Sequence diversity of the *mucABD* locus in *Pseudomonas aeruginosa* isolates from patients with cystic fibrosis

Alessandra Bragonzi,
Lutz Wiehlmann,
Jens Klockgether,
Nina Cramer,
Dieter Worlitzsch,
Gerd Döring
and Burkhard Tümmler

1Institute of Medical Microbiology and Hygiene, Universitätsklinikum Tübingen, Tübingen, Germany
2Institute for Experimental Treatment of Cystic Fibrosis, DIBIT - HS Raffaele, Milano, Italy
3Klinische Forscherguppe, OE 6710, Medizinische Hochschule Hannover, Carl-Neuberg-Str. 1, D-30625 Hannover, Germany

The *mucA* gene of the *muc* operon, which is instrumental in the control of the biosynthesis of the exopolysaccharide alginate, is a hotspot of mutation in *Pseudomonas aeruginosa*, a micro-organism that chronically colonizes the airways of individuals with cystic fibrosis (CF). The *mucA*, *mucB* and *mucD* genes were sequenced in nine environmental isolates from aquatic habitats, and in 37 *P. aeruginosa* strains isolated from 10 patients with CF, at onset or at a late stage of chronic airway colonization, in order to elucidate whether there was any association between mutation and background genotype. The 61 identified single nucleotide polymorphisms (SNPs) segregated into 18 *mucABD* genotypes. Acquired and *de novo* stop *mucA* mutations were present in 14 isolates (38%) of five *mucABD* genotypes. ΔG430 was the most frequent and recurrent *mucA* mutation detected in four genotypes. The classification of strains by *mucABD* genotype was generally concordant with that by genome-wide SpeI fragment pattern or multilocus SNP genotypes. The exceptions point to intragenic mosaicism and interclonal recombination as major forces for intraclonal evolution at the *mucABD* locus.

**INTRODUCTION**

The metabolically versatile *Pseudomonas aeruginosa* is present in soil and aquatic habitats, but it is also an important opportunistic pathogen for humans, animals and plants. Typing of strain collections by single nucleotide polymorphisms (SNPs), DNA fragment length polymorphisms and phenotypic traits indicates that the current *P. aeruginosa* population is in linkage equilibrium, and consists of a network of equivalent genotypes (termed clones), whereby a subset of clones is overrepresented due to epidemic spread (Curran et al., 2004; Morales et al., 2004; Pirnay et al., 2002a, b).

In *P. aeruginosa*, interclonal and intraclonal diversity have been reported to result from the acquisition or loss of genomic islets or islands, genome rearrangements, and recombination or point mutation (reviewed by Tümmler, 2006). Mosaic genes are a further source of genetic diversity. They contain regions with different degrees of relatedness which can be explained by recombination of two distinct alleles as a result of gene transfer events (Hakenbeck, 1998).

Evidence for a mosaic gene structure has been drawn from SNP haplotype (Spangenberg et al., 1998; Kiewitz & Tümmler, 2000) or the detection of cassettes (Spangenberg et al., 1996). Our current knowledge about mosaic genes in *P. aeruginosa* is restricted to the pyoverdine biosynthesis locus (Smith et al., 2006), *ampC* (PA4110) (Spangenberg et al., 1998), *fleP* (PA1096) (Arora et al., 2001), *fliC* (PA1092) (Spangenberg et al., 1998) and *oprD* (PA0958) (Pirnay et al., 2002b). In the case of *oprD*, sequence analysis of 55 *P. aeruginosa* isolates, collected over a period of 15 years from various, spatially separated, clinical and environmental habitats, has uncovered a microscale mosaic structure of *oprD*. All sequences fall into three main groups, which differ by 7–9% of their nucleotides. Several recombinational exchanges of DNA blocks of 100–300 bp lead to a mosaic gene structure and cause further divergence into subgroups.

