CULTURE-INDEPENDENT MULTILOCUS SEQUENCE TYPING OF *PSEUDOMONAS AERUGINOSA* FOR CROSS-INFECTION SCREENING

Philip J. Mitchelmore [a,b,*], Nicholas J. Withers [b], Christopher D. Sheldon [b], Chris J. Scotton [a], Alan R. Brown [c]

[a] Institute of Biomedical and Clinical Sciences, St Luke's Campus, University of Exeter Medical School, Exeter, UK.

[b] Department of Respiratory Medicine, Royal Devon and Exeter NHS Foundation Trust, Barrack Road, Exeter, UK.

[c] Biosciences, College of Life and Environmental Sciences, Stocker Road, University of Exeter, UK.

Running title: Culture-independent genotyping of *P. aeruginosa*

*Corresponding author. Email address: P.Mitchelmore@exeter.ac.uk

Tel: +441392 411611

ABSTRACT

The genotyping of pathogens within cystic fibrosis cohorts is an important process, enabling the detection of transmissible and clinically-important strains. Traditionally this has been via culture-dependent processes. However, culture-independent investigation of respiratory samples is becoming more common, with such approaches highlighting the limitations of culture-based methods. In this study we describe the culture-independent application of multi-locus sequence typing (MLST) for *Pseudomonas aeruginosa*, performed on DNA extracted from the sputa of cystic fibrosis patients. We compare the output to conventional culture-dependent MLST applied to the same samples and demonstrate high concordance. Culture-independent MLST enabled genotyping of culture-negative samples in patients from whom *P. aeruginosa* was intermittently isolated, and revealed the hidden presence of transmissible strains. Culture-independent MLST is also capable of highlighting samples containing multiple strains, albeit inconsistently. We conclude that culture-independent MLST can be a useful genotyping tool for screening cohorts and identifying patients that warrant further detailed investigation.

Keywords: Cystic fibrosis; *Pseudomonas aeruginosa;* MLST; sputum; genotyping; strain; infection control.

1. Introduction

Cross-infection within cystic fibrosis (CF) cohorts with significant pathogens is a well-recognised clinical risk for patients. Previous epidemiological work has revealed transmissible strains with critical clinical consequences, including the Liverpool Epidemic Strain (LES) of *Pseudomonas aeruginosa* and the ET12 strain of *Burkholderia cenocepacia (Saiman et al. 2014; Mahboubi et al. 2016).* Consequently, in many CF units patients are strictly segregated or cohorted (based on microbiological status) to reduce the risk of cross-infection. However, both patients and staff may breach these infection control standards, and prolonged viability of aerosolised pathogens may facilitate cross-infection despite strict adherence to protocols. Consequently, effective cross-infection surveillance is a critical component of the management of CF patients, and requires pro-active longitudinal genotyping of bacterial strains (Saiman et al. 2014). However, robust cross-infection studies can carry significant resource and cost implications, and there is no consensus on appropriate follow-up times. There would be considerable value in having a readily-applicable method for routine genotyping that could identify when targeted high-resolution strain genotyping within a cohort is warranted.

Strain genotyping is typically performed on cultured isolates of the organism being investigated and can utilise several methods that assess either gene-specific or genome-wide sequence variation to differentiate between strains. However, regardless of the method(s) used, the necessity for culture introduces inherent delay and bias into the process. For example, false-negative culture results from respiratory samples are widely-documented, particularly for species that are not adequately supported by routinely-used culture conditions (e.g., *Streptococcus* species) (Mahboubi et al. 2016). Even for organisms that can be readily cultured, it is not uncommon for patients to be intermittently culture-positive. An additional concern with culture-

dependent genotyping is the likelihood of missing the coexistence of multiple strains if only single representative colonies are genotyped. A failure to identify multiple strains can lead to shared strains and potential cross-infection events being missed. Indeed, in our previous study of *P. aeruginosa* within two respiratory cohorts, we identified likely cross-infection between patients harbouring multiple strains (Mitchelmore et al. 2018). With the recognised limitations of culture-dependent methods, there is increasing development and application of culture-independent methodologies that utilise DNA extracted directly from clinical samples (e.g., sputum). These culture-independent methods have been widely-applied by the research community to the analysis of microbial communities within polymicrobial infections (including those of the CF lung). These methods are likely to become common-place in clinical practise alongside traditional culture methods. However, so far they have rarely been applied to strain genotyping.

Our present study focuses on multilocus sequence typing (MLST), a well-established genotyping method that differentiates strains on the basis of sequence variation within PCR-amplified fragments of selected genes (Maiden et al. 1998). As a genotyping method, MLST does not provide the level of resolution offered by Whole Genome Sequencing (WGS), but it is a simpler approach which provides an unambiguous output that is transferrable between laboratories. MLST is therefore more appropriate as a routine screening tool, and more widely available and accessible for low-resource settings. Even in high-resource settings, MLST in an invaluable method for identifying patients harbouring shared strains that warrant further targeted investigation by WGS, thus avoiding the resource implications of WGS isolates from entire cohorts (Mitchelmore et al. 2018).

