Selection of Relevant Bacterial Strains for Novel Therapeutic Testing: a Guidance Document for Priority Cystic Fibrosis Lung Pathogens

Eshwar Mahenthiralingam[1] · Rebecca Weiser[1] · R. Andres Floto[2,3,4] · Jane C. Davies[5] · Joanne L. Fothergill[6]

Abstract

Purpose of Review People with cystic fibrosis (CF) suffer chronic lung infections with a range of antimicrobial-resistant bacterial pathogens. There is an urgent need for researchers to develop novel anti-infectives to treat these problematic infections, but how can we select bacterial strains which are relevant for robust testing and comparative research?

Recent Findings *Pseudomonas aeruginosa*, *Burkholderia cepacia* complex and *Burkholderia gladioli*, *Mycobacterium abscessus* complex, *Staphylococcus aureus, Haemophilus influenza,* and several multidrug-resistant Gram-negative species were selected as key CF infections that urgently require new therapeutics. Reference isolates and strain panels were identified, and a summary of the known genotypic diversity of each pathogen was provided.

Summary Here, we summarise the current strain resources available for priority CF bacterial pathogens and highlight systematic selection criteria that researchers can use to select strains for use in therapeutic testing.

Keywords Cystic fibrosis · Bacterial pathogens · Strain resources · Selection criteria

Introduction

Lung infection in people with cystic fibrosis (CF) is polymicrobial and comprises a mixture of multiple bacteria, fungi, viruses, and other microorganisms [1, 2]. After conventional

✉ Eshwar Mahenthiralingam
MahenthiralingamE@cardiff.ac.uk

1 Microbiomes, Microbes and Informatics Group, Organisms and Environment Division, Cardiff School of Biosciences, Cardiff University, Sir Martin Evans Building, Park Place, Cardiff, UK

2 Molecular Immunity Unit, MRC Laboratory of Molecular Biology, Department of Medicine, University of Cambridge, Cambridge, UK

3 Centre for AI in Medicine, University of Cambridge, Cambridge, UK

4 Royal Papworth Hospital, Cambridge Centre for Lung Infection, Cambridge, UK

5 Department of Paediatric Respiratory Medicine, Royal Brompton Hospital, Sydney Street, London, UK

6 Department of Clinical Infection, Microbiology and Immunology, University of Liverpool, L69 7BE Liverpool, UK

diagnosis of infection using culture-based microbiology, only specific bacterial and fungal CF pathogens are treated using antibiotics or antifungals, respectively. Multiple CF pathogens are intrinsically resistant to antibiotics, and within the environment of the CF lung, they escape killing with these antimicrobials leading to the establishment of chronic infections. The continuous or repeated use of antibiotics to suppress chronic infection may lead to increased pathogen antimicrobial resistance (AMR), and frequently limits available therapeutic options for people with CF. In addition, pathogen adaption to this antimicrobial-rich lung environment promotes diversification and phenotypic variation. The microbial molecular epidemiology of CF lung infection has been comprehensively discussed by Lipuma [2, 3•] and is audited regularly by organisations such as the UK Cystic Fibrosis within their Data Registry [4].

But from the diverse groups of isolates associated with bacterial CF lung infections, how do researchers or companies developing therapeutics select appropriate strains for analysis?

Testing of novel anti-infective agents for the treatment of CF infections should consider using microbial strains that are well characterised in the context of the respiratory disease, as well as being representative of the genetic and phenotypic diversity of species being targeted. The definition of a bacterial strain has been discussed widely and can

now be defined at several levels of resolution based on whole genome sequences [5]. For the purpose of this review, a bacterial CF strain will be taken as an isolate that is distinct from other isolates of the same species based on a genotypic feature. This is important in relation to chronic CF lung infection because sequential isolates of the same strain are frequently collected over time, and hence will bias research if not accounted for. Here, we provide a summary of strain resources for problematic bacterial CF pathogens which are treated with antibiotics in the context of the respiratory disease (Fig. 1): *Pseudomonas aeruginosa*, *Burkholderia cepacia* complex (*Bcc*) and *Burkholderia gladioli*, *Mycobacterium abscessus* complex, multi-drug-resistant Gram-negative species (*Achromobacter*, *Stenotrophomonas* and others), *Staphylococcus aureus*, and *Haemophilus influenza*. From this knowledge, systematic criteria are also proposed to help future selection strains for research and testing.

Pseudomonas aeruginosa

P. aeruginosa is arguably the most widely studied CF pathogen and is also a World Health Organisation Priority 1 (Critical) bacterium in terms of AMR and an urgent need for new treatments [6]. *P. aeruginosa* continues to be a dominant cause of chronic CF lung infection, with a prevalence of greater than 50% in large populations such as the USA [3•] and UK [4]. Aggressive antibiotic eradication therapy has been successful at limiting *P. aeruginosa* infection in children, but chronic infection eventually occurs in adolescents and adults with CF [3•, 4]. Historically, *P. aeruginosa* strains which have been widely studied in terms of antimicrobial resistance and molecular pathogenesis are:

- *P. aeruginosa* PAO1 (ATCC 15,692), a human wound isolate and arguably the prototypic *P. aeruginosa* reference strain studied for the last 30 years [7••]
- *P. aeruginosa* UCBPP-PA14, a human burn isolate [7••]
- *P. aeruginosa* PAK, a non-CF human isolate [7••]

  However, none of the above were originally isolated from CF lung infections, representing a key knowledge gap in terms of therapeutic testing and modelling disease-appropriate *P. aeruginosa* strains. In the last decade, one CF strain has been studied in more detail, providing insights into the different behaviour of isolates from chronic lung infection compared to the widely studied reference strains:

- *P. aeruginosa* LES B58, a chronic lung infection isolate from an adult with CF which is representative of a major epidemic strain, the Liverpool Epidemic Strain (LES) CF strain [8].
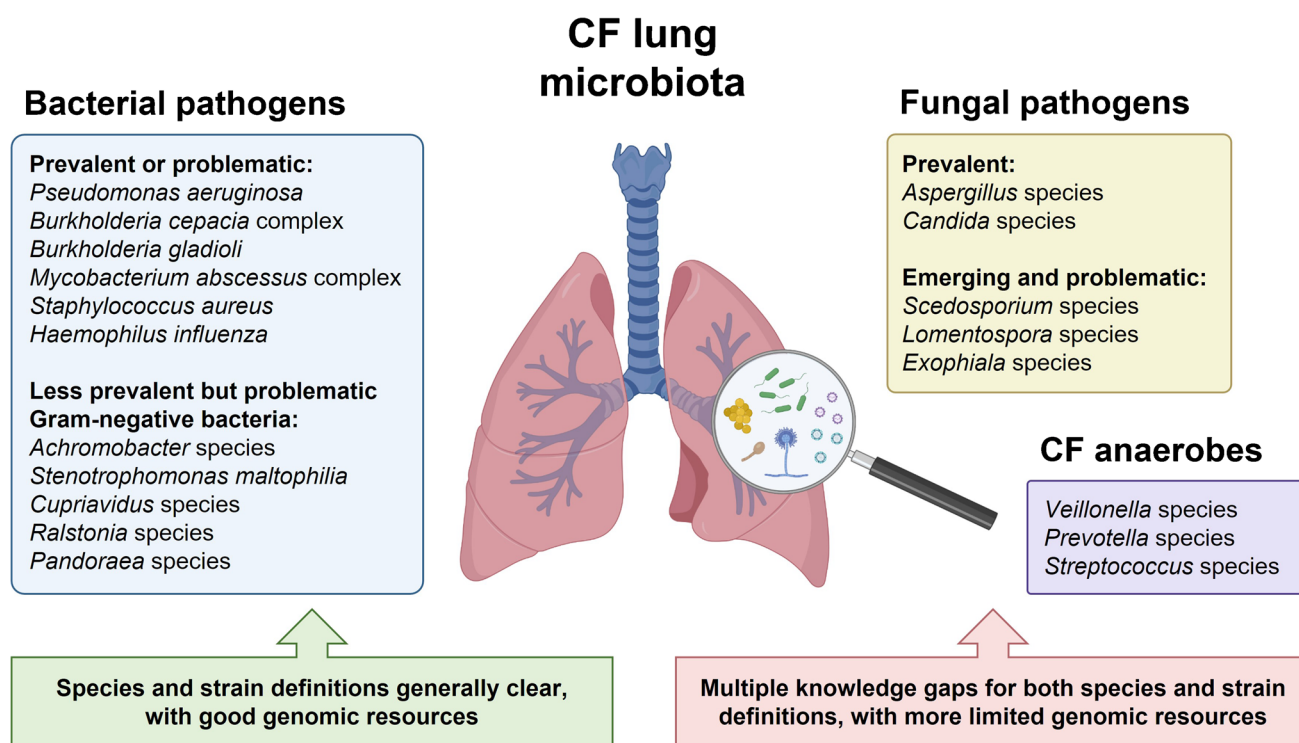

