

RESEARCH ARTICLE SUMMARY

CLINICAL MICROBIOLOGY

Evolution and host-specific adaptation of *Pseudomonas aeruginosa*

Aaron Weimann, Adam M. Dinan, Christopher Ruis, Audrey Bernut, Stéphane Pont, Karen Brown, Judy Ryan, Lúcia Santos, Louise Ellison, Emem Ukor, Arun P. Pandurangan, Sina Krokowski, Tom L. Blundell, Martin Welch, Beth Blane, Kim Judge, Rachel Bousfield, Nicholas Brown, Josephine M. Bryant, Irena Kukavica-Ibrulj, Giordano Rampioni, Livia Leoni, Patrick T. Harrison, Sharon J. Peacock, Nicholas R. Thomson, Jeff Gauthier, Jo L. Fothergill, Roger C. Levesque*, Julian Parkhill*, R. Andres Floto*

INTRODUCTION: The major human bacterial pathogen *Pseudomonas aeruginosa* causes multidrug-resistant infections, particularly in people with underlying immunodeficiencies or inflammatory lung diseases such as cystic fibrosis (CF). However, it remains unclear how *P. aeruginosa* has evolved into a highly adapted, globally disseminated pathogen.

RATIONALE: We therefore sought to understand the pathogenic evolution of *P. aeruginosa* by combining population-level genomic exploration with transcriptomic and phenotypic studies.

RESULTS: We analyzed a global collection of 9829 isolates of *P. aeruginosa*, identifying 21 major clones, which we call “epidemic” clones. These epidemic clones caused most clinical *P. aeruginosa* infections worldwide, were widely distributed across the phylogenetic tree, and had all spread globally. We estimate that these epidemic clones emerged from ancestral locations distributed around the world and then expanded nonsynchronously between the late 17th and late 20th centuries, potentially driven by changes in human population density, migration patterns, and/or air pollution. Using pan-genome analysis, we identified significant

differences between epidemic and sporadic isolate genomes in the acquisition of genes involved in specific cellular processes such as transcriptional control (see figure).

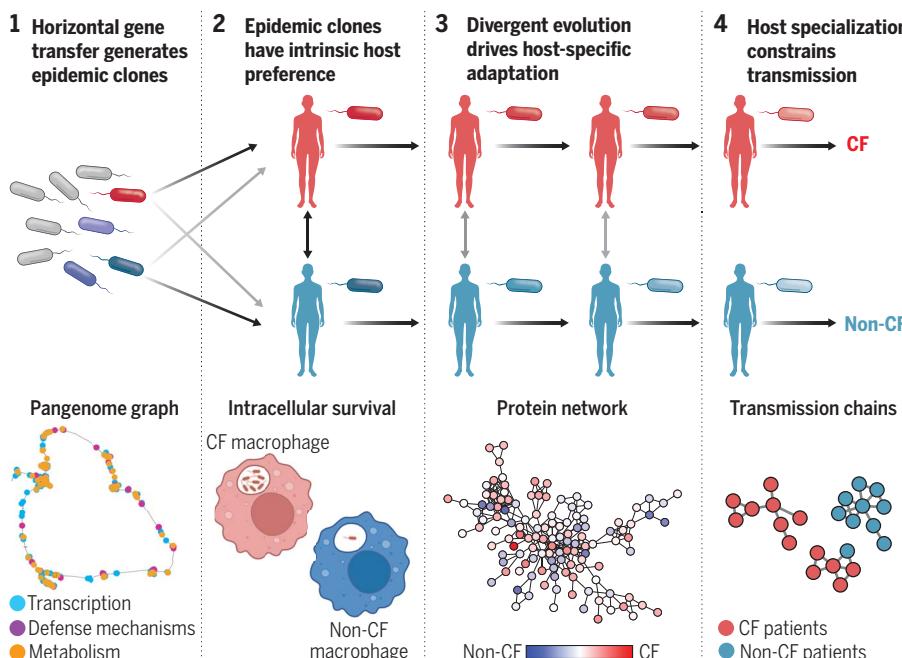
We found that epidemic clones appeared to have intrinsic preferences for CF or non-CF individuals, and we discovered a clear expression signature of genes positively and negatively associated with CF affinity. We found that high-CF-affinity clones were better able to survive within CF macrophages, in part mediated by expression of the stringent response modulator DksA1, suggesting that enhanced host innate immune evasion might explain the intrinsic success at infecting CF patients of certain epidemic clones (see figure).

Examining the recent mutation history of individual clones to understand how epidemic clones of *P. aeruginosa* have adapted to the human host through multiple rounds of within-patient evolution, we identified 224 of 5641 genes that had a higher total mutational burden than expected by chance, which we call “pathoadaptive” genes.

We found that the products of these pathoadaptive genes were tightly interconnected, indicating their likely coordinated functional roles. Many genes were more frequently mutated in either CF or non-CF isolates, suggesting that distinct functional programs were being modified as part of host-specific adaptation (see figure).

Pathoadaptive genes were frequently associated with changes in transmissibility and/or host-specific adaptation, thereby potentially driving host specialization. In support of this notion, we found strong evidence of cross-infection either between CF patients or between non-CF patients, but very little CF to non-CF transmission (see figure).

CONCLUSION: Our findings describe the key sequential steps involved in the evolution of *P. aeruginosa* from an environmental organism to a major human pathogen. These steps include saltatory evolution caused by horizontal gene transfer generating epidemic clones, varying transcriptionally driven varying preference of emerging clones for CF hosts is associated with an increased ability to survive within CF macrophages; (3) distinct trajectories of subsequent evolution of isolates infecting CF versus non-CF patients are mediated by mutations in 224 pathoadaptive genes, which influence transmissibility and/or host-specific adaptation; and (4) as a result of host specialization, transmission is constrained between CF and non-CF patients. ■



Host-specific evolution of *P. aeruginosa*. Shown are key steps in the pathogenic evolution of *P. aeruginosa*: (1) Environmental clones with epidemic potential are created through horizontal gene transfer, as revealed by comparative pan-genome graph analysis; (2) intrinsic, transcriptionally driven varying preference of emerging clones for CF hosts is associated with an increased ability to survive within CF macrophages; (3) distinct trajectories of subsequent evolution of isolates infecting CF versus non-CF patients are mediated by mutations in 224 pathoadaptive genes, which influence transmissibility and/or host-specific adaptation; and (4) as a result of host specialization, transmission is constrained between CF and non-CF patients. [Top panel created with BioRender.com]

The list of author affiliations is available in the full article online.

*Corresponding author. Email: arf27@cam.ac.uk (R.A.F.);

jp369@cam.ac.uk (J.P.); rclevesq@ibis.ulaval.ca (R.C.L.)

Cite this article as A. Weimann et al., *Science* **385**, eadi0908 (2024). DOI: 10.1126/science.adl0908

S READ THE FULL ARTICLE AT
<https://doi.org/10.1126/science.adl0908>

RESEARCH ARTICLE

CLINICAL MICROBIOLOGY

Evolution and host-specific adaptation of *Pseudomonas aeruginosa*

Aaron Weimann^{1,2,3,4}, Adam M. Dinan^{1,2,3}, Christopher Ruis^{1,2,3,4}, Audrey Bernut⁵, Stéphane Pont⁵, Karen Brown^{1,2,6}, Judy Ryan^{1,2}, Lúcia Santos⁷, Louise Ellison², Emem Ukor^{2,6}, Arun P. Pandurangan^{1,8,9}, Sina Krokowski^{1,2}, Tom L. Blundell^{1,8,9}, Martin Welch⁸, Beth Blane⁹, Kim Judge¹⁰, Rachel Bousfield^{9,11}, Nicholas Brown¹¹, Josephine M. Bryant¹⁰, Irena Kukavica-Ibrulj¹², Giordano Rampioni^{13,14}, Livia Leon¹³, Patrick T. Harrison⁷, Sharon J. Peacock^{9,11}, Nicholas R. Thomson^{10,15}, Jeff Gauthier¹², Jo L. Fothergill¹⁶, Roger C. Levesque^{12*}, Julian Parkhill^{4*}, R. Andres Floto^{1,2,3,6,9,11*}

The major human bacterial pathogen *Pseudomonas aeruginosa* causes multidrug-resistant infections in people with underlying immunodeficiencies or structural lung diseases such as cystic fibrosis (CF). We show that a few environmental isolates, driven by horizontal gene acquisition, have become dominant epidemic clones that have sequentially emerged and spread through global transmission networks over the past 200 years. These clones demonstrate varying intrinsic propensities for infecting CF or non-CF individuals (linked to specific transcriptional changes enabling survival within macrophages); have undergone multiple rounds of convergent, host-specific adaptation; and have eventually lost their ability to transmit between different patient groups. Our findings thus explain the pathogenic evolution of *P. aeruginosa* and highlight the importance of global surveillance and cross-infection prevention in averting the emergence of future epidemic clones.

P*seudomonas aeruginosa* is found widely in natural and man-made environments (1–5) and has become an increasingly important opportunistic human pathogen, causing acute nosocomial lung, soft tissue, and systemic infections (6), as well as chronic pulmonary infections in individuals with underlying inflammatory lung diseases such as chronic obstructive pulmonary disease (COPD) (7), cystic fibrosis (CF) (8), and non-CF bronchiectasis (9), in which it causes increased morbidity and mortality (8, 10). Antimicrobial resistance in *P. aeruginosa* is increasing globally, as shown by its inclusion in the World Health Organization's ESKAPE pathogen list (11), and is responsible for >500,000 deaths annually (12).

Although individuals frequently acquire *P. aeruginosa* independently from the environment, hospital-based person-to-person transmission is well recognized in people with CF, leading to strict cohort and individual segregation within clinics (8), but it has not been considered a risk in other patient cohorts (13). Nevertheless, epidemic clones of *P. aeruginosa* have been identified in both CF and non-CF infection studies (14–17), suggesting the presence of widespread transmission networks. In CF, these epidemic clones, such as the Liverpool epidemic strain (LES), are associated with accelerated inflammatory lung damage and worse clinical outcomes, demonstrating that they can emerge with increased virulence in particular patient groups.

We therefore sought to understand the pathogenic evolution of *P. aeruginosa* by defining when and how epidemic clones emerged and spread, exploring how they have adapted to specific hosts, and understanding how within-host evolution has influenced ongoing transmission patterns.

Results

Phylogenetic analysis reveals global epidemic clones

We first analyzed a globally distributed collection of 9829 human, animal, and environmental isolates of *P. aeruginosa*, including 9573 human clinical samples from 2765 patients (14, 18–27) (Fig. 1A and table S1), which we grouped into 596 genetically related clones on the basis of pairwise single-nucleotide poly-

morphism (SNP) distances and then stratified by the number of patients infected by each clone (Fig. 1B, fig. S1, and supplementary materials and methods). We identified 21 major clones, each containing isolates from at least 30 individuals, which we call “epidemic” clones and refer to by their majority multilocus sequence type (28).

We found that these epidemic clones were more likely to be detected in infected humans than in the nonbuilt environment (Fisher's exact test $P = 7.80 \times 10^{-12}$), caused 51% of all clinical *P. aeruginosa* infections worldwide, were widely distributed across the phylogenetic tree, and had all spread globally (Fig. 1C, fig. S1, and table S1).

Inferring population expansion and geography of epidemic clones over time

Because our sample collections spanned >100 years (from 1900 to 2018), we wondered whether we could explore the historical origins of epidemic clones using Bayesian temporal reconstruction (29). We estimate that epidemic clones emerged nonsynchronously between the late 17th and late 20th centuries (fig. S2) and, through Skyline demographic modeling (30) (fig. S3), infer that they have each undergone at least one major population expansion between 1850 and 2000 (Fig. 1D). Considering only extant clones, this suggests that *P. aeruginosa* has undergone relatively recent changes in host-pathogen dynamics, potentially driven by changes in human population density, migration patterns (31, 32), and/or increased susceptibility to infections caused, for example, by increased air pollution during industrialization (33–35).

For epidemic clones with geographically clustered clades (association test P value < 0.05), we implemented Bayesian phylogeographic methods (36) to infer the ancestral location of clones, accepting the limitations of our opportunistic sample collection. For some clones, such as ST235, we were able to find statistical support for the direction of intercontinental spread from South America to North America and Europe, and then subsequently Asia and Africa (fig. S4), whereas for other clones, such as ST17 and ST27, we could identify only that transmission was restricted to between Europe and North America and peaked in the second half of the 20th century (fig. S4). We conclude that epidemic clones have likely arisen from ancestral locations distributed around the world.

Horizontal gene transfer may drive emergence of epidemic clones

We next investigated why some *P. aeruginosa* clones became epidemic and investigated whether gene acquisition, through horizontal transfer, might have driven large jumps in human infectivity through saltatory evolution, defined as abrupt changes in evolutionary fitness caused by sudden large genetic changes (37), as was

¹Victor Phillip Dahdaleh Heart & Lung Research Institute, University of Cambridge, Cambridge, UK. ²University of Cambridge Molecular Immunity Unit, MRC Laboratory of Molecular Biology, Cambridge, UK. ³Cambridge Centre for AI in Medicine, University of Cambridge, Cambridge, UK.

⁴Department of Veterinary Medicine, University of Cambridge, Cambridge, UK. ⁵Laboratory of Pathogens and Host Immunity (LPHI), UMR5235, CNRS/Université de Montpellier, Montpellier, France. ⁶Cambridge Centre for Lung Infection, Royal Papworth Hospital, Cambridge, UK. ⁷Department of Physiology, Bioscience Institute, University College Cork, Cork, Ireland. ⁸Department of Biochemistry, University of Cambridge, Cambridge, UK.

