BACTERIAL TAXONOMY

Analysis of *fliC* variation among clinical isolates of *Burkholderia cepacia*

CRAIG WINSTANLEY, BARBARA A. HALES, J. ALUN W. MORGAN*, MALCOLM J. GALLAGHER†, SVITHRI D. PUTHUCHEARY‡, MOUSSA F. CISSÉ§ and C. ANTHONY HART†

Department of Biomedical Sciences, University of Bradford, Bradford, West Yorkshire BD7 1DP, *Horticulture Research International, Wellesbourne, Warwickshire CV35 9EF, †Department of Medical Microbiology and Genito-Urinary Medicine, University of Liverpool, PO Box 147, Liverpool L69 3BX, ‡Department of Medical Microbiology, University of Malaya, Kuala Lumpur 50603, Malaysia and §Department of Medical Microbiology, Hôpital d'Enfants, Albot Royer, Dakar, Senegal

PCR and restriction fragment length polymorphism (RFLP) typing of flagellin genes (*fliC*) from 57 clinical isolates of *Burkholderia cepacia* indicated that only type II flagellins were present. Twenty-two isolates previously identified as the epidemic UK cystic fibrosis strain were indistinguishable by this method, as were 11 isolates from a pseudo-outbreak in Senegal. Other clinical isolates, including 19 from disparate sources in Malaysia, were separated into nine *fliC* RFLP groups, exhibiting a large degree of divergence. When isolates were indistinguishable by *fliC* genotyping, their similarity was confirmed by whole genome macro-restriction analysis with pulsed-field gel electrophoresis following *Xba*I digestion. The variation in *fliC* sequences of *B. cepacia* was far greater than that with *B. pseudomallei*, supporting the view that '*B. cepacia*', as currently defined, may comprise several different genomic species.

Introduction

Burkholderia cepacia (formerly *Pseudomonas cepacia*) is an important opportunist pathogen in cystic fibrosis (CF). The outcome of colonisation of CF patients can vary from a rapidly fatal septicaemia to maintenance of stable respiratory function [1]. There is considerable genetic heterogeneity among CF isolates of *B. cepacia* and Vandamme *et al.* [2] reported that *B. cepacia* strains from CF patients belonged to at least five genetically distinct but phenotypically similar genomovars, although most – including the UK CF epidemic ET12 lineage [3] – belonged to genomovar III.

Motility in *B. cepacia* is achieved by means of polar flagella which consist of a basal body, hook and filament, with the latter comprising multiple flagellin molecules. Based on examination of 16 isolates, an earlier study showed that most *B. cepacia* isolates could be clustered into two types by flagellin protein

size [4]. Analysis of the *B. cepacia* flagellin genes (*fliC*) by PCR and restriction fragment length polymorphism (RFLP) enabled flagellin-type classification on the basis of product size, and revealed considerable variability in sequence. Thus, as in other organisms [5], the flagellin gene is a useful biomarker for epidemiological and phylogenetic studies [4]. This paper describes the application of *fliC* PCR/RFLP to the study of variability among 57 clinical isolates of *B. cepacia* from various sources. The results are compared with those of genotyping by pulsed-field gel electrophoresis (PFGE).

Materials and methods

B. cepacia strains

The strains used are listed in Table 1. Isolates from CF patients were cultured initially on selective media (Mast Laboratories, Bootle) and presumptively identified with the API-20NE system (bioMérieux UK, Basingstoke, Hampshire). Identification was confirmed with a range of biochemical and molecular tests, including partial 16S rRNA sequencing. Several isolates were identified as the UK CF epidemic strain (ET12 lineage) [3] by electrophoretic typing, whole protein SDS-PAGE and pyrolysis mass spectroscopy [6]. The isolates from Senegal were from wounds,

Received 24 Sept. 1998; revised version accepted 6 Nov. 1998.

Corresponding author: Dr C. Winstanley. Present address: Department of Medical Microbiology and Genito-Urinary Medicine, University of Liverpool, PO Box 147, Liverpool L69 3BX. (e-mail: C.Winstanley@bradford.ac.uk).