MLST is almost exclusively applied to the genotyping of cultured isolates, but in the present study we assess its utility as a culture-independent method for *P. aeruginosa* genotyping. Culture-independent MLST has been reported once previously, with Drevinek *et al*. reporting its use for identifying and genotyping *Burkholderia cepacia* complex (Bcc) directly from CF sputum (Drevinek et al. 2010). That study demonstrated MLST's ability to provide strain-level identification direct from sputum, including from culture-negative samples. However, culture-based MLST was not performed alongside the culture-independent approach, and thus concordance between the two could not be assessed. Furthermore, a potential strength of culture-independent MLST is its ability to reveal coexistence of multiple strains. Drevinek *et al*. were able to identify discrete allele sequences from sputum harbouring two distinct strains of *Burkholderia contaminans (Drevinek et al. 2010).* However, they were only able to do so because of the uniqueness of the alleles identified, reflecting the small number of *B. contaminans* isolates within the Bcc MLST database at the time of their study (20-30 isolates, based on date of entry). In contrast, the *P. aeruginosa* MLST database (as of January 2020) contains almost 7,000 isolates, which may hinder the ability to identify unique alleles from mixed samples.

With these issues in mind, we assessed culture-independent MLST of *P. aeruginosa* direct from sputum, comparing results with those obtained by culture-dependent MLST and assessing performance from samples containing multiple strains. In so doing, we evaluated culture-independent MLST as a tool for routine cross-infection surveillance that could help rapidly identify significant pathogenic strains and the need for further investigation by high-resolution genotyping.

2. Materials and methods

2.1. Patient recruitment

Patients were opportunistically recruited and consented to donate sputum samples to the local National Institute of Health Research (NIHR) Clinical Research Facility Tissue Bank while providing samples for standard clinical investigation. Patients were eligible for recruitment if they had a confirmed diagnosis of CF and *P. aeruginosa* had previously been isolated from their sputum. As well as this study, these patients were also recruited as part of a culture-dependent cross-infection study alongside the local non-CF bronchiectasis cohort (Mitchelmore et al. 2018). Ethical approval was obtained through the Tissue Bank (11/SW/0018).

2.2. Sputum processing

Sputum samples were processed by the microbiology department at the Royal Devon & Exeter Hospital. This involved the addition of Mucolyse™ Sputum Digestant (Pro-Lab Diagnostics) prior to sputum culture as per the standard operating procedure. If *P. aeruginosa* was cultured, 10 representative colonies were stored on Microbank™ microbial storage beads (Pro-Lab Diagnostics) at -80 °C for culture-dependent genotyping. In addition, aliquots of each sputum sample were stored at -80 °C for culture-independent genotyping.

To facilitate culture-independent genotyping, DNA was extracted from 0.4 mL of Mucolyse-diluted sputum using the QIAamp DNA Mini Kit (Qiagen) according to manufacturer's instructions. To facilitate culture-dependent genotyping, *P. aeruginosa* isolates were cultured in LB broth, with DNA subsequently extracted from 2 mL of overnight culture using the GeneJET Genomic DNA Purification Kit (ThermoFisher Scientific) according to the manufacturer's protocol for gram-negative bacteria. DNA was quantified using a NanoDrop 1000 Spectrophotometer (ThermoFisher Scientific).

2.3. Culture-dependent and culture-independent MLST

Culture-dependent MLST, performed as previously described (Curran et al. 2004), was conducted as part of our previously published cross-infection study encompassing CF and non-CF bronchiectasis patients on isolates which had unique Random Amplified Polymorphic DNA (RAPD) profiles (Mitchelmore et al. 2018). For culture-independent MLST, sputum-derived DNA was initially screened by PCR for the *acsA* and *guaA* loci. Samples that were positive for both loci were then subjected to PCR-amplification of the five remaining loci within the MLST scheme. Samples that were negative for both *acsA* and *guaA* were identified as "PCR-negative" for *P. aeruginosa*. On no occasion was a sample consistently negative for one locus but positive for the other.

Whilst all loci were readily-amplified from culture-derived DNA, successful amplification of the *nuoD* locus was inconsistent from sputum-derived DNA. Consistent amplification of the *nuoD* locus from sputum DNA was achieved after adopting some of the modifications suggested by van Mansfeld *et al*. (van Mansfeld et al. 2009), namely the addition of Q buffer (Qiagen) to the PCR reaction and the use of revised *nuoD* primers.

Prior to sequencing, all PCR products were purified using the GeneJET Gel Extraction kit (ThermoFisher Scientific). The sequencing of the purified PCR products was performed by Eurofins Genomics. Forward and reverse sequence reads were aligned using Invitrogen Vector NTI software (ThermoFisher Scientific), with the consensus sequence used to interrogate the MLST database (https://pubmlst.org/paeruginosa; (Jolley and Maiden 2010)). Sequence chromatograms were viewed using SnapGene® (GSL Biotech; available at www.snapgene.com).

2.4. Assessment of MLST from mixed samples

In addition to analysing sputum samples potentially harbouring multiple strains, DNA was extracted from two different *P. aeruginosa* isolates (of differing MLST types) and mixed in

varying ratios (10:90, 30:70, 50:50, 70:30 and 90:10) prior to use for MLST analysis. Resulting consensus sequences were analysed as described above, and sequence chromatograms reviewed for evidence of distinct alleles.

