also reanalyzed 10 genomes from Patient J from a prior study, reconstructing 16 years of within-person evolution¹⁵.

Evolutionary reconstruction indicated that Patient J was the infection source for Patients Q and R. Across a genome-wide phylogeny encompassing 927 newly sequenced and 112 previously-published isolates taken a decade prior from 14 patients, Patient Q's & R's isolates are nested within a clade of Patient J's isolates. This topology strongly supports the epidemiological association between Patient J and the new patients (Figs. 2a, S2, and S3). Yet, the phylogenetic structure does not resolve transmission between the siblings. Neither sibling's isolates are monophyletically nested within another's isolates, leaving room for multiple hypotheses: back-and-forth transmission between siblings, transmission of the same multiple genotypes from J to both siblings, or parallel evolution creating the same exact nucleotide mutation in both siblings. This complex tree topology is unlikely to emerge from recombination, as we find no strong evidence of recombination within this phylogeny ("Methods"). Patients Q and R may have been more likely to experience complex transmission because of their cohabitation. Schematics of hypothetical transmission scenarios are in Fig. 2b; each of these requires the founding of at least one patient's population by two or more independent genotypes (arrows in Fig. 2b), suggesting that *B. dolosa* infections in CF don't always hinge on a single cell transmission.

Despite this complexity, we can discern the genotype of the most recent common ancestor (MRCA) of these closely related strains—and, by proxy, the O-antigen phenotype. Our prior observations¹⁵ led us to expect an O-antigen-deficient MRCA, containing the truncated *wbaD*. This variant was present in just 1.4% of isolates in Patient J at the time of death (Fig. S4), making this scenario possible but unlikely. Contrary to our selection-based prediction, Patient Q's and R's MRCAs encode for the full length *wbaD*, and the earliest isolates from these patients expressed O-antigen. These predicted phenotypes were confirmed experimentally (Fig. S5). This observation demonstrates that O-antigen presence does not prevent transmission and thus demands another explanation for the absence of O-antigen in the early stages of the larger historical outbreak.

Selection for O-antigen disruption during early infection

Among de novo mutations accrued within Patients Q and R in the first years of infection, we observe a striking signal of convergence: many independent mutations emerge that disrupt the O-antigen pathway. Across both patients, a diverse set of 18 independent mutations were observed among 7 genes predicted to affect the O-antigen (Fig. 3a). Each isolate contained just one mutation, and each mutation was at a different nucleotide position (Supplementary Dataset S2). This concentration of mutations is statistically significant when compared to a random genomic distribution (Fig. 3b; $P < 0.01$, one-sided binomial test) and when compared to mutations accumulated in Patient J over a decade of infection and detected at the time of autopsy (Fig. 3b; $P < 0.01$, one-sided binomial proportion test). These results indicate that disruptions in O-antigen expression are strongly selected for during the first years of *B. dolosa* infection.

To probe the impact of these mutations on O-antigen expression, we stained for the O-antigen in 14 isolates with distinct mutations in this pathway, alongside 19 control isolates within the same clade without any mutations predicted to affect O-antigen synthesis. While all control isolates from Patients Q & R displayed the O-antigen, 13 out of the 14 isolates with a mutation in the O-antigen pathway (93%) showed reduced O-antigen expression—indicating that most of the identified mutations impacted O-antigen synthesis (Fig. 3c). These results provide confidence in our genotype-to-phenotype predictions.

Combining this phenotype map with the entirety of the genotypic data, we observed a strong temporal change in the prevalence of O-antigen presentation. At four months post-infection, all *B. dolosa* isolates from Patients Q & R bore O-antigen-expressing genotypes. Yet, by 20 and 38 months, 29% of Q's and 63% of R's isolates had acquired a mutation (either SNV or indel) that disrupted O-antigen expression

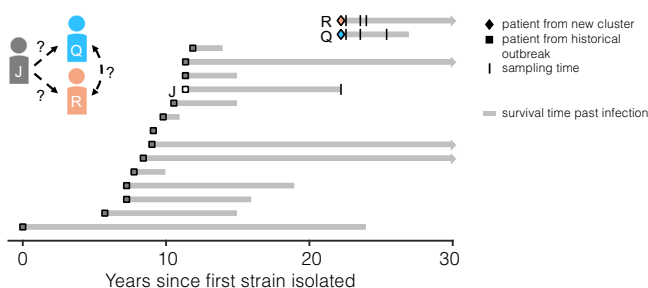


Fig. 1 | A new suspected *Burkholderia dolosa* transmission cluster among 3 adults with cystic fibrosis (CF) with connection to a historical outbreak. Two siblings with CF (Patients Q and R) were recently infected with *B. dolosa* and suspected to be part of a three-person transmission cluster (with Patient J). Patient J was previously infected as part of a historical outbreak that infected 39 people with CF, 14 of whom were profiled in a previous study and whose isolates are analyzed here. About 10 years elapsed between the last known transmission in that outbreak and the latest cluster. Here, we sequenced a total of 122 isolates cultured from the sputum of the new Patients Q and R over multiple timepoints (vertical bars). We also acquired 650 isolates from the lungs, lymph nodes, and spleen during autopsy of Patient J and 159 from blood obtained over the prior week. Gray bars indicate survival time past infection; arrowheads on the far right indicate that the patient was alive as of January 2024.

(Fig. 3d; $P < 0.01$, one-sided binomial proportion test). This signal for early O-antigen loss contrasts with the prior observation of selection for O-antigen presence late in chronic infection (Fig. 3e) and indicates an advantage for O-antigen disruption specifically during the first years of chronic infection.

Diverse O-antigen chain lengths coexist over a decade of chronic infection

To better understand O-antigen evolution in late infection, we examined the emergence of O-antigen-affecting mutations across Patient J's 809 isolates and phenotyped representative clones. This revealed four distinct O-antigen phenotype classes, each paired with a distinct genotypic signature: (1) absent, without any expression of O-antigen due to the stop codon in *wbaD*, (2) long-chain, found in isolates with the restored *wbaD* and no other mutations impacting O-antigen expression; (3) medium-chain, found in isolates with a restored *wbaD* and a 3 base pair insertion in BDAG_02328 (AK34_RS24445), encoding a different glycosyltransferase; and (4) short-chain, found in isolates with the previous two mutations plus a mutation in a third glycosyltransferase BDAG_02321 (AK34_RS24395) (Figs. S4 and S5). Genotypes corresponding to all 4 of these phenotypes were also found among isolates from Patient J collected 5–11 years earlier (Supplementary Dataset S3). To assess the generality of long-term coexistence, we reexamined isolates from other patients in the historical *B. dolosa* outbreak¹⁵. Indeed, we find multiple patient timeseries for which O-antigen-absent strains were recovered years after an O-antigen-intact strain was observed (in 4 of 13 cases). In addition, mutations that shorten the O-antigen chain were also observed in 3 of 13 patients (Figs. S6 and S7). Together, these findings indicate that diverse O-antigen phenotypes can coexist within an individual for years.

Within Patient J's lungs, the O-antigen absent phenotype was not enriched at any lung lobe, though we noted a modest difference in other O-antigen phenotypes across lobes (Figs. S8, S9, and S10). At autopsy, only 1.1% (11/809) of observed *B. dolosa* isolates had genotypes indicative of O-antigen absence, limiting our ability to detect statically significant differences in its abundance across samples. Analysis of *B. dolosa* spreading within the lung using mutational distances^{21,22} further showed only modest signal for spatial stratification overall. Given observations of large²² and modest²¹ spatial stratification for other species in CF, this weak spatial signal could be attributed to a breakdown of lung integrity following cepacia

syndrome, severe influenza, and/or mechanical ventilation. Alternatively, given the relatively large postmortem lung tissue sizes ($\sim 1\text{ cm}^2$), each sample might encompass multiple microenvironments (e.g., intracellular vs. extracellular, epithelial tissue vs. immune cells), each potentially favoring alternative O-antigen phenotypes.

Tradeoffs in murine spleen and lung colonization

What, mechanistically, can explain the time-based inversion of selection for O-antigen presentation? Expression of O-antigen is thought to provide resistance to serum complement^{23,24}, antibiotics¹⁷, and antimicrobial peptides²⁵, and is therefore expected to outcompete O-antigen-disrupted strains in most conditions. However, given its recurrent loss during early infection, O-antigen expression must be selected against in some environmental conditions.