Table 1. Description of *B. cepacia* isolates used in this study

Strain no.	Flagellin gene RFLP group*	Source/reference no.	Description*
E241 (J2315)	I	[4]	CF epidemic strain, genomovar III
E242 (A599)	II	[4]	CF isolate, genomovar II
E243 (J2543)	III	[4]	Botanical strain
E244 (J2534)	IV	[4]	Botanical strain
E245 (J2552)	V	[4]	Botanical strain, genomovar I
E195	VI	[4]	CF isolate
E196	VII	[4]	CF isolate
E197	VIII	[4]	CF isolate
PCI2	X	[4]	CF isolate
NCTC10661	IX		Type strain, soil isolate
NCTC10744	ND		Type strain
E246	V	J. Govan	Botanical strain
E517, E518, E520, E521, E526, E527, E528, E529, E532, E533, E534, E535, E537, E901, E902, E908, E918, E922, E925, E928, E941, E942, E523, E524, E899	I	This study	CF epidemic strain isolates
E917	X	This study	CF isolate
E933	CF1 CF2	This study This study	CF isolate CF isolate
E660, E661, E662, E663, E735, E736, E738, E739, E907, E909, E937	Senegal	This study	Clinical isolates from Senegal
E982, E983, E992, E993, E1000	Malaysia 1	This study	Clinical isolates from Malaysia
E984, E987, E988, E991, E994, E995, E996, E997, E998, E999	Malaysia 2	This study	Clinical isolates from Malaysia
E985	Malaysia 3	This study	Clinical isolate from Malaysia
E986, E990	Malaysia 4	This study	Clinical isolate from Malaysia
E989	Malaysia 5	This study	Clinical isolate from Malaysia

ND, not determined.

*CF UK epidemic strain confirmed by SDS-PAGE of whole protein and pyrolysis mass spectroscopy [6]. Genomovar designations were provided by J. Govan.

blood or urines of patients in a paediatric ward in Dakar. Many of these patients had responded to treatment with antibiotics to which the bacteria were resistant; moreover, environmental sampling indicated the widespread presence of the organism and it was concluded that these isolates represented a pseudo-outbreak rather than a true outbreak of infection. The isolates from Malaysia were obtained over a 10-year period from wounds, blood or urines of infected patients. None of the Malaysian isolates was epidemiologically linked. All the organisms were maintained on nutrient agar and grown at 30°C.

PCR/RFLP analysis of *B. cepacia* flagellin genes

Flagellin gene oligonucleotide primers BC4 (CTGG-TCGCACAGCAGAACCTGAAC; N-terminal) and BCR12 (ACAG/TGTTTCGCGGTTTCCTG; C-terminal) were obtained from Genosys (Cambridge, Cambs). Cells from a nutrient agar plate were suspended in

sterile distilled water and boiled for 5 min. This lysed suspension (2.5 µl) was used directly in a standard amplification mixture. Amplifications were performed in 25-µl volumes containing 2 units of DynaZyme (Flowgen Instruments, Sittingbourne, Kent), 200 nM primers (BC4 and BCR12), 1 × DynaZyme buffer and 100 µM nucleotides (dATP, dCTP, dGTP, dTTP). Thirty PCR cycles were performed, each consisting of 95°C for 1 min, 60°C for 1 min and 72°C for 2 min.

Amplified product samples (10 µl) were digested in separate experiments with the restriction endonucleases *Hae*III and *Msp*I under the conditions recommended by the supplier (Life Technologies, Renfrewshire). These digests were subjected to electrophoresis on MetaPhor agarose gels (3% w/v; Flowgen Instruments) alongside a PCR size marker (R&D Systems, Abingdon, Oxfordshire) with fragment sizes of 50, 150, 300, 500, 750, 1000, 1500 and 2000 bp.

PFGE

Bacterial cells were grown overnight at 37°C on Columbia blood agar plates, harvested by scraping, washed and suspended in a buffer comprising 10 mM Tris-HCl, 20 mM NaCl, 50 mM EDTA, pH 7.2, and then mixed with an equal volume of pulsed-field certified agarose (BioRad, Hemel Hempstead, Herts.) 2% w/v containing lysozyme 1 mg/ml. The mixtures were dispensed into 100- μ l insert moulds and allowed to solidify on ice. These plugs were incubated for 1 h at 37°C in 1 ml of lysozyme buffer (10 mM Tris-HCl, 50 mM NaCl, sodium deoxycholate 0.2% w/v, sodium lauryl sarcosine 0.5% w/v, pH 7.2) containing lysozyme 1 mg/ml. After washing in 20 mM Tris-HCl, 50 mM EDTA, pH 8.0, the plugs were incubated overnight at 50°C in 1-ml volumes of proteinase K buffer comprising 100 mM EDTA, sodium deoxycholate 0.2% w/v, sodium lauryl sarcosine 1% w/v, pH 8.0, containing proteinase K 1 mg/ml. The plugs were then washed for 30 min in 1 ml of 20 mM Tris-HCl, 50 mM EDTA, pH 8.0, containing 2 mM phenyl methyl sulphonyl fluoride. After four further washes each of 30 min in 20 mM Tris-HCl, 50 mM EDTA, pH 8.0, the plugs were incubated at 37°C for 18 h with 2 units of *Xba*I (BioRad) per block in 300 μ l of *Xba*I restriction enzyme buffer (BioRad). Restriction fragments were separated by PFGE on a CHEF DRII system (BioRad) through pulsed-field certified agarose 1% w/v in 0.5 $\times$ TBE (10 $\times$ TBE was 0.9 M Tris, 0.9 M boric acid, 32 mM EDTA). Electrophoresis was performed for 20 h in 0.5 $\times$ TBE at 14°C with initial and final pulse times of 5 s and 35 s, respectively. λ concatamers (Boehringer Mannheim, Lewes, East Sussex) were used as DNA size standards.