3. Results

3.1. Patient cohort

Thirty-two patients were recruited between September 2015 and January 2016. Relevant cohort demographics are shown in Supplementary Table 1. Sputum was obtained from 31 of the 32 patients; 22 of these sputum samples were culture-positive for *P. aeruginosa*.

3.2. Concordance between culture-dependent and culture-independent MLST

Culture-dependent MLST was successfully performed on 26 relevant isolates (based on RAPD profiles) from the 22 culture-positive sputum samples as reported previously (Mitchelmore et al. 2018). This culture-based analysis revealed a single sequence type (ST) of *P. aeruginosa* in all 22 patients (Supplementary Table 2), of which eight were novel (based on STs documented within the MLST database).

For the culture-independent MLST, DNA was successfully extracted from all 31 sputum samples, with DNA yields ranging from 1.1–76.0 µg. Culture-independent MLST data were successfully obtained from 23 of these 31 samples, including three that were culture-negative for *P. aeruginosa*. However, two culture-positive samples were PCR-negative (Table 2).

In total, we therefore had 20 patients for whom we had data via both culture-dependent and culture-independent MLST. As the MLST scheme is based on seven loci, we had 140 allele sequences to compare from these 20 patients. Of the 140 alleles, 132 (94 %) had exactly the

same sequence by both methods. When considered at a patient level, 18 of the 20 patients had complete concordance between culture-dependent and culture-independent MLST, with exact matches for all seven loci. The nature of the discrepancies for the remaining two patients are detailed in Table 2. In brief, patient PIC13 appears to harbour coexisting single locus variants of ST217 (the Manchester Epidemic Strain, MES), that were variably detected by the culture-dependent and culture-independent methods. Therefore, whilst the two MLST approaches did not provide a complete match for this patient, they did identify the same clonal complex. A more striking discrepancy between culture-dependent and culture-independent MLST arose in patient PIC32. Culture-dependent MLST identified ST195, whereas the output from culture-independent MLST consistently differed at all seven loci, with the observed allele profile matching ST217. We believe this to be a consequence of the sputum sample harbouring these two distinct coexisting strains that were variably detected by the different MLST methods (discussed further in section 3.3).

Table 2. Discrepancies between culture-dependent and culture-independent MLST.

		MLST allele								
Patient	**MLST method**	***acsA***	***aroE***	***guaA***	***mutL***	***nuoD***	***ppsA***	***trpE***	**ST**	**Relevant notes**
PIC13	Culture-dependent	28	5	11	18	4	13	271	3324[a]	Single nucleotide variants; *trpE*-271 and *trpE*-3 differ by one nucleotide
	Culture-independent	28	5	11	18	4	13	3	217[b]	
PIC32	Culture-dependent	89	30	64	26	48	24	32	195	All 7 alleles differ, suggesting coexistence of ST195 & ST217
	Culture-independent	28	5	11	18	4	13	3	217[b]	
PIC16	Culture-dependent	*Culture-negative*								
	Culture-independent	39	5	9	16	27	5	2		
PIC27	Culture-dependent	*Culture-negative*								
	Culture-independent	6	5	11	3	4	23	1	146[c]	
PIC29	Culture-dependent	*Culture-negative*								
	Culture-independent	6	5	1	1	1	12	1	395	
PIC17	Culture-dependent	39	5	11	28	4	4	63	379	
	Culture-independent	*PCR-negative*								
PIC28	Culture-dependent	6	301	11	4	3	23	272	3336[a]	
	Culture-independent	*PCR-negative*								

[a] Novel sequence type identified in the present study; [b] ST217, Manchester epidemic strain; [c] ST146, Liverpool epidemic strain. The culture-independent MLST profile of PIC16 did not match a documented sequence type (ST) within the MLST database, but is a single locus (*mutL*) variant of ST277. The observed discrepancies between the culture-dependent and culture-independent MLST for patients PIC13 and PIC32 were verified by repeat MLST analysis on all differing loci.

Via culture-dependent analysis, we had previously identified two shared strains within our CF cohort, with three patients harbouring ST27 and two (siblings) harbouring ST146 (the Liverpool epidemic strain, LES) (Mitchelmore et al. 2018). Highlighting its value as a cross-infection screening tool, the culture-independent MLST described herein resulted in the identification of further instances of shared strains within our cohort (Table 3), including an additional patient harbouring the LES and two patients harbouring the MES.

Table 3. Shared sequence types by culture-dependent and culture-independent MLST analysis.

MLST type	Number of patients with MLST type	
	Culture-dependent MLST	**Culture-independent MLST**
ST27	3	3
ST146[a]	2	3
ST217[b]	0	2
ST395	1	2

[a] Liverpool epidemic strain; [b] Manchester epidemic strain.

3.3. The impact of coexisting strains on culture-independent MLST

As highlighted above, culture-dependent and culture-independent MLST results from patient PIC32 strongly suggested the coexistence of ST195 and ST217 (MES). Accordingly, relevant sequence chromatograms were visually inspected for evidence of the coexisting strains, as judged by the presence of secondary peaks (verified by both forward and reverse sequence reads) at the expected nucleotide positions.