⁹Department of Medicine, University of Cambridge, Cambridge, UK. ¹⁰Wellcome Sanger Institute, Hinxton, UK. ¹¹Cambridge University Hospitals Trust, Cambridge, UK. ¹²Institut de Biologie Intégrative et des Systèmes (IBIS), Université Laval, Québec City, Québec, Canada. ¹³Department of Science, University Roma Tre, Rome, Italy. ¹⁴IRCCS Fondazione Santa Lucia, Rome, Italy. ¹⁵London School of Hygiene and Tropical Medicine, London, UK. ¹⁶Department of Clinical Infection, Microbiology and Immunology, University of Liverpool, Liverpool, UK.

*Corresponding author. Email: arf27@cam.ac.uk (R.A.F.); jpl369@cam.ac.uk (J.P.); rclevesq@ibis.ulaval.ca (R.C.L.)

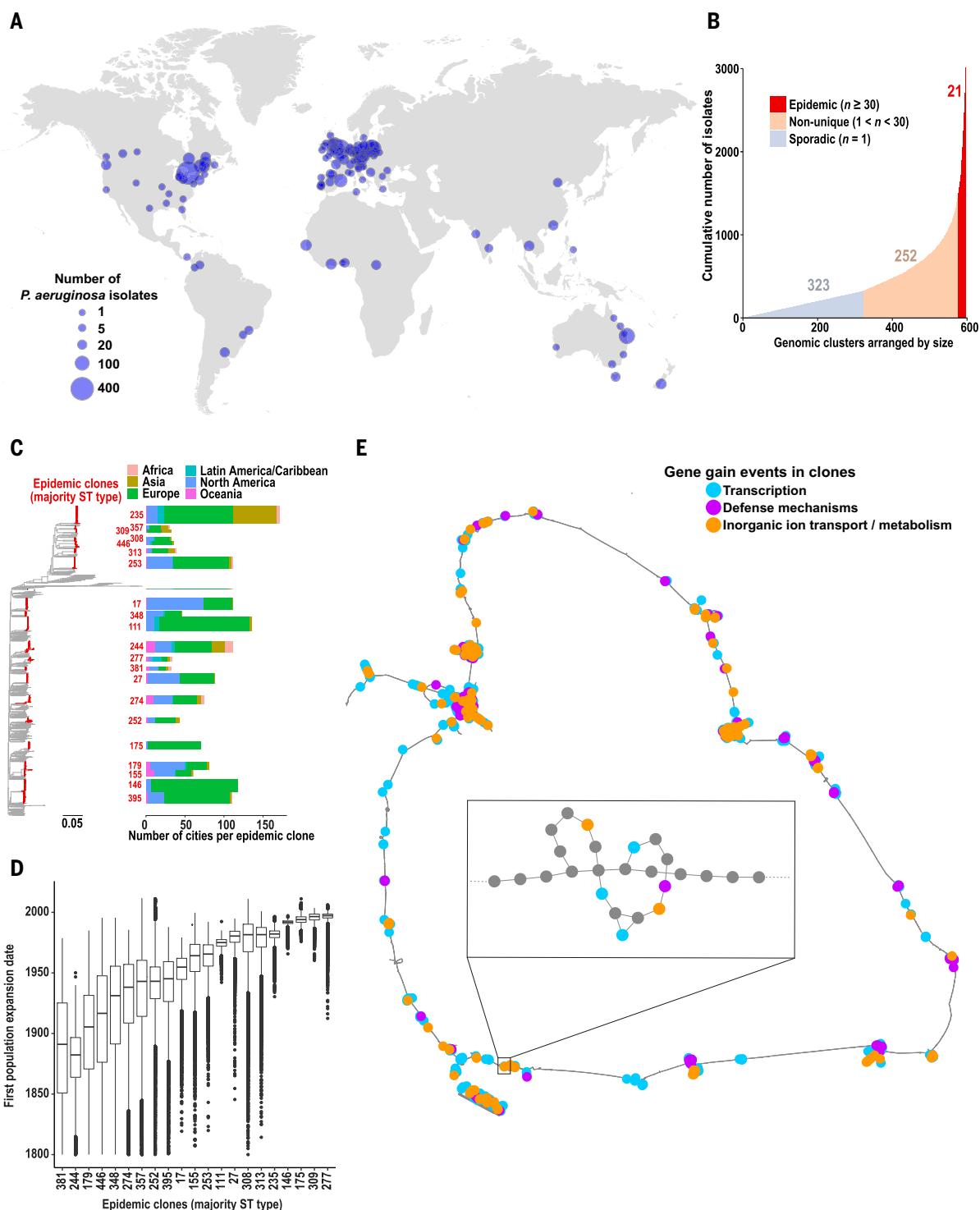


Fig. 1. The emergence of epidemic clones of *Pseudomonas aeruginosa*.

(A) Geographical location of the whole genome–sequenced *P. aeruginosa* isolates obtained from patients, animals, and environment analyzed in this study ($n = 9829$). The number of samples from each location is indicated by the size of blue dot. (B) Cumulative number of isolates across *P. aeruginosa* clones defined by clustering genomes using the unweighted pair group method with arithmetic means (see the supplementary materials and methods) and arranged by ascending number of genomes per clone and stratified into epidemic ($n \geq 30$ isolates/clone; red), non-unique ($1 < n < 30$ isolates/clone; light brown), and unique ($n = 1$ isolate/clone; blue) groups. (C) Left: maximum likelihood

phylogenetic tree generated from the genomes of all study isolates. Major epidemic clones are labeled in red. Right: bar plot representing the number of cities where each epidemic clone was found, colored by continent. (D) Estimated date of first population expansion of 21 epidemic clones predicted by Bayesian inference using BEAST (29), with the graph showing median and interquartile range (IQR; boxplots), 1.5 times the IQR range (whiskers), and data points outside of this range (black points). (E) Pan-genome graph analysis of ancestral representatives of epidemic clones ($n = 21$) and sporadic clones ($n = 80$) constructed using Panaroo (39). Nodes represent clusters of orthologous genes, and two nodes are connected by an edge if they are adjacent on a contig in

any sample from the population and define gene gain events associated with the emergence of epidemic clones (described in detail in fig. S5). Highlighted genes are involved in transcription (blue), defense mechanisms (purple), and inorganic ion

transport and metabolism (yellow). For illustration purposes, the graph has been ordered against the genome of *P. aeruginosa* PAO1. Inset: magnified section of the pan-genome graph shown to illustrate node and edge structure.

previously observed in *Mycobacterium abscessus* (38). To accurately analyze the *P. aeruginosa* accessory genome, we first inferred a pan-genome graph using Panaroo (39), with nodes as clusters of orthologous genes and two nodes linked by an edge if they were found adjacent in any contig (Fig. 1E). We then compared the accessory genomes of representatives of ancestral epidemic clones with those of sporadic isolates and found that epidemic clones had significant enrichment of genes involved in transcriptional regulation, inorganic ion transport, lipid metabolism, and protein turnover, with significant depletion of genes involved in bacterial defense and secretion [Fisher's exact test, false discovery rate (FDR) = 0.1; Fig. 1E, fig. S5, and table S2], suggesting that fundamental changes in bacterial physiology might have driven the development of epidemic clones.

Epidemic clones vary in their intrinsic host preference

We next examined the types of human infection caused by the epidemic clones, and found that ST146, also known as LES, caused infection almost exclusively in people with CF. Conversely, some clones, such as ST175 and ST309, caused infection only in non-CF individuals, and there was a range of CF and non-CF host distributions for other clones (Fig. 2A). Although our collection was opportunistically sampled, our findings were replicated when reanalyzing existing systematic UK surveillance data (40) for the frequency of CF and non-CF lung infections caused by each epidemic clone (fig. S6). We hypothesized that this apparent clone-specific host preference might relate to intrinsic differences in bacterial behavior between clones. Pan-genome analysis revealed no evidence for an association of host affinity with acquisition of genes with a particular function (fig. S7), so we reasoned that changes in gene expression could explain differences in host preference. We therefore analyzed a previous transcriptomic study of clinical *P. aeruginosa* isolates that were represented in our sample collection, nearly all of which were from non-CF patients (25), and found that clinical isolates clustered in transcriptional space based on the host preference of epidemic clones ($P < 0.001$; Fig. 2B and the supplementary materials and methods).

We next explored whether there were any patterns in gene expression associated with clone predisposition for specific human hosts and, using a negative binomial generalized linear model, identified a clear expression signature of 624 genes positively associated, and

514 genes inversely associated, with affinity for causing CF infection (Wald test, FDR = 0.05; Fig. 2C and table S3).

Increased survival within macrophages of high-CF-affinity clones

To identify a potential mechanism by which differential gene expression could alter the host preference of epidemic clones, we conducted a multidimensional phenotypic characterization of 49 representative isolates, almost all of which were from non-CF patients to minimize the contribution of secondary host-specific adaptation. These isolates ranged from epidemic clones with low (ST235 and ST111), intermediate (ST253), and high (ST17, ST27, and ST146) levels of CF affinity and were obtained from the International Pseudomonas Consortium Database (21). We initially examined isolate behavior in established assays of *Pseudomonas* virulence (biofilm formation, siderophore production, swim and twitch motility, and production of caseinase and gelatinase), but we could find no correlation with clonal host affinity (fig. S8).

We therefore decided to test the ability of representative isolates of different epidemic clones to withstand intracellular killing by macrophages that, together with neutrophils, are thought to be the first line of defense against bacterial lung infection (41, 42). We found significantly increased intracellular survival and replication of isolates from the high-CF-affinity clone (ST27) compared with isolates from the low- and intermediate-CF-affinity clones ST111 and ST235 in both wild-type (WT) and CF (F508del homozygous) isogenic macrophage cell lines (Fig. 2D and table S4). This suggested that enhanced host innate immune evasion might explain the intrinsic success of particular epidemic clones in infecting CF patients.

Host preference of epidemic strains is mediated by DksA1 expression

To further explore the bacterial mechanisms contributing to the enhanced intracellular survival of bacterial isolates from high-CF-affinity clones, we studied the differentially expressed gene set (Fig. 2C) and discovered that both the expression of the stringent response modulator DksA1 and the activation of its regulon were associated with CF affinity (Fig. 3A). Because DksA1 had previously been implicated in enhancing *P. aeruginosa* survival within mouse macrophages and increasing tolerance to H₂O₂ (43), we examined the impact of deleting DksA1 on bacterial survival by using *in vitro* and *in vivo* models of non-CF and CF infection

using *P. aeruginosa* PAO1 WT and isogenic DksA1 and DksA2 double-knockout (Δ DksA1,2) and complemented (Δ DksA1,2::DksA1) strains. Δ DksA1,2 double-knockout mutants were used because of gene redundancy (44).

We first tested bacterial survival in macrophages and found that although all three strains were effectively killed in WT THP1 macrophages, only the Δ DksA1,2 strain could be killed by isogenic CF (F508del knock-in) cells, whereas WT and complemented bacteria were able to resist macrophage killing and replicate intracellularly (Fig. 3B). Our findings reveal a number of important features of host-*Pseudomonas* interactions. First, they show that there are intrinsic defects in CF macrophages that facilitate the intracellular survival of *P. aeruginosa*, observations that are supported by previous *in vitro* reports (45) and by our *in vivo* experiments in zebrafish, in which deletion (46) or morpholino knockdown of the *cystic fibrosis transmembrane regulator* (*cfr*) gene compromised survival after intravenous infection (fig. S9). Second, our findings indicate that intracellular survival in CF macrophages is mediated by DksA1, raising intriguing mechanistic questions about the role of the stringent response in surviving the phagosomal environment. Finally, our results reveal that differences in DksA1 expression across epidemic clones may explain their observed different abilities to survive within macrophages and, potentially as a consequence, their varying intrinsic host preferences.

To explore the potential role of DksA1 further, we examined the behavior of fluorescently labeled WT, Δ DksA1,2 mutant, and complemented *P. aeruginosa* PAO1 during *in vivo* infection in zebrafish larvae (Fig. 3, C to G). We observed increased survival of both control and *cfr* morpholino-treated fish after infection with Δ DksA1,2 bacteria compared with WT and complemented strains (Fig. 3D), findings that are consistent with the observed decreased survival of Δ DksA1,2 bacteria *in vivo* (Fig. 3E).