Fig. 1 Bacterial pathogens, fungal pathogens, and the CF lung microbiota. Bacteria and fungi that are identified as pathogens and therapeutically treated in CF lung disease are listed. The wider microbiota are also illustrated, outlining prevalent bacteria (mainly anaerobes) which are not currently treated as lung pathogens. The current knowledge relating to taxonomic identity, strain definitions, and genomic resources is described. Figure graphics created with BioRender.com

The International *P. aeruginosa* Reference Panel

In 2013, Anthony De Soyza and colleagues coordinated the development of an international *P. aeruginosa* reference panel in an attempt to address the issue of studying disease-relevant strains [7••] (Table 1). The panel of 43 strains was developed to be representative of *P. aeruginosa* using a systematic range of clinical and microbiological criteria. Full historical details and source references for each panel strain were provided, and they were selected to be genetically diverse in terms of the known population biology of

P. aeruginosa at the time [7••]. The *P. aeruginosa* strain panel included all four well-studied reference strains (PAO1, UCBPP-PA14, PAK, and LES B58) and expanded on CF strains as follows:

- 8 isolates representative of transmissible CF strains including LES and from Denmark and Australia
- 8 isolates from chronic infection that were sequential to cover evolution within an individual's lung infection (3 from an adult and 5 from a childhood CF respiratory infection)

Table 1 Strain and genomic level resources for bacterial CF pathogens

Pathogen species	MLST resources for isolates and genomes (no. CF relevant):	Strain panel or published collection (useful references)	Attributes of strain panel or collection
P. aeruginosa	8251 (1352)[a] isolates 2745 (36)[a] genomes	International panel [7••] Phenotypic properties [9••] Panel genomics [12•] Population genomics [13]	43 systematically selected strains; coordinated phenotypic analysis of 42 strains from the panel; 2 major genomic groups present within species
B. cepacia complex	4451 (2092) isolates 1410 (487) genomes	Original panel [19] Updated panel [54]	30 strains (representative of first 5 named genomovars): *B. cepacia* (I), *B. multivorans* (II), *B. cenocepacia* (III), *B. stabilis* (IV) and *B. vietnamiensis* (V) Updated panel to include 4 *B. dolosa* (VI), 3 *B. ambifaria* (VII), 4 *B. anthina* (VIII), and 4 *B. pyrrocinia* isolates
B. cenocepacia	1163 (812) isolates[b] 342 (53) genomes[b]	Published genomic collections [27•]	303 strain genomes characterised; genomic support for 2 clades within species
B. gladioli	142 (53) isolates[b] 107 (25) genomes[b]	Published genomic collections and associated CF isolates [34•]	206 (194 CF) isolate genomes analysed; 5 genomic clades defined; 13% of CF isolates produce lethal bongkrekic acid toxin
Achromobacter species	1406 (1035) isolates 9 (1) genomes	Published genomic collections and associated CF isolates [42] [41]	54 isolate genomes from 26 CF patients 101 isolate genomes, 51 CF patients over 20 years
Stenotrophomonas maltophila	994 (30)[a] isolates 101 (5)[a] genomes	Published genomic collections and associated CF isolates [43]	91 isolate genomes, 10 CF individuals, and 10 years
Mycobacterium (*Mycobacteroides*) *abscessus* complex	1886 isolates 1884 genomes	Published genomic collections and associated CF isolates [35, 36, 37•]	168 isolates, 31 CF individuals over 4 years of study and subsequently 1173 isolate genomes from 526 CF individuals across UK, US, Australia and Europe
S. aureus	37887 (309)[a] isolates 27082 (17)[a] genomes	Published genomic collections and associated CF isolates [47] [46]	64 isolate genomes, 50 CF individuals from the USA Staphopia database of 43,000 genomes
H. influenzae	6798 (466)[a] isolates 2785 (76)[a] genomes	Published collection of CF isolates [48]	Extensive strain collection evaluated for antibiotic resistance but limited systematic knowledge of CF strains

[a]Source given as sputum without breakdown of being CF or other lung disease respiratory samples

[b]*B. cenocepacia* and *B. gladioli* strain information is held within the single *Bcc* MLST database

- 8 additional CF strains with specific phenotypic, virulence serotype characteristics [7••].

The wider panel encompassed 15 other *P. aeruginosa* strains from non-CF infections (burns, wounds, keratitis, acute intensive care unit infection, and chronic obstructive pulmonary disease), as well as 4 strains isolated from the environment [7••]. All the strains within the panel were deposited within an internationally recognised strain repository, the Belgium Coordinated Collection of Microorganisms/Laboratory for Microbiology Gent (BCCM/LMG; http://bccm.belspo.be/about-us/bccm-lmg). This enabled researchers to reproducibly access the *P. aeruginosa* strains from a single validated source at low cost [7••].

Phenotypic Data on the International *P. aeruginosa* Reference Panel

To follow-up the definition of the panel, Cullen and colleagues [9••] examined multiple phenotypic properties for 42 of the original 43 strains (strain NN2 was withdrawn from the analysis due to inconsistencies over its taxonomic identity). The phenotypes determined included growth rates, antibiotic susceptibility, motility, *Galleria mellonella* infection modelling, mucoidy, pyocynanin and alginate production, lipopolysaccharide properties, biofilm formation, urease activity, and antimicrobial and phage susceptibility [9••]. A key feature of the phenotypic testing was its coordination across multiple research laboratories in different geographic locations, with several phenotypes such as antibiotic susceptibility reproduced in different centres [9••]. This accounted for the variable nature of laboratory testing, and the fact that *P. aeruginosa* is a phenotypically diverse bacterium capable of significant intra-strain diversity [10]. By using the findings of the *P. aeruginosa* panel phenotype paper [9••], researchers can rapidly identify strains with the relevant virulence or antimicrobial resistance traits they wish to model.

A Genomic Understanding of the *P. aeruginosa* Reference Panel

The original *P. aeruginosa* strain panel [7••] accounted for the genetic diversity of *P. aeruginosa* by selecting strains based on their Clondiag Array Tube (AT) genotype [11]. The AT-genotyping database covered the diversity of greater than 1000 *P. aeruginosa* strains at the time, but as a molecular strain typing method, it is now been superseded by whole genome sequence analysis. In 2018, Freschi and colleagues [12•] determined the genome sequences for 33 of the strains from the *P. aeruginosa* panel [7••] and compared them to 7 reference genomes (Table 1). A key observation from

this analysis of 40 strain genomes was that *P. aeruginosa* separates into two major evolutionary lineages, designated genomic Groups 1 and 2. Well-characterised reference strains such as LESB58, PAO1, and PAK are in Group 1, while other reference strains such as UBCPP-PA14 are in Group 2.

Furthermore, genomic analysis of 1311 *P. aeruginosa* genomes identified a total of 5 genomic groups within the species [13]. This confirmed the presence of Group 1 ($n = 986$ isolates) and Group 2 ($n = 297$ isolates) and demonstrated that they are the most common *P. aeruginosa* lineages (collectively 98% of the isolates) [13]. Three additional groups, 3, 4, and 5, were defined, with Group 3 being significantly distant in terms of its average nucleotide identity (ANI; 93–94% compared to all the other *P. aeruginosa* genomes analysed) [13]. The group 3 isolates (14 in total) included the well-studied multidrug-resistant *P. aeruginosa* PA7 strain [13]. However, in the context of the current genomic taxonomy boundary of 95% ANI used to define the majority of bacterial species [14••], strain PA7 and Group 3 as a whole could be considered a novel genomic species. This suggest that strain PA7 is less relevant as a model strain when comparing to the more common and taxonomically validated, *P. aeruginosa* lineages.