Upon examination of the sequence chromatograms from two independent repeats of the culture-independent MLST of PIC32, we did indeed observe secondary peaks indicative of the two coexisting strains, albeit not consistently. This is illustrated by Fig. 1, which shows a fragment of the *aroE* locus chromatogram from the two independent repeats. The consensus sequences

obtained from both repeats match that expected of ST217 (*aroE*-5). However, the chromatogram from the 1st repeat shows secondary peaks corresponding to the coexisting *aroE*-30 allele of ST195, whilst the chromatogram from the 2nd repeat does not (Fig. 1). Across all seven MLST loci, there are 54 nucleotide differences between ST195 and ST217. When assessing these 54 nucleotide positions in both forward and reverse sequence reads from both independent repeats, discernible secondary peaks consistent with the coexistence of ST195 and ST217 were evident at 68% of positions. A consistent observation from our examination of sequence chromatograms was that those derived from culture-independent MLST had a noisier baseline than those from culture-dependent MLST (Supplementary Figure 1). This noise within the baseline frequently hindered the ability to identify discernible secondary peaks indicative of coexisting strains.

Upon review of sequence chromatograms from other patients, only one further case of coexisting strains was identified within our cohort. PIC27 (a culture-negative sample) yielded a culture-independent MLST profile matching ST146 (LES). However, the chromatograms for the *aroE* and *trpE* loci revealed coexistence of a single nucleotide variant in each locus (*aroE* C341T; *trpE* C179T), neither of which matched any documented alleles within the *P. aeruginosa* MLST database.

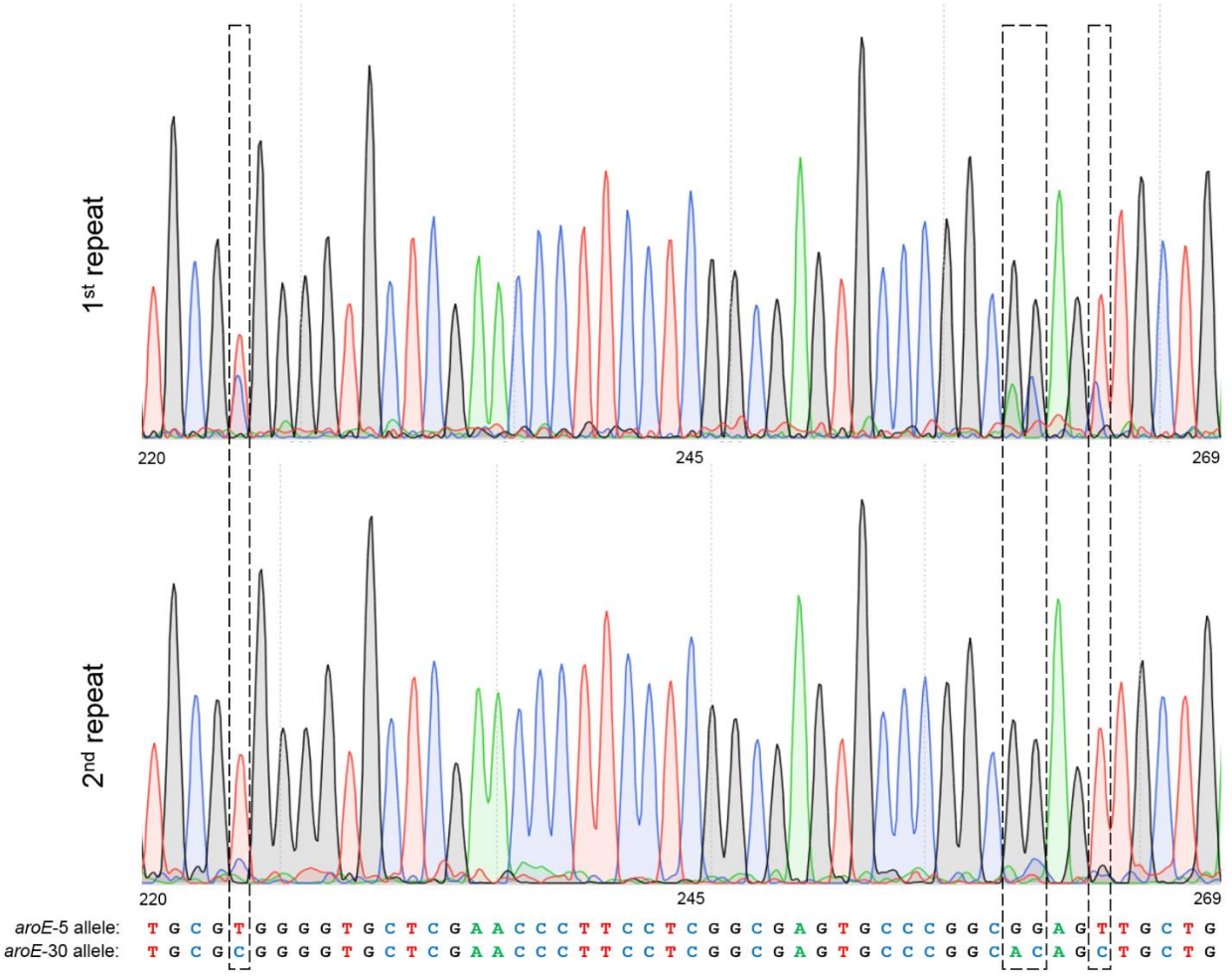

Figure 1. Representative sequence chromatograms from two independent repeats of the culture-independent MLST analysis of patient sample PIC32, showing fragments of the *aroE* locus. Nucleotide coordinates relate to their position within the *aroE* allele, as defined within the *P. aeruginosa* MLST scheme. Nucleotide positions at which the *aroE*-5 and *aroE*-30 alleles differ are highlighted. In both independent repeats, the primary chromatogram peaks (and thus the consensus sequence) exactly match the *aroE*-5 allele. However, the two independent repeats differ in their ability to resolve secondary peaks at nucleotide positions 224, 260, 261 and 264, which would be indicative of the coexisting *aroE*-30 allele.