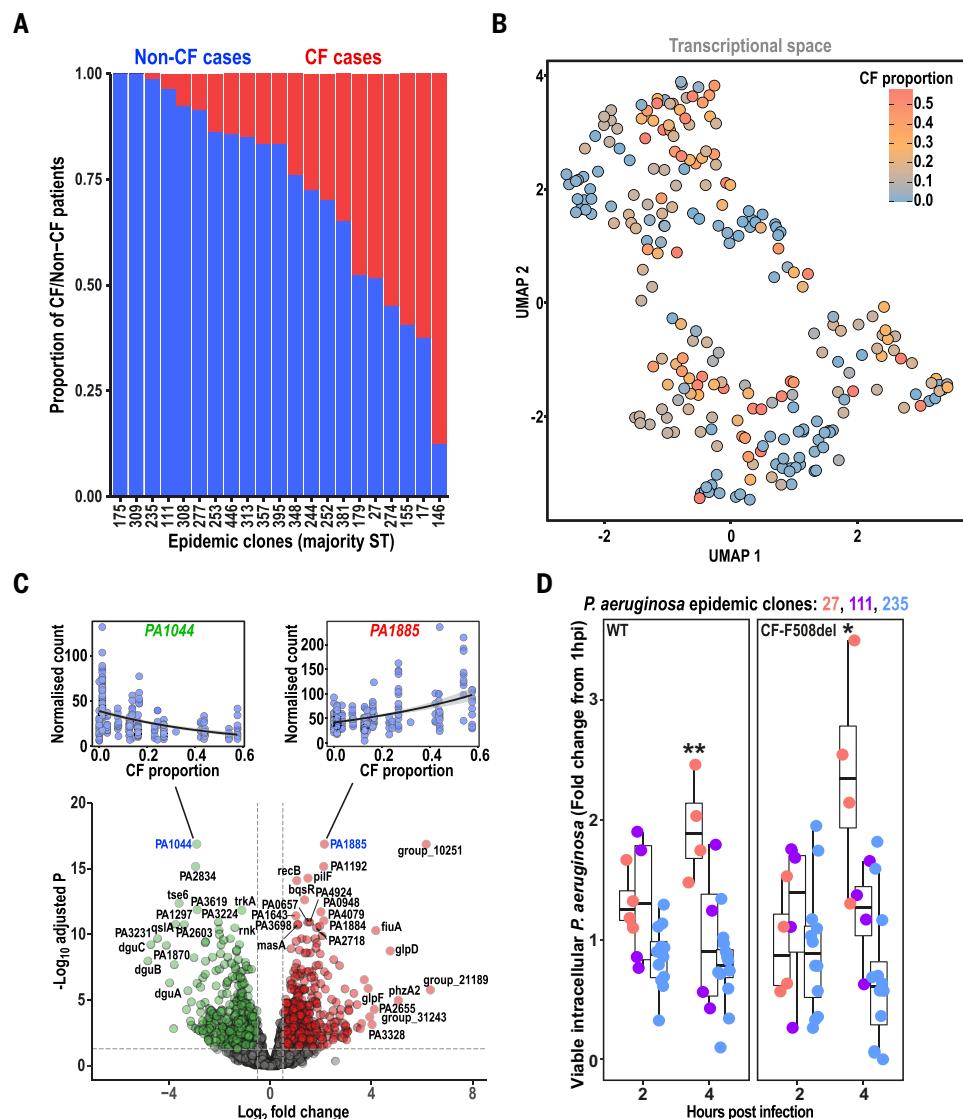
We then exploited the optical transparency of zebrafish larvae to track the interaction of macrophages with *P. aeruginosa* after intramuscular infection using intravital confocal microscopy. Using a fluorescent macrophage reporter fish line, *Tg(mpeg1:mcherry-F)ump2* (47), we were able to clearly identify macrophages and distinguish extracellular from intracellular fluorescent bacteria (Fig. 3F). We found no difference in the mobilization of macrophage to the site of infection or in the proportion of infected macrophages in control and

Fig. 2. Variable intrinsic host preference of epidemic *P. aeruginosa* clones.

(A) Proportion of infections caused by epidemic clones labeled by their majority multilocus sequence type (ST) in CF (red) and non-CF (blue) patients. (B) Uniform manifold approximation and projection (UMAP) projection of transcriptomes from representative isolates of epidemic clones (25), color coded by the CF affinity of each clone. Expression data were pseudo-aligned to strain-specific gene indices to produce estimates of gene transcript abundance. (C) Transcriptome-wide association of gene expression with CF affinity. Transcript abundances were modeled as a response to the proportion of CF infections caused by each epidemic clone using a negative binomial generalized linear model. Volcano plot visualization of the log₂-fold expression change with CF proportion for every gene in the 99% core genome of *P. aeruginosa* (center). Genes with an adjusted *P* value < 0.05 and a log₂-fold change <-0.5 are shown in green; genes with a log₂-fold change >0.5 are shown in red. The coefficients for gene models were assessed using the Wald test (FDR = 0.05). Normalized expression counts versus CF proportions per epidemic clone with a trendline for the two genes with the lowest and highest log₂ fold change are shown above (top left and top right, respectively). Bulk RNA-seq data were analyzed from 241 clinical isolates of epidemic clones (25) included in our strain collection. (D) Survival of epidemic clones within WT or isogenic F508del knock-in THP1 macrophages at 2 and 4 hours after infection, expressed as fold change from 1 hour after infection showing median and IQR (boxplots) and 1.5 times IQR range (whiskers).

Experiments were performed (at least in

duplicate) by exposing THP1 macrophages to pooled isolates of 51 clinical isolates at an MOI of <1. Viable bacteria were isolated from macrophages at the time points indicated and grown on solid media. Isolate abundance was quantified using sequence-based deconvolution. Strains with <1% abundance at the 1-hour time point were excluded from the analysis. A difference in the abundance of ST27 strains versus ST111 and ST235 strains at the 4-hour time point was assessed using a two-tailed *t* test. **P* < 0.05, ***P* < 0.01.



cfr morpholino fish infected with WT, ΔDksA1,2 mutant, or complemented *P. aeruginosa* PAO1 bacteria (fig. S9 and Fig. 3G). However, we did observe a clear reduction in macrophage bacterial burden, suggesting reduced intracellular replication, after ΔDksA1,2 infection of both control and *cfr* morpholino fish lines (Fig. 3G), confirming the critical role of DksA1 for intracellular survival during *in vivo* *P. aeruginosa* infection.

Taken together, our data indicate that intrinsic elevations in DksA1 expression levels in some epidemic clones may have enabled them to exploit potential innate immune defects in CF and adopt the specific evolutionary strategy of replicating within macrophages.

Convergent host-specific adaptation of *P. aeruginosa*

We next investigated how, once they are selected from the environment, epidemic clones of *P. aeruginosa* have adapted to the human host through multiple rounds of within-patient evolution by analyzing the recent mutation history of individual clones.

By reconstructing mutations that had likely occurred since the emergence of each clone, we found strong evidence for convergent molecular evolution, identifying 224 out of 5641 genes that had a higher total mutational burden than would be expected by chance (Poisson test, FDR = 0.05; Fig. 4A and table S5), which we refer to as “pathoadaptive” genes.

Mutations in these pathoadaptive genes were more likely to be nonsynonymous and deleterious, as shown by variant effect annotation (48), than those found in other genes (Fisher's exact test *P* < 1.0×10^{-16}), and were predicted to be more likely to cause protein dysfunction, as estimated by both sequence conservation methods [SIFT (49); Wilcoxon rank sum test, *P* = 9.04×10^{-15}] and structural modeling approaches [FoldX (50); Wilcoxon rank sum test, *P* = 1.34×10^{-6}] (fig. S10), suggesting that pathoadaptation is largely driven by loss-of-function mutations. We next explored the functional impact of pathoadaptive mutations experimentally by using existing RNA-seq datasets (25) to examine

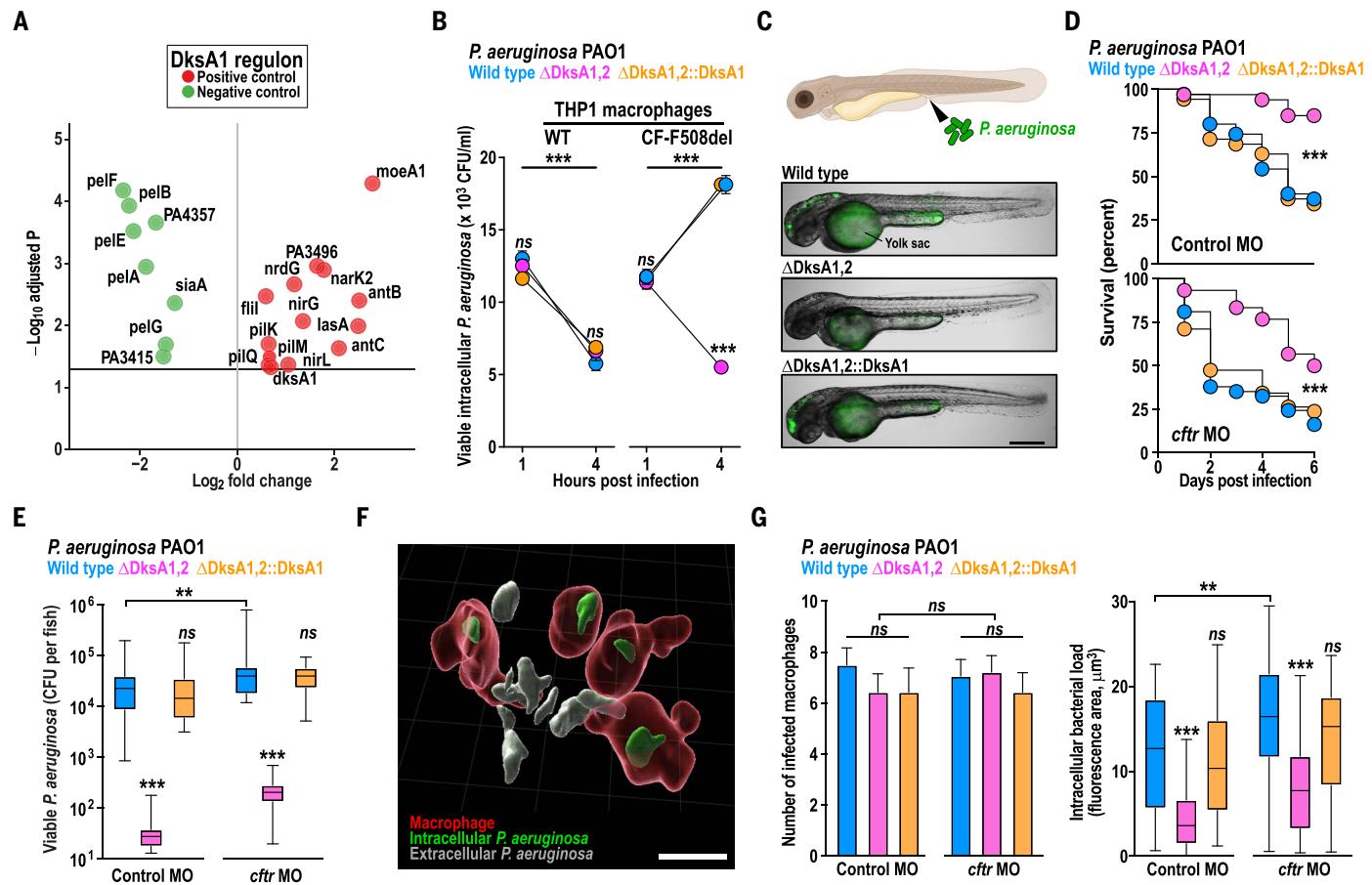


Fig. 3. Activation of the DksA1 regulon contributes to CF host preference of *P. aeruginosa* clones. (A) Volcano plot visualization of the log₂-fold expression change with CF proportion for genes positively controlled (red) and negatively controlled (green) within the DksA1 regulon as defined by Fortuna et al. (44). Bulk RNA-seq data were analyzed from 241 clinical isolates of epidemic clones (24) included in our strain collection. (B) DksA1 promotes survival of *P. aeruginosa* within CF macrophages. Viable intracellular *P. aeruginosa*, quantified through enumeration of cell-associated CFUs, were measured at 1 and 4 hours after infection of differentiated WT and isogenic F508del homozygous knock-in (CF-F508del) THP1 cells with WT (blue), isogenic DksA1-DksA2 double-knockout (Δ DksA1,2; pink), and knockout complemented with DksA1 (Δ DksA1,2::DksA1; yellow) *P. aeruginosa* PAO1. Data (mean \pm SEM) are representative of at least three independent experiments performed in at least triplicate. *** P < 0.001; ns, not significant (two-tailed Student's *t* test). (C) Top: cartoon of zebrafish illustrating the injection site for GFP-labeled fluorescent *P. aeruginosa*. Bottom: representative fluorescence and differential interference contrast (DIC) images of whole infected zebrafish larvae at 1 day after infection. Scale bar, 150 μ m. The labeled yolk sac is autofluorescent. [Figure created with BioRender.com] (D) Survival analysis of control (top) and cftr morphant (cftr MO; bottom) zebrafish larvae infected intravenously (250 to 350 CFU) with *P. aeruginosa* PAO1 WT (blue), Δ DksA1,2 knockout (pink), and Δ DksA1,2::DksA1 complemented (yellow) fluorescent strains plotted as the percentage of surviving animals over

6 days. Data are the average of two independent experiments, with n = 30 to 38 fish for each condition (*** P < 0.001, Mantel-Cox log-rank test). (E) Viable *P. aeruginosa* in zebrafish larvae at day 1 after infection with *P. aeruginosa* PAO1 WT (blue), Δ DksA1,2 knockout (pink), and Δ DksA1,2::DksA1 complemented (yellow) fluorescent strains plotted as mean \pm IQR CFUs per fish of at least three independent experiments and n = 15 to 20 larvae per condition (*** P < 0.001; ns not significant, two-way ANOVA with Tukey's post test). (F and G) Control and cftr morphant zebrafish larvae with mCherry-labeled macrophages [*Tg(mpeg1:mcherry-F)ump2* (45)] were intramuscularly infected with 250 to 350 GFP-labeled *P. aeruginosa* PAO1 WT, Δ DksA1,2, or Δ DksA1,2::DksA1 strains, and the infection was tracked using real-time intravital confocal microscopy. (F) Representative three-dimensional reconstruction of confocal imaging showing macrophages (red) and automatic classification of extracellular (gray) and intracellular (green) *P. aeruginosa*. Scale bar, 10 μ m. (G) Quantification of the number of infected macrophages at the site of injection (left) and the level of intracellular bacterial load calculated by the volume of bacteria-associated fluorescence observed within each macrophage at 6 hours after infection with *P. aeruginosa* PAO1 WT (blue), Δ DksA1,2 knockout (pink), and Δ DksA1,2::DksA1 complemented (yellow) fluorescent strains. Data are shown as mean \pm IQR of at least 54 cells per condition from n = 4 to 6 larvae recorded from two independent experiments (** P < 0.01; *** P < 0.001; ns not significant, two-way ANOVA with Tukey's post test).

the effect of transcription factor variants on the expression of their previously characterized regulons (51–54). We found that clinical isolates with pathoadaptive variants in several transcription factors had statistically significant shifts in regulon expression levels compared with controls (two-tailed *t* test with adjusted P value < 0.00014; fig. S10), supporting the concept of a general loss-of-function evolutionary process driving *P. aeruginosa* pathoadaptation.