A study of 103 *P. aeruginosa* genomes [15] confirmed the Groups 1 and 2 population biology as encompassing the majority of strains [13], and also defined the preservative tolerance of multiple panel strains. *P. aeruginosa* is intrinsically resistant to both antibiotics and other antimicrobials such as preservatives, disinfectants, and biocides [15]. This analysis also showed that preservative tolerant strains from the contamination of nonsterile industrial products possessed genomes that were larger (7 Mb) than clinical strains (6.6 Mb) primarily due to the presence of a 0.5-Mb megaplasmid. The same family of megaplasmids was subsequently identified in multidrug-resistant non-CF clinical *P. aeruginosa* isolates from Thailand [16]. Overall, given the clear presence of 2 common genomic lineages within *P. aeruginosa*, researchers should seek to understand how strains from each behave in relation to novel therapeutics. At minimum, studies should include *P. aeruginosa* PAO1 (Group 1) and UBCPP-PA14 (Group 2) as references for each lineage and potentially expand analysis to include wider CF or other strains from sequenced panel [12•], or more recent genetic studies which characterise AMR megaplasmid encoding strains [15, 16].

Burkholderia cepacia Complex Bacteria and *Burkholderia gladioli*

Burkholderia are intrinsically antimicrobial-resistant Gram-negative bacteria that emerged as highly virulent and transmissible CF pathogens in the 1980s [2, 17]. From the onset of their recognition as CF pathogens, there were difficulties

associated with their identification and taxonomy, with 5 species groups shown to make up isolates of *B. cepacia* [18]. Hence, to help researchers study them, they were collectively designated as the *B. cepacia* (Bcc) complex [18], with a representative strain panel assembled in 2000 to facilitate comparative research [19] (Table 1). The original *Bcc* panel included 30 strains from the 5 species at the time, and accounted for basic genetic strain diversity as determined by macrorestriction and pulsed-field gel electrophoresis fingerprinting [19]. Multiple follow-up publications have examined these original panel strains, and in 2003, an expansion to the panel was made to include *B. dolosa* (designated genomovar VI at the time), *B. ambifaria*, *B. anthina*, and *B. pyrrocinia* (20) (Table 1). As with the *P. aeruginosa* panel [7••], all strains within the *B. cepacia* complex panels are deposited within the BCCM/LMG (http://bccm.belspo.be/about-us/bccm-lmg) to enable access.

With the ongoing taxonomic classification, by 2010, there were 17 *Bcc* species defined [2]. Using whole genome sequence, a total of 22 known species and 14 further novel genomic species have recently been identified [20]. In a more extensive analysis of 4000 *Burkholderia* genomes, 26 novel *Bcc* genomic species groups have been recently defined [21•], indicating that the taxonomic complexity of this group of bacteria will continue to grow. In the context of CF lung infections, *B. multivorans* and *B. cenocepacia* are the most dominant species encountered in people with CF [3•, 22], and hence should be the focus of therapeutic development. Although the prevalence of *B. cepacia* complex infections in CF is low (between 2 and 5% in different populations [3•, 22]), they are important as a target for therapeutic development for a number of reasons. Firstly, there are limited therapeutic options to suppress or eradicate infection. Secondly, as a result of the unpredictable clinical outcome and severe lung disease associated with *Bcc* CF lung disease, infected individuals are frequently excluded from clinical trials of other novel drugs and lung transplantation, further limiting options available for this highly vulnerable group.

B. cenocepacia

This species name was proposed in 2003 [23] and known to encompass considerable CF strain diversity. At least four genetic lineages of *B. cenocepacia* were defined based on the *recA* gene sequence (III-A, III-B, III-C, and III-D) [23, 24]. *B. cenocepacia* isolates from the ET12 strain have been widely studied and include model strains J2315 [25] and K56-2 [19]. They are representative of the highly virulent and transmissible CF *B. cenocepacia recA* III-A lineage [26]. *B. cenocepacia recA* III-B lineage strains are prevalent within the US CF population [2]. Within the original strain panel [19], strain PC184 is a suitable *recA* III-B strain and member of the US Mid-West epidemic clone.

Recent genomic analysis by Wallner and colleagues [27•] validated the major genetic split within *B. cenocepacia* originally observed by the analysis of the *recA* gene [24]. They also proposed that the III-B grouping represent a novel genomic species, proposing the name *Burkholderia servocepacia* for this group [27•]. This new species name has not been taxonomically validated. It is also the subject of controversy because it misrepresents that this *B. cenocepacia* genomic species encompasses strains which are mainly plant associated [27•]. This completely contrasts the known CF infection epidemiology, with both III-A and III-B *B. cenocepacia* strains causing devastating CF lung infections and epidemic outbreaks of infection [2]. Future studies on *B. cenocepacia* as a CF pathogen should examine strains from both the III-A and III-B genetic lineages. Also, with genomic characterisation of multiple *B. cenocepacia* isolates completed [21•], systematic selection of future strain panels for this species is fully enabled.

B. multivorans

Unfortunately, there is no well-characterised CF strain of this species, even though it is now the most dominant *Bcc* species seen in CF [2, 3•, 22]. The soil isolate *B. multivorans* ATCC 17,616 has been well studied [19], with multilocus sequence typing (MLST) identifying 2 CF strains of the same sequence type (ST-21) as this environmental isolate [28]. Hence, strain ATCC 17,616 can be considered a model that is genetically representative of isolates that are capable of causing CF infection.

There have been two detailed studies following the genomic evolution of single *B. multivorans* strains during chronic CF lung infection. Silva et al. [29] followed the evolution of single strain through a switch from a mucoid to nonmucoid phenotype associated with 20 years of chronic CF infection. Cabellero et al. [30] performed another genomic study of the evolution of a single *B. multivorans* strain which chronically infected a CF adult for 10 years prior to lung transplantation. The isolate possessed the MLST sequence type, ST-783 [30], and current analysis of this sequence type at the MLST database shows that it is a sequence type shared by other CF (US infection), non-CF (Belgium), and environmental isolates (Belgium). Given the depth of clinical and genomic information that accompanies these 2 *B. multivorans* CF strains [29, 30], they are worthy of inclusion within future test panels.

B. gladioli

This species is not a member of the *Bcc*, but it is the third most common *Burkholderia* species seen in the US [2] and UK CF population [22]. Clinical outcomes of infection with *B. gladioli* may also be variable and problematic as seen

with the *Bcc*. A useful model strain for this group is *B. gladioli* BCC0238 (LMG-P 26202), a CF isolate recovered in 1996 from a paediatric patient in Minnesota, USA [31]. The complete genome sequence for *B. gladioli* BCC0238 is available [31, 32], and although it has not been characterised in detail as a CF pathogen, the isolate has been studied because it produces multiple antibiotics, such as gladiolin which is active against tuberculosis [31, 33].

The population biology of *B. gladioli* has recently been determined [34•] by genomic comparison of 206 isolates, of which the majority were from CF infection (*n* = 194). The analysis demonstrated the presence of 3 distinct genomic groups and 5 evolutionary clades, with strain BCC0238 residing in *B. gladioli* group 3 [34•]. Strains from CF infection were found within all 5 evolutionary clades, and the study also uniquely identified that 13% of the *B. gladioli* CF isolates can produce the toxin, bongkrekic acid. This virulence factor could potentially drive clinical disease in CF as it is known to be lethal to non-CF individuals following ingestion of bongkrekic acid via food that had been contaminated with *B. gladioli* [34•]. The antibiotic trimethoprim suppressed *B. gladioli* bongkrekic acid production in vitro, and hence could be useful to prevent rapid clinical decline due to toxin poisoning in CF [34•]. *B. gladioli* should be included as a test organism for CF therapeutics that target *Burkholderia* CF infection, with BCC0238 or other genomically characterised strains selected as models [34•].

