Correspondence
Philippe Lanotte
philippe.lanotte@med.univ-tours.fr

Received 27 May 2003
Accepted 4 August 2003

Genetic features of Pseudomonas aeruginosa isolates from cystic fibrosis patients compared with those of isolates from other origins

Philippe Lanotte, Stephane Watt, Laurent Mereghetti, Nathalie Dartiguelongue, Aziz Rastegar-Lari, Alain Goudeau and Roland Quentin

Département de Microbiologie Médicale et Moléculaire, EA 3250, Unité de Bactériologie, Faculté de Médecine, Tours, France

In order to improve our understanding of the colonization of the pulmonary tract of cystic fibrosis (CF) patients by Pseudomonas aeruginosa, 162 isolates from five different ecological origins were studied. The genetic features of each isolate were determined by random amplification of polymorphic DNA (RAPD) and by searching for eight virulence genes (six known virulence genes, algD, lasB, toxA, plcH, plcN and exoS, and two genes encoding putative neuraminidases, nan1 and nan2). Five RAPD groups were identified. Most of the CF isolates were distributed equally in three of these groups (RA, RB and RC). The CF isolates in RB were related to isolates from a wide variety of origins. The CF isolates in RA were related to a population composed of 65 % of the non-CF isolates from pulmonary tract infections. RC was mainly composed of CF isolates that were related to 30 % of isolates from plants. All genes except exoS and nan1 were present in all isolates. The exoS and nan1 virulence factor genes were most prevalent in CF isolates. exoS, which encodes exoenzyme S, was present in 94 % of CF isolates but also in 80 % of non-CF isolates from pulmonary tract infections. nan1, which encodes a putative neuraminidase, was found in 82.5 % of the isolates from group RC, which was composed largely of CF isolates. In conclusion, three major genogroups of P. aeruginosa isolates, each of which exhibits peculiar genetic features, are able to colonize CF patients. This may have different consequences on the outcome of pulmonary disease.

INTRODUCTION

Pseudomonas aeruginosa is a ubiquitous micro-organism that can infect immunocompromised individuals and is responsible for nosocomial infections (Van Delden & Iglewski, 1998). Patients with cystic fibrosis (CF) are particularly susceptible to chronic P. aeruginosa infection of the airways. This bacterium is rarely eradicated from CF patients after colonization, regardless of the antibiotherapy used (Govan & Deretic, 1996).

P. aeruginosa possesses a large number of cell-associated and extracellular virulence factors, which are tightly regulated by cell-to-cell signalling systems (Van Delden & Iglewski, 1998). The formation of mucoid colonies of P. aeruginosa composed of alginates, involving algD genes, protects the bacterium from the host’s immune response and from antibiotics and thus contributes to chronic pulmonary inflammation (Govan & Deretic, 1996). Other virulence factors can cause pulmonary damage by different mechanisms. Exoenzyme S, encoded by the exoS gene, is an ADP-ribosyltransferase that is secreted by a type-III secretion system directly into the cytosol of epithelial cells (Yahr et al., 1995). Exotoxin A, encoded by the toxA gene, inhibits protein biosynthesis. LasB elastase, a zinc metalloprotease encoded by the lasB gene, has an elastolytic activity on lung tissue (Jaffar-Bandjee et al., 1995). In addition, the phospholipids contained in pulmonary surfactants may be hydrolysed by two phospholipases C encoded by plcH and plcN (PLC-H and PLC-N, respectively) (Konig et al., 1997; Ostroff et al., 1990). An extracellular neuraminidase is thought to play an important role in implantation of the bacterium (Cacalano et al., 1992; Davies et al., 1999), but the genetic basis of this process is still unknown.