We were able to functionally annotate most of these pathoadaptive genes using previously published information (55), identifying many of them as having established roles in recognized pathogenic processes, including biofilm formation, antibiotic resistance, and lipopolysaccharide modification (fig. S11). The number of genes with an established function was much

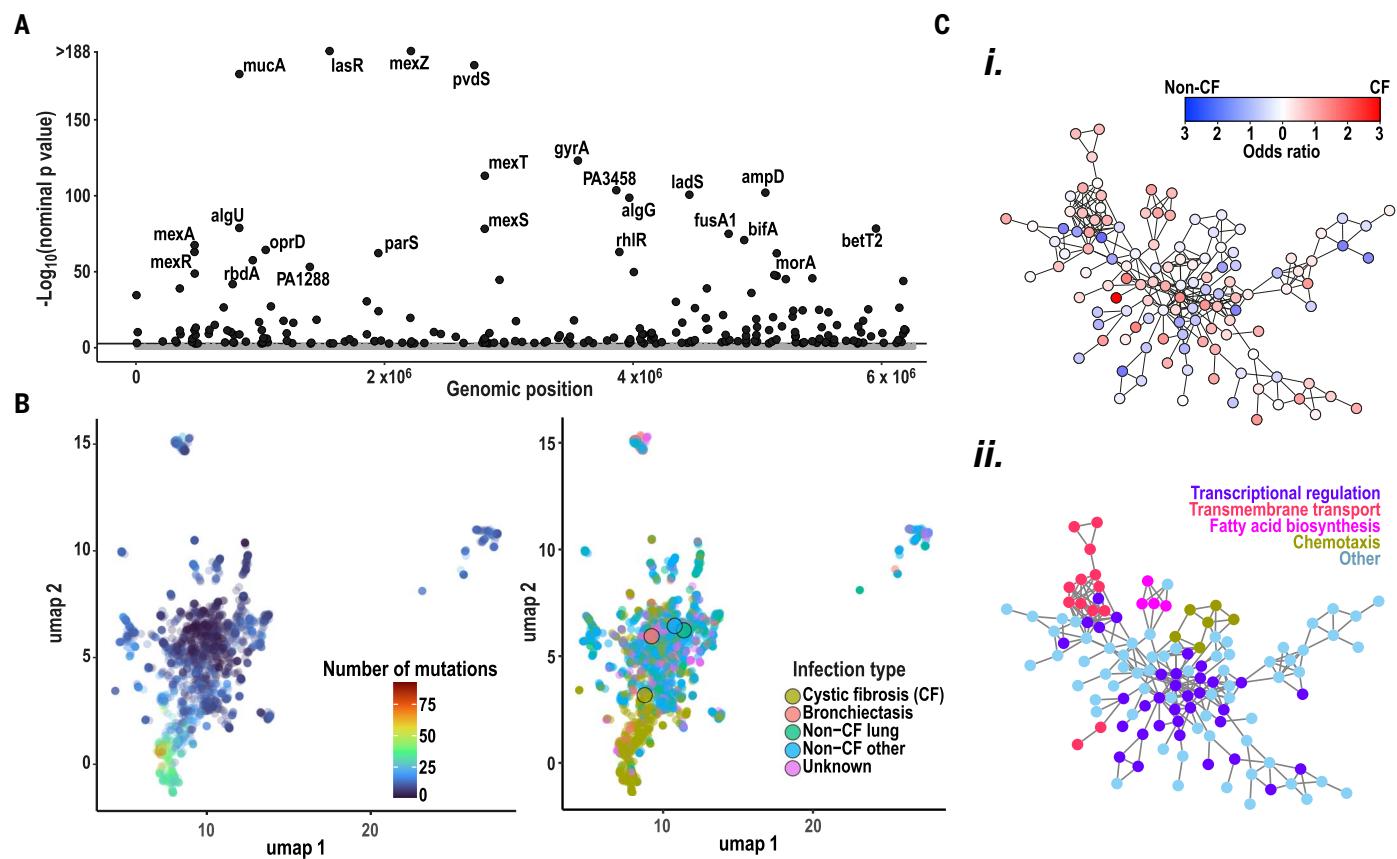


Fig. 4. Host-specific pathoadaptation of *P. aeruginosa*. (A) Manhattan plot showing nominal P values ($-\log_{10}$) from genome-wide mutational burden test across all genes in *P. aeruginosa* PAO1. Significance was assessed using a Poisson test comparing the expected and observed number of mutations in each gene accounting for the proportion of genomes in which that gene was found in the pan-genome (FDR = 0.1). Genes with a significant mutational burden (pathoadaptive genes) are shown in black; others are shown in gray. (B) UMAP projections of host adaptation of isolates based on acquired mutations in pathoadaptive genes color-coded by number of pathoadaptive mutations (left) and type of infection (right). Centroids are denoted by larger dots. Isolates without any pathoadaptive mutations were removed from the analysis. (C) Protein-protein

interaction network for the pathoadaptive genes extracted from the STRING database. Only main connected components are shown; full graphs are shown in fig. S12 (56). Genes are shown as nodes connected by an edge if they had an interaction reported in STRING (confidence > 0.7). Top: To estimate host-specific pathoadaption, the numbers of CF versus non-CF mutations, determined by stratifying mutations in pathoadaptive genes on terminal branches by the infection type of isolates, were compared using a Fisher's exact test (FDR = 0.1) and are expressed as an odds ratio for each gene. Bottom: Gene nodes are color coded by class of functional annotation based on overrepresented pathways using Gene Ontology (89) biological process enrichment analysis with TopGO (57) (for CF, transmembrane transport and fatty acid biosynthesis; for non-CF, transcriptional regulation and chemotaxis).

higher among pathoadaptive genes than in other genes (Fisher's exact test $P < 1.0 \times 10^{-16}$; fig. S11), potentially reflecting their central role in *P. aeruginosa* pathobiology. We also characterized the function of 41 pathoadaptive genes experimentally by de novo screening relevant transposon mutants in a series of functional assays to quantify virulence traits (see the supplementary materials and methods and fig. S11).

We next examined the nature of host adaptation achieved by individual *P. aeruginosa* isolates by using the profile of their pathoadaptive gene mutations to map them in evolutionary space, as defined by the presence or absence of mutations in the 224 pathoadaptive genes. We found that CF isolates clustered separately from others and had accumulated more mutations, suggesting that the CF lung is a distinct niche with different selective pres-

sures compared with other lung or nonlung environments (Fig. 4B).

We found that the products of these pathoadaptive genes were tightly interconnected, with more protein-protein interactions than would be expected by chance [STRING database (56); $P < 1 \times 10^{-16}$; Fig. 4C and fig. S12], indicating their likely coordinated functional roles. We observed that 54 genes were more frequently mutated in CF isolates, whereas 47 genes were more commonly mutated in non-CF isolates (Fisher's exact test, FDR = 0.05). Among genes that were more commonly mutated in CF or non-CF, we found several overrepresented pathways using Gene Ontology biological pathway enrichment analysis with TopGO (57), suggesting that distinct functional programs were being modified as part of host-specific adaptation (Fig. 4C). For example, CF isolates were more likely to have mutated AlgU, a key

regulator of mucoidy (58) with mutations occurring predominantly at the interface between this sigma-H factor and its negative regulator protein, MucA, and PcnA, a putative nicotinamide with mutations found within the protein core or at sites of protein-protein interaction (fig. S13). By contrast, non-CF isolates were more likely to have mutated LadS, a calcium-responsive histidine kinase (59) with mutations concentrated in the N-terminal (sensor) and transmembrane domains, and in the putative choline transporter BetT2 (60) with helix-breaking mutations found within the transmembrane domain (fig. S13).

Distinct evolutionary trajectories lead to host specialization

We then used ancestral state reconstruction to determine the order of acquisition of each pathoadaptive mutation and thereby recreate

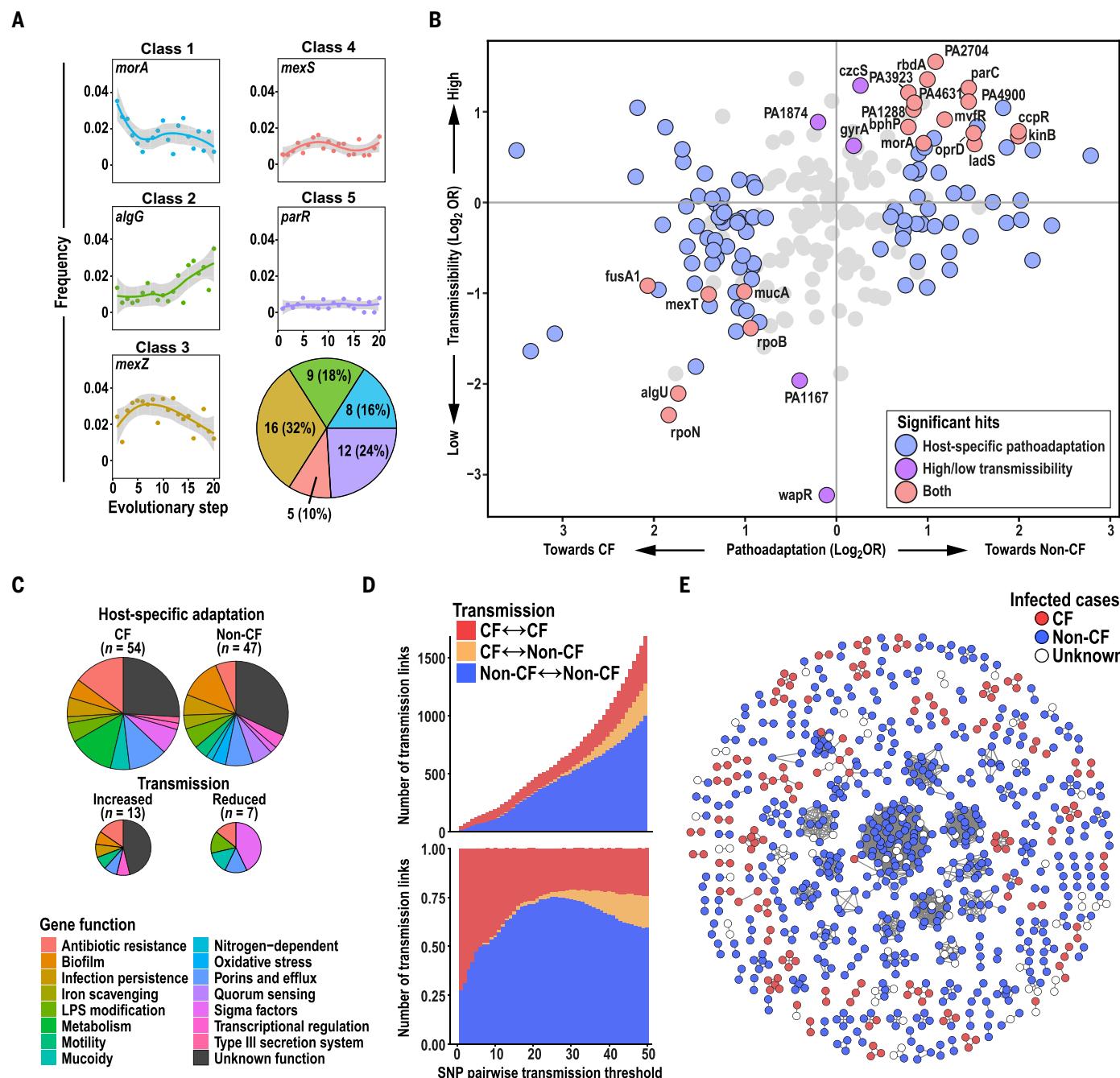


Fig. 5. Evolutionary trajectories of *P. aeruginosa* during pathoadaptation.