Mycobacterium abscessus Complex

A range of nontuberculous mycobacteria (NTM) may cause lung infection in people with CF, including isolates from the *M. abscessus* complex and *M. avium* complex [2]. Currently, the most problematic and prevalent group within a number of global CF populations is *M. abscessus*. In 2013, Bryant and colleagues [35] published the first extensive genomic evidence that shared strains exist and transmission of isolates may occur between individuals with CF. Over 4 years, they assembled a collection of 168 isolates from 31 individuals attending a single CF UK treatment centre, and used genome sequencing and phylogenomic analysis to characterise this collection [35].

In the last decade, it has become clear from genomic studies characterizing over 1000 clinical isolates that *M. abscessus* CF infections are acquired frequently through transmission leading to the emergence of globally dominant clones [36]. Furthermore, high-resolution genomic analyses of 1173 isolates from 526 CF patients were used to demonstrate the presence of 3 dominant circulating clones of *M. abscessus*, and model their pathogenic evolution within selected isolates such as strain BIR1049 [37•]. These large collections of well-characterised *M. abscessus* strains [35, 36, 37•] provide highly valuable resources including information on antibiotic resistance and molecular pathogenesis in CF. A sub-selection of limited numbers of model strains such as BIR1049, and those which account for each of the dominant circulating clones, will be required to enable straightforward CF therapeutic testing against *M. abscessus*.

Multiresistant Gram-Negative CF Pathogens

In addition to *P. aeruginosa* and *Burkholderia* species, multiple intrinsically antibiotic resistant Gram-negative species have emerged as CF pathogens [2], which can also dominate the lung microbiota in CF adults [38]. These include *Ralstonia* species, *Pandoraea* species, *Stenotrophomonas maltophilia*, *A. xylosoxidans*, and *Inquilinus limosus* [2]. There are currently no highly characterised strains, but a small panel of representative species was assembled as part of a European Union funded project, EuroCareCF. The panel of CF isolates and other reference strains (Table 2; *n* = 25) were used for a quality assurance trial of bacteriology laboratories in different countries [39]. This EuroCareCF panel is also available from the BCCM/LMG (http://bccm.belspo.be/about-us/bccm-lmg), and although biased towards *Burkholderia* species (14 of 25 strains), it forms a systematic collection of other multiresistant species useful for therapeutic testing (*Ralstonia*, *n* = 3; *Pandoraea*, *n* = 4; and one isolate of *S. maltophilia*, *A. xylosoxidans*, *Cupriavidus (Ralstonia) respiriculi*, and *Inquilinus limosus*; Table 2).

Since the assembly of the EuroCareCF strain panel [39], it has become clear that multiple species of *Achromobacter* may infect people with CF [40], including *A. xylosoxidans*, *A. ruhlandii*, *A. insuavis*, *A. aegrifaciens*, *A. dolens*, *A. insolitus*, and other novel genomic species [41, 42]. Two recent genomic studies of 101 [41] and 54 [42] CF isolates of *Achromobacter* provide excellent collections from which to draw diverse isolates for testing. *S. maltophilia* is another multiresistant Gram-negative pathogen of increasing concern in CF lung infection, which in the USA had an overall prevalence of 13% in 2012 [3•]. Few genomic studies of *S. maltophilia* CF isolates exist, but Esposito et al. [43] characterised a collection of 91 isolates from 10 individuals with CF, looking at their phenotypic and genomic evolution over a 12-year period. This study showed that the *S. maltophilia* strains infecting the 10 individuals were genetically diverse, with many of them representing novel MLST sequence types [43].

Staphylococcus aureus

As a Gram-positive bacterium, *S. aureus* is as synonymous with CF lung infection as *P. aeruginosa*. *S. aureus* is found in CF patients of all ages but is at high prevalence in the paediatric CF population (often greater than 60%) [3•]. The increased isolation of methicillin-resistant *S. aureus* (MRSA) in CF is also worrying [3•], as it is WHO 'Priority

Table 2 The EuroCareCF panel of Gram-negative CF pathogens

Species	BCCM identifier	Pseudonym or other names	Source:	Comments (useful reference)
B. cepacia	LMG 1222 [T]	ATCC 25416	Plant rot (onion), US	Species type strain, Bcc panel [24]
B. multivorans	LMG 13010 [T]	ATCC BAA-247	CF, UK	Species type strain, Bcc panel [24]
B. multivorans	LMG 16660	C1576	CF, UK	Glasgow epidemic strain, Bcc panel [24]
B. multivorans	LMG 16775	95–14886	CF, Sweden	Strain from an adult with CF
B. cenocepacia	LMG 16656 [T]	J2315	CF, UK	Species type strain, epidemic *recA* III-A lineage, ET12 strain, Bcc panel[24]
B. cenocepacia	LMG 12614	NCTC 13010	CF, UK	Epidemic *recA* III-A lineage, ET12 strain
B. cenocepacia	LMG 18829	PC184	CF, USA	US Mid-West Epidemic strain, *recA* III-B lineage, Bcc panel [24]
B. cenocepacia	LMG 18830	CEP511	CF, Australia	*recA* III-B lineage, Bcc panel [24]
B. cenocepacia	LMG 16654	J415	CF, UK	III-B lineage, Bcc panel [24]
B. stabilis	LMG 14294 [T]	ATCC BAA-67	CF, Belgium	Species type strain, Bcc panel [24]
B. vietnamiensis	LMG 18835	PC259	CF, USA	Bcc panel [24]
B. dolosa	LMG 18943 [T]	AU0645	CF, USA	Species type strain, Boston epidemic strain, Updated Bcc panel
B. pyrrocinia	LMG 21824	C1469	CF, UK	Updated Bcc panel [54]
B. gladioli	LMG 18157	CC40, BCC1651	CF, USA	*B. gladioli* clade 3, genome available [34•]
Ralstonia mannitolylitica	LMG 18103	CCUG 38756	CF, USA	
Cupriavidus (Ralstonia) respiraculi	LMG 21510 [T]	AU3313	CF, USA	Species type strain
Ralstonia pickettii	LMG 18088	CEP0488	CF, USA	
Ralstonia insidiosa	LMG 18101	C6317	CF, Canada	
Pandoraea apista	LMG 16407	ATCC BAA-61	CF, Denmark	Species type strain
Pandoraea sputorum	LMG 18819 [T]	ATCC BAA-64	CF, USA	Species type strain
Pandoraea pulmonicola	LMG 18106 [T]	FC330	CF, Canada	Species type strain
Pandoraea pnomenusa	LMG 18087 [T]	C1513	CF, UK	Species type strain
Stenotrophomonas maltophilia	LMG 958 [T]	ATCC 13637	Human infection (mouth)	Species type strain
Achromobacter xylosoxidans	LMG 1863 [T]	ATCC 27061	Human infection (ear), Japan	Species type strain
Inquilinus limosus	LMG 20952 [T]	AU0476	CF, USA	Species type strain

2 (high)' AMR pathogen [6]. The prevalence of chronic MRSA lung infection is increasing in several CF populations [3•, 4, 44]. Outside of CF lung infection, *S. aureus* strain collections and knowledge of genomic diversity are considerably advanced primarily because it is a major AMR human pathogen [6]. An extensive database of over 36,000 isolates and 26,000 genomes is maintained at the *S. aureus* MLST database [45••]. The recently established Staphopia analysis pipeline, database, and web application programming interface also provide a unique suite of tools to assemble and analyse *S. aureus* genomes [46]. The developers of Staphopia used the platform to analyse over 43,000 publicly available *S. aureus* genomes, exploring its genetic diversity and selecting high-quality genomes with robust metadata as a reference subset for the species [46].