To discern a potential advantage for *B. dolosa* O-antigen disruption across body compartments, we utilized a murine pneumonia model ("Methods"). We selected 3 independent, near-isogenic pairs of strains that differ in O-antigen presentation, each containing an O-antigen-disrupted strain and near-isogenic strain with an intact O-antigen (Fig. S11; up to 1 additional nonsynonymous mutation; see "Methods"). For each pair, we performed in vivo competitions by labeling the O-antigen-intact strain with a transposon that constitutively expresses *lacZ*. This labeled strain was mixed with its unlabeled O-antigen-disrupted partner and used to intranasally infect mice (Fig. 4a).

After 7 days, the ratio of O-disrupted to O-antigen-intact strains was significantly greater in the spleen than in the lungs for all pairs (Fig. 4a; 1.7 to 2.4-fold, paired t test, $P < 10^{-3}$), though both phenotypes disseminated to all profiled organs. We also performed label-swap experiments and saw similar enrichment (Fig. S12). Slight variations were seen among pairs of isolates, potentially due to differences in the O-antigen-disrupting mutation or other genomic variations. Nevertheless, the organ-based tradeoff was consistent across isolate pairs and label-swaps, suggesting a causal role for O-antigen disrupting mutations. This increase in O-antigen-disrupted strains in the spleen could reflect higher rates of hematogenous dissemination, higher rates of survival in the blood, or higher uptake by immune cells.

We therefore hypothesized that O-antigen might affect *B. dolosa*'s ability to be internalized or persist within phagocytic cells²⁶. The intracellular niche is known to be critical to establishment of infection in the broader *Burkholderia cepacia* complex (*Bcc*)^{27,28}. We infected THP-1-derived macrophage-like cells with O-antigen-intact and disrupted strains ("Methods"). In line with our expectation, the number of intracellular bacteria was 1.7–8.2 fold higher for O-antigen-disrupted strains after two hours of macrophage infection (Fig. 4b and S13, $P < 0.03$ for 8/9 comparisons, two-sided Wilcoxon rank sum). To distinguish between increased invasion or increased intracellular survival, we performed time-course assays and confirmed that O-antigen disruption affected internalization but not intracellular survival ($P < 0.05$, Fig. S14), consistent with microscopy-based observations in other *Bcc* species²⁹. Combined with the mouse data, these results suggest that O-antigen-disrupted strains have a higher ability to enter the immune cell niche, which may be important to dissemination within the lung.

Together, these results are consistent with a model in which O-antigen disruption is selected for during the first years of infection in the CF lung. We propose a mechanism for this selection: O-antigen-disruption advantages growth in or uptake by phagocytic cells, which is superseded by changing environmental pressures during the course of longer chronic infection. Changing pressures may include diminished oxygen availability with increasing lung damage³⁰, increased antibiotic pressures, changing phage pressure³¹, or co-evolving adaptive immune responses³² (Fig. 5). However, we cannot rule out the possibility that other mechanistic forces, such as the higher propensity of O-antigen-disrupted strains to adhere to host cells²⁹ or interactions with complement, could also be influencing the selective landscape.

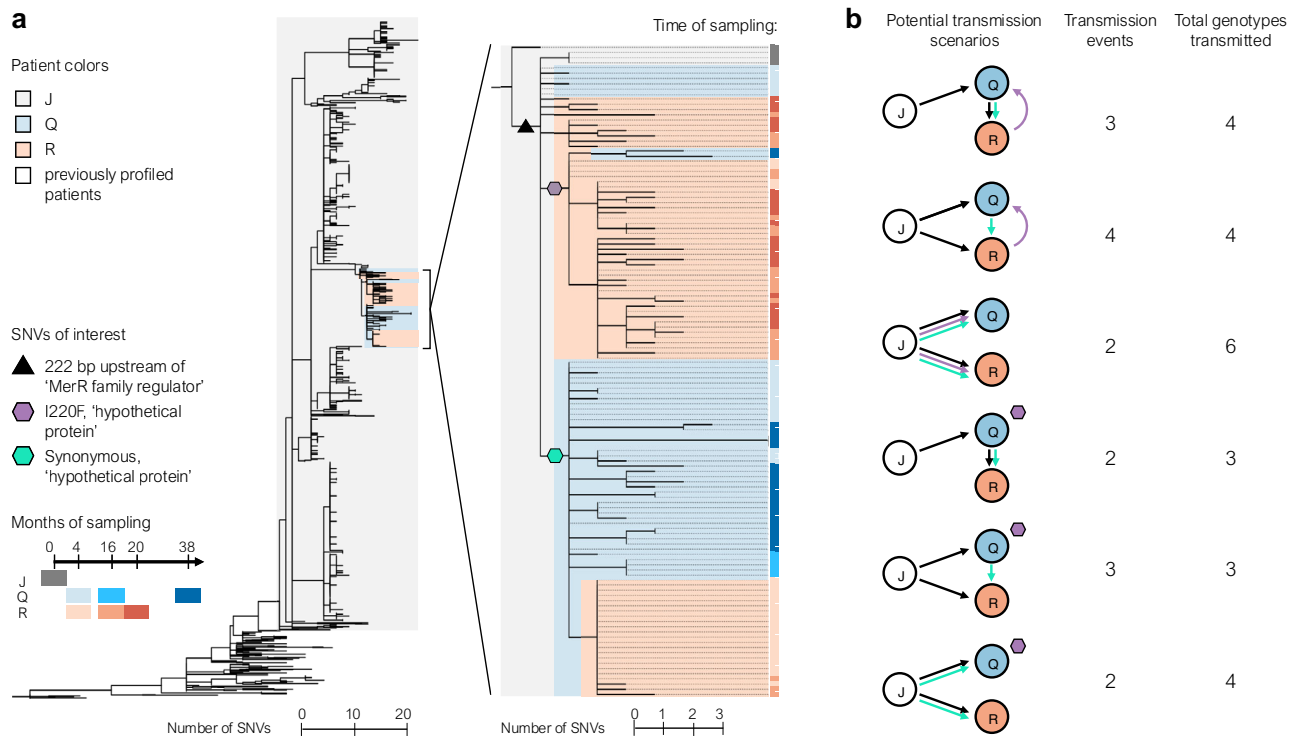


Fig. 2 | In early infection establishment, multiple *B. dolosa* clones are transmitted between CF patients. **a** A maximum parsimony SNV phylogeny was built from whole-genome sequencing of 805 *B. dolosa* isolates from autopsy samples of the suspected index case (Patient J), 122 from serial sputum samples of Patients Q and R, and 112 previously published sequences of isolates from Patient J and 13 other patients a decade prior¹⁵. All isolates from Patient Q and R are descended from a subset of Patient J's diversity; a high-resolution phylogeny of this clade ("Methods") is shown to the right. The time of sampling of each isolate is indicated

by color in the rightmost vertical bar, clades within the tree are shaded by patient, and SNVs of interest for transmission inference are indicated by colored shapes. **b** Potential transmission scenarios between Patients J, Q, and R are diagrammed, showing the need for either more than two transmission events or parallel nucleotide evolution to explain the observed diversity. Arrows point in the direction of *B. dolosa* transfer and are colored by the SNV in (a) that precedes a transmission.

Discussion

In this study, we show that *B. dolosa* alternates O-antigen phenotypes via point mutations in response to shifting environmental pressures within individual people. O-antigen absence is selected for within immune cells and in early infection, while its expression is selected for within the lung and during late infection. In response to these trade-offs, *B. dolosa* cells switch O-antigen expression on and off through a seemingly permanent solution—mutation—only to revert to the previous phenotype through additional mutation.

Our results help to clarify previously observed O-antigen variation among members of the *Bcc*. Conflicting observations of loss^{20,33,34} and acquisition of O-antigen in species^{15,20} can be rationalized when viewed through the lens of context-dependent mutations. This illumination was enabled by a combination of deep and longitudinal population sequencing, phenotypic profiling, and in vitro and in vivo models of infection. Further work will be needed to understand why mutations that enhance immune cell internalization are selected for only in early chronic infection. If immune internalization is indeed a dominant selective pressure driving O-antigen evolution, we might cautiously speculate that lymphatic spread may be critical for initial within-lung dissemination, as has been hypothesized for *M. tuberculosis*^{35,36}. We do not anticipate O-antigen mutations to be required for dissemination during cepacia syndrome, given the observations here (Fig. S9) and previously¹⁵ that disseminated genotypes reflect the spectrum of diversity in the lung.

