Porphyromonas, a potential predictive biomarker of Pseudomonas aeruginosa pulmonary infection in cystic fibrosis

Marlène Keravec,1 Jérôme Mounier,1 Charles-Antoine Guilloux,2 Marie-Sarah Fangous,2 Stanislas Mondot,3 Sophie Vallet,2 Stéphanie Gouriou,2 Rozenn Le Berre,2 Gilles Rault,3 Claude Férec,2 Georges Barbier,1 Patricia Lepage,5 Geneviève Héry-Arnaud

ABSTRACT

Introduction Pseudomonas aeruginosa pulmonary infections are the primary cause of morbidity-mortality in patients with cystic fibrosis (CF). In this cohort study, the objective was to identify candidate biomarkers of P. aeruginosa infection within the airway microbiota.

Methods A 3-year prospective multicentre study (PYOMUCO study) was conducted in Western France and included patients initially P. aeruginosa free for at least 1 year. A 16S-targeted metagenomics approach was applied on iterative sputum samples of a first set of patients (n=33). The composition of airway microbiota was compared according to their P. aeruginosa status at the end of the follow-up (colonised vs non-colonised), and biomarkers associated with P. aeruginosa were screened. In a second step, the distribution of a candidate biomarker according to the two groups of patients was verified by qPCR on a second set of patients (n=52) coming from the same cohort and its load quantified throughout the follow-up.

Results Porphyromonas (mainly P. catoniae) was found to be an enriched phylotype in patients uninfected by P. aeruginosa (p<0.001). This result was confirmed by quantitative PCR. Conversely, in patients who became P. aeruginosa-positive, P. catoniae significantly decreased before P. aeruginosa acquisition (p=0.014).

Discussion Further studies on replication cohorts are needed to validate this potential predictive biomarker, which may be relevant for the follow-up in the early years of patients with CF. The identification of infection candidate biomarkers may offer new strategies for CF precision medicine.

Key messages

Could the airway microbiota be a source of predictive biomarkers of early Pseudomonas aeruginosa pulmonary infection?

The airway microbiota might be more or less permissive to P. aeruginosa in cystic fibrosis (CF) airways.

Based on a cohort study, we found the first candidate biomarker predictive of P. aeruginosa pulmonary infection in CF.

INTRODUCTION

Respiratory polymicrobial infections play a major role in cystic fibrosis (CF) progression and the acquisition of bacterial pathogens during the course of the disease is now well described. After ~25 years old, the establishment of CF pathogens is usually completed, P. aeruginosa being the most predominant species in CF lung. P. aeruginosa has a negative impact on pulmonary function promoting more frequent acute exacerbations. P. aeruginosa is linked to worsened prognosis for patients with CF; they have a decreased life expectancy and experience a more rapid decline in pulmonary function compared with non-colonised patients. At the stage of chronic infection, the eradication of the pathogen is impossible; this places all the hopes of treatment at the early stage of the infection. Indeed, the chances to successfully eradicate P. aeruginosa are the most important in the early stages of P. aeruginosa colonisation.

Improvement in the median survival of patients with CF is correlated with early antibiotic therapy in patients colonised with P. aeruginosa, the eradication success being essentially dependent on how early P. aeruginosa is detected. As early P. aeruginosa pulmonary infection is completely asymptomatic in most cases, the monitoring is based on systematic microbiological analysis. In this aim, P. aeruginosa quantitative PCR (qPCR) detection was shown to be relevant in patients' follow-up. For the coming years, one challenge is to decipher the factors involved in the early onset of P. aeruginosa colonisation. Demographic and environmental factors were shown to increase the risk of P. aeruginosa acquisition, but are not precise enough to predict the risk of P. aeruginosa early colonisation.

In the framework of the present study, we hypothesised that lung commensal micro- biota could be associated with P. aeruginosa...
Figure 1  Two-step approach of the study. Samples were issued from the PYOMUCO cohort study whose patients (n=96), initially (T0) all Pseudomonas aeruginosa (PA) free for at least 1 year, were separated into two groups (group 1 and group 2) according to their P. aeruginosa status at the end of the follow-up (Tf). Group 1 patients remained negative, whereas group 2 patients became positive. In a first step carried out in a first set of patients (n=33), bacterial biomarkers associated with P. aeruginosa were screened by 16S-targeted metagenomics; a candidate biomarker (Porphyromonas catoniae) was revealed. In a second step, distribution of the candidate biomarker according to the two groups of patients was verified by quantitative PCR (qPCR) on a second set of patients (n=52) coming from the same cohort. CF, cystic fibrosis.