(A) Normalized frequency of mutations over evolutionary time in specific pathoadaptive genes. The trajectories of the 50 most commonly mutated genes were manually assigned to one of five classes (fig. S15), based on the shape of their mutation frequency curves. The relative size of each class and representative examples with trendlines from locally weighted smoothing are shown. (B) The relative transmissibility and host-specific adaptation of pathoadaptive genes. To estimate host-specific pathoadaptation, the number of CF versus non-CF mutations was determined by stratifying mutations in pathoadaptive genes on terminal branches by the infection type of isolates, compared using a Fisher's exact test (FDR = 0.1), and expressed as an odds ratio. To assess the transmissibility of pathoadaptive changes, the numbers of mutations that had been observed in at least two isolates were compared with

mutations that had only been observed once using a Fisher's exact test (FDR = 0.1). Genes are color coded if they showed significant host-specific adaptation (blue), changes in transmissibility (purple), or both (pink). Genes with zero or infinite odds ratios are not shown. (C) Functional annotation of pathoadaptive genes associated in host-specific adaptation (top) and changes in transmissibility (bottom) (FDR = 0.05). (D) The number (top) and proportion (bottom) of transmission links across a range of pairwise SNP thresholds for CF to CF (red), CF to non-CF (yellow), and non-CF to non-CF (blue) individuals. Data were down-sized to contain equal numbers of CF and non-CF infections. (E) Transmission clusters involving patients with CF (red), non-CF (blue), or unknown status (white). Nodes representing isolates were connected by edges if pairwise SNP distances were 26 SNPs or less. This cutoff represents the 95th percentile of the within-host genetic diversity analyzed in 81 patients.

the evolutionary trajectory of each isolate. We found that, on average, CF isolates had longer trajectories than non-CF isolates, with 20.5 compared with 11.2 steps (Wilcoxon signed-rank test $P < 1.0 \times 10^{-16}$; fig. S14). By looking at the frequency of mutations in pathoadaptive genes at each evolutionary time step, we were able to cluster genes into five groups with distinct temporal signatures (Fig. 5A and fig. S15), suggesting that mutations in specific genes may be important at different stages of evolution [as noted previously, specifically for the evolution of antimicrobial resistance (61)].

Given the above results, we wondered whether specific sets of pathoadaptive genes might be driving distinct evolutionary processes such as host-specific adaptation, person-to-person transmission, or both. To investigate this, we inferred the impact of each pathoadaptive gene on host-specific pathoadaptation by examining the relative frequency of gene mutations occurring in CF compared with non-CF lung isolates and on bacterial transmissibility based on the frequency of specific gene mutations being found in isolates from at least two patients. We then used this information to create a map of the contribution of each pathoadaptive gene to each of these evolutionary processes (Fig. 5B), annotating each gene by previously known or experimentally derived function or the type of temporal mutation signature observed (Fig. 5C and fig. S16).

We found that whereas some pathoadaptive genes were associated with changes in either transmissibility or host-specific pathoadaptation, many were implicated in both processes (fig. S16). For example, mutations in several genes, such as *mvfR* and *morA*, occurred early in the evolutionary trajectories of isolates and were associated with both adaptation to the non-CF host and increased transmissibility, whereas mutations in other genes, such as in *fusA1* and *algU*, occurred late in evolutionary journeys and were associated with adaptation to the CF host and decreased transmissibility (fig. S16).

Because we observed a likely negative impact of several pathoadaptive mutations on transmissibility, we investigated whether pathoadaptation might lead to host specialization and result in reduced transmission of isolates between CF and non-CF individuals. To explore this possibility, we used the genomic relatedness of isolates to plot the number and proportion of transmission links over a range of SNP pairwise thresholds representing transmission chains of various lengths. We found strong evidence for CF-to-CF patient transmission and non-CF to non-CF patient cross-infection, but very little CF to non-CF transmission (Fig. 5D). Additionally, we reconstructed transmission clusters at a specific SNP threshold (26 SNPs) on the basis of the measured genetic diversity within individual

patients (see the supplementary materials and methods) and found transmission clusters of variable sizes but very few containing both CF and non-CF patients (Fig. 5E). Thus, it seems likely that host-specific pathoadaptation of epidemic *P. aeruginosa* clades limits transmission between different hosts.

Discussion

Our findings describe the key sequential steps involved in the evolution of *P. aeruginosa* from an environmental organism to a major human pathogen. We identified horizontal gene acquisition as a likely driver for the emergence of epidemic clones from the environment through saltational evolution [as previously described for *M. abscessus* (38)] and infer their spatiotemporal spread, which suggests an increasing rate of new epidemic clone expansions over time even though only extant clones are considered. We identified an intrinsic and variable host-specific affinity across epidemic clones with CF preference potentially causally associated with improved intracellular survival in macrophages. We then show how deleterious mutations in a discrete set of functionally interrelated genes likely mediate further host specialization through multiple rounds of within-patient adaptation and onward transmission, thereby plausibly explaining the observed lack of person-to-person transmission between CF and non-CF patients.

Our work highlights the importance of preventing pathogenic evolution by minimizing cross-infection, not just within CF cohorts, in which infection control measures are well established, but also between non-CF patients, and emphasizes how global surveillance and targeted monitoring of high-risk patient groups will be needed to detect expansion, pathoadaptation, and transmission of new and extant epidemic *P. aeruginosa* clones.

Methods summary

Genomic datasets and clone assignment

We collated *P. aeruginosa* genome datasets from studies of antibiotic resistance (18, 23–25) from individuals with cystic fibrosis (14, 26) and non-CF bronchiectasis (22); from the International *Pseudomonas* Consortium (21); and from studies targeting high-risk clones (19, 20, 27). Newly sequenced genomes from the TeleCF study, which involved adults with CF ($n = 15$) who underwent home monitoring for 6 months and were chronically infected with *Pseudomonas*; from bacteremia infections ($n = 365$) as part of the UK BSAC bacteremia resistance program (62); and from patients attending hospitals in Cambridgeshire, UK, were also included. DNA was extracted using QIAxtractor (QIAGEN), and samples were sequenced on the Illumina HiSeq 2000, 2500, and X10 platforms.

Variants were called by mapping reads against the *P. aeruginosa* PAO1 reference genome (acces-

sion number AE004091.2) using the multiple_mappings_to_bam 1.6 pipeline with default parameters (<https://github.com/sanger-pathogens/bact-gen-scripts>) using BWA (63) for mapping, followed by stringent quality-control filtering and removing samples with an excess number of minority variants. Ariba 2.14.6 (64) was used for multilocus sequencing typing (28). FastTree (2.1.10) was used to infer a global phylogenetic tree (65).

Clones were assigned by first grouping samples on the basis of pairwise SNP distances using the ultrametric pairwise group method with arithmetic means (UPGMA) and then applying a cutoff of 7000 SNPs. SNP-sites was used to infer a clone-specific alignment of variable sites (66). Gubbins version 2.4.1 (67) was used to remove recombination for individual clones with at least four available genomes.

Dating and phylogeography

Molecular dating was performed for all 21 epidemic clones separately. Potential hypermutators (distorting the temporal signal) were removed by identifying samples with an unusual ratio of transition and transversion mutations. The temporal signal was assessed with TempEst (68) by comparing collection dates with root-to-tip distances using undated phylogenetic trees inferred with RAxML 8.2.12 (69). The significance of the signal was assessed using a permutation test with a custom script (https://github.com/chrisruis/tree_scripts/blob/main/bootstrap_TempEst_rttd_date.R). Clones with a significant temporal signal in this test ($P < 0.05$) were then used for molecular dating with BEAST 2.6.6 (29). We modeled the population history using the coalescent Bayesian skyline population prior. Convergence was assessed with Tracer 1.7.1 (70) with 10% burn-in. For clones that did not pass the bootstrap randomization test ($n = 9$), a uniform prior for the substitution rate was set, informed by the above clones.

For clones that passed the initial test, we ran a more thorough date randomization test as described previously (71). The estimated median substitution rates and most recent common ancestor dates for randomized BEAST runs ($n = 10$) did not overlap with those of the runs using real collection dates, indicating a significant temporal signal. To test whether each epidemic clone had undergone a historical population expansion, we analyzed Bayesian skyline plot estimates of relative genetic diversity across the posterior distribution.

The association index was computed to find evidence of geographic clustering within clone phylogenies (72). We identified clones for further spatiotemporal analysis in which fewer than 5% of randomizations had a higher association index than the nonpermuted dataset. Asymmetric phylogeographic discrete trait reconstructions of the isolate continents were

then performed using the BEAST classic 1.9.0 package of BEAST 2.6.6 (29). Subsampling to account for overrepresentation of certain continents was repeated five times, and results were compared between subsamples. Spread 0.9.7.1 (73) was used to identify candidate migration routes between continents (Bayes factor ≥ 3).

Pan-genome analysis

Genomes were assembled from short-read data, and Panaroo 1.2.8 (39) was used to cluster the gene sequences from all samples into gene families and to infer a graphical pan-genome, which was reduced, ordered against the *P. aeruginosa* PAO1 genome, pruned of long-range connections, and then visualized (see the supplementary materials and methods for details). Parsimony ancestral character state reconstruction was then used to infer gene gains and losses on the branches of the rooted tree leading to the ancestral epidemic and sporadic clones. Gene functions were annotated using EggNog-mapper 2.1.6 (74), with the number of genes gained and annotated within a specific COG functional category compared using a Fisher's exact test (adjusted P value < 0.05).

Macrophage infection experiments

Isogenic F508del homozygous THP1 cells were created from WT THP1 cells (obtained from the American Type Culture Collection) using CRISPR-Cas9 editing and confirmed by Sanger sequencing (see the supplementary materials and methods). WT and F508del THP-1 monocytes were cultured, seeded at 200,000 cells/ml, and differentiated into macrophages as previously described (75) before being exposed to pooled clinical isolates of *P. aeruginosa* at a multiplicity of infection (MOI) of 1:1. The cells were then incubated at 37°C for 1 hour before the supernatant was removed and cells were either lysed at the 1-hour time point or incubated in fresh medium for a further 2 or 4 hours before supernatant removal, cell lysis, and DNA extraction and sequencing. Strain abundance was quantified using the mSWEET 1.4.0 sequence-based deconvolution method (76). Strains with <1% abundance at the 1-hour time point were excluded from the analysis.

Transcriptomic analysis

Gene expression data for clinical *P. aeruginosa* strains and the UCBPP-PA14 WT control strain were obtained (25) and pseudoaligned to strain-specific gene indices to produce abundance estimates using Kallisto (77). Length-scaled abundance estimates were size-factor normalized by the median ratio method and modeled as a response to CF proportion per genomic cluster (as defined by the number of CF versus non-CF patients and environmental samples) using a negative binomial generalized linear model with DESeq2 (78). The coefficients for

gene models were assessed using the Wald test (adjusted P value < 0.05). To assess the distribution and clustering of transcriptional diversity of strains with respect to CF proportion, we used k-means clustering ($k = 20$) on the principal components of the gene expression data and then calculated the mean SD (σ) of the CF proportion by cluster (mean $\sigma = 0.135$). A permutation test was used to assess significance (see the supplementary materials and methods for details)

Zebrafish infection models

The following zebrafish lines were used (see the supplementary materials and methods for details): the WT AB line, a knockout *cfr* sh540 mutant (46), and the Tg(mpeg1:mcherry-F)ump2 line (47). The morpholinos for *cfr* knockdown (5'-GACACATTGAGACTCACACAA-3') were prepared and injected into the one-cell stage as previously described (79). Systemic infections were achieved by microinjection of green fluorescent protein (GFP)-expressing *P. aeruginosa* strains into the caudal vein of zebrafish embryos 30 hours after fertilization as previously described (80), with survival after infection assessed daily and viable *in vivo* *P. aeruginosa* quantified by colony-forming units (CFUs) at 1 day after infection. Macrophage responses were examined by intramuscular injection of anesthetized Tg(mpeg1:mcherry-F)ump2 larvae at 3 days after fertilization with GFP-expressing fluorescent *P. aeruginosa* as previously described (79, 81). Macrophage chemotaxis, phagocytosis, and intracellular *P. aeruginosa* burden were quantified by confocal microscopy (see the supplementary materials and methods for details).

Mutational burden analysis

Treetime 0.8.1 (82) was used to reconstruct ancestral character states of every nucleotide position in every clone. A pipeline (83) was then implemented to identify single-nucleotide changes and to annotate variant effect in their phylogenetic context using the gene annotation from <https://pseudomonas.com/> (PAO1 107) and ancestral character state reconstructions (55). Parsimony ancestral character state reconstruction was used to infer ancestral insertions and deletions using SNPeff (48) for variant effect annotation.

We assessed the mutational burden of every gene on the basis of the number of non-synonymous variants across all clones (using a Poisson test, adjusted P value < 0.05). The 224 genes passing the adjusted P -value threshold were used to query the STRING 11.5 database (56) of protein-protein interactions. Pathoadaptive genes were assigned to 17 functional categories on the basis of the gene products description on <https://pseudomonas.com/> (55) (fig. S1). A Fisher's exact test was used to compare the number of assigned with the number

of unassigned genes among pathoadaptive genes and nonhits.

Impact of amino acid changes on protein stability and structural analysis

All amino acid changes were analyzed with SIFT 4G 6.2.1 (49) and FoldX 5 (50) (see the supplementary materials and methods for details). A two-tailed t test was used to compare the averaged scores per gene and protein scores between mutational burden test hits and nonhits. Mutational frequencies were mapped on the structural models of the identified hotspot genes in *P. aeruginosa* using the Chimera molecular modeling package (84). Models were downloaded from the Protein Data Bank and UniProt (85).

Phenotyping of pathoadaptive gene mutations

PAO1 mutants with transposon insertions in 154 pathoadaptive genes selected from the Manoil library (86) were arrayed in 96-well plates and imaged using the Phenobooth Imager (Singer Instruments) to quantify the following phenotypic traits: swimming motility, twitching motility, siderophore production, caseinase activity, gelatinase activity, and rhamnolipid production (see the supplementary materials and methods for details).