To our knowledge, this is the most direct observation of repeated phenotypic switching via conventional mutation outside of a laboratory setting. Even though prior in vitro studies³⁷ and controlled mouse experiments¹⁰ observed mutation-mediated switching, many models

of evolution consider this phenomena irrelevant in natural settings^{8,9}. Instead, changing selective pressures are considered more likely to favor the emergence of a compensatory mutation¹², the migration of cells with ancestral phenotypes³⁸, or the resurgence of a persistent, low-frequency ancestral genotype^{39,40}. Mutation-driven switching is commonly relegated only to genomic elements with elevated mutation rates (e.g., simple sequence repeats)^{41–43}. Our work challenges this expectation by presenting an in situ case of bacteria undergoing mutation-mediated phenotypic switching and calls for theoretical models that consider mutation-mediated reversion.

Mutation-mediated switching and reversion may well be common among bacteria with within-host population sizes exceeding the inverse of the per-nucleotide mutation rate ($\sim 10^{-9}$ mutations/generation for most species⁴⁴), including other CF lung pathogens and commensal members of the gut microbiome^{2,11,45}. Mutation-driven adaptations could be considered less likely for commensals and other organisms with prolonged tenure in mammalian environments, which have had ample time to evolve more complex regulatory mechanisms^{46,47}. However, recent observations suggest that mutation-mediated phenotypic switching might be common even for species with ample time to evolve gene regulation: *Escherichia coli* undergoes mutation-mediated switching in laboratory mice¹⁰ and *Bacteroides fragilis*'s in-human adaptive mutations target conserved genes⁴⁸, despite these same genes possessing complex regulatory mechanisms⁴².

Altogether, our finding of phenotypic alternation via repeated point mutation suggests that the speed and dynamism of bacterial adaptations may be underappreciated. While mutations that change bacterial cell envelopes have been observed previously in vivo^{37,49–51}, it has been speculated that within-host adaptations to these changing

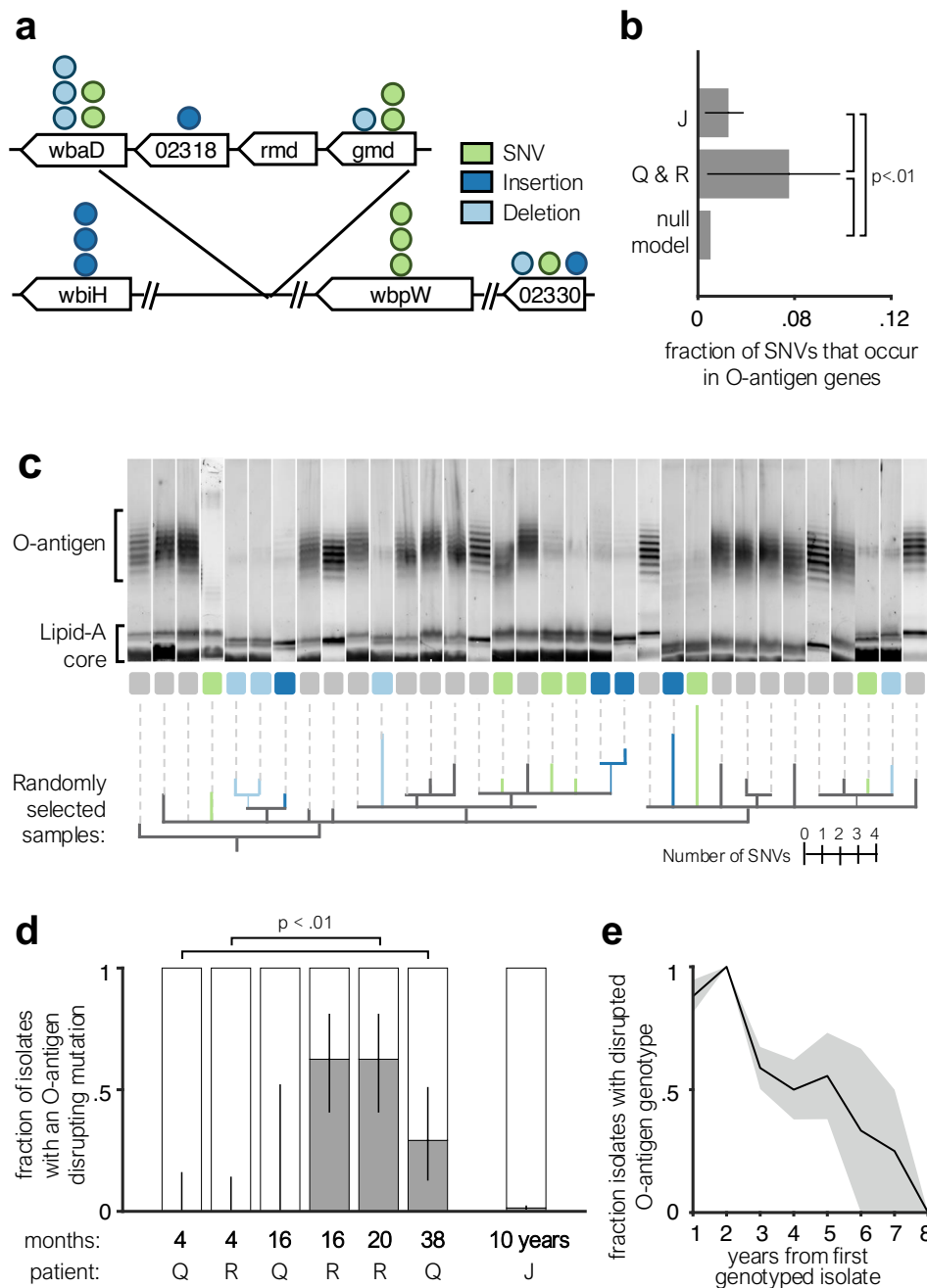
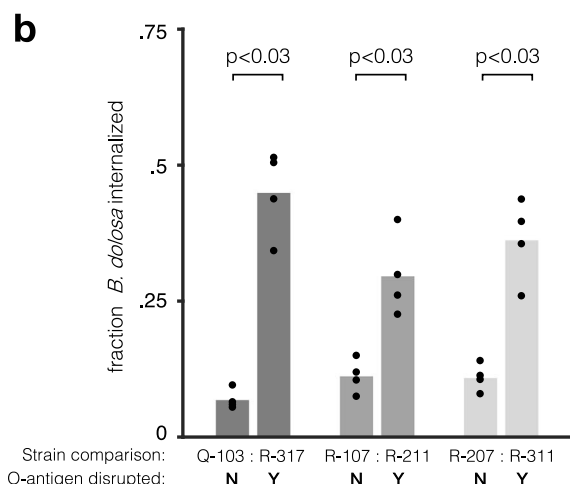
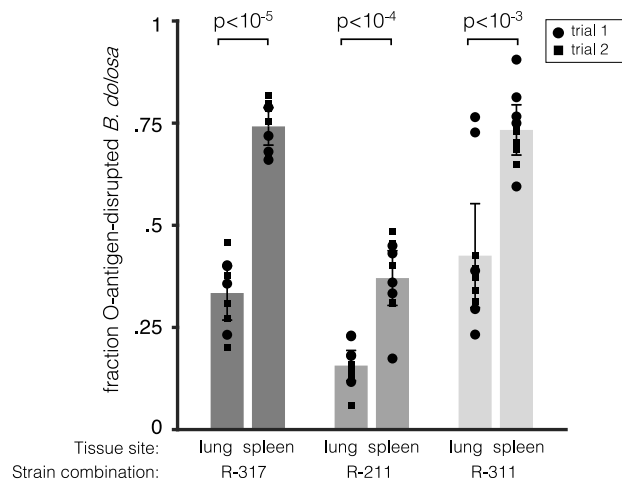
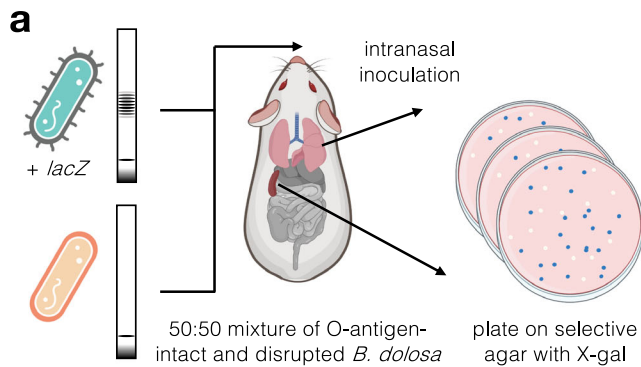