Most information about the evolution of intraclonal diversity of *P. aeruginosa* has been obtained from retrospective cross-sectional and longitudinal analyses of isolates recovered from the atypical habitat of the lungs of patients with cystic fibrosis (CF) (Ernst et al., 2003; Kresse et al., 2003; Larbig et al., 2002; Smith et al., 2006). A hallmark of

**Abbreviations:** CF, cystic fibrosis; SNP, single nucleotide polymorphism.
the adaptation of *P. aeruginosa* to the atypical niche of the CF lung is the emergence of mucoid alginate-overexpressing morphotypes that are typically associated with stop mutations in the *mucA* gene (Martin et al., 1993a).

The activation of alginate genes is complex (reviewed by Ramsey & Wozniak, 2005). The *algU* gene, which is essential for alginate production, encodes the alternative sigma factor σA. *MucA* is an anti-sigma factor that sequesters σA (Mathee et al., 1996), and MucB is a periplasmic negative regulator (Martin et al., 1993b). MucD is a homologue of the *Escherichia coli* periplasmic serine protease HtrA, and is apparently both a chaperone and a regulator of alginate production (Wood & Ohman, 2006).

This study reports on the sequence diversity of the *mucABD* operon in nine environmental and 37 CF *P. aeruginosa* isolates. The association between *mucABD* SNP haplotype and chromosomal genotype has been inferred from SpeI restriction fragment patterns (Ro¨mling et al., 1994) and multilocus SNP genotypes (Morales et al., 2004). The CF strains were isolated from 10 patients at onset and at a chronic late stage of airway colonization. In living patients, the late strains were isolated from 10 patients at onset and at a chronic stage of airway colonization. In deceased patients, the last isolates prior to death were studied. Sequence analysis of sequential isolates has so far only been performed on strains from a young CF patient who acquired *P. aeruginosa* during infancy, and who was monitored for a period of 8 years thereafter (Spencer et al., 2003; Smith et al., 2006).

Since *mucA* mutations are common in CF isolates (Martin et al., 1993a, b), the sequence analysis provided the opportunity to investigate the diversity of the genetic background in a mutation hotspot in the *P. aeruginosa* genome. The comparative study of sequential isolates allowed us to differentiate between acquired and *de novo* mutations. Sequence and genotype analysis provided direct evidence for recurrent *mucA* mutations and the generation of mosaic *muc* genes in *P. aeruginosa* that persists in CF lungs.

**METHODS**

**CF patients, bacterial strains and growth conditions.** *P. aeruginosa* strains were isolated from 10 pancreatic-insufficient CF patients attending the CF clinic of the Medizinische Hochschule Hannover. Thirty-seven *P. aeruginosa* strains were analysed from spu·ta or throat swabs from the 10 CF patients (Table 2). Twenty *P. aeruginosa* strains were isolated from the CF lung, and from the 10 CF patients (patients 6–10), the last isolate collected prior to death 5–14 years after acquisition of *P. aeruginosa* was analysed. Strains were stored at −70 °C, and were not subcultured. In addition, the completely sequenced strains *P. aeruginosa* PA01 and PA14, and nine environmental isolates from aquatic habitats (SG17M, PT2, PT6, PT12, PT20, PT22 and DSM3939), mushroom (ATCC 33818) and a fuel tank (ATCC 33988) were studied.

**Genotyping by PFGE.** Samples were analysed as described by Römling et al. (1994), with minor modifications. *P. aeruginosa* strains were grown overnight at 37 °C in LB medium, centrifuged for 10 min at 1500 g and suspended in 75 mM NaCl, 25 mM EDTA, pH 7-4, to a concentration of 5 × 10⁶ cells ml⁻¹. The cell suspension was mixed 1:1 with 2 % (w/v) low-melting-point agarose to prepare the agarose plugs. The embedded cells were digested for 48 h at 56 °C with proteinase K (1 mg ml⁻¹) in 0-5 M EDTA, pH 9-5, 1 %, v/v, N-lauryl sarcosine. Agarose plugs were equili-

---

**Multilocus SNP genotyping.** Boiled colonies of each strain were subjected to PCR with appropriate primers to amplify DNA fragments containing the SNP, as described by Morales et al. (2004). Amplified DNA was digested with the restriction enzyme that discriminated each SNP (Table 1), and the fragments generated were ana-