Methods
Patient cohort, inclusion criteria and global data
A 3-year prospective multicentre study (PYOMUCO study) was conducted in Western France to assess the time saved in the detection of P. aeruginosa in patients with CF by qPCR compared with culture detection methods. Only patients P. aeruginosa free for at least 1 year were included. The cohort was divided into two groups at the end of the follow-up; group 1 contained patients who remained free of P. aeruginosa while patients from group 2 became positive in culture for P. aeruginosa during the follow-up. For each patient, sputum was collected every 3 months up to the first P. aeruginosa positivity in culture. For this ancillary study, 33 patients with CF of the PYOMUCO cohort were selected as follows: 20 patients from the 36 who became P. aeruginosa-positive at the end of the follow-up, and 13 patients from the 28 who remained P. aeruginosa-negative both in qPCR and culture. The analysis of airway microbiota was performed retrospectively on spontaneous sputum samples collected at two time points: at enrolment (T0) and at the end of the follow-up (Tf) (figure 1). Overall, 75.7% of the samples originated from a paediatric population (<18 years old at sampling time). Clinical and biological data were collected at each sampling time (see online supplementary table S1). The majority of patients were homozygous (n=21, 63.6%) or heterozygous (n=12, 36.4%) for the F508del-CFTR mutation. Four clinical states were defined: baseline clinical state, pulmonary exacerbation, treatment for exacerbation and recovery (BETR categories). Sputum sample quality was verified by cytological examination as previously described. P. aeruginosa was quantified using qPCR and a culture-based method, as previously described. In order to confirm Porphyromonas distribution with respect to P. aeruginosa colonisation throughout the follow-up, P. catoniae absolute quantification was carried out with qPCR on 52 additional patients of the PYOMUCO cohort.

Targeted metagenomics and P. catoniae absolute quantification in sputum samples
Total DNA was extracted using the QIAamp DNA Mini Kit (QIAGEN, Courtabœuf, France) as previously described. For bacterial diversity assessment, barcoded high-throughput 454 pyrosequencing was performed on the amplified V3 and V4 hypervariable regions of the 16S rRNA gene, and data analysed as previously described (Bioproject PRJNA445243). The absolute quantification of P. catoniae was performed using a validated qPCR.

PYOMUCO cohort study
96 patients; 3y-follow-up

Figure 1  Two-step approach of the study. Samples were issued from the PYOMUCO cohort study whose patients (n=96), initially (T0) all Pseudomonas aeruginosa (PA) free for at least 1 year, were separated into two groups (group 1 and group 2) according to their P. aeruginosa status at the end of the follow-up (Tf). Group 1 patients remained negative, whereas group 2 patients became positive. In a first step carried out in a first set of patients (n=33), bacterial biomarkers associated with P. aeruginosa were screened by 16S-targeted metagenomics; a candidate biomarker (Porphyromonas catoniae) was revealed. In a second step, distribution of the candidate biomarker according to the two groups of patients was verified by quantitative PCR (qPCR) on a second set of patients (n=52) coming from the same cohort. CF, cystic fibrosis.

Methods
Patient cohort, inclusion criteria and global data
A 3-year prospective multicentre study (PYOMUCO study) was conducted in Western France to assess the time saved in the detection of P. aeruginosa in patients with CF by qPCR compared with culture detection methods. Only patients P. aeruginosa free for at least 1 year were included. The cohort was divided into two groups at the end of the follow-up; group 1 contained patients who remained free of P. aeruginosa while patients from group 2 became positive in culture for P. aeruginosa during the follow-up. For each patient, sputum was collected every 3 months up to the first P. aeruginosa positivity in culture. For this ancillary study, 33 patients with CF of the PYOMUCO cohort were selected as follows: 20 patients from the 36 who became P. aeruginosa-positive at the end of the follow-up, and 13 patients from the 28 who remained P. aeruginosa-negative both in qPCR and culture. The analysis of airway microbiota was performed retrospectively on spontaneous sputum samples collected at two time points: at enrolment (T0) and at the end of the follow-up (Tf) (figure 1). Overall, 75.7% of the samples originated from a paediatric population (<18 years old at sampling time). Clinical and biological data were collected at each sampling time (see online supplementary table S1). The majority of patients were homozygous (n=21, 63.6%) or heterozygous (n=12, 36.4%) for the F508del-CFTR mutation. Four clinical states were defined: baseline clinical state, pulmonary exacerbation, treatment for exacerbation and recovery (BETR categories). Sputum sample quality was verified by cytological examination as previously described. P. aeruginosa was quantified using qPCR and a culture-based method, as previously described. In order to confirm Porphyromonas distribution with respect to P. aeruginosa colonisation throughout the follow-up, P. catoniae absolute quantification was carried out with qPCR on 52 additional patients of the PYOMUCO cohort.