Fig. 3 | *B. dolosa* lipopolysaccharide (LPS) O-antigen expression is convergently disrupted by de novo mutation in early CF lung infection, contrasting with trends in longer-term colonization. **a** Mutations in genes suspected to be involved in O-antigen synthesis are represented by circles; each mutation occurred within Patient Q and R on an independent branch of the tree and each isolate contained at most one such mutation. Mutations are colored by the type of mutation and gene icon length is proportional to gene size. **b** The proportion of observed SNVs that fall in the O-antigen synthesis is significantly greater for Patients Q and R (recently infected) than for Patient J (infected for nearly a decade; $P < 0.01$ one-sided binomial proportion test). Both sets of mutations occur at a greater frequency than expected by random mutation across the O-antigen pathway ($P < 0.01$, one-sided binomial test, see “Methods”). Error bars indicate 95% confidence intervals. **c** Selected isolates from Patient Q and R, as well as near-

isogenic samples from Patient J, were profiled for O-antigen banding pattern. Isolates are ordered by a phylogeny of SNVs and O-antigen-affecting indels (see “Methods”) and branches are colored by the occurrence of an O-antigen-affecting SNV, insertion, or deletion. Of 14 isolates with an O-antigen-affecting mutation, 13 show a reduction of O-antigen expression. **d** The proportion of O-antigen-disrupted isolates is significantly increased from the first to last sampled time point for both Patients Q and R ($P < 0.01$, one-sided binomial proportion test), but negligible in Patient J after 10 years. Lines represent 95% confidence intervals. **e** In a reanalysis of 112 isolate genomes from the original *B. dolosa*¹⁵ outbreak collected over a greater duration of infection, we identify a trend during extended chronic infection that contrasts with early infection. See Fig. S6 for more details. The shaded region indicates standard error of the mean. Source data are provided as a Source Data file.

fitness landscapes represent ‘dead-ends’ that do not transmit to others^{9,38}. However, the observation of nucleotide-level and phenotype-level reversions seen here add to recent theoretical work¹¹ suggesting that additional studies with deep sampling along a

transmission chain are needed to fully observe reversions. If mutation-mediated phenotypic alternation is indeed widespread, the extent to which within-human bacterial evolution drives phenotypic change may be vastly underestimated.



Methods

Study cohort, isolate sampling, and genome sequencing

The clinical research conducted in this study complies with all relevant ethical regulations, and the study protocol was approved by the Institutional Review Board of Boston Children's Hospital. Informed consent was obtained for sample use/collection and medical record review. Starting in 1998, 39 CF patients were infected with *B. dolosa* during an outbreak in Boston Children's Hospital, Boston, MA. No new *B. dolosa* infections were reported in the Boston area between 2005 and 2013, and the outbreak was presumed to have ended. In 2014, a pair of adult siblings with CF (Patient Q and Patient R) were confirmed

Fig. 4 | Disruption of *B. dolosa* LPS O-antigen expression leads to tradeoffs in lung versus spleen infection in vivo. **a** Mice were infected with a 50:50 mixture of O-antigen-intact and disrupted phenotypes, with one *B. dolosa* strain marked with a *lacZ* cassette conferring blue colony color on X-gal-containing Bcc-selective agar. Lungs and spleens were excised to determine the bacterial load of O-antigen-intact and O-antigen-disrupted strains. Two trials of this experiment were performed for each of three near-isogenic pairs. The mean fraction of O-antigen-disrupted bacteria in mouse spleens and lungs is shown with error bars representing 95% confidence intervals. The O-antigen-disrupted fraction is higher in spleens relative to lungs for all pairs ($P < 10^{-3}$, two-sided paired t test). Macrophages were infected for 2 hours with near-isogenic *B. dolosa* strains with and without an O-antigen-disrupting mutation and then incubated for 2 hours with kanamycin, which kills extracellular bacteria. The number of intracellular bacteria is compared to the total bacteria obtained from an identical culture without kanamycin treatment, revealing an advantage for O-antigen-disrupted bacteria within macrophages ($P < 0.03$, two-sided Wilcoxon rank sum). Bars represent averages across four concurrent technical replicates. **a** was created in BioRender. Poret, A. (2025) <https://BioRender.com/ig94dd7>.

to have newly acquired *B. dolosa* infections, and a likely index adult patient was identified. The suspected index patient of this new cluster was infected during the known previous outbreak, and this patient's *B. dolosa* population was described in two previous studies: Patient J in ref. 15 and Patient 2 in ref. 16; referred to as Patient J in this study).

Sputum samples were obtained during normal care under IRB approved protocol 05-02-014 R (with written informed consent) at Boston Children's Hospital (BCH) from Patients Q and R at 3 timepoints following initial diagnosis: 4 months, 16 months, and either 20 months (Patient R) or 38 months (Patient Q) (Fig. 1 and Supplementary Dataset S1). Sputum samples were diluted in 1% proteose peptone and then plated on oxidation/fermentation-polymyxin-bacitracin-lactose (OFBPL) media (BD Biosciences) to select single *B. dolosa* colonies. From every sample, up to 24 colonies were randomly picked and frozen in 15% glycerol; in the event of low CFU density, we sampled all present colonies. This resulted in 123 total isolates, with 5–24 collected per patient timepoint.

Patient J died from cepacia syndrome that developed during hospitalization for influenza shortly after identification of the new *B. dolosa* cases (Patients Q and R). Blood cultures grew *B. dolosa* over the 7 days prior to death. Blood culture bottles found to be growing *B. dolosa* were obtained under the same IRB protocol as above. Contents of blood culture bottles were diluted in 1% proteose peptone and then plated on OFBPL plates for colony picking. We collected 170 total blood colonies, ranging from 16 to 29 isolates per time point. During Patient J's autopsy, we obtained lung (599), spleen (47), and lymph node (48) isolates from 38 different sites (Fig. S8) under the same IRB protocol as above. Tissue samples were taken using sterile technique, homogenized, and frozen in glycerol as described in ref. 21. Serial dilutions of lung, spleen, and lymph node tissue were plated on OFBPL media. Up to 24 *B. dolosa* colonies were grown from each sample for 24 h in Luria broth (LB) and then frozen in glycerol.

Bacteria were cultured to mid-log in LB for DNA extraction, library preparation, and DNA sequencing. The standard Illumina Nextera protocol was used in a modified method for library construction using a microfluidic device that enabled parallel and resource-efficient DNA library preparation. Microfluidic methods were based on ref. 52 and further updated, as described in a forthcoming publication. The resultant libraries were sequenced to a median depth of 87.7 on an Illumina HiSeq X Ten with 2×150 bp dual indexed reads.

Mutation detection and phylogenetic inference

Adaptors were trimmed by cutadapt v1.18⁵³ and then filtered via sickle v1.33⁵⁴ (filters: “-q 20 -l 50 -x -n”). Filtered reads were aligned by bowtie2⁵⁵ (v2.2.6; bowtie2 -X 2000 -no-mixed -dovetail -very-sensitive -n-ceil 0,0.01) to the reference genome AU0158, a *B. dolosa*

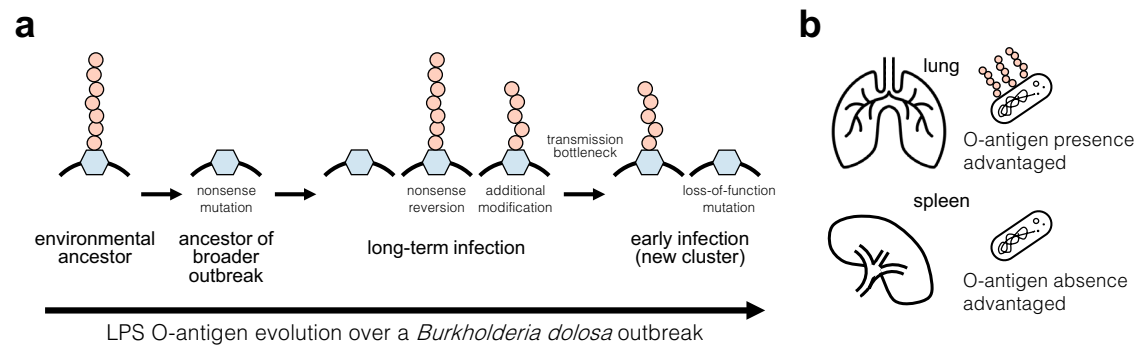


Fig. 5 | Navigation of selective tradeoffs on O-antigen presentation via de novo mutation. **a** Inferred natural history of *B. dolosa* LPS O-antigen presentation during an outbreak in people with CF. The initial *B. dolosa* outbreak was initiated by a strain lacking O-antigen due to a premature stop codon in a glycosyltransferase gene. This premature stop mutation likely occurred shortly before or during outbreak initiation¹⁵. As the outbreak spread and persisted in chronic infections, independent reversions restored *B. dolosa*'s O-antigen in multiple patients. Within each patient, strains with variable O-antigen presentation were recoverable and

coexisted for years. Two patients recently infected by a survivor of this outbreak were initially colonized by strains expressing O-antigen. Remarkably, O-antigen-disrupted phenotypes re-emerged independently within each of these patients via a variety of de novo mutations in the O-antigen pathway. **b** Our in vivo experimental results suggest that O-antigen presence is advantageous in the lung while O-antigen absence is favored in the spleen. Combined with the historical observations, these results suggest that survival in the spleen or immune cells may be important during the early years of chronic *B. dolosa* lung infection.

sample isolated from the index patient of the original Boston outbreak five years into infection (Genbank: GCA_000959505.1). SAMtools v1.5³⁶, mpileup (-q30 -t SP -d3000), bcftools call (-c), and bcftools view (-v snps -q.75) was used to create consensus sequences and call candidate single nucleotide variants (SNVs) across all isolates. All code used to call these commands was aggregated in a MATLAB v2015b and run on the high performance computing cluster, "Commonwealth Computational Cloud for Data Driven Biology" (C3DDB).