**Sequence analysis at the muc locus.** One loop of a single *P. aeruginosa* colony, grown on blood agar, was processed for DNA extraction using a commercial DNA isolation kit (Qiagen), according to the instructions of the manufacturer. PCR amplification of the entire mucA, mucB and mucD genes was carried out using Taq DNA polymerase (Qiagen). The following primers (MWG-Biotech) were used: mucA, 5'-CTC GTG AAGCAA TCG AGA AA-3' and 5'-AAA AGC AAC AGG GAG GTG GT-3' for mucA; mucB, 5'-ATC CGC CGT CAG TGG TAC AG-3' and 5'-CGA GCA GGA CGA GCA GGT AC-3', and mucD, 5'-GTC CGA TTC GGC GTG AGT CT-3' and 5'-ACG CAG GTA ACG GAT TGA CG-3'. The amplified DNA samples were sequenced by standard automated DNA sequence technology employing the primers described above and the following additional internal primers for mucB and mucD: mucB int1, 5'-CAG TGG TCC TTG CAG TGA AA-3' and mucD int1, 5'-GAT CAA CCA GGG TTA CTC CG-3' and mucD int2, 5'-AGA TCT GCG AGT TGA CG-3'. The sequence results were compared with the strain PAO1 sequence (www.pseudomonas.com) by the BLAST program at the NCBI data-

---

The relatedness of strains by *mucABD* or multilocus genotype was calculated by parsimony analysis with the program PARS of the phylogenetic software package PHYLIP version 3.6 (alpha 3) (Felsenstein, 2002).
Table 1. Multilocus SNP genotyping of *P. aeruginosa*

The sequences of the *P. aeruginosa* strains PA103 (*exoU*), PAK (*fliC* a type) and PAO1 (all other genes) were taken as reference.

<table>
<thead>
<tr>
<th>Locus</th>
<th>Restriction enzyme</th>
<th>SNP or presence of gene (see Fig. 1)</th>
<th>Designation</th>
</tr>
</thead>
<tbody>
<tr>
<td>oriC</td>
<td>AlwNI</td>
<td>T267C</td>
<td>I</td>
</tr>
<tr>
<td>oprL</td>
<td>BstUI</td>
<td>T396C</td>
<td>II</td>
</tr>
<tr>
<td><em>fliC</em> a</td>
<td>MspA1I</td>
<td>A672T</td>
<td>III</td>
</tr>
<tr>
<td><em>alkB2</em></td>
<td>MspA1I</td>
<td>G471A</td>
<td>IV</td>
</tr>
<tr>
<td><em>citS</em></td>
<td>AciI</td>
<td>A533G</td>
<td>V</td>
</tr>
<tr>
<td></td>
<td>MnII</td>
<td>G896C</td>
<td>VI</td>
</tr>
<tr>
<td><em>oprI</em></td>
<td>AciI</td>
<td>T195C</td>
<td>VII</td>
</tr>
<tr>
<td><em>ampC</em></td>
<td>MspI</td>
<td>G789A</td>
<td>VIII</td>
</tr>
<tr>
<td></td>
<td>BatUI</td>
<td>C1060T</td>
<td>IX</td>
</tr>
<tr>
<td></td>
<td>Hinfl</td>
<td>G1150A</td>
<td>X</td>
</tr>
<tr>
<td></td>
<td>AlwNI</td>
<td>G1166C</td>
<td>XI</td>
</tr>
<tr>
<td></td>
<td>Agel</td>
<td>T1585C</td>
<td>XII</td>
</tr>
<tr>
<td></td>
<td>BsrI</td>
<td>C1639A</td>
<td>XIII</td>
</tr>
<tr>
<td><em>fliC</em></td>
<td>a type</td>
<td></td>
<td>XIV</td>
</tr>
<tr>
<td><em>fliC</em></td>
<td>b type</td>
<td></td>
<td>XV</td>
</tr>
<tr>
<td><em>exoS</em></td>
<td></td>
<td></td>
<td>XVI</td>
</tr>
<tr>
<td><em>exoU</em></td>
<td></td>
<td></td>
<td>XVII</td>
</tr>
</tbody>
</table>