In order to further investigate the impact of multiple strains on MLST outcomes, DNA from two *P. aeruginosa* isolates of differing MLST types was mixed in different ratios and MLST performed from the resulting mixed template. Representative chromatograms from the analysis of the *trpE* locus from these mixed samples are shown in Fig. 2. The constituent samples had been identified as harbouring *trpE*-7 and *trpE*-47 allele sequences, which differ at 13 nucleotide positions. The ability to discern these two alleles was inconsistent across the different template ratios tested. Sequencing the *trpE* locus from a 10:90 mixture of *trpE*-7:*trpE*-47 resulted in a consensus sequence that did not consistently match either allele (Fig. 2A). In contrast, analysis of a 30:70 *trpE*-7:*trpE*-47 mixture resulted in a consensus sequence that fully matched *trpE*-7 but with secondary chromatogram peaks consistent with *trpE*-47 (Fig. 2B), whilst a 50:50 template mixture resulted in a *trpE*-7 consensus with no evidence of secondary peaks at any nucleotide positions (Fig. 2C). Overall, upon examination of forward and reverse sequence reads from all template ratios tested, secondary peaks indicative of the coexisting alleles were observed at only 40% of the expected nucleotide positions.

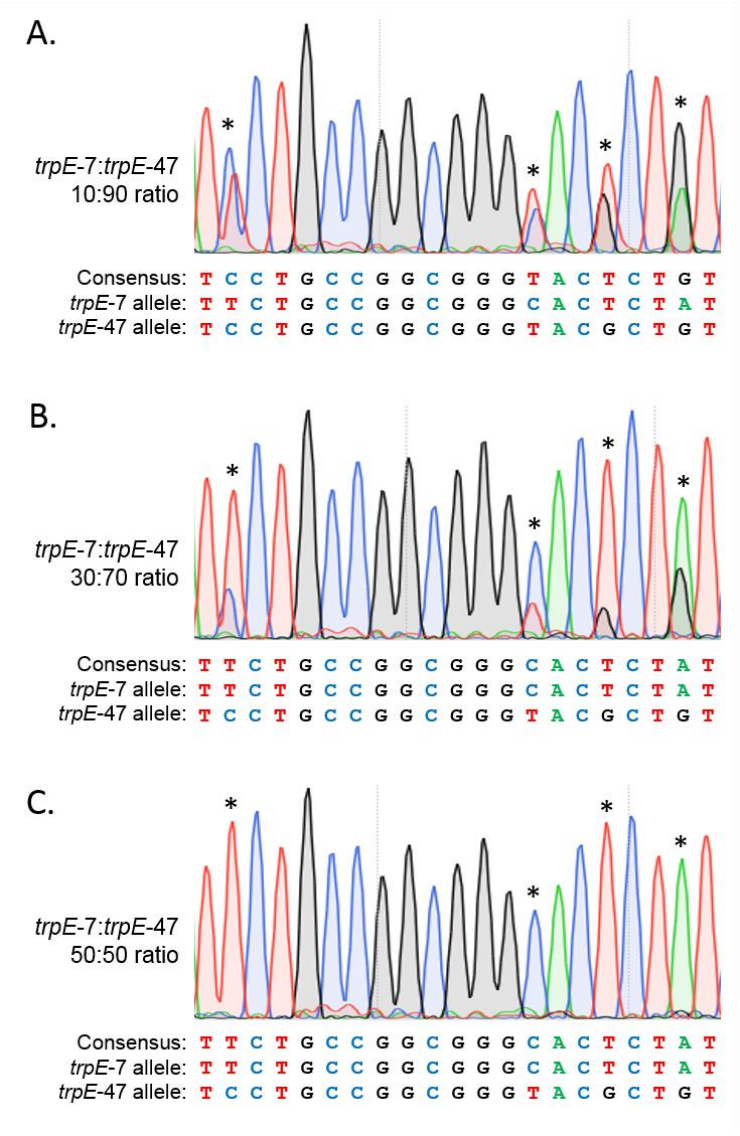

Figure 2. Representative sequence chromatograms from MLST analysis of mixed template DNA harbouring known differences in the *trpE* locus. Genomic DNA from two *P. aeruginosa* isolates of differing MLST types (incorporating *trpE*-7 and *trpE*-47 alleles) were mixed in various ratios prior to use as template in MLST analysis of the *trpE* locus. Consensus sequences obtained are shown alongside the corresponding *trpE*-7 and *trpE*-47 sequences. The region depicted corresponds to nucleotides 330-350 of the *trpE* allele sequence, as defined within the MLST scheme. Asterisks indicate nucleotide positions where the *trpE*-7 and *trpE*-47 alleles vary.