Genetic methods have been used to explore the genetic diversity of various populations of P. aeruginosa strains and to evaluate the rate of genetic recombination in the species. Most of these studies have included few or no CF isolates (Denamur et al., 1993; Kiewitz & Tummler, 2000; Lomholt et al., 2001; Pirnat et al., 2002; Ruimy et al., 2001) or have
specifically explored the genome of P. aeruginosa isolates from CF patients, especially for epidemiological purposes (Boukadida et al., 1993; Hoogkamp-Korstanje et al., 1995). Fewer studies have demonstrated a link between isolates from CF patients and isolates from environmental habitats (Romling et al., 1994). Therefore, the genetic relationships between CF isolates and populations of isolates from patients with other diseases remain unexplored.

We have previously studied enzyme polymorphism in P. aeruginosa isolates recovered from French CF patients by multilocus enzyme electrophoresis (MLEE) (Martin et al., 1999). The data suggested that isolates involved in pulmonary disease in these patients belong to three populations. The purpose of the present study was to determine the extent to which populations of French CF isolates are genetically linked with P. aeruginosa isolates from other pathological origins. For this, genetic features were determined by random amplification of polymorphic DNA (RAPD) and by assessing variations in the prevalence of eight virulence genes for five populations of P. aeruginosa isolates from five different origins, including CF patients. The RAPD analysis identified groups of CF isolates that were strongly correlated with the different populations of CF isolates previously defined by MLEE. The use of isolates from origins other than CF added three pieces of information: (i) P. aeruginosa isolates are not distributed randomly in the RAPD groups according to their origin, (ii) two-thirds of CF isolates belong to one of two RAPD groups that mostly contain respiratory tract isolates from CF and non-CF patients and (iii) one of the RAPD groups is specifically associated with the colonisation of the airways of CF patients and exhibits a particular virulence gene set.

**METHODS**

**CF patients and bacterial isolates.** One hundred and sixty-two P. aeruginosa isolates were studied. Eighty-one isolates were recovered from the sputum of 81 patients with CF (44 males and 37 females) examined in France. The CF patients were from 19 towns scattered across France. Strain 118 was deposited in the Collection de l’Institut Pasteur (Paris) with the number CIP 106484.

A modified version of the Shwachman–Kulczycki–Khaw (SKK) scoring system was used to evaluate the clinical status of CF patients at the time of sputum collection (Shwachman, 1965). Four categories were scored: chest X-ray, general activity, nutrition and physical condition. The maximum score for each category was 25, giving a maximum total score of 100. Patients were then assigned to one of five groups on the basis of their total score: excellent (score of 86–100), good (71–85), moderate (56–70), poor (41–55) and weak (<40).

Non-CF P. aeruginosa isolates were from urine (n = 27), lungs (spumatum or distal bronchial specimens) (n = 25) and wounds (n = 17) from patients from 14 different hospitals in France. Ten P. aeruginosa reference strains isolated from plants were studied: NCPPB 292, NCPPB 1224, NCPPB 2195, NCPPB 2650, NCPPB 2652, NCPPB 2653, ICMP 8649, ICMP 7844, ICMP 7843 and ICMP 12558. P. aeruginosa strains PAO1 and ATCC 10145 were also tested as reference strains. All the bacteria were stored at −80°C in vitamin K3 broth (bioMérieux) containing 20% (v/v) glycerol. Each strain was identified as being P. aeruginosa on the basis of colony morphology, oxidase reaction and the appropriate response pattern in the ID 32 GN system (bioMérieux).

**Preparation of bacterial DNA.** Each strain was subcultured aero-bically on two nutrient-agar plates for 18–24 h at 37°C. Cultures were harvested in 15 ml TE buffer (40 mM Tris/HCl, 2 mM EDTA, pH 8.0) and lysed in 220 µl of a 25% (w/v) aqueous solution of SDS and 30 µl Pronase (Sigma). The mixture was incubated overnight at 37°C to allow cell lysis. DNA was extracted as described previously (Brenner et al., 1982). The DNA obtained was resuspended in 1 ml 1X TE.