Candidate SNVs and samples were filtered based on coverage, nucleotide identity, and SAMtools produced FQ score in MATLAB v2021a according to the below filters. Samples with a mean candidate SNV coverage of <12× across candidate positions were discarded (17/987 samples). SNV positions were removed if they had a median depth of coverage across samples of <40× or if more than 20% of sequence calls across samples were marked as ambiguous, where ambiguity is defined as having a major allele frequency of less than 0.9, a per-strand coverage of less than 12×, or a FQ score below 0 in a given sample. This resulted in a list of 541 trusted positions.

For each sample's nucleotide call at a given position was defined as the major allele across reads, or labeled as N if the major allele frequency was <0.8 or had no reads (Supplementary Dataset S4, S5, and S6). Any sample with more than 10 trusted positions called as N was removed from further analysis (39/987 samples). In total, 931 samples passed this filtering protocol. The resulting allele calls were used to construct a maximum-parsimony phylogenetic tree of Patients J, Q, and R using DNAPars (Phylip v3.69)⁵⁷. The tree was rooted with the reference genome AU0158 and can be seen in Fig. S2.

We also identified short insertions and deletions (indels) in and nearby genes affecting the O-antigen presentation (Supplementary Dataset S7). Indels are difficult to call with high fidelity, and both false-positive and false-negative mutation calls can impact phylogenetic inference. We used Breseq v0.30.0⁵⁸ and its associated program gdttools to identify indels relative to the reference genome AU0158. Variants around homopolymers and variants for which >0.05 of samples were called as unknown or deleted by Breseq were discarded (Supplementary Dataset S8). All remaining indels within 100,000 base pairs of an O-antigen-affecting gene as defined by KEGG's gene ontology⁵⁹ were queried via PaperBlast⁶⁰ to deduce whether a mutation in that gene could impact *B. dolosa*'s O-antigen (KEGG: 00541, Supplementary Dataset S9 and S10). After curating all indels that could plausibly impact O-antigen presentation, we created a second maximum-parsimony phylogenetic tree of the new infection cluster,

including one additional SNV for each indel in the input file to DNAPars. The inclusion of these indels proved useful in understanding the spread of O-antigen-affecting mutations across the phylogeny, but did not change the overall shape of the phylogeny. An indel-inclusive tree is shown in Fig. S4; a subset of this tree is drawn in Fig. 3d.

To determine whether the mutations identified above could have emerged from recombination or some other non-SNV event (e.g., multi-nucleotide lesion), we screened for pairs of nearby SNVs (<1000 bp) that were perfectly correlated in the presence on the phylogeny. This identified 5 small multi-variant blocks (1–15 bp in size) affecting 11 SNPs, which are labeled as such in Supplementary Dataset S11. Given the small size of these blocks, these variants may emerge from recombination or vertical multi-nucleotide lesions. Each set of these mutations were monophyletic; thus, if these blocks emerge from recombination they came from some undetected genotype. None of these loci are implicated in O-antigen variation. Only 1 of these blocks, affecting 2 SNVs, was found in subjects Q or R (in isolates Q-305, R-311), and it does not affect the transmission analysis performed in Fig. 2.

Reanalysis of 112 *B. dolosa* isolates from a past outbreak

We created a phylogeny that combines the 112 previously-sequenced isolates from ref. 15 with the 987 new isolates from Patients J, Q and R. All of the 112 historical isolates were processed using the same procedure described above, with the following exceptions: (1) we used looser sickle v1.33 filters to account for differences in sequencing quality ("q 0 -l 0 -x -n") (2) reads were aligned to reference genome AU0158 using bowtie2 using settings modified for single reads (-phred64 -sensitive -n-ceil 0,0.01). Using the aforementioned SAMtools v1.5, bcftools call, and bcftools view settings, we created consensus sequences and called candidate SNVs across all 1099 total isolates.

To identify loci of interest, we analyzed each set of samples separately; this minimizes the impact of coverage and error bias between sequencing types. We use the same procedure described above to reanalyze and filter Patient J, Q, and R's samples. We similarly identify SNVs among Lieberman and Michel et al.'s historical samples, with the following modifications: (1) Each allele call at a given position was defined as ambiguous if it had a major allele frequency of less than 0.9 or coverage of less than 20 reads; (2) No samples were excluded based on coverage; (3) Positions with a median coverage across samples below 20× were removed.

We combine these two candidate loci sets into one trusted list, and define each sample's nucleotide call at a given position as above. New isolates from Patients J, Q, and R with greater than 10 trusted positions called as N were removed from further analysis. These quality control steps resulted in 927 filter-passing isolates; no such cutoff was implemented for historical samples (which overall had lower sequencing depth and quality). A phylogenetic tree was constructed as above to better understand transmission within the scope of the broader *B. dolosa* epidemic (Fig. 2A).

Additionally, we ran Breseq (–predict-polymorphisms) on these 112 historical samples and searched for O-antigen-affecting mutations using the process defined above (Supplementary Dataset S3 and S12). Since these isolates are sequenced to a lower depth than those from patients Q, R, and J, we utilize the –predict-polymorphisms setting to increase sensitivity. As such we do not construct a new, indel-aware phylogeny that combines historical samples with our new, differently-processed sequenced data.

O-antigen phenotyping and quantification

Frozen stocks of *B. dolosa* were used to inoculate 1–5 ml of overnight LB cultures. These were normalized to an OD₆₀₀ of 0.25, pelleted using a microcentrifuge at 10,600 × *g* for 10 min, and stored at –20 °C. The next day pellets were thawed and LPS was extracted according to Davis and Goldberg⁶¹, stopping after the addition of proteinase K (step six) to store overnight at –20 °C. Subsequently, 15 µl of each sample were loaded into each well of a 12% Mini-PROTEAN TGX gel (Bio-Rad) along with a CandyCane glycoprotein ladder (ThermoFisher). We additionally loaded the same two *B. dolosa* samples with different LPS phenotypes (Q-103; medium and R-221; disrupted) into each gel to assess phenotype repeatability. We run SDS-PAGE on each gel and stain the separated LPS bands using Pro-Q Emerald 300 Lipopolysaccharide Gel Stain Kit (ThermoFisher) according to the manufacturer's instructions, modifying the initial fixation step to be repeated twice and each washing step three times. Images were taken using 300 nm light in a Syngene G:Box Mini-9 imager. The resulting images were processed to quantify the expression and size of the O-antigen.

Creation of *lacZ* mutant *B. dolosa* strains

To differentiate between O-antigen mutant and intact strains in competition experiments, a *lacZ* cassette was inserted one strain of several pairs of near isogenic *B. dolosa* isolates: R-317 and Q-103, R-211 and R-107, R-311 and R-207. All *lacZ* cassettes were inserted into the O-antigen-intact strain (Q-103, R-107, and R-207). An inverse, R-317 *lacZ* and Q-103 strain pair was engineered to ensure the *lacZ* insertion did not impart any costly fitness effects.

Strain R-317 is near isogenic to Q-103, differing by a Y438* (non-sense) SNV in *wbpW* (BDAG_02323/AK34_RS24410), a “mannose-1-phosphate guanylyltransferase” projected to impact O-antigen expression, an intergenic G → T SNV at position 1,284,152 on chromosome 2, and a synonymous SNV in BDAG_01365 (AK34_RS18940), a “porin.” Strain R-211 differs from R-107 by a P97R SNV in *gmd* (BDAG_02320/AK34_RS24390), a “GDP-mannose 4,6-dehydratase” projected to impact O-antigen expression and no other known SNVs. Strain R-311 differs from R-207 differ by a +G insertion in *ubiH* (BDAG_02309/AK34_RS24330) projected to impact the O-antigen, and SI30A in a hypothetical protein AK34_RS02570 (no BDAG identification number).