To assess the association between genetic variants and the expression of transcription factor regulons, gene expression data from (25) were pseudoaligned to strain-specific gene sets, and the normalized expression levels of tissue factor regulons were compared between strains with and without genetic variants using Welch's two-sample t tests (adjusted P value < 0.05) (see the supplementary materials and methods for details).

Transmission and host selectivity of pathoadaptive mutations

To assess the transmissibility of pathoadaptive changes, the number of mutations that had been observed in at least two isolates from different patients was compared with hitherto untransmitted mutations using a Fisher's exact test (adjusted P value < 0.1). TopGO 2.4.6 was used for functional enrichment analysis of the host-specific Gene Ontology biological pathway annotation compared with background (57) using annotations from <https://pseudomonas.com/> (55) (P value < 0.05).

Mutations in pathoadaptive genes were stratified by the (ancestral) infection type (CF or non-CF) of every branch based on outgroup-rooted clone trees. To assess host-specific pathoadaptation, the number of CF versus non-CF mutations were compared using Fisher's exact test (adjusted P value < 0.1). Mutations on branches with nonconcordant ancestral infection types were discarded.

Trajectories were inferred as the sequence of mutations in pathoadaptive genes since the

emergence of the clone ancestor as implied by the PAO1-rooted tree stratified by CF and non-CF infection types. Mutation frequencies were position normalized, and the frequency plots of the 40 genes with the lowest *P* value from the mutational-burden test were manually assigned into five groups of genes with similar frequency-curve shapes. Trendlines were generated by locally weighted smoothing.

A relatedness cutoff was established to define potential transmission links using pairwise SNP differences between pairs of isolate genomes from the same patient (*n* = 81 patients). Potential transmission events were then identified as isolates from the same clone sampled from different patients that differed by 26 SNPs or fewer and visualized using Cytoscape.

REFERENCES AND NOTES

- J. L. Whitby, A. Rampling, *Pseudomonas aeruginosa* contamination in domestic and hospital environments. *Lancet* **1**, 15–17 (1972). doi: [10.1016/S0140-6736\(72\)90006-2](https://doi.org/10.1016/S0140-6736(72)90006-2); pmid: [4108816](https://pubmed.ncbi.nlm.nih.gov/4108816/)
- S. K. Green, M. N. Schroth, J. J. Cho, S. K. Komino, V. B. Vitanza-jack, Agricultural plants and soil as a reservoir for *Pseudomonas aeruginosa*. *Appl. Microbiol.* **28**, 987–991 (1974). doi: [10.1128/am.28.6.987-991.1974](https://doi.org/10.1128/am.28.6.987-991.1974); pmid: [4217591](https://pubmed.ncbi.nlm.nih.gov/4217591/)
- J.-P. Pirnay et al., Global *Pseudomonas aeruginosa* biodiversity as reflected in a Belgian river. *Environ. Microbiol.* **7**, 969–980 (2005). doi: [10.1111/j.1462-2920.2005.00776.x](https://doi.org/10.1111/j.1462-2920.2005.00776.x); pmid: [15946293](https://pubmed.ncbi.nlm.nih.gov/15946293/)
- L. M. Ringen, C. H. Drake, A study of the incidence of *Pseudomonas aeruginosa* from various natural sources. *J. Bacteriol.* **64**, 841–845 (1952). doi: [10.1128/jb.64.6.841-845.1952](https://doi.org/10.1128/jb.64.6.841-845.1952); pmid: [1301158](https://pubmed.ncbi.nlm.nih.gov/1301158/)
- S. Crone et al., The environmental occurrence of *Pseudomonas aeruginosa*. *APMIS* **128**, 220–231 (2020). doi: [10.1111/apm.13010](https://doi.org/10.1111/apm.13010); pmid: [31709616](https://pubmed.ncbi.nlm.nih.gov/31709616/)
- European Centre for Disease Prevention and Control, *Healthcare-Associated Infections Acquired in Intensive Care Units. Annual Epidemiological Report for 2017* (ECDC, 2019).
- T. F. Murphy et al., *Pseudomonas aeruginosa* in chronic obstructive pulmonary disease. *Am. J. Respir. Crit. Care Med.* **177**, 853–860 (2008). doi: [10.1164/rccm.200709-1413OC](https://doi.org/10.1164/rccm.200709-1413OC); pmid: [18202344](https://pubmed.ncbi.nlm.nih.gov/18202344/)
- S. Rajan, L. Saiman, Pulmonary infections in patients with cystic fibrosis. *Semin. Respir. Infect.* **17**, 47–56 (2002). doi: [10.1053/srin.2002.31690](https://doi.org/10.1053/srin.2002.31690); pmid: [11891518](https://pubmed.ncbi.nlm.nih.gov/11891518/)
- S. Finch, M. J. McDonnell, H. Abo-Leyah, S. Alberti, J. D. Chalmers, A Comprehensive analysis of the impact of *Pseudomonas aeruginosa* colonization on prognosis in adult bronchiectasis. *Ann. Am. Thorac. Soc.* **12**, 1602–1611 (2015). pmid: [26356317](https://pubmed.ncbi.nlm.nih.gov/26356317/)
- M. R. Kosorok et al., Acceleration of lung disease in children with cystic fibrosis after *Pseudomonas aeruginosa* acquisition. *Pediatr. Pulmonol.* **32**, 277–287 (2001). doi: [10.1002/ppul.2009](https://doi.org/10.1002/ppul.2009); pmid: [11568988](https://pubmed.ncbi.nlm.nih.gov/11568988/)
- E. Tacconelli et al., Discovery, research, and development of new antibiotics: The WHO priority list of antibiotic-resistant bacteria and tuberculosis. *Lancet Infect. Dis.* **18**, 318–327 (2018). doi: [10.1016/S1473-3099\(17\)30753-3](https://doi.org/10.1016/S1473-3099(17)30753-3); pmid: [29276051](https://pubmed.ncbi.nlm.nih.gov/29276051/)
- GBD 2019 Antimicrobial Resistance Collaborators, Global mortality associated with 33 bacterial pathogens in 2019: A systematic analysis for the Global Burden of Disease Study 2019. *Lancet* **400**, 2221–2248 (2022). doi: [10.1016/S0140-6736\(22\)01285-7](https://doi.org/10.1016/S0140-6736(22)01285-7); pmid: [36423648](https://pubmed.ncbi.nlm.nih.gov/36423648/)
- P. J. Mitchelmore et al., Molecular epidemiology of *Pseudomonas aeruginosa* in an unsegregated bronchiectasis cohort sharing hospital facilities with a cystic fibrosis cohort. *Thorax* **73**, 677–679 (2017). doi: [10.1136/thoraxjnl-2016-209889](https://doi.org/10.1136/thoraxjnl-2016-209889); pmid: [28844058](https://pubmed.ncbi.nlm.nih.gov/28844058/)
- P. J. Stapleton et al., *Pseudomonas aeruginosa* strain sharing in early infection among children with cystic fibrosis. *Clin. Infect. Dis.* **73**, e2521–e2528 (2021). doi: [10.1093/cid/ciaa788](https://doi.org/10.1093/cid/ciaa788); pmid: [32544950](https://pubmed.ncbi.nlm.nih.gov/32544950/)
- G. A. Tramper-Stranders et al., *Pseudomonas aeruginosa* diversity in distinct paediatric patient groups. *Clin. Microbiol. Infect.* **14**, 935–941 (2008). doi: [10.1111/j.1469-0961.2008.02054.x](https://doi.org/10.1111/j.1469-0961.2008.02054.x); pmid: [18752596](https://pubmed.ncbi.nlm.nih.gov/18752596/)
- T. E. Woo et al., Epidemiology and natural history of *Pseudomonas aeruginosa* airway infections in non-cystic fibrosis bronchiectasis. *ERJ Open Res.* **4**, 00162-2017 (2018). doi: [10.1183/23120541.00162-2017](https://doi.org/10.1183/23120541.00162-2017); pmid: [29930949](https://pubmed.ncbi.nlm.nih.gov/29930949/)
- A. AbdulWahab et al., Genetic relatedness and host specificity of *Pseudomonas aeruginosa* isolates from cystic fibrosis and non-cystic fibrosis patients. *Infect. Drug Resist.* **7**, 309–316 (2014). pmid: [25429232](https://pubmed.ncbi.nlm.nih.gov/25429232/)
- V. N. Kos et al., The resistome of *Pseudomonas aeruginosa* in relationship to phenotypic susceptibility. *Antimicrob. Agents Chemother.* **59**, 427–436 (2015). doi: [10.1128/AAC.03954-14](https://doi.org/10.1128/AAC.03954-14); pmid: [25367914](https://pubmed.ncbi.nlm.nih.gov/25367914/)
- M. P. Moore et al., Transmission, adaptation and geographical spread of the *Pseudomonas aeruginosa* Liverpool epidemic strain. *Microb. Genom.* **7**, mgen000511 (2021). doi: [10.1099/mgen.0.000511](https://doi.org/10.1099/mgen.0.000511); pmid: [33720817](https://pubmed.ncbi.nlm.nih.gov/33720817/)
- C. López-Causapi et al., Evolution of the *Pseudomonas aeruginosa* mutational resistome in an international Cystic Fibrosis clone. *Sci. Rep.* **7**, 5555 (2017). doi: [10.1038/s41598-017-05621-5](https://doi.org/10.1038/s41598-017-05621-5); pmid: [28717172](https://pubmed.ncbi.nlm.nih.gov/28717172/)
- L. Freschi et al., The *Pseudomonas aeruginosa* pan-genome provides new insights on its population structure, horizontal gene transfer and pathogenicity. *Genom. Biol. Evol.* **11**, 109–120 (2019). doi: [10.1093/gbe/evy259](https://doi.org/10.1093/gbe/evy259); pmid: [30496396](https://pubmed.ncbi.nlm.nih.gov/30496396/)
- Y. Hilliam et al., *Pseudomonas aeruginosa* adaptation and diversification in the non-cystic fibrosis bronchiectasis lung. *Eur. Respir. J.* **49**, 1602018 (2017). doi: [10.1183/13993003.02108-2016](https://doi.org/10.1183/13993003.02108-2016); pmid: [28446558](https://pubmed.ncbi.nlm.nih.gov/28446558/)
- J. Chilam et al., Genomic surveillance of *Pseudomonas aeruginosa* in the Philippines, 2013–2014. *Western Pac. Surveill. Response J.* **12**, 4–18 (2021). doi: [10.5365/wpsar.2020.11.1.006](https://doi.org/10.5365/wpsar.2020.11.1.006); pmid: [34540307](https://pubmed.ncbi.nlm.nih.gov/34540307/)
- E. del Barrio-Tofiño et al., Genomics and susceptibility profiles of extensively drug-resistant *Pseudomonas aeruginosa* isolates from Spain. *Antimicrob. Agents Chemother.* **61**, e01589-17 (2017). doi: [10.1128/AAC.01589-17](https://doi.org/10.1128/AAC.01589-17); pmid: [28874376](https://pubmed.ncbi.nlm.nih.gov/28874376/)
- A. Khaledi et al., Predicting antimicrobial resistance in *Pseudomonas aeruginosa* with machine learning-enabled molecular diagnostics. *EMBO Mol. Med.* **12**, e10264 (2020). doi: [10.15252/emmm.201910264](https://doi.org/10.15252/emmm.201910264); pmid: [32048461](https://pubmed.ncbi.nlm.nih.gov/32048461/)
- R. L. Marvig, L. M. Sommer, S. Molin, H. K. Johansen, Convergent evolution and adaptation of *Pseudomonas aeruginosa* within patients with cystic fibrosis. *Nat. Genet.* **47**, 57–64 (2015). doi: [10.1038/ng.3148](https://doi.org/10.1038/ng.3148); pmid: [25401299](https://pubmed.ncbi.nlm.nih.gov/25401299/)
- J. F. Turton et al., High-resolution analysis by whole-genome sequencing of an international lineage, (sequence type 11) of *Pseudomonas aeruginosa* associated with metallo-carbapenemases in the United Kingdom. *J. Clin. Microbiol.* **53**, 2622–2631 (2015). doi: [10.1128/JCM.00505-15](https://doi.org/10.1128/JCM.00505-15); pmid: [26041902](https://pubmed.ncbi.nlm.nih.gov/26041902/)
- B. Curran, D. Jonas, H. Grundmann, T. Pitt, C. G. Dowson, Development of a multilocus sequence typing scheme for the opportunistic pathogen *Pseudomonas aeruginosa*. *J. Clin. Microbiol.* **42**, 5644–5649 (2004). doi: [10.1128/JCM.42.12.5644-5649.2004](https://doi.org/10.1128/JCM.42.12.5644-5649.2004); pmid: [15583294](https://pubmed.ncbi.nlm.nih.gov/15583294/)
- R. Bouckaert et al., BEAST 2: A software platform for Bayesian evolutionary analysis. *PLOS Comput. Biol.* **10**, e1003537 (2014). doi: [10.1371/journal.pcbi.1003537](https://doi.org/10.1371/journal.pcbi.1003537); pmid: [24722319](https://pubmed.ncbi.nlm.nih.gov/24722319/)
- A. J. Drummond, A. Rambaut, B. Shapiro, O. G. Pybus, Bayesian coalescent inference of past population dynamics from molecular sequences. *Mol. Biol. Evol.* **22**, 1185–1192 (2005). doi: [10.1093/molbev/msi103](https://doi.org/10.1093/molbev/msi103); pmid: [15703244](https://pubmed.ncbi.nlm.nih.gov/15703244/)
- W. H. McNeill, Human migration in historical perspective. *Popul. Dev. Rev.* **10**, 1–18 (1984). doi: [10.2307/1973159](https://doi.org/10.2307/1973159)
- A. P. Dobson, E. R. Carper, Infectious diseases and human population history. *Bioscience* **46**, 115–126 (1996). doi: [10.2307/1312814](https://doi.org/10.2307/1312814)
- F. J. Kelly, J. C. Fussell, Air pollution and airway disease. *Clin. Exp. Allergy* **41**, 1059–1071 (2011). doi: [10.1111/j.1365-2222.2011.03776.x](https://doi.org/10.1111/j.1365-2222.2011.03776.x); pmid: [21623970](https://pubmed.ncbi.nlm.nih.gov/21623970/)
- A. J. Chauhan, S. L. Johnston, Air pollution and infection in respiratory illness. *Br. Med. Bull.* **68**, 95–112 (2003). doi: [10.1093/bmb/lgd022](https://doi.org/10.1093/bmb/lgd022); pmid: [14757711](https://pubmed.ncbi.nlm.nih.gov/14757711/)
- M. Kampa, E. Castanas, Human health effects of air pollution. *Environ. Pollut.* **151**, 362–367 (2008). doi: [10.1016/j.envpol.2007.06.012](https://doi.org/10.1016/j.envpol.2007.06.012); pmid: [17646040](https://pubmed.ncbi.nlm.nih.gov/17646040/)
- P. Lemey, A. Rambaut, A. J. Drummond, M. A. Suchard, Bayesian phylogeography finds its roots. *PLOS Comput. Biol.* **5**, e1000520 (2009). doi: [10.1371/journal.pcbi.1000520](https://doi.org/10.1371/journal.pcbi.1000520); pmid: [19779555](https://pubmed.ncbi.nlm.nih.gov/19779555/)
- M. I. Katsnelson, Y. I. Wolf, E. V. Koonin, On the feasibility of saltational evolution. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 21068–21075 (2019). doi: [10.1073/pnas.1909031116](https://doi.org/10.1073/pnas.1909031116); pmid: [31570621](https://pubmed.ncbi.nlm.nih.gov/31570621/)
- J. M. Bryant et al., Stepwise pathogenic evolution of *Mycobacterium abscessus*. *Science* **372**, eabb8699 (2021). doi: [10.1126/science.abb8699](https://doi.org/10.1126/science.abb8699); pmid: [33926925](https://pubmed.ncbi.nlm.nih.gov/33926925/)
- G. Tonkin-Hill et al., Producing polished prokaryotic pangenomes with the Panaroo pipeline. *Genome Biol.* **21**, 180 (2020). doi: [10.1186/s13059-020-02090-4](https://doi.org/10.1186/s13059-020-02090-4); pmid: [32698896](https://pubmed.ncbi.nlm.nih.gov/32698896/)
- K. Martin et al., Clusters of genetically similar isolates of *Pseudomonas aeruginosa* from multiple hospitals in the UK. *J. Med. Microbiol.* **62**, 988–1000 (2013). doi: [10.1099/jmm.0.054841-0](https://doi.org/10.1099/jmm.0.054841-0); pmid: [23558134](https://pubmed.ncbi.nlm.nih.gov/23558134/)
- J. Day, A. Friedman, L. S. Schlesinger, Modeling the immune rheostat of macrophages in the lung in response to infection. *Proc. Natl. Acad. Sci. U.S.A.* **106**, 11246–11251 (2009). doi: [10.1073/pnas.0904846106](https://doi.org/10.1073/pnas.0904846106); pmid: [19549875](https://pubmed.ncbi.nlm.nih.gov/19549875/)
- A. Craig, J. Mai, S. Cai, S. Jayaselvan, Neutrophil recruitment to the lungs during bacterial pneumonia. *Infect. Immun.* **77**, 568–575 (2009). doi: [10.1128/IAI.00832-08](https://doi.org/10.1128/IAI.00832-08); pmid: [19015252](https://pubmed.ncbi.nlm.nih.gov/19015252/)
- A. Fortuna et al., The *Pseudomonas aeruginosa* DksA1 protein is involved in H₂O₂ tolerance and within-macrophages survival and can be replaced by DksA2. *Sci. Rep.* **12**, 10404 (2022). doi: [10.1038/s41598-022-14635-7](https://doi.org/10.1038/s41598-022-14635-7); pmid: [35729352](https://pubmed.ncbi.nlm.nih.gov/35729352/)
- A. Fortuna, H. Bähré, P. Visca, G. Rampioni, L. Leoni, The two *Pseudomonas aeruginosa* DksA stringent response proteins are largely interchangeable at the whole transcriptome level and in the control of virulence-related traits. *Environ. Microbiol.* **23**, 5487–5504 (2021). doi: [10.1111/1462-2920.15693](https://doi.org/10.1111/1462-2920.15693); pmid: [34327807](https://pubmed.ncbi.nlm.nih.gov/34327807/)
- K. B. Turton, R. J. Ingram, M. A. Valvano, Macrophage dysfunction in cystic fibrosis: Nature or nurture? *J. Leukoc. Biol.* **109**, 573–582 (2021). doi: [10.1002/JLB.B4RU0620-245R](https://doi.org/10.1002/JLB.B4RU0620-245R); pmid: [32678926](https://pubmed.ncbi.nlm.nih.gov/32678926/)
- A. Bernut, C. A. Loynes, R. A. Floto, S. A. Renshaw, Deletion of *cfr* Leads to an Excessive Neutrophilic Response and Defective Tissue Repair in a Zebrafish Model of Sterile Inflammation. *Front. Immunol.* **11**, 1733 (2020). doi: [10.3389/fimmu.2020.01733](https://doi.org/10.3389/fimmu.2020.01733); pmid: [32849617](https://pubmed.ncbi.nlm.nih.gov/32849617/)
- A. Bernut et al., *Mycobacterium abscessus* cording prevents phagocytosis and promotes abscess formation. *Proc. Natl. Acad. Sci. U.S.A.* **111**, E943–E952 (2014). doi: [10.1073/pnas.1321390111](https://doi.org/10.1073/pnas.1321390111); pmid: [24567393](https://pubmed.ncbi.nlm.nih.gov/24567393/)
- P. Cingolani et al., A program for annotating and predicting the effects of single nucleotide polymorphisms. *SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w118; iso-2; iso-3. Fly* **6**, 80–92 (2012). doi: [10.4161/fly.19695](https://doi.org/10.4161/fly.19695); pmid: [22728672](https://pubmed.ncbi.nlm.nih.gov/22728672/)
- R. Vaser, S. Adusumalli, S. N. Leng, M. Sikic, P. C. Ng, SIFT missense predictions for genomes. *Nat. Protoc.* **11**, 1–9 (2016). doi: [10.1038/nprot.2015.123](https://doi.org/10.1038/nprot.2015.123); pmid: [26633127](https://pubmed.ncbi.nlm.nih.gov/26633127/)
- J. Delgado, L. G. Radusky, D. Cianferoni, L. Serrano, FoldX 5.0: Working with RNA, small molecules and a new graphical interface. *Bioinformatics* **35**, 4168–4169 (2019). doi: [10.1093/bioinformatics/btz184](https://doi.org/10.1093/bioinformatics/btz184); pmid: [30874800](https://pubmed.ncbi.nlm.nih.gov/30874800/)
- S. Westbrook-Wadman et al., Characterization of a *Pseudomonas aeruginosa* efflux pump contributing to aminoglycoside impermeability. *Antimicrob. Agents Chemother.* **43**, 2975–2983 (1999). doi: [10.1128/AAC.43.12.2975](https://doi.org/10.1128/AAC.43.12.2975); pmid: [10582892](https://pubmed.ncbi.nlm.nih.gov/10582892/)
- K. Poole et al., Expression of the multidrug resistance operon *mexA-mexB-oprM* in *Pseudomonas aeruginosa*: *mexR* encodes a regulator of operon expression. *Antimicrob. Agents Chemother.* **40**, 2021–2028 (1996). doi: [10.1128/AAC.40.9.2021](https://doi.org/10.1128/AAC.40.9.2021); pmid: [8878574](https://pubmed.ncbi.nlm.nih.gov/8878574/)
- L. Rust, E. C. Pesci, B. H. Iglesias, Analysis of the *Pseudomonas aeruginosa* elastase (las) regulatory region. *J. Bacteriol.* **178**, 1134–1140 (1996). doi: [10.1128/JB.178.4.1134-1140.1996](https://doi.org/10.1128/JB.178.4.1134-1140.1996); pmid: [8576049](https://pubmed.ncbi.nlm.nih.gov/8576049/)
- T. L. Yahr, D. W. Frank, Transcriptional organization of the trans-regulatory locus which controls exoenzyme S synthesis in *Pseudomonas aeruginosa*. *J. Bacteriol.* **176**, 3832–3838 (1994). doi: [10.1128/JB.176.13.3832-3838.1994](https://doi.org/10.1128/JB.176.13.3832-3838.1994); pmid: [8021164](https://pubmed.ncbi.nlm.nih.gov/8021164/)
- G. L. Winsor et al., Enhanced annotations and features for comparing thousands of *Pseudomonas* genomes in the *Pseudomonas* genome database. *Nucleic Acids Res.* **44**, D646–D653 (2016). doi: [10.1093/nar/gkv1227](https://doi.org/10.1093/nar/gkv1227); pmid: [26578582](https://pubmed.ncbi.nlm.nih.gov/26578582/)
- D. Szklarczyk et al., STRING v11: Protein-protein association networks with increased coverage, supporting functional discovery in genome-wide experimental datasets. *Nucleic Acids Res.* **47**, D607–D613 (2019). doi: [10.1093/nar/gky131](https://doi.org/10.1093/nar/gky131); pmid: [30476243](https://pubmed.ncbi.nlm.nih.gov/30476243/)
- A. Alexa, J. Rahnenführer, T. Langbeuer, Improved scoring of functional groups from gene expression data by decorrelating GO graph structure. *Bioinformatics* **22**, 1600–1607 (2006). doi: [10.1093/bioinformatics/bti140](https://doi.org/10.1093/bioinformatics/bti140); pmid: [16606683](https://pubmed.ncbi.nlm.nih.gov/16606683/)
- D. W. Martin, M. J. Schurr, H. Yu, V. Deretic, Analysis of promoters controlled by the putative sigma factor AlgU