In the context of CF, Bernady et al. [47] recently established a collection of 64 lung infection isolates, mapped a range of their phenotypes, and used genome sequencing and comparison via Staphopia to place them within the known *S. aureus* population biology. The CF strains were mapped to 8 of the 66 clonal complexes defined in the Staphobia study, with the most common being CC5 and CC8 MRSA strains [47]. The study also compared the CF strains to the well-known non-CF reference MRSA strain, *S. aureus* JE2, which

is representative of a globally distributed strain known as USA300 [47]. *S. aureus* phenotypes such as toxin production and nonmucoid variants were shown to be retained during chronic lung infection in the MRSA and antibiotic susceptible CF strains that were characterised [47]. This study [47] together with greater use of representative clonal complex strains within the MLST [45••] (Table 1) or Staphopia databases [46] can provide an excellent basis to select *S. aureus* strain panels for therapeutic testing.

Haemophilus influenzae

Nontypeable *H. influenzae* isolates are common colonisers of children with CF, with an overall prevalence of 15% in the USA, that peaks at 32% among children aged 2 to 5 years [3•]. Although frequently treated as a pathogen in CF, *H. influenzae* represents a poorly studied bacterial infection in this context. In contrast, as an invasive pathogen of children that is capable of causing meningitis, extensive isolate (>6000) and genomic collections (>2000) exist for *H. influenzae* at the MLST database (Table 1) [45••]. Searching the MLST database shows that a small proportion of isolates may represent CF infection but proving this is limited by the fact that the source for *H. influenzae* isolates is generically recorded as sputum (Table 1). Without robust sample metadata, it is unknown whether the sputum is from an individual with CF or another form of respiratory disease. In addition, CF may actually be under-represented given most infection occurs in children, who generally cannot expectorate sputum.

Ebbing et al. [48] characterised the antibiotic resistance of both *H. influenzae* and *H. parainfluenzae* in isolates collected over 15 years within the Australian CF community. *H. parainfluenza* is considered part of the normal microbiota within the human oral and laryngeal cavity but can occasionally cause lung in infections and endocarditis. With 518 *H. influenzae* and 1020 *H. parainfluenzae* isolates analysed, it represents one of the largest studies in the context of CF. The isolates were representative of the infection seen in 349 CF individuals, and it was demonstrated that overall antibiotic resistance increased by 46% for *H. influenza* and 61% for *H. parainfluenza* over the 15-year study. The striking frequency at which *Haemophilus* species are isolated during childhood respiratory infection [3•], their increasing antimicrobial resistance in CF [48], and the status of *H. influenzae* as a WHO 'Priority 3 (Medium)' pathogen warrants that it should be the subject of much more systematic study in CF. Given the assumption that microbial species that infect the CF lung early and any antibiotics administered against them will alter the lung environment, it is likely that early *H. influenzae* infections may play a significant role in influencing the future course of lung disease.

Systematic Criteria for the Selection of Bacterial CF Pathogen Strains

It is clear from the current state of knowledge that the ability to systematically select bacterial strains for testing novel therapeutics varies for different CF pathogens. However, multiple core principles come through in considering how strains for testing and reproducible CF research should be selected in the future. In particular, the development of the *P. aeruginosa* international strain panel [7••] and its characterisation over the last decade [9••, 12•] prompts the following selection criteria and key questions for consideration when selecting strains in future CF research (Fig. 2):

I. A systematic strain selection process. Initially, there was a systematic selection of *P. aeruginosa* strains based on relevant strain diversity, underlying disease and other criteria [7••]. These criteria and the strains selected were discussed widely among multiple groups of interdisciplinary researchers [7••].

II. Detailed phenotypic characterisation. Phenotypic analysis of *P. aeruginosa* strains was subsequently performed, reproduced, and validated to understand their behaviour across a range of standard laboratory models including antibiotic susceptibility analysis [9••]. A clear understanding of multidrug resistance within a testing strain panel is also highly relevant in the context of new therapeutics being able to overcome existing AMR [6].

III. Understanding of the species and strain population biology. The panel strains were initially placed within the overall population biology of *P. aeruginosa* as a bacterial species using AT-genotyping [7••]. This accounted for their genetic diversity and representation of *P. aeruginosa* as a pathogen in CF, other diseases, and as a free-living environmental bacterial species. The ultimate resolution of genetic analysis, whole genome sequencing, and the comparison of isolates at a phylogenomic scale revealed further diversity into two major evolutionary groups [13]. Although we do not fully understand the clinical or biological significance of *P. aeruginosa* genomic Group 1 or Group 2 strains, inclusion of both within future research should be carried out to identify if systematic differences exist.

IV. Understanding within-strain evolution and adaption to CF lung infection. The within-strain variation of *P. aeruginosa* as a CF pathogen has also been extensively characterised [10, 49, 50]. For example, the international *P. aeruginosa* reference panel includes a mucoid and nonmucoid isolates of the same strain (IST 27) to account

Fig. 2 Core principles and relevant questions to aid the design of systematic CF pathogen strain panels. Consideration of the design and testing of strains within the international *P. aeruginosa* panel, and the current state of knowledge of multiple other bacterial CF pathogens was made. The five core principles (boxes) and key questions within them (ovals) that derive from considering what is needed in future studies are illustrated. Figure graphics created with BioRender.com
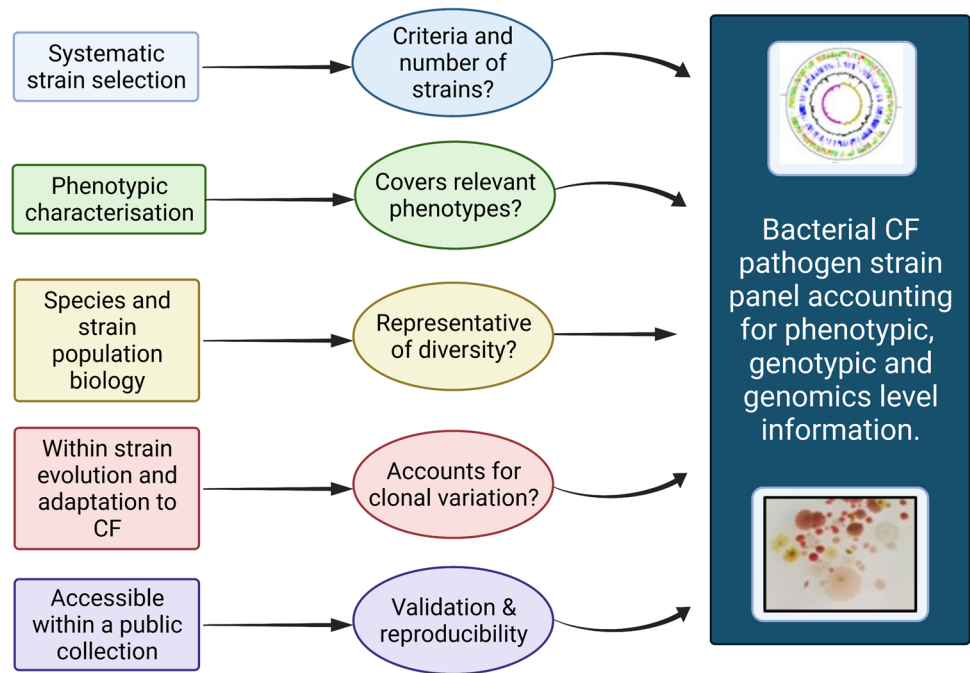
for this important CF phenotype [7••]. The transition to a hypermutator for *P. aeruginosa* [50], or a small colony variant in *S. aureus* [51], are also important sub-strain phenotypes in CF. Consideration should be given to which types of phenotypic strain derivatives should be included in any therapeutic testing, particularly if new compounds are directed, for example, at virulence factors.

V. Accessibility within a public collection. A final key consideration is that any strains that are used widely in therapeutic testing must be made available from curated and accessible microbial collections. The establishment of the international *P. aeruginosa* reference panel and its central deposition in a recognised microbial resource repository [7••] enabled testing by multiple laboratories to ensure data validation and reproduction [9••]. Given the within strain phenotypic and genetic variance of multiple CF pathogens, using a single curated source rather than passing a strain from research laboratory to research laboratory is vital to obtain robust data.