Alginate determination. For the quantitative determination of alginate production by *P. aeruginosa* patient isolates, the carbazole assay was used (Lam et al., 1980). Briefly, *P. aeruginosa* was first cultured aerobically in trypticase soy broth (Oxoid) overnight at 37°C, then plated on *Pseudomonas isolation agar* containing 1% glycerol, v/v, as a carbon source, and incubated for 24 h at 37°C, followed by determination of alginate production. For uronic acid detection, growth from three different plates was pooled. The uronic acid content was normalized by measuring the protein content (Bio-Rad).

RESULTS

Sequence diversity of the *mucABD* locus

Sequencing of the *mucA*, *mucB* and *mucD* genes in 37 sequential *P. aeruginosa* strains from 10 CF patients (Table 2) and in nine environmental isolates revealed 61 SNPs (Table 3) corresponding to 2% sequence diversity at this locus, which is significantly higher than the mean sequence diversity of 0.5% in the *P. aeruginosa* genome (Spencer et al., 2003). The overrepresentation of transitions (83%) and synonymous substitutions (80%) is similar to that found for evolutionarily neutral SNPs in other loci (Kiewitz & Tümler, 2000) or in the bulk genome (Spencer et al., 2003). The 61 SNPs segregated into 18 *mucABD* genotypes, one of which split into two variants (Table 2, Figs 1 and 2). The largest divergence of the *mucABD* sequence was observed between the completely sequenced reference strains PAO1 and PA14, the latter represented in our panel by one environmental and five CF isolates (Fig. 1).

Of the 14 non-synonymous nucleotide substitutions, the two common valine–isoleucine polymorphisms I137V and V441I and the sequence variant Q225E in *mucD* were found in more than one SNP genotype. All other amino acid substitutions were specific for a single SNP genotype, including the three non-conservative substitutions S113F in *mucB* and W171R and S363N in *mucD*. A representation of Dayhoff’s mutation odds matrix (Dayhoff et al., 1978) by projecting the matrix using multidimensional scaling (French & Robson, 1983; Taylor, 1986) indicated that these three amino acid replacements at least were not neutral; however, it is not yet known whether any substitution modulates the function of the gene product.

Comparison of the *mucABD* single locus genotype with the genome-wide SNP and SpeI fragment genotype: evidence for intragenic mosaicism and interclonal recombination

The sequential isolates from individual patients belonged either to one clone (three patients) or to two clones (seven patients) (Table 2). Four clones (A, C, D and L; Table 2) were detected in more than one patient. Environmental and CF isolates of the abundant clones C (Dinesh et al., 2003) and D (strain PA14; Liberati et al., 2006) were identical in their *mucABD* sequences (Fig. 1).

The classification by *mucABD* genotype corresponded for most strains with that by genome-wide multilocus SNP and anonymous macrorestriction fragment genotypes (Table 2, Fig. 1). Differences were taken as evidence for intra- or interclonal exchange of sequence at the *muc* locus. First, unrelated clones could share the same *mucABD* sequence. One set was represented by clones I and L, the other set by
Table 2. P. aeruginosa sequential isolates from CF patients