Targeted metagenomics and P. catoniae absolute quantification in sputum samples
Total DNA was extracted using the QIAamp DNA Mini Kit (QIAGEN, Courtabœuf, France) as previously described. For bacterial diversity assessment, barcoded high-throughput 454 pyrosequencing was performed on the amplified V3 and V4 hypervariable regions of the 16S rRNA gene, and data analysed as previously described (Bioproject PRJNA445243). The absolute quantification of P. catoniae was performed using a validated qPCR.

PYOMUCO cohort study
96 patients; 3y-follow-up

Figure 1  Two-step approach of the study. Samples were issued from the PYOMUCO cohort study whose patients (n=96), initially (T0) all Pseudomonas aeruginosa (PA) free for at least 1 year, were separated into two groups (group 1 and group 2) according to their P. aeruginosa status at the end of the follow-up (Tf). Group 1 patients remained negative, whereas group 2 patients became positive. In a first step carried out in a first set of patients (n=33), bacterial biomarkers associated with P. aeruginosa were screened by 16S-targeted metagenomics; a candidate biomarker (Porphyromonas catoniae) was revealed. In a second step, distribution of the candidate biomarker according to the two groups of patients was verified by quantitative PCR (qPCR) on a second set of patients (n=52) coming from the same cohort. CF, cystic fibrosis.

Methods
Patient cohort, inclusion criteria and global data
A 3-year prospective multicentre study (PYOMUCO study) was conducted in Western France to assess the time saved in the detection of P. aeruginosa in patients with CF by qPCR compared with culture detection methods. Only patients P. aeruginosa free for at least 1 year were included. The cohort was divided into two groups at the end of the follow-up; group 1 contained patients who remained free of P. aeruginosa while patients from group 2 became positive in culture for P. aeruginosa during the follow-up. For each patient, sputum was collected every 3 months up to the first P. aeruginosa positivity in culture. For this ancillary study, 33 patients with CF of the PYOMUCO cohort were selected as follows: 20 patients from the 36 who became P. aeruginosa-positive at the end of the follow-up, and 13 patients from the 28 who remained P. aeruginosa-negative both in qPCR and culture. The analysis of airway microbiota was performed retrospectively on spontaneous sputum samples collected at two time points: at enrolment (T0) and at the end of the follow-up (Tf) (figure 1). Overall, 75.7% of the samples originated from a paediatric population (<18 years old at sampling time). Clinical and biological data were collected at each sampling time (see online supplementary table S1). The majority of patients were homozygous (n=21, 63.6%) or heterozygous (n=12, 36.4%) for the F508del-CFTR mutation. Four clinical states were defined: baseline clinical state, pulmonary exacerbation, treatment for exacerbation and recovery (BETR categories). Sputum sample quality was verified by cytological examination as previously described. P. aeruginosa was quantified using qPCR and a culture-based method, as previously described. In order to confirm Porphyromonas distribution with respect to P. aeruginosa colonisation throughout the follow-up, P. catoniae absolute quantification was carried out with qPCR on 52 additional patients of the PYOMUCO cohort.

Targeted metagenomics and P. catoniae absolute quantification in sputum samples
Total DNA was extracted using the QIAamp DNA Mini Kit (QIAGEN, Courtabœuf, France) as previously described. For bacterial diversity assessment, barcoded high-throughput 454 pyrosequencing was performed on the amplified V3 and V4 hypervariable regions of the 16S rRNA gene, and data analysed as previously described (Bioproject PRJNA445243). The absolute quantification of P. catoniae was performed using a validated qPCR.
scheme with the standard curve method. The qPCR was set up on the ABI 7500 Fast Real-Time PCR system (Applied Biosystems, Foster City, California, USA) with SYBR Green, and original primers (sense: 5′-GTGTCGTCGTGCATCTA-3′; antisense: 5′-GCATTCCTCAGGGTGCAGTC-3′) targeting the rplb gene. PCR reactions were carried out in a total volume of 25 µL with 12.5 µL of Select Mastermix (Applied Biosystems), and a temperature profile of 50°C for 2 min, 95°C for 10 min, followed by 40 cycles at 95°C for 15 s, 60°C for 60 s, 95°C for 30 s and 60°C for 15 s.