**RAPD fingerprinting.** The RAPD PCR mixture (25 µl) consisted of buffer (10 mM Tris/HCl, 50 mM KCl, 2.5 mM MgCl2, pH 8.3), 250 µM of each dNTP (Boehringer Mannheim), 40 pmol oligonucleotide, 1 U AmpliTaq DNA polymerase (Perkin Elmer) and 40 ng genomic DNA. The oligonucleotides used, primers 208 and 272, and the amplification protocol were as described previously (Mahenthiralingam et al., 1996). The PCRs were carried out in a Cetus 9600 DNA Thermal Cycler (Perkin Elmer). A 1-kb ladder (Gibco-BRL) was used as a molecular size standard. Amplified products were stained with ethidium bromide and detected by UV transillumination.

**Computerized analysis.** Photographs of RAPD gels were digitized with a video camera connected to a microcomputer (Bioprofil; Vilber-Lourmat). The Taxotron 2000 package (Institut Pasteur), including Restrictoscan, Restrictotyper, Adanson and Dendrogram, was used for numerical analysis. The size of each fragment was calculated from the distance migrated by use of the reciprocal method of Schaffer & Sederoff (1981). The distance matrix was used to calculate the Dice coefficient complement for each pair of isolates. For a given strain, the RAPD type was defined as the combination of patterns obtained with the two oligonucleotides. Relationships between types were calculated by use of the neighbour-joining method (Saitou & Nei, 1987; Sneath & Sokal, 1973).

**Strategy for the identification of putative neuraminidase genes in P. aeruginosa.** The sequence of the P. aeruginosa PAO1 genome was published recently (accession no. AE004091; Stover et al., 2000). The strategy used to detect putative neuraminidase genes was based on the observation that all described bacterial sialidases have four or five copies of an aspartate box (‘Asp-box’) motif, Ser–X–Asp–X–Gly–X–Thr–Trp, where X is any amino acid (Roggentin et al., 1989; Vinm, 1994). By using this peculiarity and information from the Pseudomonas Genome Project (http://www.pseudomonas.com), two genes containing four Asp-boxes were identified in the P. aeruginosa PAO1 genome; we named them nan1 and nan2.

**Detection of virulence genes by PCR and sequencing.** The prevalence of virulence genes encoding alginate (algD), elastase (lasB), haemolytic phospholipase C (plcH), non-haemolytic phospholipase C (plcN), exoenzyme S (exoS) and exotoxin A (toxA) and of the putative neuraminidase genes (nan1 and nan2) was determined by PCR. The genes were amplified with primers selected on the basis of the published PAO1 sequence (Table 1) (Stover et al., 2000). The PCR mixture contained PCR buffer (10 mM Tris/HCl, 50 mM KCl, 1.5 mM MgCl2, pH 8.3), 200 µM of each dNTP (Boehringer), 12.5 pmol of each primer, DMSO at a final concentration of 4%, 1 U AmpliTaq DNA polymerase (Perkin Elmer) and 25 ng DNA template. The DNA was amplified in a Cetus 9600 DNA Thermal Cycler (Perkin Elmer) using the following protocol: 94°C for 3 min, 30 cycles of 94°C for 30 s, 55°C for 1 min and 72°C for 1 min 30 s, and 72°C for 5 min. Each gene was amplified separately. PCR products were separated in a 1% agarose gel for 1 h at 100 V, stained with ethidium bromide and detected by UV transillumination. Amplified genes were identified on the basis of fragment size. Each of the amplified products for strain CIP 106484, a CF strain from...
our collection, was sequenced using the Thermo Sequenase dye terminator cycle sequencing premix according to the manufacturer's recommendations (Amersham Life Sciences) and an ABI Prism 377 DNA sequencer.

Southern blotting. When a virulence gene was not detected by PCR, Southern blotting was performed (Southern, 1975). Probes for genes were synthesized by PCR using the corresponding primers (Table 1) and DNA from strain PAO1. The probe was labelled with alkaline phosphatase using the AlkPhos Direct labelling kit (Amersham Pharmacia Biotech). Four micrograms of DNA were digested overnight with 20 U EcoRI for nan1 and with 20 U HindIII for exoS. Restriction fragments were resolved on a 1 % agarose gel for 3 h at constant voltage (30 V). The DNA was transferred onto a positively charged nylon membrane (Boehringer Mannheim) by use of a vacuum system (Amersham Pharmacia Biotech) with 20× SSC (0.3 M trisodium citrate, 3 M NaCl, pH 7). The DNA on the membrane was then hybridized with the probe at 55°C in hybridization buffer [1 % SDS, 1 M NaCl, 50 mM Tris/HCl, pH 7.5, 1 % (w/v) blocking reagent]. The hybridized probe was detected by a dioxetane chemiluminescence system (CDP Star; Amersham Pharmacia Biotech) and light emission was recorded on Hyperfilm MP (Amersham Pharmacia Biotech).