<table>
<thead>
<tr>
<th>Patient</th>
<th>Age at onset of colonization (years)</th>
<th>Strain</th>
<th>Colonization time (years)</th>
<th>Alginate [μg uronic acid (μg protein)^−1]*</th>
<th>Multilocus SNP genotype</th>
<th>mucABD genotype [mucA] stop mutation</th>
<th>PFGE genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>Patient 1</td>
<td>7.4</td>
<td>RP1</td>
<td>0</td>
<td>0.02</td>
<td>A</td>
<td>a</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td></td>
<td>RP2</td>
<td>0</td>
<td>0.02</td>
<td>A</td>
<td>a</td>
<td>No</td>
</tr>
<tr>
<td>Patient 2</td>
<td>12.5</td>
<td>SG1</td>
<td>0</td>
<td>0.01</td>
<td>A</td>
<td>a</td>
<td>No</td>
</tr>
<tr>
<td>Patient 3</td>
<td>10.6</td>
<td>NN1</td>
<td>0</td>
<td>0.01</td>
<td>A</td>
<td>c</td>
<td>No</td>
</tr>
<tr>
<td>Patient 4</td>
<td>13.8</td>
<td>BT1</td>
<td>0</td>
<td>0.03</td>
<td>D</td>
<td>d</td>
<td>C523T</td>
</tr>
<tr>
<td>Patient 5</td>
<td>19.5</td>
<td>KB1</td>
<td>0</td>
<td>0.01</td>
<td>D</td>
<td>d</td>
<td>T412C C523T</td>
</tr>
<tr>
<td>Patient 6</td>
<td>3.9</td>
<td>AA1</td>
<td>0</td>
<td>0.01</td>
<td>H</td>
<td>h</td>
<td>No</td>
</tr>
<tr>
<td>Patient 7</td>
<td>7.5</td>
<td>TR1</td>
<td>0</td>
<td>0.01</td>
<td>I</td>
<td>i</td>
<td>ΔA358</td>
</tr>
<tr>
<td>Patient 8</td>
<td>13.2</td>
<td>MF1</td>
<td>0</td>
<td>0.01</td>
<td>J</td>
<td>j</td>
<td>No</td>
</tr>
<tr>
<td>Patient 9</td>
<td>17.0</td>
<td>KK1</td>
<td>0</td>
<td>0.01</td>
<td>L</td>
<td>i</td>
<td>No</td>
</tr>
<tr>
<td>Patient 10</td>
<td>8.1</td>
<td>BST1</td>
<td>0</td>
<td>0.02</td>
<td>M</td>
<td>m</td>
<td>No</td>
</tr>
</tbody>
</table>

*The alginate contents of the PAO1 strain and of its isogenic PDO300 mutant (Mathee et al., 1999) harbouring the mucA stop mutation Δ430 were 0.01 and 2.7 μg uronic acid (μg protein)^−1, respectively.

Clone H and the environmental isolate PT6. In other words, the same mucABD genotype was observed in different genomic backgrounds. Second, the clones B and H represented by strains RP73 and RP74, and AA2, AA43, AA44 and PT6, respectively, shared an identical SNP pattern in mucA, mucB and the 5′ portion of mucD, but differed in their SNP pattern 3′ of SNP32 (Fig. 1), indicating that these two mucABD genotypes emerged by intragenic recombination.

Evidence for intragenic recombination at the muc locus was derived not only from interclonal, but also from intraclonal sequence comparison. The first case was the strains KB1 and KB2 which were retrieved from the same patient’s lungs within the first year of colonization with P. aeruginosa. The two isolates were perfectly matched in their genome-wide SNP and SpeI fragment genotypes, but differed in their mucA SNP genotype, with the mucBD genotypes being identical (Fig. 1). Since it is extremely unlikely that three
successive mutations at common SNP positions occurred de novo, one strain must have incorporated its mucA gene from another P. aeruginosa clone. The second case was the clone G strains AA1 and PT20 which had been isolated in 1985 and 1992, at a distance of 300 km from one another, from CF airways and a sanitary facility, respectively. The two strains were identical in their genomic SNP genotype and most muc sequences, but differed in the outermost SNPs at the 5' end of mucA and the 3' end of mucD. The genomically related mushroom isolate ATCC 33818 shared the mucA sequence with strain PT20 and the mucD sequence with strain AA1 (Fig. 1). Hence, two recombinations in the muc locus can explain the emergence of three related mosaic muc genotypes.