- regulating conversion to mucoidy in *Pseudomonas aeruginosa*: Relationship to sigma E and stress response. *J. Bacteriol.* **176**, 6688–6696 (1994). doi: [10.1128/JB.176.21.6688-6696.1994](https://doi.org/10.1128/JB.176.21.6688-6696.1994); pmid: [7961422](#)
59. U. N. Broder, T. Jaeger, U. Jenal. LacS is a calcium-responsive kinase that induces acute-to-chronic virulence switch in *Pseudomonas aeruginosa*. *Nat. Microbiol.* **2**, 16184 (2016). doi: [10.1038/nmicrobiol.2016.184](https://doi.org/10.1038/nmicrobiol.2016.184); pmid: [27775685](#)
60. C. Chen, G. A. Beattie, *Pseudomonas syringae* BetT is a low-affinity choline transporter that is responsible for superior osmoprotection by choline over glycine betaine. *J. Bacteriol.* **190**, 2717–2725 (2008). doi: [10.1128/JB.01585-07](https://doi.org/10.1128/JB.01585-07); pmid: [18156257](#)
61. I. Santi, P. Manfredi, E. Maffei, A. Egli, U. Jenal. Evolution of antibiotic tolerance shapes resistance development in chronic *Pseudomonas aeruginosa* infections. *mBio* **12**, e03482-20 (2021). doi: [10.1128/mBio.03482-20](https://doi.org/10.1128/mBio.03482-20); pmid: [33563834](#)
62. R. Reynolds, R. Hope, L. Williams; BSAC Working Parties on Resistance Surveillance, Survey, laboratory and statistical methods for the BSAC Resistance Surveillance Programmes. *J. Antimicrob. Chemother.* **62**, ii15–ii28 (2008). doi: [10.1093/jac/dkn349](https://doi.org/10.1093/jac/dkn349); pmid: [18819976](#)
63. H. Li. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM. [arXiv:1303.3997](https://arxiv.org/abs/1303.3997) [q-bio.GN] (2013).
64. M. Hunt et al., ARIBA: Rapid antimicrobial resistance genotyping directly from sequencing reads. *Microb. Genom.* **3**, e000131 (2017). doi: [10.1099/mgen.0.000131](https://doi.org/10.1099/mgen.0.000131); pmid: [29170789](#)
65. M. N. Price, P. S. Dehal, A. P. Arkin, FastTree 2—Approximately maximum-likelihood trees for large alignments. *PLoS ONE* **5**, e9490 (2010). doi: [10.1371/journal.pone.0009490](https://doi.org/10.1371/journal.pone.0009490); pmid: [20224823](#)
66. A. J. Page et al., SNP-sites: Rapid efficient extraction of SNPs from multi-FASTA alignments. *Microb. Genom.* **2**, e000056 (2016). doi: [10.1099/mgen.0.000056](https://doi.org/10.1099/mgen.0.000056); pmid: [28348851](#)
67. N. J. Croucher et al., Rapid phylogenetic analysis of large samples of recombinant bacterial whole genome sequences using Gubbins. *Nucleic Acids Res.* **43**, e15 (2015). doi: [10.1093/nar/gku196](https://doi.org/10.1093/nar/gku196); pmid: [25414349](#)
68. A. Rambaut, T. L. Lam, L. Max Carvalho, O. G. Pybus, Exploring the temporal structure of heterochronous sequences using TempEst (formerly Path-O-Gen). *Virus Evol.* **2**, vew007 (2016). doi: [10.1093/ve/vew007](https://doi.org/10.1093/ve/vew007); pmid: [27774300](#)
69. A. Stamatakis, RAxML version 8: A tool for phylogenetic analysis and post-analysis of large phylogenies. *Bioinformatics* **30**, 1312–1313 (2014). doi: [10.1093/bioinformatics/btu033](https://doi.org/10.1093/bioinformatics/btu033); pmid: [24451623](#)
70. A. Rambaut, A. J. Drummond, D. Xie, G. Baele, M. A. Suchard, Posterior summarization in Bayesian phylogenetics using Tracer 1.7. *Syst. Biol.* **67**, 901–904 (2018). doi: [10.1093/sysbio/syy032](https://doi.org/10.1093/sysbio/syy032); pmid: [29718447](#)
71. F. Menardo, S. Duchêne, D. Brites, S. Gagneux. The molecular clock of Mycobacterium tuberculosis. *PLOS Pathog.* **15**, e1008067 (2019). doi: [10.1371/journal.ppat.1008067](https://doi.org/10.1371/journal.ppat.1008067); pmid: [31513651](#)
72. J. Parker, A. Rambaut, O. G. Pybus, Correlating viral phenotypes with phylogeny: Accounting for phylogenetic uncertainty. *Infect. Genet. Evol.* **8**, 239–246 (2008). doi: [10.1016/j.meegid.2007.08.001](https://doi.org/10.1016/j.meegid.2007.08.001); pmid: [17921073](#)
73. F. Bielejec, A. Rambaut, M. A. Suchard, P. Lemey, SPREAD: Spatial phylogenetic reconstruction of evolutionary dynamics. *Bioinformatics* **27**, 2910–2912 (2011). doi: [10.1093/bioinformatics/btr481](https://doi.org/10.1093/bioinformatics/btr481); pmid: [21911333](#)
74. J. Huerta-Cepas et al., Fast Genome-Wide Functional Annotation through Orthology Assignment by eggNOG-Mapper. *Mol. Biol. Evol.* **34**, 2115–2122 (2017). doi: [10.1093/molbev/msw148](https://doi.org/10.1093/molbev/msw148); pmid: [28460117](#)
75. J. M. Bryant et al., Emergence and spread of a human-transmissible multidrug-resistant nontuberculous mycobacterium. *Science* **354**, 751–757 (2016). doi: [10.1126/science.aaf8156](https://doi.org/10.1126/science.aaf8156); pmid: [27846606](#)
76. T. Mäklin et al., High-resolution sweep metagenomics using fast probabilistic inference. *Wellcome Open Res.* **5**, 14 (2021). doi: [10.12688/wellcomeopenres.156391](https://doi.org/10.12688/wellcomeopenres.156391); pmid: [34746439](#)
77. N. L. Bray, H. Pimentel, P. Melsted, L. Pachter, Near-optimal probabilistic RNA-seq quantification. *Nat. Biotechnol.* **34**, 525–527 (2016). doi: [10.1038/nbt.3519](https://doi.org/10.1038/nbt.3519); pmid: [27043002](#)
78. M. Love, S. Anders, W. Huber, Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **15**, 550 (2014). doi: [10.1186/s13059-014-0550-8](https://doi.org/10.1186/s13059-014-0550-8); pmid: [25516281](#)
79. A. Bernut et al., CFTR protects against *Mycobacterium abscessus* infection by fine-tuning host oxidative defenses. *Cell Rep.* **26**, 1828–1840.e4 (2019). doi: [10.1016/j.celrep.2019.01.071](https://doi.org/10.1016/j.celrep.2019.01.071); pmid: [30759393](#)
80. C. Belon et al., A macrophage subversion factor is shared by intracellular and extracellular pathogens. *PLOS Pathog.* **11**, e1004969 (2015). doi: [10.1371/journal.ppat.1004969](https://doi.org/10.1371/journal.ppat.1004969); pmid: [26080006](#)
81. V. Le Moigne et al., Serological biomarkers for the diagnosis of *Mycobacterium abscessus* infections in cystic fibrosis patients. *J. Cyst. Fibros.* **21**, 353–360 (2022). doi: [10.1016/j.jcf.2021.08.019](https://doi.org/10.1016/j.jcf.2021.08.019); pmid: [34511392](#)
82. P. Sagulenko, V. Puller, R. A. Neher, TreeTime: Maximum-likelihood phylodynamic analysis. *Virus Evol.* **4**, vex042 (2018). doi: [10.1093/ve/vex042](https://doi.org/10.1093/ve/vex042); pmid: [29340210](#)
83. A. Weimann, C. Ruis, "PhyloEffects," Zenodo (2024); doi: [10.5281/zenodo.10606175](https://doi.org/10.5281/zenodo.10606175)
84. E. F. Pettersen et al., UCSF Chimera—A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004). doi: [10.1002/jcc.20084](https://doi.org/10.1002/jcc.20084); pmid: [15264254](#)
85. B. E. Supek, H. Huang, P. McGarvey, R. Mazumder, C. H. Wu, UniRef: Comprehensive and non-redundant UniProt reference clusters. *Bioinformatics* **23**, 1282–1288 (2007). doi: [10.1093/bioinformatics/btm098](https://doi.org/10.1093/bioinformatics/btm098); pmid: [17379688](#)
86. K. Held, E. Ramage, M. Jacobs, L. Gallagher, C. Manoil, Sequence-verified two-allele transposon mutant library for *Pseudomonas aeruginosa* PAO1. *J. Bacteriol.* **194**, 6387–6389 (2012). doi: [10.1128/JB.01479-12](https://doi.org/10.1128/JB.01479-12); pmid: [22984262](#)
87. Code for: A. Weimann et al., Evolution and host-specific adaptation of *Pseudomonas aeruginosa*, Zenodo (2024); doi: [10.5281/zenodo.10625500](https://doi.org/10.5281/zenodo.10625500)
88. Data for: A. Weimann et al., Evolution and host-specific adaptation of *Pseudomonas aeruginosa*, Zenodo (2024); doi: [10.5281/zenodo.10600286](https://doi.org/10.5281/zenodo.10600286)
89. M. Ashburner et al., Gene ontology: Tool for the unification of biology. *Nat. Genet.* **25**, 25–29 (2000). doi: [10.1038/75556](https://doi.org/10.1038/75556); pmid: [10802651](#)