Statistical methods. The distribution of virulence genes with respect to genomic groups or strain origin was compared using the χ² test.

RESULTS

RAPD analysis and genetic link between P. aeruginosa isolates

RAPD analysis with primer 208 generated 132 patterns, each containing three to 20 DNA fragments of between 0-2 and 4-2 kbp. RAPD analysis with primer 272 generated 127 patterns, each containing five to 23 DNA fragments of between 0-3 and 4-3 kbp. The combination of RAPD patterns obtained with the two primers identified 162 distinct RAPD types.

The genetic link between each P. aeruginosa strain of the studied population, determined by the neighbour-joining method, is represented as a dendogram (Fig. 1). This analysis identified five primary lineages that allowed the differentiation of five major RAPD groups of isolates (RA, RB, RC, RD and RE) at a dissimilarity distance of 0-43. These groups respectively contained 37, 40, 40, 31 and 14 isolates. Each of the five RAPD groups appeared to be equally as diverse, given that the branches appeared to be the same length in all of the groups.

Distribution of P. aeruginosa isolates in RAPD groups according to their ecological origins

Three of the five RAPD groups were mostly composed of isolates originating from lung disease (Table 2). RA (37 isolates) contained 31 respiratory tract isolates (83.8 %). RB (40 isolates) contained 30 pulmonary tract isolates (75 %). RC (40 isolates) contained 34 pulmonary tract isolates (85 %). RD (31 isolates) contained 22 isolates of non-pulmonary origin (71 %), including urine isolates (7 isolates, 22.6 % of the RD isolates), wound isolates (10 isolates, 32.2% of the RD isolates) and plant strains (5 strains, 16.1 % of the RD isolates). RE (14 isolates) appeared to be a ‘urinary tract group’, because it contained 12 urinary tract isolates (85.7 %). Therefore, P. aeruginosa isolates from respiratory tract (CF and non-CF patients), urine, wounds and plants are not distributed randomly among the five RAPD groups (P < 0.0001). More precisely, this significant difference is observed because some of the lung disease (95 of 106 pulmonary isolates) and urine (12 of 27 urinary isolates) isolates belonged to particular RAPD groups; RA, RB and RC for pulmonary isolates and RE for urinary isolates. Nevertheless, the remaining isolates of each origin appeared to be closely related in RD (31 isolates), which contained 9 of the 106 pulmonary isolates, 7 of the 27 urinary isolates, 10 of the 17 wound isolates and 5 of the 10 plant strains.