Computer analyses

For computer analysis of *fliC* RFLP groups, pairwise similarity between isolates was calculated with a Jaccard score assigning equal weight to each fragment length for which at least one of the isolates produced a fragment and assigning a score of one where both isolates produced the fragment and zero where only one isolate did so [7]. A hierarchical cluster analysis was performed with the UPGMA algorithm in the package GENSTAT 5 (GENSTAT 5 Committee, Statistics Department, Rothamstead Experimental Station, Harpenden, Herts.) and a dendrogram was plotted. PFGE gels were analysed with the BioRad Image Analysis System with Molecular Analyst PC Fingerprinting software (version 1.12). Cluster analysis was performed with the UPGMA algorithm within the package Clusterbases (BioRad).

Results

PCR/RFLP genotyping of flagellin genes from *B. cepacia* isolates

PCR amplification of *fliC* from the 57 clinical isolates of *B. cepacia* revealed that only type II flagellins were

present, as indicated by an amplified product size of 1.0 kb. A botanical strain (E246) contained a type I flagellin, indicated by an amplified product size of 1.4 kb. Examples of different restriction patterns observed with *Msp*I are shown in Fig. 1. Combinations of *Msp*I and *Hae*III profiles were used to separate the clinical isolates into nine *fliC* RFLP groups, three of which – groups I, X and CF2 – were reported in a previous study [4]. Group CF2 was indistinguishable from RFLP group III. Isolates belonging to a single RFLP group but from different sources were given different designations. The 22 isolates of the UK epidemic CF strain all belonged to *fliC* RFLP group I, confirming previous observations [4]. The 11 clinical isolates from the Senegal pseudo-outbreak were indistinguishable from each other, but the clinical isolates from Malaysia were separated into five *fliC* RFLP groups. A dendrogram (Fig. 2a) illustrates the relationships between the *fliC* RFLP groups identified in this and the previous study [4].

Dissimilarity among the banding patterns made a meaningful analysis of relationships difficult, but it was possible to identify three broad clusters: two of the type I *fliC* RFLP groups (RFLP groups V and VI), together with one strain possessing an abnormally large flagellin (RFLP group VIII), clustered apart from the third type I *fliC* RFLP group (RFLP group II), which itself clustered with three type II *fliC* RFLP groups (RFLP groups III, CF2 and VII). The rest of the type II *fliC* RFLP groups formed one large cluster. The three broad clusters contained representatives of genomovars I, II and III, as previously defined [2].

PFGE genotyping of *B. cepacia* isolates

PFGE genotyping was performed on 19 isolates, including strains representing each of the *fliC* RFLP groups (Fig. 2b). Members of type I and type II *fliC* RFLP groups were not readily distinguishable by PFGE analysis. Strains representing the three type I *fliC*

Fig. 1. RFLP patterns of PCR amplified flagellin genes. Examples of *Msp*I digests of *fliC* amplified products representing RFLP groups: lane 1, Malaysia 5; 2, I; 3, Senegal; 4, Malaysia 4; 5, X; 6, III; 7, VII; 8, Malaysia 2; 9, Malaysia 3; 10, IX; M, PCR size marker.

Fig. 2. Dendrograms indicating inferred relationships among strains of *B. cepacia*. (a) Relationships among the *fliC* RFLP groups; (b) relationships among PFGE patterns, including those for strains representing different *fliC* RFLP groups.

RFLP groups (RFLP groups II, V and VI) were widely divergent from each other; on the other hand, strains representing the *fliC* RFLP groups II and Malaysia 4 clustered together by PFGE, despite their very divergent *fliC* RFLP patterns. Likewise, *fliC* RFLP groups VII and VIII – although widely divergent by *fliC* RFLP – clustered together by PFGE. Representatives of *fliC* RFLP groups CF1 and Malaysia 2, which gave identical *fliC* RFLP patterns, were similar by PFGE; likewise, the two methods indicated relatedness of *fliC* RFLP groups Senegal and Malaysia 1.