Strain Collections, Useful Databases, and Prospective Clinical Isolates

Multiple internationally recognised strain repositories exist from which reference strains may be ordered and hence enable greater reproducibility across global studies. The BCCM/LMG collection (http://bccm.belspo.be/about-us/bccm-lmg) already forms an excellent repository for multiple CF pathogen species as outlined (Table 2). The Bacterial Diversity Metadatabase (Bac*Dive*; > 81,000 strain) is also a highly useful online resource for researchers and companies to explore culture collections [52•]. Bac*Dive* enables rapid searches of taxonomy, physiology, isolation source and location, phenotypic data, and genomic resources for multiple bacterial strains [52•]. The database is very useful for providing unifying information on bacterial strains which may have been published under multiple pseudonyms (e.g., ATCC, LMG, or other identifiers; see Table 2).

Multiple researchers and industrials developing therapeutics also use currently circulating clinical isolates to test novel anti-infectives. This approach is relevant if the basic strain selection can account for the first four criteria noted above. Working directly with CF treatment centres may also be beneficial because of the relative sparsity of available clinical data that accompanies strains requested from curated culture collections. CF lung infection also represents a highly complex disease, where several factors impact the success of therapeutic approaches and there are multiple knowledge gaps (Table 3). Clinical CF centres can provide rich clinical data to accompany bacterial isolates such as the rates of decline in lung function, frequency of exacerbations, co-infecting pathogens, and other drug treatments.

Table 3 Knowledge gaps in relation to the selection and testing of relevant CF pathogen strains

Pathogen or knowledge area	Knowledge gaps	Useful reviews and references
Fungal CF pathogens • *Aspergillus* species • *Candida* species • *Scedosporium* species • *Lomentospora* species • *Exophiala* species	• A critical limit of available antifungal antibiotics • Questions on their pathogenicity and transmissibility • Variable growth phenotypes and modelling systems • Relevant model strains • Taxonomy and strain classification • Genomic resources	CF fungal pathogens [55] Fungal classification [56]
Gram-negative bacterial CF pathogens • *H. influenzae* • *Stenotrophomonas* species • *Achromobacter* species • *Cupriavidus* species • *Ralstonia* species • *Pandoraea* species	• Questions on their pathogenesis and correlation to poor clinical outcome • Multidrug resistance and limited antibiotics as per priority CF pathogens • Taxonomic and genomic diversity within certain species groups • An ability to form dominant lung infection in certain CF individuals	*H. influenzae* [48] Dominant microbiota [38]
Prevalent CF anaerobes • *Veillonella* species • *Prevotella* species • *Streptococcus* species	• Difficulties in laboratory growth • Identification and taxonomy • Strain level understanding • Resistance to antibiotics directed against aerobic pathogens • Limited relevant strain collections and genomic resources • Questions on their pathogenesis and contribution to infection	(1) [57]
The CF lung microbiome	• Complexity of modelling the microbial and host diversity associated with lung infection • What pathogens co-occur • How is AMR influenced by the microbiota and interactions present • The disconnect between in vitro antibiotic susceptibility testing (on a single target organism) and predicting clinical efficacy for anti-infectives to work in a polymicrobial lung infection	(1) [58]
Appropriate therapeutic testing models	• Single target pathogens versus polymicrobial testing and co-infecting pathogens • Artificial sputum versus standard growth media • Ex vivo versus in vivo models • Choice of model systems from invertebrates through to mammals • Delivery of therapeutic and systemic versus inhaled or respiratory targeting • Impact of CFTR modulation and clinical relevance of models	[59] [60] [53]

The last few decades have witnessed improved health outcomes and decreasing rates of chronic infection with several priority CF pathogens [4]. The revolutionary CFTR modulator drugs have the ability to restore a substantial amount of CFTR function, and are now available for the vast majority of the CF population in many countries [53]. Registry data have demonstrated a modest impact on infection prevalence for certain bacterial pathogens, even in populations commencing treatment in adulthood. It is anticipated that the impacts on infection acquisition and evolution to chronicity will be even more marked in CF children commencing CFTR modulators before airway disease is established. With this changing landscape in bacterial CF lung infections, there is a strong case to continue prospective collection and characterisation of CF infection isolates to join up current and historical information, and account for the clear differences in the clonal population biology of problematic pathogens such as *P. aeruginosa*, *Bcc*, and *M. abscessus*:

How can researchers help provide context to clinical isolates or other testing strains with limited genotypic characterisation?

Given that genome sequencing is now cost-effective, carrying this out for uncharacterised testing isolates is highly recommended. A draft bacterial genome can be uploaded to the SpeciesID tool (https://pubmlst.org/species-id) available at the PubMLST database, which hosts strain genotyping resources for multiple CF and other pathogens (Table 1) [45••]. The tool will carry out ribosomal multi-locus sequence typing (rMLST) [45••] directly on the draft genome and provide both taxonomic and strain level matches to the query isolate sequence against the databases. The provenance data on closely related strains within pubMLST (Table 1) can then be used to provide additional context to the uncharacterized testing isolate. If the genome sequence of the isolate is subsequently deposited in the DNA archives, this will help researchers with future comparative analysis and set a baseline for what AMR resistance factors may have present in the testing isolate.

Conclusions

The criteria and core principles on strain selection (Fig. 2) can be applied to multiple priority CF pathogens beyond *P. aeruginosa*. Phenotypic and genomic knowledge of *Bcc* species, *B. gladioli*, *M. abscessus, S. aureus*, and *Achromobacter* species in CF is reaching a point where there is now sufficient understanding to enable systematic strain selection (Table 1). In light of this, an additional consideration becomes:

What is the optimum number of strains of a given CF pathogen species that is required for any testing panel to be relevant?

Given that we know multiple strain lineages and core genomic groups exist within *P. aeruginosa* [13] and *B. cenocepacia* [27•], initial testing of at least 5 to 10 strains should enable basic coverage of such population biology diversity. Going beyond this will depend on multiple variables including the logistics of the testing required, the cost, how many well-characterised strains are easily accessible, and whether panels are representative of current and emerging AMR. The strain coverage within the *P. aeruginosa* international reference panel and its use [9••] suggests that 40 to 50 strains can be handled as a relevant number for investigations on multiple traits, at least for broader screening.

This review of strain choice for bacterial pathogens also highlights multiple knowledge gaps for CF lung infection which researchers should seek to fill and include (Table 3): fungal CF pathogens, *H. influenzae*, and emerging AMR Gram-negative bacterial CF pathogens, anaerobic CF bacteria, the wider lung microbiome, and CF relevant models for therapeutic testing. Outlining relevant bacterial strain resources for the major well-characterised CF pathogens (Tables 1 and 2) and core principles for selecting relevant strains (Fig. 2) provides a basis from which to consider the challenge of filling these knowledge gaps (Table 3) in future.

Declarations

References

Papers of particular interest, published recently, have been highlighted as:
• Of importance
•• Of major importance

1. Filkins LM, O'Toole GA. Cystic fibrosis lung infections: polymicrobial, complex, and hard to treat. PLoS Pathog. 2015;11(12):e1005258.

2. Lipuma JJ. The changing microbial epidemiology in cystic fibrosis. Clin Microbiol Rev. 2010;23(2):299–323.

3.• Salsgiver EL, Fink AK, Knapp EA, LiPuma JJ, Olivier KN, Marshall BC, et al. Changing epidemiology of the respiratory bacteriology of patients with cystic fibrosis. Chest. 2016;149(2):390–400. **This study provides a current overview of the epidemiology of CF lung infection in the US CF population**.

4. UK Cystic Fibrosis 2020 Registry Annual Data Report [Internet]. 2020. Available from: https://www.cysticfibrosis.org.uk/the-work-we-do/uk-cf-registry/reporting-and-resources. Accessed: 11 August 2022.

5. Simar SR, Hanson BM, Arias CA. Techniques in bacterial strain typing: past, present, and future. Curr Opin Infect Dis. 2021;34(4):339–45.

6. Tacconelli E, Carrara E, Savoldi A, Harbarth S, Mendelson M, Monnet DL, et al. Discovery, research, and development of new antibiotics: the WHO priority list of antibiotic-resistant bacteria and tuberculosis. Lancet Infect Dis. 2018;18(3):318–27.