Table 1. Primers used for PCR amplification of virulence factors

<table>
<thead>
<tr>
<th>Gene</th>
<th>Primer sequence (5’–3’)</th>
<th>Product (bp)</th>
<th>Position on PAO1 chromosome</th>
<th>G+C content (mol%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>algD</td>
<td>ATGGCAATCACGATCTTTTGCTACCGAACGATGGCTCCCGGC</td>
<td>1310</td>
<td>PA3540</td>
<td>62.9</td>
</tr>
<tr>
<td>lasB</td>
<td>GGAATGAACGAAGCTTTCATGTCGAGCAGCTCCCGTGTGG</td>
<td>300</td>
<td>PA3724</td>
<td>64.3</td>
</tr>
<tr>
<td>toxA</td>
<td>GTTAAACCCGCTCAGCCACATGACTCTCAGGTTGGA</td>
<td>352</td>
<td>PA1148</td>
<td>69.5</td>
</tr>
<tr>
<td>plcH</td>
<td>GAAGCCATGGGCTACCTTCAAGAGTGAACGGAGGCGGTAG</td>
<td>307</td>
<td>PA0844</td>
<td>65.6</td>
</tr>
<tr>
<td>plcN</td>
<td>GGTATCCGAAACCCCGCCTTACAGTTCTGGAACATAGTCCGTAAG</td>
<td>466</td>
<td>PA3319</td>
<td>67.3</td>
</tr>
<tr>
<td>exoS</td>
<td>CTGGAAGGCCGCTCAGACAGGTTCAGGCCGTAAT</td>
<td>504</td>
<td>PA3841</td>
<td>64.0</td>
</tr>
<tr>
<td>nan1</td>
<td>AGGATGATAATTTCATTGCTGTAATACTGTCCTGAGCAGCAGGGTTA</td>
<td>1316</td>
<td>PA2794</td>
<td>48.4</td>
</tr>
<tr>
<td>nan2</td>
<td>ACAGAAACACGGAGGAGGTTATGTTTGCTGATGCTGGTTCA</td>
<td>1161</td>
<td>PA3080</td>
<td>69.4</td>
</tr>
</tbody>
</table>
Fig. 1. Dendrogram showing genetic relationships between 162 P. aeruginosa isolates. The results were obtained by RAPD analysis using two primers. The dendrogram was constructed by the Taxotron program using the neighbour-joining method. Five groups (RA, RB, RC, RD and RE) were identified at a dissimilarity distance of 0.43. The presence of exoS or nan1 is indicated by shading. All the isolates possess algD, lasB, plcH, plcN, toxA and nan2.
Table 2. Distribution of the 162 P. aeruginosa isolates in five genomic groups with respect to ecological origin

Values are numbers (percentages calculated according to the origin of the isolates) of isolates. Isolates are not distributed randomly in the five genomic groups (P < 0.0001; 2 test).

<table>
<thead>
<tr>
<th>Genomic group</th>
<th>Respiratory tract</th>
<th>Urine (n = 27)</th>
<th>Wound (n = 17)</th>
<th>Plant (n = 10)</th>
<th>Reference (n = 2)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>CF isolates (n = 81)</td>
<td>Non-CF isolates (n = 25)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RA (n = 37)</td>
<td>20 (24.7)</td>
<td>11 (44)</td>
<td>3 (11.1)</td>
<td>1 (5.9)</td>
<td>1 (50)</td>
</tr>
<tr>
<td>RB (n = 40)</td>
<td>23 (28.4)</td>
<td>7 (28)</td>
<td>4 (14.8)</td>
<td>4 (23.5)</td>
<td>1 (10)</td>
</tr>
<tr>
<td>RC (n = 40)</td>
<td>31 (38.3)</td>
<td>3 (12)</td>
<td>1 (3.7)</td>
<td>2 (11.7)</td>
<td>3 (30)</td>
</tr>
<tr>
<td>RD (n = 31)</td>
<td>5 (6.2)</td>
<td>4 (16)</td>
<td>7 (25.9)</td>
<td>10 (58.8)</td>
<td>5 (50)</td>
</tr>
<tr>
<td>RE (n = 14)</td>
<td>2 (2.7)</td>
<td>–</td>
<td>12 (44.4)</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Link between P. aeruginosa isolates from CF airways and isolates from other ecological origins

CF P. aeruginosa isolates were equally distributed in the three 'pulmonary RAPD groups', RA, RB and RC (Table 2). Indeed, 20/81 CF isolates (24.7%) were in group RA, 23/81 CF isolates (28.4%) in group RB and 31/81 CF isolates (38.3%) in group RC. However, the link between CF isolates and populations of P. aeruginosa from other origins varied for each of these three RAPD groups. RC was mainly composed of CF isolates (31/40 isolates, 77.5%). RA (37 isolates) contained 20 CF isolates (54%), which were closely related to isolates that are mostly implicated in lung disease in non-CF patients (11 of the 17 remaining isolates). RB (40 isolates) contained 23 CF isolates (57%), which were related to isolates from a wide variety of origins (Table 2). In contrast, CF isolates only accounted for 16.1% (5/31 isolates) of the isolates in group RD and for 14.3% (2/14 isolates) of the isolates in group RE.