Discussion

In an earlier study of 16 clinical and environmental isolates of *B. cepacia*, three strains with type I flagellin and one strain (E197) with an abnormally large flagellin conforming to neither type I nor type II, were reported [4]. The present study mainly identified strains with type II flagellins, suggesting that these are far more numerous among clinical isolates. The reproducibility of *fliC* genotyping was confirmed by recognition of all the representatives of the UK CF epidemic strain as belonging to a single *fliC* RFLP group (group I). Likewise, the 11 isolates from Senegal – all from the same clinic and indistinguishable by PFGE – were indistinguishable by *fliC* genotyping, although they clustered with five epidemiologically unrelated Malaysian isolates, which also resembled the Senegal isolates in PFGE analysis. More generally, the Malaysian clinical isolates were highly variable as a group, reflecting the fact that they were isolated from different patients and different sites, and at differing times. The Malaysian isolates were responsible for most novel *fliC* RFLP groups reported in this study, although some were indistinguishable from one of the UK CF isolates (RFLP groups CF1 and Malaysia 2).

A comparison of *fliC* RFLPs from single strains representing genomovars I (RFLP group V), II (RFLP group II) and III (RFLP group I) is consistent with their subdivision into genomovars; thus there were three major branches on the dendrogram (Fig. 2a), and these corresponded with the split of the species into genomovars I, II and III. However, these three branches shared <20% similarity, indicating that *fliC* was highly variable among strains currently designated as *B. cepacia*. This variation is such that no real distinction can be made between type I and type II flagellins, despite the different sizes of their PCR products. This variation in *B. cepacia* contrasts markedly with the lack of variance observed among clinical isolates of *B. pseudomallei* (although some variation was seen between clinical and environmental isolates of this latter organism) [8]. As a species, *B. pseudomallei* seems far more genetically homogeneous than *B. cepacia*.

The diversity of *B. cepacia* did not reflect their environmental sources: isolates from botanical sources

fell into each of the three branches of the dendrogram. In a previous study, Butler *et al.* [9] examined several environmental *B. cepacia* strains that were phenotypically different from the CF epidemic strain and, by PFGE analysis, found not only that they were readily distinguishable from the CF epidemic strain, but also from each other. Following multilocus linkage disequilibrium analysis of a lotic population of *B. cepacia*, Wise *et al.* [10] concluded that much recombination had occurred, and that the population was non-clonal. The authors suggested that such recombination might be typical of environmental bacteria because these would frequently encounter divergent strains. In some cases clinical strains may be more homogeneous: using iso-enzyme analysis, Carson *et al.* [11] reported identical profiles for 31 nosocomial *B. cepacia* isolates obtained during an outbreak. It was suggested that this apparent contrast between clinical and environmental *B. cepacia* populations indicated that the bacterial species might exhibit different types of population structures depending on environmental conditions [10]. However, the study of Vandamme *et al.* [2] and our observations indicate considerable variation among clinical *B. cepacia* and that clinical and environmental isolates do not readily divide into two discrete clusters.

No overall agreement was apparent between the dendrograms generated from *fliC* RFLP and PFGE. The PFGE dendrogram divided into four main clusters, three of which contained individual strains of genomovars I, II and III. As the genomovar subgroups were delineated on the basis of DNA-DNA hybridisation experiments [2], correspondence with delineation by PFGE (which is based upon separation of restriction enzyme-digested genomic DNA) might reasonably be expected. However, as only three of the strains used in the analysis have been assigned to a genomovar, it is not certain that the clusters as a whole correspond with genomovar designations. The fourth cluster did not contain any strains of known genomovar. The lack of general correlation between the two dendrograms is probably caused by the extreme variability exhibited by the *fliC* RFLP patterns. Nevertheless, despite this variability, where the *fliC* RFLPs of different isolates were the same, their relatedness was confirmed by PFGE; thus, no strains indistinguishable by *fliC* RFLP exhibited >10% variation in PFGE analysis.