mucA stop mutations in CF isolates: frequent recurrent mutations

Unequivocal loss-of-function stop mutations were identified in the mucA gene of 14 of our 37 CF isolates (38 %), but in none of the nine environmental isolates that were sequenced in the mucABD genes. No loss-of-function mutations were detected in mucB or mucD. mucA stop mutations were associated with five mucABD genotypes (Fig. 2), and were observed with similar frequency in strains recovered from both healthier and more severely ill patients (Table 2) (chi-square test: \( P = 1 \)). Two C-to-T transitions, numerous single nucleotide deletions and two double mutations were observed (Table 2). Four of the six different mutations were located in stretches of thermodynamically stable double strands with > 80 % GC content.
mucA mutations have been reported to arise in *P. aeruginosa* during the stage of chronic colonization, when the strains become mucoid (Spencer *et al.*, 2003). However, in our cohort, initial isolates from the first *P. aeruginosa*-positive sputum from patients 4, 7 and 9 carried mucA mutations (Table 2). These three patients became positive for *P. aeruginosa* after discharge from a stay at a CF ward, or after return from a CF rehabilitation centre, indicating that their mucA mutations were acquired with the initially colonizing strain from an external CF-related source. This suspicion of nosocomial acquisition is supported by the phenotypes of the initial isolates that are typically observed in samples retrieved from chronically colonized CF lungs. The strains from these three patients were mucoid and/or resistant to antipseudomonal agents.

In patients 2, 3, 5, 6 and 10, mucA stop mutations were detected in just one isolate per patient, which strongly suggests that the mutations had emerged *de novo* in the CF lung habitat. Half of all the mutations consisted of a deletion of a G in a string of five G residues located at positions 426–430 in a GC-rich stretch of the mucA gene. This ΔG430 deletion occurred in four mucABD haplotypes (Table 2, Fig. 2), which is proof of recurrent mutations on diverse genetic backgrounds.

Recent sequence analyses of the mucA gene in large collections of CF isolates from North America (Spencer *et al.*, 2003; Yoon *et al.*, 2006), Australia (Anthony *et al.*, 2002) and Europe (this study) have identified insertions or deletions in 18 strains, and 21 different stop mutations in 92 strains. More than 80% of the frameshift or nonsense mutations that give rise to a premature stop signal were clustered within the 100 bp stretch from nt 340–439 of the 585 bp mucA gene (Fig. 3). The most frequent ΔG430 mutation accounted for 40% of all stop mutations. The
recurrent emergence of the ΔG430 mutation in independent CF lung habitats and unrelated mucABD genetic backgrounds (Table 2, Fig. 2) is in agreement with in vitro studies showing that this homopolymeric tract is prone to single nucleotide deletion upon exposure of P. aeruginosa biofilms to reactive oxygen species, a scenario envisaged to occur in the CF lung (Mathee et al., 1999).

Stop mutations in mucA are known to lead to alginate overproduction and conversion to a stable mucoid phenotype in P. aeruginosa (Martin et al., 1993b). We were able to confirm the causative role of the loss-of-function mutation in mucA for mucoidy in the reference strain PAO1 and its isogenic single mutant ΔG430 mucA (footnote in Table 2). This association, however, was not so evident for the CF isolates. Just five of the 14 mucA stop mutants were overexpressing alginate (Table 2). During the chronic colonization of CF airways, P. aeruginosa is continuously accumulating mutations in its genome (Smith et al., 2006); therefore, secondary site mutations probably caused the reversion to a non-mucoid morphotype in the majority of mucA mutant strains.