There was no significant correlation between the distribution of CF isolates in the five RAPD groups and SKK score (data not shown). Nevertheless, the lower the SKK score, the higher the prevalence of isolates belonging to group RC.

Prevalence of virulence genes detected by PCR and Southern blotting

We used PCR to assess the prevalence of eight virulence genes. Any isolates that gave negative PCR results were tested by Southern blotting. Each of the amplified products obtained for strain CIP 106484 was sequenced. The identity of each gene was confirmed by comparison with the PAO1 sequence in the GenBank database. PCR detected algD, lasB, toxA, plcH, plcN and nan2 in all of the 162 isolates studied. exoS was only detected in 137 isolates (84.5%), and Southern blotting was negative for all PCR-negative isolates. nan1 was detected in 86 isolates (53%), 80 by PCR and six by Southern blotting for PCR-negative isolates.

The prevalence of exoS was significantly higher in CF respiratory isolates than in non-CF isolates (P = 0.002) and ranged from 64.7% for wound isolates to 93.8% for CF isolates (Table 3).

The distribution of nan1 was significantly related to strain origin and RAPD group (Table 3). The prevalence of nan1 was higher in CF isolates: 50 of the 81 CF isolates (61.7%) were positive for nan1 compared with only 36 of the 81 non-CF isolates (44.4%) (P = 0.038). Consequently, the prevalence of nan1 was highest in group RC, which is composed mainly of CF isolates: 33 of the 40 isolates (82.5%) in group RC contained nan1 (P = 0.0003). The prevalence of nan1 was lowest in group RD (32.2%; 10/31 isolates).

The prevalence of this gene in CF isolates tended to increase as the clinical status worsened. Indeed, nan1 was detected in 57% (23/40 isolates) of isolates from patients with an

Table 3. Prevalence of exoS and nan1 according to ecological origin of P. aeruginosa isolates

Values are numbers (percentages) of isolates.

<table>
<thead>
<tr>
<th>Virulence gene</th>
<th>CF isolates</th>
<th>Non-CF isolates</th>
<th>P value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sputum (n = 81)</td>
<td>Urine (n = 27)</td>
<td>Lung (n = 25)</td>
</tr>
<tr>
<td>exoS*</td>
<td>76 (93.8)</td>
<td>19 (70.4)</td>
<td>20 (80)</td>
</tr>
<tr>
<td>nan1*</td>
<td>50 (61.7)</td>
<td>11 (40.7)</td>
<td>12 (48)</td>
</tr>
</tbody>
</table>

*Calculated by 2 test for the difference between prevalence in CF and non-CF isolates.
excellent or good clinical status, in 63% (17/27) of isolates from patients with a moderate status and in 71% (10/14 isolates) of isolates from patients with a poor or weak clinical status. Nevertheless, this variation is not statistically significant.

**DISCUSSION**

To improve our understanding of the genetic link between CF isolates and isolates from other origins, we studied 162 *P. aeruginosa* isolates from five different ecological origins. The relationships between isolates were calculated by numerical analysis of genetic features determined by RAPD (Wang et al., 1993). This molecular approach has been used successfully, for example, for *P. aeruginosa*, to determine the link between pathogenic and non-pathogenic isolates or between pathogenic isolates and isolates originating from various ecological sites (Ruimy et al., 2001). RAPD is an arbitrary method that does not necessarily explore the totality of the genome. The ability of such typing methods to establish the phylogeny of bacterial populations that are widely distributed in the environment is a matter of debate (Denamur et al., 1993; Maynard Smith et al., 1993). Nevertheless, the CF isolates we used here have previously been analysed by MLEE (Martin et al., 1999), an approach that has been largely validated for genetic purposes (Selander et al., 1986). Interestingly, the RAPD method allowed us to identify RAPD groups that correlate strongly with MLEE populations (Table 4). This suggests that the RAPD method used in this study is valid for the detection of genetic links between the various populations of *P. aeruginosa* isolates studied.