ACKNOWLEDGMENTS

We thank Sanger Pathogen Informatics for providing informatics support; H. Gardner and R. E. McLaughlin for sharing their unpublished short-read sequencing data of their previously published genomes ([18](#)); D. Aanensen and S. Häubler for sharing their genomic and transcriptomic data in the public domain ([23](#), [25](#)); BSAC and the laboratory staff at the clinical microbiology laboratory at Cambridge University Hospitals for their support in isolating the hospital strains; A.-B. Blanc-Potard for VLM support and the gift of pMF230; M. Nguyen-Chi for support with access to animal experimentation; S. Renshaw for support and helpful discussions; the ZEFIX-LPHI (ZEBrafish and *Xenopus* platform, Lphi, University of Montpellier) and the Bateson Center (University of Sheffield, Sheffield) Aquarium teams for zebrafish maintenance and care; the imaging facility MRI (Montpellier Ressources Imagerie) and E. Jublanc and V. Diakou for assistance; and the University of Montpellier and the University of Sheffield for support. **Funding:** This work was supported by The Wellcome Trust (grant 107032AIA to R.A.F., A.W., C.R., and A.M.D.; grant 226602/Z/22/Z to R.A.F. and A.W.; grant 10224/Z/15/Z to J.M.B.; grant 098051 to J.P.; and grant 220472/Z/20/Z to S.K.); the Botnar Foundation (grant 6063 to R.A.F., A.W., T.L.B., C.R., and J.P.); the UK Cystic Fibrosis Trust (Innovation Hub Grant 001 to R.A.F., T.L.B., J.P., A.W., C.R., and L.E.); the NIHR Cambridge Biomedical Research Centre (R.A.F. and K.B.); Health Enterprise East (to K.B., J.R., and R.A.F.); the Royal Papworth Hospital Research Innovation Fund (R.A.F.); the Canadian Institute of Health Research (grant 168006), Cystic Fibrosis Canada (grant 2610), and Genome Quebec and Genome Canada (all to R.C.L.); the Italian Ministry of University and Research (grant 314-337 Legge 232/2016 to G.R. and L.L.); the Italian Cystic Fibrosis Research Foundation (grant FFC 10/2023 to G.R.); Horizon 2020 Research (grant H2020-MSCA-IF-2016 to A.B.); the Marie Skłodowska-Curie Innovation Framework Program (CFZEBRA grant 751977 to A.B.); the FC3R through the ZFishforCFcare project (grant 22FC3R-017 to A.B.); and the French Cystic Fibrosis Foundation Vaincre La Mucoviscidose (VLM grant RF20220503060 to S.P.). **Author**

Contributions: A.W., J.P., and R.A.F. conceived the project. A.W., R.C.L., J.P., and R.A.F. designed the experiments and wrote the manuscript. A.W. performed the bioinformatic analyses. A.M.D. and A.W. performed the differential expression analysis. A.W. and C.R. performed the Bayesian phylodynamic analysis. A.P.P. and T.L.B. performed the computational structural modeling. L.E. and M.W. performed the virulence phenotyping assays. L.S. and P.T.H. made the THP1 F508 knock-in cell line. K.B., J.M.B., S.K., and A.W. performed the THP1 in vitro screen, sequencing, and analysis. E.U., K.B., J.R., J.M.B., and R.A.F. undertook the TeleCF study, sample collection, DNA extraction, and analysis. B.B., K.J., R.B., and S.J.P. conducted the bacteraemia sample collection and processing. G.R. and L.L. generated knockout bacterial lines and contributed to experimental design. A.B. and S.P. conceived and designed experiments and analyzed data for the in vivo testing in zebrafish models. S.P. generated fluorescent *P. aeruginosa* strains. N.R.T. provided feedback on the project outline and access to the Sanger computing facilities. J.G., J.L.F., and R.L. provided clinical isolates and contributed to analysis. R.A.F. and J.P. provided supervisory support. **Competing interests:** The authors declare no competing interests. National ethical approval for the TeleCF study (www.ClinicalTrials.gov number NCT01877707) was granted from the National Research Ethics Service (NRES) Committee of Hertfordshire, UK (REC 12/EE/0462). The study design was peer reviewed by the Cambridge Centre for Lung Infection (CCL) and the Cambridge Institute for Medical Research (CIMR). All zebrafish (*Danio rerio*) procedures described in the present study were performed by authorized staff and conducted by following the “3 Rs” (replacement, reduction, and refinement) principles in compliance with the European Union guidelines for handling of laboratory animals to local standards set approved by the UK Home Office under Animal Welfare and Ethical Review Body and by the Direction Sanitaire et Vétérinaire de l’Hérault et Comité d’Ethique pour l’Expérimentation Animale de la région Languedoc Roussillon and the French Ministry of Agriculture (authorization number APAFIS 36309-202204014222432). Breeding and maintenance of adult zebrafish were performed at the ZEFIX (Lphi, UMR5294, University of Montpellier, France) and the Bateson Centre (University of Sheffield, UK) Home Office-approved aquaria, according to the local animal welfare standards (license numbers CEEA-LR-B3417237 and P1A447A5E). The number of animals used for each procedure was guided by pilot experiments. **Data and materials availability:** Short-read DNA and RNA-seq data for the clinical isolates were downloaded from the European Nucleotide Archive (ENA). THP1 pooled infection assay DNA sequencing data were uploaded to the ENA under PRJEB20836 (see table S4). Short-read DNA sequencing data for newly sequenced isolates genomes (or where only assemblies were previously available) were uploaded to the ENA under PRJEB22089, IPC: PRJNA325248, AZ: PRJEB66158, LES: PRJEB69223, Bronchiectasis UK PRJEB69219. ENA run accessions can be found in table S1. The analysis codes and the PhyloEffects software were made available on GitHub and snapshots provided on Zenodo ([83](#), [87](#)). Intermediate and additional data were made available in a Zenodo repository ([88](#)). **License information:** Copyright © 2024 the authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original US government works. <https://www.science.org/about/science-licenses-journal-article-reuse>. In the interest of rapid dissemination of results with immediate public health relevance and because this research was funded in whole or in part by the Wellcome Trust (grant #107032AIA), a cOAition S organization, the author will make the Author Accepted Manuscript (AAM) version available under a CC BY public copyright license.

SUPPLEMENTARY MATERIALS

science.org/doi/10.1126/science.ad0908

Materials and Methods

Figs. S1 to S16

Tables S1 to S7

References ([90–105](#))

MDAR Reproducibility Checklist

Submitted 3 April 2023; resubmitted 4 December 2023

Accepted 2 May 2024

10.1126/science.ad0908