DISCUSSION

Chromosomal mosaic genes are a major source of genetic diversity, and are common in naturally transformable bacteria such as neisseriae and streptococci (Hakenbeck, 1998). In the non-transformable P. aeruginosa, mosaicism is known in the pyoverdine biosynthesis locus (Smith et al., 2006), ampC (PA4110) (Spangenberg et al., 1998), ffrP (PA1096) (Arora et al., 2001), fliC (PA1092) (Spangenberg et al., 1998) and oprD (PA0958) (Pirnay et al., 2002b). This report adds the muc operon to the list of mosaic genes. Compared to the number of analysed clones and isolates, the frequency of intragenic recombination may appear to be high; however, this does not imply that this locus is subject to evolutionary forces additional to those of the bulk of the genome. Bacteria such as the ubiquitous and metabolically versatile P. aeruginosa are believed to have very large population sizes (Lynch & Conery, 2003), and correspondingly are expected to evolve under purifying or negative natural selection, i.e. natural selection acting to decrease the frequency of deleterious alleles. The quantitative criterion is a low ratio of non-synonymous \( d_N \) over synonymous \( d_S \) nucleotide substitutions, and indeed, low \( d_N/d_S \) ratios of 0-14 are observed when large datasets of phylogenetically independent pairs of genes are compared between closely related pairs of bacterial species (Friedman et al., 2004; Jordan et al., 2002). A significant excess of \( d_N \) over \( d_S \) nucleotide substitutions is interpreted as evidence for the action of positive selection. The \( d_N/d_S \) ratio of 0-27 for the muc locus, however, is not significantly different from the global \( d_N/d_S \) ratio of \( \sim 0.16 \) for the P. aeruginosa core genome (Spencer et al., 2003). Hence, we conclude that the muc locus, like the core genome, is subject to ongoing purifying selection.
Although at first glance the observed intragenic recombination rate seems to be high, it is in complete accordance with earlier reports that recombination is frequent in *P. aeruginosa*, and that correspondingly the population has an epidemic non-clonal population structure (Kiewitz & Tümmler, 2000; Pirnay et al., 2002a). All genotypes, each of which is characterized by non-random association of alleles, are in linkage equilibrium to each other. In other words, the frequency of recombination is high enough to prevent linkage disequilibrium among genes and hitchhiking effects on the phenotype. Previous publications on the population genetics of *P. aeruginosa* have focused on the diversity of unrelated clones. By sequencing several strains of the same clone, we now demonstrate that interclonal recombination contributes to intraclonal variation and the generation of mosaic genes. The colonized CF lung is characterized by high numbers of *P. aeruginosa* and only a few other bacterial competitors (Breitenstein et al., 1997). Such a scenario may facilitate the exchange of homologous genes between two or more *P. aeruginosa* clones.

Sequential isolates from CF lungs have been extensively characterized in one index case who became chronically colonized with *P. aeruginosa* very early, by the age of 6 months (Smith et al., 2006). Complete genome sequencing of clonally related month 6 and month 96 isolates uncovered the fixation of 68 mutations in the chromosome of the late isolate. No stop mutation in the *mucA* gene was observed in the month 96 isolate. A previous analysis on 36 sequential isolates recovered from the same patient during the first 30 months of airway colonization, however, identified eight mucoid strains, six of which carried either the ΔG357 or the ΔG430 stop mutation in *mucA* (Spencer et al., 2003). Lineages harbouring *mucA* mutations have not been detected in month 60 and month 96 isolates (see Fig. 1 in Smith et al., 2006). The disappearance of *mucA* mutants from the lungs of this early colonized patient is in accordance with our observation in the older patients 2 and 9 of our cohort (Table 2), who also lost *mucA* mutants during colonization. Our study on initial and late isolates from 10 patients moreover demonstrates that loss-of-function mutations in *mucA* either emerge de novo in the CF lung (see Table 2, patients 2, 3, 5, 6 and 10) or are acquired by nosocomial transmission (patients 4, 7 and 9).

In summary, the spectrum of loss-of-function mutations in the *mucA* gene of CF isolates is dominated by the major mutation ΔG430 and a cluster of further stop mutations within 100 bp upstream of ΔG430. The dominant mutations are recurrent mutations on different *mucABD* haplotypes. No mutations were detected in the negative regulator *mucB* or in *mucD*, which plays a dual role in the regulation of alginate production, as well as in temperature resistance (Wood & Ohman, 2006). Sequencing of sequential isolates moreover provided direct experimental evidence for intragenic recombination as a means of generating intraclonal diversity in *P. aeruginosa*.

ACKNOWLEDGEMENTS

The study was supported in part by a Marie Curie Fellowship to A. B. (MCFI-2001-51061), and by the Mukoviszidose eV, Bonn, Germany, and the Italian Cystic Fibrosis Foundation. J. K. and N. C. are members of the Deutsche Forschungsgemeinschaft (DFG)-sponsored Research Training Group ‘Pseudomonas: Pathogenicity and Biotechnology’.

REFERENCES


P. aeruginosa muc sequence diversity