**Bioinformatic and statistical analyses**

Sequences were analysed with the standard UPARSE pipeline according to Edgar’s instructions as previously described5 (see online supplementary file 1).

Statistical comparison between groups was performed with the Mann-Whitney U and Kruskal-Wallis tests, and linear discriminant analysis effect size (LEfSe) was used to elucidate bacterial taxa associated with group 1 or group 2 patients. The false discovery rate was calculated to correct for multiple hypothesis testing. Principal component analysis and clustering analysis were computed on different distance matrix to document the presence of enterotype-like clusters in airway CF microbiota.7 These clusters were named pulmotypes.

**RESULTS**

The cohort samples clustered into three pulmotypes (p=0.001, Kruskal-Wallis test) driven by the differences in relative abundance of three dominant genera, *Streptococcus*, *Haemophilus* and *Staphylococcus*, as well as other co-occurring genera (see online supplementary figure S1). Overall, 11 predominant genera (relative abundance ≥1%) were found, including *Porphyromonas* which abundance varied between 2.5% and 6.9% depending on the pulmotype. Random forest analysis revealed a significant connexion between relative abundance of *Pseudomonas* and *Porphyromonas* (figure 2A).

Interestingly, in group 1 patients who remained uninfected by *P. aeruginosa* during the follow-up (figure 2B), Porphyromonas relative abundance at T0 was significantly higher than in group 2 patients (p<0.001, Mann-Whitney U test) (figure 2C).

As *Porphyromonas* reads mainly corresponded to *P. catoniae* species, we focused on this bacterial species. In order to check the distribution of *P. catoniae* according to the patient group, we quantified *P. catoniae* by qPCR in another set of patients from the PYOMUCO cohort. For group 1 patients, we did not observe any statistical difference in *P. catoniae* population between the first (T0) and last sample (Tf) (p=0.41, t-test). Conversely, group 2 patients showed a significant drop in *P. catoniae* population (p=0.039, t-test) (figure 2D). Then, we compared

---

**Figure 2** Inverse correlation between *Pseudomonas aeruginosa* and *Porphyromonas* revealed by both 16S rRNA microbiota study and quantitative PCR study in the two groups of patients with cystic fibrosis: group 1 patients (G1) remained *P. aeruginosa*-negative from the initial visit (T0) to the last one (Tf); group 2 patients (G2), initially negative at T0, became *P. aeruginosa*-positive at Tf. (A) Results of random forest analysis showing the 15 taxa that contributed the most to each group based on the measure of mean decrease in accuracy. (B) Normalised abundance of *P. aeruginosa* in each group at the two time points. (C) Normalised abundance of *Porphyromonas* in each group at the two time points. (D) Absolute quantification of *P. catoniae* in each group (G1, G2) at the two time points (T0, Tf); a third time point was added (Tx) that corresponded to the visit which preceded the last one (Tf).
patients according to their \textit{P. catoniae} population before \textit{P. aeruginosa} colonisation. Group 1 had a significantly higher initial \textit{P. catoniae} absolute quantity than group 2 ($p=0.026$). Finally, we tested the predictive power of \textit{P. catoniae}. To do this, we analysed the susceptibility of patients with CF to acquire \textit{P. aeruginosa} according to the presence or absence of \textit{P. catoniae} in the penultimate sputum (Tx) (figure 2D). We observed that 40.7\% of patients without \textit{P. catoniae} developed a \textit{P. aeruginosa} infection the visit after (3 months later), while only 24\% of patients positive for \textit{P. catoniae} developed the infection (table 1).

**DISCUSSION**  
Bacteria from the \textit{Porphyromonas} genus are anaerobic commensals of the core pulmonary microbiota in healthy people\textsuperscript{8} and also described as part of the CF pulmonary core microbiota.\textsuperscript{5,6,9–12} Characterisation of bronchoalveolar lavages’ microbiota in infants with asymptomatic CF retrieved \textit{Porphyromonas} as one of the six highest abundant taxa.\textsuperscript{10}

Looking into details at the taxonomic affiliation of \textit{Porphyromonas} reads, the vast majority of them were affiliated to \textit{P. catoniae}, which is in agreement with other culture-dependent and culture-independent studies.\textsuperscript{11,12} Interestingly, in another study, the abundance of \textit{Porphyromonas} was significantly lower in sputa of patients with \textit{P. aeruginosa} infection the visit after (3 months later), while only 24\% of patients positive for \textit{P. catoniae} developed the infection (table 1).