4. Discussion

Monitoring for cross-infection in CF cohorts is crucial. Episodes put patients at risk and demand the re-assessment of infection control standards. However, robust cross-infection studies can be costly and time-consuming to perform when applied to whole cohorts, particularly if using multiple methodologies (that increasingly include WGS). There is considerable value in adopting a suitable genotyping method as a cohort-wide screening tool for the identification of patients harbouring shared strains that would benefit from further investigation by high-resolution methods. Culture-independent MLST may lend itself to such an approach. Herein we have provided the first report of culture-independent MLST genotyping of *P. aeruginosa* direct from sputum. This culture-independent approach showed good concordance with culture-dependent MLST of the same sputa, and avoids the need for culture and the associated bias when picking representative colonies from selective plates. We believe that MLST applied in this manner is a valuable tool for the initial genotypic screening of *P. aeruginosa* within cohorts, enabling the rapid identification of patients harbouring unique or shared strains.

Through the culture-independent analysis described herein, we identified patients harbouring shared strains that had been missed by traditional culture-based analysis. These included the LES (ST146) and the MES (ST217), both of which are associated with increased healthcare burden, morbidity and/or mortality (Aaron et al. 2010; Al-Aloul et al. 2004; Jones et al. 2001; Jones et al. 2010). In our CF unit, a strict segregation policy is employed. Consequently, no patients should come into direct contact with each other or use the same clinical space within any one hospital session. The presence of transmissible strains in multiple patients raises obvious concerns about infection control standards. However, two of the three LES-infected patients in our cohort are close household contacts, and all three LES-infected patients had *P. aeruginosa* isolated from

their sputum in other CF units before having their care transferred to us. Similarly, the two MES-infected patients in our cohort have long-harboured *P. aeruginosa*, and both patients previously received care at the same CF unit in another city. The other shared strains observed herein were ST27 (n = 3) and ST395 (n = 2). Our previous WGS-based analysis of these sequence types (which incorporated all three ST27-infected patients and one of the ST395-infected patients) ruled out cross-infection, favouring instead independent acquisition (Mitchelmore et al. 2018). The identification of an additional case of ST395 infection in a culture-negative sample justifies high-resolution genotyping of this patient's strain upon next isolation.

Whilst culture-independent MLST has previously been reported for Bcc (Drevinek et al. 2010), our studies represent the first comparison of culture-dependent and culture-independent analyses. Nineteen of the 20 patients for whom we acquired both culture-dependent and culture-independent MLST data had either complete sequence concordance across all seven loci (18 patients) or had discrepancies within only a single locus (and thus would be considered of the same clonal complex; one patient). Only in patient PIC32 were discordant outcomes obtained by the two MLST approaches, reflecting the coexistence of at least two distinct strains. The culture-independent analysis also enabled genotyping information to be obtained from two culture-negative patients who had a prior history of *P. aeruginosa* infection, highlighting the value of this culture-independent approach for those in whom the pathogen is only intermittently isolated from. Conversely, culture-independent MLST failed to yield results from two patients who were culture-positive, likely as a consequence of low yield of template DNA. One of the PCR-negative culture-positive samples (PIC28) had the joint lowest sputum DNA yield (1.1 µg) of all samples processed. The other PCR-negative culture-positive sample (PIC17) yielded 8.2 µg of DNA (closer to the median yield of 12.4 µg). However, as sputum-extracted DNA is a complex

300 mixture of mammalian and microbial DNA, *P. aeruginosa*-derived DNA may be a minor
301 component of the overall yield obtained (Maiden et al. 1998).

302 A major aim of our study was to assess how culture-independent MLST results are influenced by
303 the coexistence of multiple strains within a single sample, and to assess the ability of culture-
304 independent MLST to detect such samples. A full appreciation of the strain diversity harboured
305 by individual patients is important; without such knowledge, cross-infection and/or super-
306 infection events may be missed. From our studies, it is evident that culture-independent MLST
307 can reveal the coexistence of multiple strains, but not consistently. The absence of secondary
308 peaks within sequence chromatograms cannot exclude the possibility of multiple strains being
309 present, and when secondary peaks are observed, it is not possible to discern (with confidence)
310 the identity of the coexisting strains. In part, this is due to the fact that it is unlikely that all
311 nucleotide variants between coexisting strains will be represented as secondary peaks; across our
312 combined analysis of PIC32 and the mixed *trpE* allele samples, only 57% of the expected
313 nucleotide variants were identifiable as secondary peaks. Furthermore, 'hybrid' sequences can be
314 generated from samples containing mixed alleles as shown by our artificial mixing of two strains
315 (*trpE* alleles; Fig. 2A). Despite these limitations, culture-independent MLST can certainly aid the
316 identification of sputum samples harbouring multiple strains that may be missed by culture-based
317 methods. Detection of coexisting strains by culture demands genotyping of multiple
318 representative colonies. The challenge of this is strikingly illustrated in the present study, where
319 culture-dependent analysis of 10 colonies from PIC32 only revealed ST195, whereas two
320 independent repeats of culture-independent MLST from the same sputum consistently identified
321 ST217. In this culture-independent analysis, the primary peaks in the sequence chromatograms
322 were entirely consistent with ST217 alleles, whilst the secondary peaks (when present)