7.•• De Soyza A, Hall AJ, Mahenthiralingam E, Drevinek P, Kaca W, Drulis-Kawa Z, et al. Developing an international *Pseudomonasaeruginosa* reference panel. MicrobiologyOpen. 2013;2(6):1010–23. **This outlines how a panel of *P. aeruginosa* strains was selected for robust and systematic analysis**.

8. Winstanley C, Langille MG, Fothergill JL, Kukavica-Ibrulj I, Paradis-Bleau C, Sanschagrin F, et al. Newly introduced genomic prophage islands are critical determinants of in vivo competitiveness in the Liverpool epidemic strain of *Pseudomonas aeruginosa*. Genome Res. 2009;19(1):12–23.

9.•• Cullen L, Weiser R, Olszak T, Maldonado RF, Moreira AS, Slachmuylders L, et al. Phenotypic characterization of an international Pseudomonas aeruginosa reference panel: strains of cystic fibrosis (CF) origin show less in vivo virulence than non-CF strains. Microbiology. 2015;161(10):1961–77. **This work**

characterizes multiple phenotypes including the antibiotic susceptibility of the *P. aeruginosa* strain panel.

10. Mowat E, Paterson S, Fothergill JL, Wright EA, Ledson MJ, Walshaw MJ, et al. *Pseudomonas aeruginosa* population diversity and turnover in cystic fibrosis chronic infections. Am J Respir Crit Care Med. 2011;183(12):1674–9.

11. Wiehlmann L, Wagner G, Cramer N, Siebert B, Gudowius P, Morales G, et al. Population structure of *Pseudomonas aeruginosa*. Proc Natl Acad Sci U S A. 2007;104(19):8101–6.

12.● Freschi L, Bertelli C, Jeukens J, Moore MP, Kukavica-Ibrulj I, Emond-Rheault JG, et al. Genomic characterisation of an international *Pseudomonas aeruginosa* reference panel indicates that the two major groups draw upon distinct mobile gene pools. FEMS Microbiol Lett. 2018;365(14). https://doi.org/10.1093/femsle/fny120 **This study describes the genome sequencing and presence of two major genomic groups within *P. aeruginosa*.**

13. Freschi L, Vincent AT, Jeukens J, Emond-Rheault JG, Kukavica-Ibrulj I, Dupont MJ, et al. The *Pseudomonas aeruginosa* pan-genome provides new insights on its population structure, horizontal gene transfer, and pathogenicity. Genome Biol Evol. 2019;11(1):109–20.

14.●● Parks DH, Chuvochina M, Chaumeil PA, Rinke C, Mussig AJ, Hugenholtz P. A complete domain-to-species taxonomy for bacteria and archaea. Nat Biotechnol. 2020;38(9):1079–86. **This study outlines the principles of large-scale genomic taxonomy and provides a key database for understanding current bacterial diversity**.

15. Weiser R, Green AE, Bull MJ, Cunningham-Oakes E, Jolley KA, Maiden MCJ, et al. Not all *Pseudomonas aeruginosa* are equal: strains from industrial sources possess uniquely large multireplicon genomes. Microb Genom. 2019;5(7). https://doi.org/10.1099/mgen.0.000276

16. Cazares A, Moore MP, Hall JPJ, Wright LL, Grimes M, Emond-Rheault JG, et al. A megaplasmid family driving dissemination of multidrug resistance in *Pseudomonas*. Nat Commun. 2020;11(1):1370. https://doi.org/10.1038/s41467-020-15081-7.

17. Mahenthiralingam E, Urban TA, Goldberg JB. The multifarious, multireplicon *Burkholderiacepacia* complex. Nat Rev Microbiol. 2005;3(2):144–56.

18 Vandamme P, Holmes B, Vancanneyt M, Coenye T, Hoste B, Coopman R, et al. Occurrence of multiple genomovars of Burkholderia cepacia in cystic fibrosis patients and proposal of Burkholderia multivorans sp. nov. Int J Syst Bacteriol. 1997;47(4):1188–200.

19. Mahenthiralingam E, Coenye T, Chung JW, Speert DP, Govan JR, Taylor P, et al. Diagnostically and experimentally useful panel of strains from the *Burkholderiacepacia* complex. J Clin Microbiol. 2000;38(2):910–3.

20. Jin Y, Zhou J, Zhou J, Hu M, Zhang Q, Kong N, et al. Genome-based classification of *Burkholderiacepacia* complex provides new insight into its taxonomic status. Biol Direct. 2020;15(1):6.

21.● Mullins AJ, Mahenthiralingam E. The hidden genomic diversity, specialised metabolite capacity, and revised taxonomy of *Burkholderia* sensu lato. Front Microbiol. 2021. https://doi.org/10.3389/fmicb.2021.726847. **This study demonstrates the presence of multiple novel genomic species groups within *Burkholderia* by examining over 4000 genome sequences.**

22. Kenna DTD, Lilley D, Coward A, Martin K, Perry C, Pike R, et al. Prevalence of *Burkholderia* species, including members of *Burkholderiacepacia* complex, among UK cystic and non-cystic fibrosis patients. J Med Microbiol. 2017;66(4):490–501.

23. Vandamme P, Holmes B, Coenye T, Goris J, Mahenthiralingam E, LiPuma JJ, et al. *Burkholderiacenocepacia* sp nov — a new twist to an old story. Res Microbiol. 2003;154(2):91–6.

24. Mahenthiralingam E, Bischof J, Byrne SK, Radomski C, Davies JE, Av-Gay Y, et al. DNA-based diagnostic approaches for identification of *Burkholderiacepacia* complex, *Burkholderiavietnamiensis*, *Burkholderiamultivorans*, *Burkholderiastabilis,* and *Burkholderiacepacia* genomovars I and III. J Clin Microbiol. 2000;38(9):3165–73.

25. Holden MT, Seth-Smith HM, Crossman LC, Sebaihia M, Bentley SD, Cerdeno-Tarraga AM, et al. The genome of *Burkholderiacenocepacia* J2315, an epidemic pathogen of cystic fibrosis patients. J Bacteriol. 2009;191(1):261–77.

26. Drevinek P, Mahenthiralingam E. *Burkholderiacenocepacia* in cystic fibrosis: epidemiology and molecular mechanisms of virulence. Clin Microbiol Infect. 2010;16(7):821–30.

27.● Wallner A, King E, Ngonkeu ELM, Moulin L, Bena G. Genomic analyses of *Burkholderia cenocepacia* reveal multiple species with differential host-adaptation to plants and humans. BMC Genomics. 2019;20(1). https://doi.org/10.1186/s12864-019-6186-z. **This report correctly describes a novel genomic species group within *B. cenocepacia*, but its conclusions and proposal of the species name "*B. servocepacia*" are controversial.**

28. Baldwin A, Mahenthiralingam E, Drevinek P, Vandamme P, Govan JR, Waine DJ, et al. Environmental *Burkholderiacepacia* complex isolates in human infections. Emerg Infect Dis. 2007;13(3):458–61.

29. Silva IN, Santos PM, Santos MR, Zlosnik JE, Speert DP, Buskirk SW et al. Long-term evolution of *Burkholderia multivorans* during a chronic cystic fibrosis infection reveals shifting forces of selection. mSystems. 2016;1(3). https://doi.org/10.1128/mSystems.00029-16

30. Diaz Caballero J, Clark ST, Wang PW, Donaldson SL, Coburn B, Tullis DE, et al. A genome-wide association analysis reveals a potential role for recombination in the evolution of antimicrobial resistance in *Burkholderiamultivorans*. PLoS Pathog. 2018;14(12):e1007453.

31. Song L, Jenner M, Masschelein J, Jones C, Bull MJ, Harris SR, et al. Discovery and biosynthesis of gladiolin: A *Burkholderia gladioli* Antibiotic with Promising Activity against *Mycobacterium tuberculosis*. J Am Chem Soc. 2017;139(23):7974–81.