Previous studies have suggested that there is no correlation between *P. aeruginosa* clones and diseases or environmental habitats (Foght et al., 1996; Kiewitz & Tummler, 2000; Pirnay et al., 2002; Romling et al., 1994). The use of one or a few strains to represent one disease or environmental habitat may have limited the conclusions of these studies. Our data are nevertheless in accordance with this finding. This is particularly valuable for CF isolates, which were almost equally distributed in three of the five RAPD groups (RA, RB and RC) that also contained about 30% of the urinary isolates, 40% of the wound isolates and 50% of the plant strains. This indicates that the biological conditions in the CF respiratory tract allow *P. aeruginosa* strains from a wide range of ecological origins to be pathogenic. Thus, as previously suggested, CF patients should be isolated when they are hospitalized, and meetings of CF patients, for example in cure centres or holiday camps, may not be advisable (Kosorok et al., 1998; Ojeniyi et al., 2000; Tummler et al., 1991).

Although the *P. aeruginosa* isolates found in CF patients are genetically linked to isolates from patients with a number of other diseases, two of the five RAPD subpopulations of *P. aeruginosa* isolates (RA and RC) mostly contained isolates from the pulmonary tract of CF and non-CF patients. The existence of subpopulations that may have greater potential to colonize or infect the pulmonary tract of CF patients and favour other chronic obstructive airway diseases was suggested previously by an analysis of the incidence of common pyocin types (Tredgett et al., 1990).

Due to the high prevalence of *P. aeruginosa* CF isolates in RAPD group RC (77.5%), this population may have a particular ability to inhabit the pulmonary tract of CF patients and to spread in CF patients. The probability of this type of subpopulation really existing is reinforced by the fact that RAPD group RC corresponds strongly to the ET2 group of isolates previously identified by MLEE (Martin et al., 1999). The emergence of the RAPD group RC may have resulted from a combination of many factors. Although the CF patients included in this study were from 19 different centres, cross-colonization by strains of this RAPD group cannot be excluded, because some of them may have previously been hospitalized together (Hoogkamp-Korstanje et al., 1995; Kosorok et al., 1998; Ojeniyi et al., 2000). However, RAPD group RC, which contained 38% of all CF isolates, also contained 30% of the isolates from plants (Table 2). Therefore, strains of this group may also have emerged from the environment. This is in accordance with the conclusion of another European study, which found that 28% of CF isolates were related to 21% of *P. aeruginosa* isolates from an aquatic habitat (Romling et al., 1994).

The differences in the distributions of virulence factor genes in the populations strengthen the probability that some *P. aeruginosa* strains are better adapted to the pulmonary conditions found in CF patients. The prevalence of exoS was highest in CF pulmonary isolates (94%; Table 3), in agreement with previous studies (Dacheux et al., 2000; Feltman et al., 2001). The prevalence of exoS was similar in isolates from plants (90%; Table 3) and from soil isolates (Ferguson et al., 2001). Nevertheless, 80% of all pulmonary isolates harboured exoS, indicating that this gene plays a major role in all pulmonary infections. The most remarkable genetic feature of CF isolates was the high prevalence of nan1 (62% of the CF isolates versus 44% of non-CF isolates; Table 3). This characteristic more specifically marked isolates in

---

**Table 4. Distribution of electrophoretic types defined by MLEE of *P. aeruginosa* CF isolates with respect to the five genomic groups defined by RAPD**

The distribution of electrophoretic types (ET) defined by MLEE (Martin et al., 1999) correlated very well with the five genomic groups defined by RAPD ($P < 0.0001$; $\chi^2$ test).