To constitutively express *lacZ*, mutant *B. dolosa* strains were generated using pCElacZ⁶² reporter, which utilizes a mini-Tn7 based transposon for stable integration into the chromosome without the need for antibiotic selection. This reporter plasmid was conjugated into each *B. dolosa* strain with the helper plasmid pRK2013 and integration-helped plasmid pTNS3 as described by ref. 63. Conjugates were selected by plating the colonies on LB agar containing trimethoprim (1 mg/ml) and gentamicin (50 µg/ml). Insertions into the attTn7 site downstream of BDAG_04221 (AK34_RS08675) were

confirmed by PCR (Forward-PTn7L- ATTAGCTTACGACGCTACACCC and Reverse-bdag-4221 GCGTTCCTGCACCGAACATG).

Modeling spatial spread of *B. dolosa* via intranasal murine infection

We implemented the murine infection model described in ref. 64 to compare how intact (“medium” phenotype) and disrupted O-antigen phenotypes impact infection establishment. All animal experiments were approved by the Boston Children's Hospital Institutional Animal Care and Use Committee under assurance number A3303-01 and protocol number 1241. All protocols are compliant with the NIH Office of Laboratory Animal Welfare, the Guide for the Care and Use of Laboratory Animals, the US Animal Welfare Act, and the PHS Policy on Humane Care and Use of Laboratory Animals.

To prepare *B. dolosa* inoculum, overnight cultures of *lacZ*-marked intact and disrupted O-antigen strains were grown in tryptic soy broth (TSB). After subculturing each strain to log-phase, each culture was diluted with PBS to a final concentration of 2 × 10⁶ CFU/ml. Before inoculation, samples were plated on PC Agar (*Pseudomonas cepacia* agar, BD biosciences) containing 100 mg/ml of X-gal (5-Bromo-4-chloro-3-indolyl β-D-galactopyranoside) (ThermoFisher) to confirm CFU/ml counts.

Mice (6-to-8-week-old C57BL/6 female mice from Taconic Labs) were anesthetized with ketamine (66.6 mg/kg) and xylazine (13.3 mg/kg) given intraperitoneally. Mice were housed in a biosafety level 2 facility under specific pathogen-free conditions at 20–23 °C, with 35–70% humidity and a 12-h light/dark cycle. Female mice were used since, in our experience, lower size variation in female mice compared to male mice leads to less variation in lung deposition after intranasal inoculation of female mice compared to male mice. To infect mice with a 50:50 mixture of *B. dolosa* phenotypes, 10 µl of inoculum was inserted into each nostril of mice held in dorsal recumbency. After 7 days, mice were euthanized by CO₂ overdose, and the lungs and spleens were then aseptically removed. Each organ was weighed, placed into 1 ml of 1% proteose peptone in water, homogenized, and serially diluted and plated on custom PC-X-gal agar. After 36 h, blue and white colonies were counted to determine the phenotype ratio in each lung. This experiment was performed twice for each of the three strain pairs. We additionally competed the pair Q-103 and R-317 *lacZ* and pair Q-103 *lacZ* and R-317 against each other once to account for potential effects of the *lacZ* cassette (Fig. S12).

Kanamycin exclusion assays to measure *B. dolosa* phagocytosis and survival in macrophages

To understand how different O-antigen phenotypes survived within macrophages, we used a kanamycin exclusion assay from ref. 65. Human THP-1 monocytes obtained from ATCC were grown in RPMI 1640 medium (Gibco) containing 2 mM L-glutamine, 10 mM HEPES, 1 mM sodium pyruvate, 4500 mg/liter glucose, 1500 mg/liter sodium bicarbonate supplemented with 10% heat-inactivated fetal bovine serum (Gibco), and 0.05 mM 2-mercaptoethanol at 37 °C with 5% CO₂. For differentiation into macrophages, THP-1 cells (ATCC cat. # TIB-202) were diluted to 7 × 10⁵ cells/mL in fresh media and treated with 200 nM phorbol 12-myristate 13-acetate (PMA); 1 mL/well of cells were plated into 24-well plates. Cells were incubated for 3 days as described above. Cells were washed 2× with PBS (containing calcium and magnesium), and RPMI 1640 medium, lacking PMA, was replaced before being incubated in similar conditions for an additional 24 h.

Cultures of each *B. dolosa* strain used in this analysis, R-317, Q-103, R-211, R-107, R-311, R-207, as well as Q-101, R-318, and R-221, were concurrently streaked out onto OFPBL petri dishes and then incubated overnight in LB while shaking at 200 rpm at 37 °C. After subculturing and dilution, log-phase cultures of each bacterial strain were washed in RPMI three times and ~2 × 10⁶ CFU were added to each well containing THP-1 macrophages (multiplicity of infection of ~10:1). To synchronize

infection, plates were spun at $500 \times g$ for 5 min and then incubated for 2 h at 37°C with 5% CO_2 . In a near-identical time course experiment using strains R-211 and R-311, this incubation time was modified to 15, 30, 45, and 60 min (Fig. S14b). After incubation, the total number of bacteria in the wells were determined by adding 100 μL of 10% Triton X-100 lysis buffer to each well with a final concentration of 1%. Having burst the macrophages with Triton, we serially diluted and plated the resultant mixtures on tryptic soy agar (TSA) plates to determine the number of CFU.

To determine the number of invaded bacteria in macrophages, after the aforementioned 2 h incubation period each well of parallel infected cells were washed with $2\times$ with cell culture grade PBS containing calcium and magnesium (+/+ PBS). RPMI media was then replaced in each well along with 1.0 mg/mL kanamycin. The resultant mixture was incubated for another 2 h under similar conditions and washed $3\times$ with +/- PBS. In parallel experiments, the kanamycin incubation period was extended to 4 or 6 h (Fig. S14c). To lyse the macrophages, 1% Triton X-100 lysis buffer was similarly added to each well. The plate was then shaken for 15 min to detach cells, and the resultant bacteria were diluted, plated on TSA, incubated for 1–2 days, and counted.

To ensure our results are not confounded by differences in kanamycin resistance, we tested the kanamycin resistance of *B. dolosa* strains R-317, Q-103, R-211, R-107, R-311, R-207, and their *lacZ*-modified counterparts (Fig. S15). Strains were grown overnight in 5 mL of LB, then diluted 10^{-2} -fold, followed by serial tenfold dilutions ranging from 10^{-4} to 10^{-10} . A 3 μL volume of each dilution was drop-spotted onto LB agar plates containing varying concentrations of kanamycin (0, 0.5, 0.7, 0.9, 1.1, and 1.3 mg/mL). Each drop spot therefore represents a total dilution of 3×10^{-5} -fold as well as 3×10^{-7} to 3×10^{-13} -fold. Plates were incubated at 37°C for 2 days, after which colony growth was imaged. Colonies were counted at the 3×10^{-9} and 3×10^{-10} dilutions on each plate, as these dilutions produce well-isolated colonies on the control plate (0 mg/mL kanamycin), minimizing the influence of contact-dependent bacterial behavior. The minimum inhibitory concentration was defined as the highest kanamycin concentration at which no visible colonies appeared in both dilutions.

Statistics and reproducibility. Statistical analyses using Mann–Whitney *U*-test (ranksum) and *t* tests (ttest2) were conducted using built-in packages in MATLAB (v2021a). Code used to implement all statistical tests is available at GitHub. No statistical method was used to predetermine sample size for the clinical study. No data were excluded from the analyses. As this was an observational study, random allocation of patients is not applicable to this study. We conducted a random sampling of *B. dolosa* colonies from each cultured plate of blood or tissue. To ensure random selection of colonies when 24 or more colonies were present on the plate, we marked a piece of paper cut to the size of the Petri dish with 24–29 “x” marks. We taped this paper to the back of each plate and picked the colony closest to each “x” mark. In the event of low CFU numbers (<24) on a plate, we sampled all present colonies.

For in vitro experiments and mouse experiments, the investigators were blinded to group allocations when performing CFU counts. For characterizing clinically relevant phenotypes of mutants, we allocated isolates as control vs. experimental based on their genotype. All in vitro experiments were performed 2–3 times, and all findings were replicated successfully.

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

Sequences from *B. dolosa* samples obtained in this study are uploaded under BioProject PRJNA1063312. All strains and plasmids used in this

study are available upon request. Source data are provided with this paper.

Code availability

All code used in this project is published under https://github.com/ajporet/b_dolosa_evolution.