32. Mullins AJ, Jones C, Bull MJ, Webster G, Parkhill J, Connor TR, et al. Genomic assemblies of members of *Burkholderia* and related genera as a resource for natural product discovery. Microbiol Resour Announc. 2020;9(42): https://doi.org/10.1128/MRA.00485-20

33. Jenner M, Jian X, Dashti Y, Masschelein J, Hobson C, Roberts DM, et al. An unusual *Burkholderia gladioli* double chain-initiating nonribosomal peptide synthetase assembles "fungal" icosalide antibiotics. Chem Sci. 2019;10(21):5489–94.

34.● Jones C, Webster G, Mullins AJ, Jenner M, Bull MJ, Dashti Y, et al. Kill and cure: genomic phylogeny and bioactivity of *Burkholderia gladioli* bacteria capable of pathogenic and beneficial lifestyles. Microb Genom. 2021. Vol. 7, issue (1). https://doi.org/10.1099/mgen.0.000515. **This study shows multiple phylogenomic groups and strains which produce the toxin bongkreckic acid occur within CF isolates of *B. gladioli*.**

35. Bryant JM, Grogono DM, Greaves D, Foweraker J, Roddick I, Inns T, et al. Whole-genome sequencing to identify transmission of *Mycobacterium abscessus* between patients with cystic fibrosis: a retrospective cohort study. Lancet. 2013;381(9877):1551–60.

36. Bryant JM, Grogono DM, Rodriguez-Rincon D, Everall I, Brown KP, Moreno P, et al. Emergence and spread of a human-transmissible multidrug-resistant nontuberculous mycobacterium. Science. 2016;354(6313):751–7.

37 ●. Bryant JM, Brown KP, Burbaud S, Everall I, Belardinelli JM, Rodriguez-Rincon D et al. Stepwise pathogenic evolution of *Mycobacterium abscessus*. Science. 2021;372(6541). https://doi.org/10.1126/science.abb8699. **This study provides the most current knowledge of how *M. abscessus* strains evolve during CF lung infection.**

38. Flight WG, Smith A, Paisey C, Marchesi JR, Bull MJ, Norville PJ, et al. Rapid detection of emerging pathogens and loss of microbial diversity associated with severe lung disease in cystic fibrosis. J Clin Microbiol. 2015;53(7):2022–9.

39. Hogardt M, Ulrich J, Riehn-Kopp H, Tummler B. EuroCareCF quality assessment of diagnostic microbiology of cystic fibrosis isolates. J Clin Microbiol. 2009;47(11):3435–8.

40. Spilker T, Vandamme P, LiPuma JJ. Identification and distribution of *Achromobacter* species in cystic fibrosis. J Cystic Fibrosis. 2013;12(3):298–301.

41. Gabrielaite M, Nielsen FC, Johansen HK, Marvig RL. *Achromobacter* spp. genetic adaptation in cystic fibrosis. Microb Genom. 2021;7(7). https://doi.org/10.1099/mgen.0.000582

42. Veschetti L, Sandri A, Patuzzo C, Melotti P, Malerba G, Lleo MM. Genomic characterization of *Achromobacter* species isolates from chronic and occasional lung infection in cystic fibrosis patients. Microb Genom. 2021;7(7). https://doi.org/10.1099/mgen.0.000606

43. Esposito A, Pompilio A, Bettua C, Crocetta V, Giacobazzi E, Fiscarelli E, et al. Evolution of *Stenotrophomonasmaltophilia* in cystic fibrosis lung over chronic infection: a genomic and phenotypic population study. Front Microbiol. 2017;8:1590.

44. Akil N, Muhlebach MS. Biology and management of methicillin resistant *Staphylococcus aureus* in cystic fibrosis. Pediatr Pulmonol. 2018;53(S3):S64–74. https://doi.org/10.1002/ppul.24139

45. ●● Jolley KA, Bray JE, Maiden MCJ. Open-access bacterial population genomics: BIGSdb software, the PubMLST.org website and their applications. Wellcome Open Res. 2018;3:124. https://doi.org/10.12688/wellcomeopenres.14826.1.**This study describes the genomic database and the suite of tools such as MLST freely available for international bacterial strain analysis**.

46. Petit RA 3rd, Read TD. *Staphylococcus aureus* viewed from the perspective of 40,000+ genomes. PeerJ. 2018;6:e5261. https://doi.org/10.7717/peerj.5261.

47. Bernardy EE, Petit RA, 3rd, Raghuram V, Alexander AM, Read TD, Goldberg JB. Genotypic and phenotypic diversity of *Staphylococcus aureus* isolates from cystic fibrosis patient lung infections and their interactions with *Pseudomonas aeruginosa*. mBio. 2020;11(3). https://doi.org/10.1128/mBio.00735-20.

48. Ebbing R, Robertson CF, PJ R. *Haemophilus influenzae* and *Haemophilus parainfluenza* in Cystic Fibrosis: J Medical Microbiol Diag. 2015;S5:004. https://doi.org/10.4172/2161-0703.S5-004.

49. Malhotra S, Hayes D Jr, Wozniak DJ. Mucoid *Pseudomonas aeruginosa* and regional inflammation in the cystic fibrosis lung. J Cystic Fibrosis. 2019;18(6):796–803.

50. Oliver A, Mena A. Bacterial hypermutation in cystic fibrosis, not only for antibiotic resistance. Clin Microbiol Infect. 2010;16(7):798–808.

51. Wolter DJ, Onchiri FM, Emerson J, Precit MR, Lee M, McNamara S, et al. Prevalence and clinical associations of *Staphylococcus aureus* small-colony variant respiratory infection in children with cystic fibrosis (SCVSA): a multicentre, observational study. Lancet Respir Med. 2019;7(12):1027–38.

52.● Reimer LC, SardaCarbasse J, Koblitz J, Ebeling C, Podstawka A, Overmann J. BacDive in 2022: the knowledge base for standardized bacterial and archaeal data. Nucleic Acids Res. 2022;50(D1):D741–6. **This study overviews the BacDive database as a useful resource to search for bacterial strains deposited within recognised repositories.**

53. King JA, Nichols AL, Bentley S, Carr SB, Davies JC. This study overviews the BacDive database as a useful resource to search for bacterial strains deposited within recognised repositories. An update on CFTR modulators as new therapies for cystic fibrosis. Paediatr Drugs. 2022;24(4):321–33.

54. Coenye T, Vandamme P, LiPuma JJ, Govan JR, Mahenthiralingam E. Updated version of the *Burkholderiacepacia* complex experimental strain panel. J Clin Microbiol. 2003;41(6):2797–8.

55. Schwarz C, Eschenhagen P, Bouchara JP. Emerging fungal threats in cystic fibrosis. Mycopathologia. 2021. https://doi.org/10.1007/s11046-021-00574-w

56. Hibbett D, Abarenkov K, Koljalg U, Opik M, Chai B, Cole J, et al. Sequence-based classification and identification of Fungi. Mycologia. 2016;108(6):1049–68.

57. Scott JE, O'Toole GA. The Yin and Yang of *Streptococcus* lung infections in cystic fibrosis: a model for studying polymicrobial interactions. J Bacteriol. 2019;201(11). https://doi.org/10.1128/JB.00115-19

58. Waters VJ, Kidd TJ, Canton R, Ekkelenkamp MB, Johansen HK, LiPuma JJ, et al. Reconciling antimicrobial susceptibility testing and clinical response in antimicrobial treatment of chronic cystic fibrosis lung infections. Clin Infect Dis. 2019;69(10):1812–6.

59. O'Toole GA, Crabbe A, Kummerli R, LiPuma JJ, Bomberger JM, Davies JC, et al. Model systems to study the chronic polymicrobial infections in cystic fibrosis: current approaches and exploring future directions. mBio. 2021;12(5):e0176321.

60. Kirchner S, Fothergill JL, Wright EA, James CE, Mowat E, Winstanley C. Use of artificial sputum medium to test antibiotic efficacy against *Pseudomonas aeruginosa* in conditions more relevant to the cystic fibrosis lung. J Vis Exp. 2012;64:e3857.