<table>
<thead>
<tr>
<th>Genomic group</th>
<th>ET1 ($n = 35$)</th>
<th>ET2 ($n = 35$)</th>
<th>Other ETs ($n = 11$)</th>
</tr>
</thead>
<tbody>
<tr>
<td>RA ($n = 20$)</td>
<td>15</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td>RB ($n = 23$)</td>
<td>13</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>RC ($n = 31$)</td>
<td>3</td>
<td>27</td>
<td>1</td>
</tr>
<tr>
<td>RD ($n = 5$)</td>
<td>3</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>RE ($n = 2$)</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
</tbody>
</table>
RAPD group RC, which mainly contained CF isolates, in which the prevalence of nan1 was 82.5%. In addition, the prevalence of nan1 increased as the clinical state of CF patients worsened. These data suggest a possible role of Nan1 in CF pulmonary disease evolution. The nature of the protein encoded by this gene in P. aeruginosa PA01 is unknown (Stover et al., 2000). Several arguments suggest that nan1 encodes a sialidase, an enzyme theoretically able to release sialic acid from sialylated gangliosides, thus increasing the amount of asialoGM1, a major receptor for adherence to the respiratory tract (Bryan et al., 1998; de Bentzmann et al., 1996; Imundo et al., 1995; Saiman et al., 1992; Saiman & Prince, 1993). The first argument is that the deduced bacterial protein encoded by nan1 possesses four Asp-boxes, a characteristic of bacterial sialidases (Taylor, 1996; Vimr, 1994). Secondly, the deduced protein contains conserved amino acids at key sites that are probably part of the catalytic site (Crennell et al., 1993). Although there is a low degree of similarity between Nan1 and previously identified bacterial sialidases, this is usual for bacterial sialidase genes (Taylor, 1996). Nevertheless, part of the deduced protein encoded by nan1 has 26% identity to part of the Clostridium tertium sialidase protein (Y08695). Interestingly, among non-CF isolates, nan1 was most prevalent in strains from plants, and was found in all strains from plants belonging to RAPD group RC (Fig. 1). This reinforces the conclusion that a group of CF isolates may have originated specifically from the environment. The G+C content of nan1 (48.2 mol%) differs notably from the G+C content of the whole PAO1 genome (66.7 mol%) (Stover et al., 2000). Thus, the gene was probably acquired by horizontal transfer, as observed previously for some other bacterial sialidase genes (Roggentin et al., 1993). The prevalence of nan1 in plant strains, the stability of the genotype of P. aeruginosa isolates from the pulmonary tract of CF patients (Romling et al., 1994; Renders et al., 1996) and the strong correlation between RAPD and MLEE results suggest that such genetic events probably originated in environmental populations of strains rather than in CF populations.

In conclusion, our data suggest that P. aeruginosa strains from a wide range of ecological origins are able to colonize the respiratory tract of CF patients. Nevertheless, three major populations of CF isolates that exhibit a link with isolates of other origins can be identified. One population had no specific characteristics. Isolates in the second population are closely related to isolates from non-CF pulmonary infectious diseases. Isolates in the third population appear to be more particularly associated with the CF pulmonary tract and were related to 30% of isolates from plants. With regard to the virulence gene patterns, the major trait of the CF populations is the high prevalence of nan1. The protein encoded by this gene is unknown, but there is evidence to suggest that it is a neuraminidase. This remains to be confirmed, and the role of this protein in the physiopathology of CF pulmonary disease has yet to be clarified. This is important at a time when neuraminidase inhibitors are being developed for therapeutic purposes.

ACKNOWLEDGEMENTS

We are very grateful for help from the Pseudomonas Genome Project. R.Q. was supported by a grant from the Anjou Mucoviscidose association. We would also like to thank our colleagues from microbiology laboratories for supplying non-CF P. aeruginosa isolates: J. Caillon (Nantes), G. Chambrehieu (La Rochelle sur Yon), F. Geffroy (Quimper-Concarneau), D. Jan (Laval), S. Jarraud (Lyon), C. Martin (Limoges), P. Poudras (Vannes), A. Prandini (Guigamp), H. Senechal (Quimperlé), H. Silvestre (Pontivy), D. Tandé (Brest), J. Vaucel (Saint-Brieuc) and J. F. Ygout (Lorient).

REFERENCES