The analysis of *fliC* variation has been applied to study populations of several pathogenic bacteria including *Escherichia coli* [12], *Salmonella* spp. [13], *Campylobacter* spp. [14], *Helicobacter pylori* [15], *Borrelia burgdorferi* [16] and *Listeria monocytogenes* [17]. Recombination and lateral gene transfer are thought to have contributed to the flagellar antigenic variation observed in a number of bacterial pathogens [5]. Thus, although flagellin proteins contain conserved N- and C-terminal domains, the

central domain can vary greatly without an adverse effect on function. The variability between *B. cepacia* *fliC* RFLP groups indicates considerable divergence, possibly due to such recombinational events. Sequencing of *fliC* genes from three different *B. cepacia* strains confirms this variability [4], but any meaningful insight into the evolution of *B. cepacia* *fliC* genes will require analysis of many more such sequences. In conclusion, it is clear that the flagellin genes of *B. cepacia*, as currently defined, are highly variable. This supports the suggestion that '*B. cepacia*' may represent a number of different genomic species [2].

This work was supported by an award to C.W. from The Wellcome Trust (044249/PMG/VW) and by the Biotechnology and Biological Sciences Research Council.

References

1. Govan JRW, Deretic V. Microbial pathogenesis in cystic fibrosis: mucoid *Pseudomonas aeruginosa* and *Burkholderia cepacia*. *Microbiol Rev* 1996; **60**: 539–574.
2. Vandamme P, Holmes B, Vancanneyt M *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**: 1188–1200.
3. Govan JRW, Hughes JE, Vandamme P. *Burkholderia cepacia* – medical, taxonomic and ecological issues. *J Med Microbiol* 1996; **45**: 395–407.
4. Hales BA, Morgan JAW, Hart CA, Winstanley C. Variation in the flagellin genes and proteins of *Burkholderia cepacia*. *J Bacteriol* 1998; **180**: 1110–1118.
5. Winstanley C, Morgan JAW. The bacterial flagellin gene as a biomarker for detection, population genetics and epidemiological analysis. *Microbiology* 1997; **143**: 3071–3084.
6. Corkill JE, Sisson PR, Smyth A *et al.* Application of pyrolysis mass spectroscopy and SDS-PAGE in the study of the epidemiology of *Pseudomonas cepacia* in cystic fibrosis. *J Med Microbiol* 1994; **41**: 106–111.
7. Jaccard P. Etude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bull Soc Vaudoise Sci Nat* 1901; **44**: 223–270.
8. Winstanley C, Hales BA, Corkill JE, Gallagher MJ, Hart CA. Flagellin gene variation between clinical and environmental isolates of *Burkholderia pseudomallei* contrasts with the invariance among clinical isolates. *J Med Microbiol* 1998; **47**: 689–694.
9. Butler SL, Doherty CJ, Hughes JE, Nelson JW, Govan JRW. *Burkholderia cepacia* and cystic fibrosis: do natural environments present a potential hazard? *J Clin Microbiol* 1995; **33**: 1001–1004.
10. Wise MG, Shimkets LJ, McArthur JV. Genetic structure of a lotic population of *Burkholderia (Pseudomonas) cepacia*. *Appl Environ Microbiol* 1995; **61**: 1791–1798.
11. Carson LA, Anderson RL, Panlilio AL, Beck-Sague CM, Miller JM. Isoenzyme analysis of *Pseudomonas cepacia* as an epidemiologic tool. *Am J Med* 1991; **91**: 252S–255S.
12. Schoenhals G, Whitfield C. Comparative analysis of flagellin sequences from *Escherichia coli* strains possessing serologically distinct flagellar filaments with a shared complex surface pattern. *J Bacteriol* 1993; **175**: 5395–5402.
13. Masten BJ, Joys TM. Molecular analyses of the *Salmonella* g... flagellar antigen complex. *J Bacteriol* 1993; **175**: 5359–5365.
14. Alm RA, Guerry P, Trust TJ. Distribution and polymorphism of the flagellin genes from isolates of *Campylobacter coli* and *Campylobacter jejuni*. *J Bacteriol* 1993; **175**: 3051–3057.
15. Forbes KJ, Fang Z, Pennington TH. Allelic variation in the *Helicobacter pylori* flagellin genes *flaA* and *flaB*: its consequences for strain typing schemes and population structure. *Epidemiol Infect* 1995; **114**: 257–266.
16. Fukunaga M, Koreki Y. A phylogenetic analysis of *Borrelia burgdorferi* sensu lato isolates associated with lyme disease in Japan by flagellin gene sequence determination. *Int J Syst Bacteriol* 1996; **46**: 416–421.
17. Rasmussen OF, Skouboe P, Dons L, Rossen L, Olsen JE. *Listeria monocytogenes* exists in at least three evolutionary lines: evidence from flagellin, invasive associated protein and listeriolysin O genes. *Microbiology* 1995 **141**: 2053–2061.