consistently matched ST195 alleles. It is clear from our wider studies that this clear ST-specific distinction between primary and secondary peaks, and thus our ability to identify ST217 as the likely co-infecting strain, was entirely fortuitous. Given the complete match of all seven loci to ST217, it seems unlikely that hybrid sequences are confounding our culture-independent analysis of PIC32. Whilst further culture-based investigation of this patient is clearly required to confirm the coexistence of ST217, plausibly a very large number of colonies may need to be picked from plates and at intervals to confirm this. This highlights a potential bias of the culture-based approach to strains which are readily cultivatable in the laboratory.

Whilst culture-independent MLST was inconsistent in its ability to detect coexisting strains and has the potential to yield hybrid sequences, we believe it remains a useful screening tool to guide further investigation but not as a definitive technique in potentially complex samples. While higher-resourced settings have greater availability and capacity for WGS-based studies, culture-independent MLST within these settings still offers potential benefits for certain scenarios. This would include the interrogation of patients from whom a pathogen is only intermittently isolated from, conceivably due to long-term suppressive therapies such as inhaled colistin and aminoglycosides. Higher-resourced settings are also more likely to see culture-independent microbiological investigation being performed alongside tradition culture of sputum become more common, and therefore have the template of extracted DNA available for rapid screening. Additional uses of this approach could also be used for longitudinal assessment of strain retention or clearance. This would be of particular interest in researching eradication attempts and the aetiology of subsequent re-infection. In lower-resourced settings, culture-independent MLST can be adapted further to efficiently utilise capacity. For example, by initially

interrogating a limited number of loci, ongoing examination could be focused on those with shared (or potentially clinically significant) sequences without needing to sequence all loci.

Many different genotyping techniques exist, and the selection of the technique used will be dependent on resources, the skills of the investigators, and the research question being asked. Other commonly used techniques include Pulsed Field Gel Electrophoresis (PFGE) and Multiple Locus Variable-number Tandem Repeat Analysis (MLVA or VNTR). However, MLST has a higher turnover of specimens and requires less laboratory skills and experience than PFGE or MLVA (Tümmler 2020). The ease of extraction of MLST from WGS data also allows compatible alignment with an increasingly common high-resolution technique and will remain relevant in the future. Whilst other techniques may be more discriminatory than MLST, this is less of an issue when being used to screen for potential shared strains in the manner we describe.

5. Conclusion

In summary, we conclude that culture-independent MLST is a useful screening tool for the initial genotypic characterization of *P. aeruginosa* within CF cohorts, allowing identification of individual patients that warrant further investigation with higher resolution culture-based methods. Unhindered by the bias of culture, culture-independent MLST can aid the identification of samples containing multiple distinct strains, and can yield clinically-important genotype data from culture-negative samples. With MLST schemes established for many notable CF pathogens (Enright et al. 2000; Kim et al. 2013; Meats et al. 2003; Spilker et al. 2009), there is clearly scope for the wider incorporation of such methods into routine clinical practice as DNA extraction from sputum becomes more commonplace.

Acknowledgements

The authors would like to thank all the patients who participated in this study; Hilary Mortimer, Sarah Murray, Rachel Rogers, Jayne Trott, Roseanne Lowless and Miriam Green for their help in obtaining samples; the Microbiology Department at the Royal Devon and Exeter Hospital, particularly Nigel Richardson, Dr Jo Randall and Dr Cressida Auckland, and Dr Matthew Robinson for laboratory guidance. The project was supported by NIHR Exeter Clinical Research Facility, and the authors thank Chloe Slade and Kathryn Shears for their assistance in managing samples and data.

Funding: This study was funded though charitable donations and a Small Grants Award by the Research and Development department at the Royal Devon and Exeter NHS Foundation Trust. The funding sources had no input into this study or its submission for publication.

Conflict of Interest Statement

No conflicts of interest to declare.

References

Aaron, S. D., K. L. Vandemheen, K. Ramotar, T. Giesbrecht-Lewis, E. Tullis, A. Freitag, N. Paterson, M. Jackson, M. D. Lougheed, C. Dowson, V. Kumar, W. Ferris, F. Chan, S. Doucette, and D. Fergusson. Infection with transmissible strains of Pseudomonas aeruginosa and clinical outcomes in adults with cystic fibrosis. Jama 2010; 304: 2145-53. doi: 10.1001/jama.2010.1665.

Al-Aloul, M., J. Crawley, C. Winstanley, C. A. Hart, M. J. Ledson, and M. J. Walshaw. Increased morbidity associated with chronic infection by an epidemic Pseudomonas aeruginosa strain in CF patients. *Thorax* 2004; 59: 334-6. doi: 10.1136/thx.2003.014258.

Curran, B., D. Jonas, H. Grundmann, T. Pitt, and C. G. Dowson. Development of a multilocus sequence typing scheme for the opportunistic pathogen Pseudomonas aeruginosa. J Clin Microbiol 2004; 42: 5644-9. doi: 10.1128/JCM.42.12.5644-5649.2004.