References

1. Ferenci, T. Trade-off mechanisms shaping the diversity of bacteria. *Trends Microbiol.* **24**, 209–223 (2016).
2. Didelot, X., Walker, A. S., Peto, T. E., Crook, D. W. & Wilson, D. J. Within-host evolution of bacterial pathogens. *Nat. Rev. Microbiol.* **14**, 150–162 (2016).
3. Cooper, V. S. & Lenski, R. E. The population genetics of ecological specialization in evolving *Escherichia coli* populations. *Nature* **407**, 736–739 (2000).
4. McNally, A. et al. Combined analysis of variation in core, accessory and regulatory genome regions provides a super-resolution view into the evolution of bacterial populations. *PLOS Genet.* **12**, e1006280 (2016).
5. Juhas, M. Horizontal gene transfer in human pathogens. *Crit. Rev. Microbiol.* **41**, 101–108 (2015).
6. Zhou, K., Aertsen, A. & Michiels, C. W. The role of variable DNA tandem repeats in bacterial adaptation. *FEMS Microbiol. Rev.* **38**, 119–141 (2014).
7. Teotónio, H. & Rose, M. R. Perspective: reverse evolution. *Evolution* **55**, 653–660 (2001).
8. Tan, L., Serene, S., Chao, H. X. & Gore, J. Hidden randomness between fitness landscapes limits reverse evolution. *Phys. Rev. Lett.* **106**, 198102 (2011).
9. Shopsis, B. et al. Mutations in *agr* do not persist in natural populations of methicillin-resistant *Staphylococcus aureus*. *J. Infect. Dis.* **202**, 1593–1599 (2010).
10. Sousa, A. et al. Recurrent reverse evolution maintains polymorphism after strong bottlenecks in commensal gut bacteria. *Mol. Biol. Evol.* **34**, 2879–2892 (2017).
11. Torrilho, P. A. & Lieberman, T. D. Reversions mask the contribution of adaptive evolution in microbiomes. *eLife* **13**, e93146 (2024).
12. Pennings, P. S., Ogbunugafor, C. B. & Hershberg, R. Reversion is most likely under high mutation supply when compensatory mutations do not fully restore fitness costs. *G3 Genes|Genomes|Genet.* **12**, jkac190 (2022).
13. Biddick, R., Spilker, T., Martin, A. & LiPuma, J. J. Evidence of transmission of *Burkholderia cepacia*, *Burkholderia multivorans* and *Burkholderia dolosa* among persons with cystic fibrosis. *FEMS Microbiol. Lett.* **228**, 57–62 (2003).
14. Kalish, L. A. et al. Impact of *Burkholderia dolosa* on Lung Function and Survival in Cystic Fibrosis. *Am. J. Respir. Crit. Care Med.* **173**, 421–425 (2006).
15. Lieberman, T. D. et al. Parallel bacterial evolution within multiple patients identifies candidate pathogenicity genes. *Nat. Genet.* **43**, 1275–1280 (2011).
16. Lieberman, T. D. et al. Genetic variation of a bacterial pathogen within individuals with cystic fibrosis provides a record of selective pressures. *Nat. Genet.* **46**, 82–87 (2014).
17. Maldonado, R. F., Sá-Correia, I. & Valvano, M. A. Lipopolysaccharide modification in Gram-negative bacteria during chronic infection. *FEMS Microbiol. Rev.* **40**, 480–493 (2016).
18. Hassan, A. A. et al. Structure of O-antigen and hybrid biosynthetic locus in *Burkholderia cenocepacia* clonal variants recovered from a cystic fibrosis patient. *Front. Microbiol.* **8**, 1027 (2017).
19. Hassan, A. A., Vitorino, M. V., Robalo, T., Rodrigues, M. S. & Sá-Correia, I. Variation of *Burkholderia cenocepacia* cell wall morphology and mechanical properties during cystic fibrosis lung infection, assessed by atomic force microscopy. *Sci. Rep.* **9**, 16118 (2019).

20. Silva, I. N. et al. Long-term evolution of *Burkholderia multivorans* during a chronic cystic fibrosis infection reveals shifting forces of selection. *mSystems* **1**, e00029-16 (2016).
21. Chung, H. et al. Global and local selection acting on the pathogen *Stenotrophomonas maltophilia* in the human lung. *Nat. Commun.* **8**, 14078 (2017).
22. Jorth, P. et al. Regional isolation drives bacterial diversification within cystic fibrosis lungs. *Cell Host Microbe* **18**, 307–319 (2015).
23. Syed, I. & Wooten, R. M. Interactions between pathogenic *Burkholderia* and the complement system: a review of potential immune evasion mechanisms. *Front. Cell. Infect. Microbiol.* **11**, 701362 (2021).
24. Butler, S. L., Nelson, J. W., Poxton, I. R. & Govan, J. R. Serum sensitivity of *Burkholderia* (*Pseudomonas*) *cepacia* isolates from patients with cystic fibrosis. *FEMS Immunol. Med. Microbiol.* **8**, 285–292 (1994).
25. Campos, M. A. et al. Capsule polysaccharide mediates bacterial resistance to antimicrobial peptides. *Infect. Immun.* **72**, 7107–7114 (2004).
26. Saldías, M. S. & Valvano, M. A. Interactions of *Burkholderia cenocepacia* and other *Burkholderia cepacia* complex bacteria with epithelial and phagocytic cells. *Microbiol. Read. Engl.* **155**, 2809–2817 (2009).
27. Martin, D. W. & Mohr, C. D. Invasion and intracellular survival of *Burkholderia cepacia*. *Infect. Immun.* **68**, 24–29 (2000).
28. Saini, L. S., Galsworthy, S. B., John, M. A. & Valvano, M. A. Intracellular survival of *Burkholderia cepacia* complex isolates in the presence of macrophage cell activation. *Microbiol. Read. Engl.* **145**, 3465–3475 (1999).
29. Saldías, M. S., Ortega, X. & Valvano, M. A. Y. *Burkholderia cenocepacia* O antigen lipopolysaccharide prevents phagocytosis by macrophages and adhesion to epithelial cells. *J. Med. Microbiol.* **58**, 1542–1548 (2009).
30. Worlitzsch, D. et al. Effects of reduced mucus oxygen concentration in airway *Pseudomonas* infections of cystic fibrosis patients. *J. Clin. Invest.* **109**, 317–325 (2002).
31. Jayaraman, J. et al. Variation at the common polysaccharide antigen locus drives lipopolysaccharide diversity within the *Pseudomonas syringae* species complex. *Environ. Microbiol.* **22**, 5356–5372 (2020).
32. Kronborg, G., Fomsgaard, A., Galanos, C., Freudenberg, M. A. & Høiby, N. Antibody responses to lipid A, core, and O sugars of the *Pseudomonas aeruginosa* lipopolysaccharide in chronically infected cystic fibrosis patients. *J. Clin. Microbiol.* **30**, 1848–1855 (1992).
33. Guadalupe Cabral, M., et al. Using dendritic cells to evaluate how *Burkholderia cenocepacia* clonal isolates from a chronically infected cystic fibrosis patient subvert immune functions. *Med. Microbiol. Immunol.* **206**, 111–123 (2017).
34. Hassan, A. A., Coutinho, C. P. & Sá-Correia, I. *Burkholderia cepacia* complex species differ in the frequency of variation of the lipopolysaccharide O-antigen expression during cystic fibrosis chronic respiratory infection. *Front. Cell. Infect. Microbiol.* **9**, 273 (2019).
35. McMurray, D. N. Hematogenous reseeding of the lung in low-dose, aerosol-infected guinea pigs: unique features of the host–pathogen interface in secondary tubercles. *Tuberculosis* **83**, 131–134 (2003).
36. Ssengooba, W., de Jong, B. C., Joloba, M. L., Cobelens, F. G. & Meehan, C. J. Whole genome sequencing reveals mycobacterial microevolution among concurrent isolates from sputum and blood in HIV infected TB patients. *BMC Infect. Dis.* **16**, 371 (2016).
37. Nucci, A., Rocha, E. P. C. & Rendueles, O. Adaptation to novel spatially-structured environments is driven by the capsule and alters virulence-associated traits. *Nat. Commun.* **13**, 4751 (2022).
38. Culyba, M. J. & Tyne, D. V. Bacterial evolution during human infection: adapt and live or adapt and die. *PLoS Pathog.* **17**, e1009872 (2021).
39. Chung, H. et al. Rapid expansion and extinction of antibiotic resistance mutations during treatment of acute bacterial respiratory infections. *Nat. Commun.* **13**, 1231 (2022).
40. Snitkin, E. S. et al. Genomic insights into the fate of colistin resistance and *Acinetobacter baumannii* during patient treatment. *Genome Res.* **23**, 1155–1162 (2013).
41. Krinos, C. M. et al. Extensive surface diversity of a commensal microorganism by multiple DNA inversions. *Nature* **414**, 555–558 (2001).
42. Jiang, X. et al. Invertible promoters mediate bacterial phase variation, antibiotic resistance, and host adaptation in the gut. *Science* **363**, 181–187 (2019).
43. Mrázek, J., Guo, X. & Shah, A. Simple sequence repeats in prokaryotic genomes. *Proc. Natl. Acad. Sci. USA* **104**, 8472–8477 (2007).
44. Dillon, M. M., Sung, W., Lynch, M. & Cooper, V. S. The rate and molecular spectrum of spontaneous mutations in the GC-rich multichromosome genome of *Burkholderia cenocepacia*. *Genetics* **200**, 935–946 (2015).
45. Lieberman, T. D. Detecting bacterial adaptation within individual microbiomes. *Philos. Trans. R. Soc. B Biol. Sci.* **377**, 20210243 (2022).
46. Groussin, M. et al. Elevated rates of horizontal gene transfer in the industrialized human microbiome. *Cell* **184**, 2053–2067.e18 (2021).
47. Wiser, M. J., Ribeck, N. & Lenski, R. E. Long-term dynamics of adaptation in asexual populations. *Science* **342**, 1364–1367 (2013).
48. Zhao, S. et al. Adaptive evolution within the gut microbiomes of healthy people. *Cell Host Microbe* **25**, 656–667.e8 (2019).
49. Ernst, C. M. et al. Adaptive evolution of virulence and persistence in carbapenem-resistant *Klebsiella pneumoniae*. *Nat. Med.* **26**, 705–711 (2020).
50. Key, F. M. et al. On-person adaptive evolution of *Staphylococcus aureus* during treatment for atopic dermatitis. *Cell Host Microbe* **31**, 593–603.e7 (2023).
51. Ma, K. C. et al. Adaptation to the cervical environment is associated with increased antibiotic susceptibility in *Neisseria gonorrhoeae*. *Nat. Commun.* **11**, 4126 (2020).
52. Kim, S. et al. High-throughput automated microfluidic sample preparation for accurate microbial genomics. *Nat. Commun.* **8**, 13919 (2017).
53. Martin, M. Cutadapt removes adapter sequences from high-throughput sequencing reads. *EMBnet. J.* **17**, 10–12 (2011).
54. Joshi, N. A. & Fast, J. N. Sickle: a sliding-window, adaptive, quality-based trimming tool for FastQ files (Version 1.33) [Software]. <https://github.com/najoshi/sickle>.
55. Langmead, B. & Salzberg, S. L. Fast gapped-read alignment with Bowtie 2. *Nat. Methods* **9**, 357–359 (2012).
56. Li, H. et al. The sequence alignment/map format and SAMtools. *Bioinformatics* **25**, 2078–2079 (2009).
57. Felsenstein, J. PHYLIP (Phylogeny Inference Package) version 3.6. <https://evolution.genetics.washington.edu/phylip.html> (2005).
58. Deatherage, D. E. & Barrick, J. E. Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. *Methods Mol. Biol.* **1151**, 165–188 (2014).
59. KEGG: Kyoto Encyclopedia of Genes and Genomes. <https://www.genome.jp/kegg/>.
60. Price, M. N. & Arkin, A. P. PaperBLAST: text mining papers for information about homologs. *mSystems* **2**, e00039-17 (2017).
61. Davis, M. R. & Goldberg, J. B. Purification and visualization of lipopolysaccharide from Gram-negative bacteria by hot aqueous-phenol extraction. *J. Vis. Exp.* 3916 <https://doi.org/10.3791/3916> (2012).
62. Ellis, C. N. & Cooper, V. S. Experimental adaptation of *Burkholderia cenocepacia* to onion medium reduces host range. *Appl. Environ. Microbiol.* **76**, 2387–2396 (2010).