<table>
<thead>
<tr>
<th>Patient group*</th>
<th>Negative</th>
<th>Positive</th>
</tr>
</thead>
<tbody>
<tr>
<td>Group 1</td>
<td>16 (59.3%)</td>
<td>19 (46%)</td>
</tr>
<tr>
<td>Group 2</td>
<td>11 (40.7%)</td>
<td>6 (24%)</td>
</tr>
<tr>
<td>Total</td>
<td>27</td>
<td>25</td>
</tr>
</tbody>
</table>

\*Group 1, patients who remained \textit{P. aeruginosa}-negative; group 2, patients who became \textit{P. aeruginosa}-positive; Hazard Ratio 1.7.

†Detection carried out on the penultimate sputum sample.

no \textit{P. catoniae} within their airway microbiota showed 1.7-fold risk of acquiring \textit{P. aeruginosa} later.

Taken together, these results suggest that \textit{P. catoniae} may be considered as a favourable prognostic biomarker in CF. Further prospective studies on replication cohorts are needed to define the benefit provided by \textit{P. catoniae} quantification in identifying patients with a higher risk of \textit{P. aeruginosa} infection. In case of confirmation, we suggest to perform molecular quantifications of both \textit{P. catoniae} and \textit{P. aeruginosa} as part of the CF diagnosis toolbox in order to enlarge the ‘window of opportunity’ in the management of \textit{P. aeruginosa} infection.

To conclude, this study showed the crucial importance of microbiota data in the management of patients with CF. In the personalised and precision medicine era, microbiota-based study could identify signatures that could be useful in predicting the CF progression. Identification of new bacteria of interest opens the possibility of using them as diagnostic biomarkers or as companion diagnostic test. Further cohort studies are needed to validate these findings and to address the question of causality. The influence of the input microbiota on CF progression during early life has also to be investigated.

In the not-too-distant future, study of both biochemical and microbial signatures will constitute new approaches to understand CF microbiology.

**Table 1** Table of contingency showing the distribution of patients with cystic fibrosis (n=52) according to their \textit{Porphyromonas catoniae} colonisation status during the follow-up with respect to \textit{Pseudomonas aeruginosa} status at the end of the follow-up

**Author affiliations**  
1EA3882-Laboratoire de Biodiversité et Ecologie Microbienne (LUBEM), Université de Bretagne Occidentale, Brest, France  
2UMR1078 ‘Génétique, Génomique Fonctionnelle et Biotechnologies’, INSERM, EFS, Université de Bretagne, IBSAM, CHU de Brest, University of Brest, Brest, France  
3Micalis Institute, INRA, AgroParisTech, University Paris-Saclay, Jouy-en-Josas, France  
4Centre de Ressources et de Compétences de la Mucoviscidose, Illids, Roscoff, France  
5Micalis Institute, INRA, AgroParisTech, University Paris-Saclay, Jouy-en-Josas, France

**Acknowledgements** The authors gratefully acknowledge the patients with CF who participated in the study. The authors thank Iain MacGill for English revision of the manuscript.

**Contributors** GH-A and JM conceived the study. SG, C-AG and MK performed DNA extraction, qPCR and pyrosequencing. MK, JM, PL and GB performed pyrosequencing data and statistical analysis. MK, C-AG, M-SF and SG performed multiplex PCR. GH-A, RLB, CF and GR contributed to data collection. JM, PL, SV and SM contributed to design. MK, GH-A, JM and PL wrote the manuscript. All authors participated in data analysis and editing the manuscript for final approval. GH-A and JM take overall responsibility for the content of the manuscript.

**Funding** This work was supported by a grant from the French Cystic Fibrosis Associations ‘Vaincre la Mucoviscidose’ and ‘Grégory Lemarchal’ (contract no. RC20170501971).

**Competing interests** None declared.

**Patient consent for publication** Not required.

**Ethics approval** Two review boards, the local Comité de Protection des Personnes VI-Ouest and the institutional review board of the Brest University Hospital Centre, approved the protocol.

**Provenance and peer review** Not commissioned; externally peer reviewed.

**Data sharing statement** No additional data are available.

**Author note** This work was supported by a grant from the French Cystic Fibrosis Associations ‘Vaincre la Mucoviscidose’ and ‘Grégory Lemarchal’ (contract no. RC20170501971).

**Copyright** © 2019 The Author(s). This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (http://creativecommons.org/licenses/by-nc-nd/4.0/), which permits non-commercial re-use, distribution, and reproduction in any medium, provided the original work is unaltered and is properly cited. The moral rights of the named individual(s) have been asserted.
REFERENCES


2. ECFS. The early cystic fibrosis years de Boeck and K. Southern, 2018.