Drevinek, P., S. Vosahlikova, K. Dedeckova, O. Cinek, and E. Mahenthiralingam. Direct culture-independent Strain typing of Burkholderia cepacia complex in sputum samples from patients with cystic fibrosis. J Clin Microbiol 2010; 48: 1888-91. doi: 10.1128/JCM.02359-09.

Enright, M. C., N. P. Day, C. E. Davies, S. J. Peacock, and B. G. Spratt. Multilocus sequence typing for characterization of methicillin-resistant and methicillin-susceptible clones of Staphylococcus aureus. J Clin Microbiol 2000; 38: 1008-15. doi: 10.1128/JCM.38.3.1008-1015.2000.

Jolley, K. A., and M. C. Maiden. BIGSdb: Scalable analysis of bacterial genome variation at the population level. BMC Bioinformatics 2010; 11: 595. doi: 10.1186/1471-2105-11-595.

Jones, A. M., M. E. Dodd, J. Morris, C. Doherty, J. R. Govan, and A. K. Webb. Clinical outcome for cystic fibrosis patients infected with transmissible pseudomonas aeruginosa: an 8-year prospective study. Chest 2010; 137: 1405-9. doi: 10.1378/chest.09-2406.

Jones, A. M., J. R. Govan, C. J. Doherty, M. E. Dodd, B. J. Isalska, T. N. Stanbridge, and A. K. Webb. Spread of a multiresistant strain of Pseudomonas aeruginosa in an adult cystic fibrosis clinic. Lancet 2001; 358: 557-8. doi: 10.1016/s0140-6736(01)05714-2.

Kim, S. Y., Y. A. Kang, I. K. Bae, J. J. Yim, M. S. Park, Y. S. Kim, S. K. Kim, J. Chang, and S. H. Jeong. Standardization of multilocus sequence typing scheme for Mycobacterium abscessus and Mycobacterium massiliense. Diagn Microbiol Infect Dis 2013; 77: 143-9. doi: 10.1016/j.diagmicrobio.2013.06.023.

Mahboubi, M. A., L. A. Carmody, B. K. Foster, L. M. Kalikin, D. R. VanDevanter, and J. J. LiPuma. Culture-Based and Culture-Independent Bacteriologic Analysis of Cystic Fibrosis Respiratory Specimens. J Clin Microbiol 2016; 54: 613-9. doi: 10.1128/JCM.02299-15.

Maiden, M. C., J. A. Bygraves, E. Feil, G. Morelli, J. E. Russell, R. Urwin, Q. Zhang, J. Zhou, K. Zurth, D. A. Caugant, I. M. Feavers, M. Achtman, and B. G. Spratt. Multilocus sequence typing: a portable approach to the identification of clones within populations of pathogenic microorganisms. Proc Natl Acad Sci U S A 1998; 95: 3140-5. doi: 10.1073/pnas.95.6.3140.

Meats, E., E. J. Feil, S. Stringer, A. J. Cody, R. Goldstein, J. S. Kroll, T. Popovic, and B. G. Spratt. Characterization of encapsulated and noncapsulated Haemophilus influenzae and determination of phylogenetic relationships by multilocus sequence typing. J Clin Microbiol 2003; 41: 1623-36. doi: 10.1128/jcm.41.4.1623-1636.2003.

Mitchelmore, P. J., J. Randall, M. J. Bull, K. A. Moore, P. A. O'Neill, K. Paszkiewicz, E. Mahenthiralingam, C. J. Scotton, C. D. Sheldon, N. J. Withers, and A. R. Brown. Molecular epidemiology of Pseudomonas aeruginosa in an unsegregated bronchiectasis cohort sharing hospital facilities with a cystic fibrosis cohort. Thorax 2018; doi: 10.1136/thoraxjnl-2016-209889.

Saiman, L., J. D. Siegel, J. J. LiPuma, R. F. Brown, E. A. Bryson, M. J. Chambers, V. S. Downer, J. Fliege, L. A. Hazle, M. Jain, B. C. Marshall, C. O'Malley, S. R. Pattee, G. Potter-Bynoe, S. Reid, K. A. Robinson,

K. A. Sabadosa, H. J. Schmidt, E. Tullis, J. Webber, and D. J. Weber. Infection prevention and control guideline for cystic fibrosis: 2013 update. Infect Control Hosp Epidemiol 2014; 35 Suppl 1: S1-s67. doi: 10.1086/676882.

Spilker, T., A. Baldwin, A. Bumford, C. G. Dowson, E. Mahenthiralingam, and J. J. LiPuma. 2009. 'Expanded multilocus sequence typing for burkholderia species', *J Clin Microbiol*, 47: 2607-10.

Tümmler, B. Molecular epidemiology in current times. *Environ Microbiol* 2020. doi: 10.1111/1462-2920.15238.

van Mansfeld, R., R. Willems, R. Brimicombe, H. Heijerman, F. T. van Berkhout, T. Wolfs, C. van der Ent, and M. Bonten. Pseudomonas aeruginosa genotype prevalence in Dutch cystic fibrosis patients and age dependency of colonization by various P. aeruginosa sequence types. J Clin Microbiol 2009; 47: 4096-101. doi: 10.1128/JCM.01462-09.