63. Choi, K.-H., DeShazer, D. & Schweizer, H. P. mini-Tn7 insertion in bacteria with multiple glmS-linked attTn7 sites: example *Burkholderia mallei* ATCC 23344. *Nat. Protoc.* **1**, 162–169 (2006).
64. Schaefer, M. M. et al. Evolution towards virulence in a *Burkholderia* two-component system. *mBio* **12**, e01823-21 (2021).
65. Schaefer, M. M. et al. An oxygen-sensing two-component system in the *Burkholderia cepacia* complex regulates biofilm, intracellular invasion, and pathogenicity. *PLoS Pathog.* **13**, e1006116 (2017).

Acknowledgements

We thank Robert Fowler, Jonathan Greenberg, Elizabeth Carpino, and Craig Gerard for assistance in obtaining samples. We thank Chris Mancuso, Laura Markey for feedback on the manuscript, and all members of Lieberman and Priebe labs for advice and support. This work was funded in part by the Richard A. and Susan F. Smith President's Innovation Award (to G.P.P.) and by funds for the Translational Research for Infection Prevention in Pediatric Anesthesia and Critical Care (TRIPPACC), Program from the Department of Anesthesiology, Critical Care and Pain Medicine at Boston Children's Hospital (to G.P.P.), and U.S. National Institutes of Health grants DP2-GM140922 (to T.D.L.), and the Burroughs Wellcome Fund via a Career Award at the Scientific Interface, Broad Institute Startup Funds, an Emerging Technologies Opportunity Program award from the Department of Energy Joint Genome Institute, and the National Institute for Allergy and Infectious Disease (AI121932) (to P.C.B.).

Author contributions

T.D.L., R.K., and G.P.P. conceptualized the project. T.D.L. and G.P.P. designed, and jointly supervised the work. G.P.P., T.D.L., and S.O.V. collected autopsy samples. A.J.P., C.M., G.K.L. P.C.B. prepared isolates and sequencing libraries. A.J.P., C.M., M.M.S., K.E.M., C.D.A., A.H., A.R.C., J.B.G., conducted and interpreted phenotypic assays. A.J.P., M.M.S., and K.E.M. generated mutant constructs and conducted in vivo experiments. A.Z.U., A.J.M. assisted with patient recruitment and clinical descriptions. A.J.P., T.D.L., M.M.S., and G.P.P. conducted all analyses and interpreted results. A.J.P. and T.D.L. wrote the manuscript with M.M.S. and G.P.P. All authors provided feedback on the manuscript.

Competing interests

P.C.B. is a consultant to or holds equity in 10x Genomics, General Automation Lab Technologies/Isolation Bio, Celsius Therapeutics, Next

Gen Diagnostics, Cache DNA, Concerto Biosciences, Stately, Ramona Optics, Bifrost Biosystems, and Amber Bio. His laboratory has received research funding from Calico Life Sciences, Merck, and Genentech for unrelated work. The remaining authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-025-61168-4>.

Correspondence and requests for materials should be addressed to Tami D. Lieberman or Gregory P. Priebe.

Peer review information *Nature Communications* thanks Paul Planet and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. A peer review file is available.

Reprints and permissions information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2025

¹Institute for Medical Engineering and Sciences, Massachusetts Institute of Technology, Cambridge, MA, USA. ²Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. ³Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA, USA. ⁴Department of Anesthesiology, Critical Care and Pain Medicine, Division of Critical Care Medicine, Boston Children's Hospital, Boston, MA, USA. ⁵Department of Anaesthesia, Harvard Medical School, Boston, MA, USA. ⁶Broad Institute of MIT and Harvard, Cambridge, MA, USA. ⁷Department of Pediatrics, Division of Pulmonary, Asthma, Cystic Fibrosis, and Sleep, Emory University School of Medicine, Atlanta, GA, USA. ⁸Faculty of Biology and Faculty of Computer Science, Technion-Israel Institute of Technology, Haifa, Israel. ⁹Department of Pediatrics, Division of Respiratory Diseases, Boston Children's Hospital, Boston, MA, USA. ¹⁰Adult CF Program, Brigham and Women's Hospital, Boston, MA, USA. ¹¹Department of Pediatrics, Harvard Medical School, Boston, MA, USA. ¹²Department of Laboratory Medicine, Boston Children's Hospital, Boston, MA, USA. ¹³Department of Pathology, Harvard Medical School, Boston, MA, USA. ¹⁴Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA, USA. ¹⁵Department of Pediatrics, Division of Infectious Diseases, Boston Children's Hospital, Boston, MA, USA. ¹⁶These authors jointly supervised this work: Tami D. Lieberman, Gregory P. Priebe.

✉ e-mail: tami@mit.edu; Gregory.Priebe@childrens.harvard.edu