



HAL
open science

Genomic Diversity in Porphyromonas: Evidence of Porphyromonas catoniae Commensality in Lungs

Lourdes Velo-Suarez, Yann Moalic, Charles-Antoine Guilloux, Jacky Ame, Claudie Lamoureux, Stephanie Gouriou, Rozenn Le Berre, Clemence Beauruelle, Genevieve Hery-Arnaud

► **To cite this version:**

Lourdes Velo-Suarez, Yann Moalic, Charles-Antoine Guilloux, Jacky Ame, Claudie Lamoureux, et al.. Genomic Diversity in Porphyromonas: Evidence of Porphyromonas catoniae Commensality in Lungs. 2024. hal-04780096

HAL Id: hal-04780096

<https://hal.science/hal-04780096v1>

Preprint submitted on 13 Nov 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1 **Genomic Diversity in *Porphyromonas*:**

2 **Evidence of *P. catoniae* Commensality in Lungs**

3

4 **1.1 Author names**

5 Lourdes Velo-Suarez¹, Yann Moalic¹, Charles-Antoine Guilloux², Jacky Ame², Claudie
6 Lamoureux^{2,3}, Stéphanie Gouriou², Rozenn Le Berre^{2,4}, Clémence Beauruelle^{2,3}, Geneviève
7 Héry-Arnaud^{1,2,3} ORCID 0000-0002-2089-3825

8 **1.2 Affiliation(s)**

- 9 1. Centre Brestois d'Analyse du Microbiote (CBAM), CHRU de Brest, Brest, France
10 2. INSERM, EFS, Univ Brest, UMR 1078, GGB, 29200 Brest, France.
11 3. Department of Infectious Agents, Brest University Hospital, Boulevard Tanguy Prigent,
12 29200 Brest, France
13 4. Department of Internal Medicine and Pulmonology, Brest University Hospital, Boulevard
14 Tanguy Prigent, 29200 Brest, France

15 **1.3 Corresponding author and email address**

16 Geneviève Héry-Arnaud. e-mail: hery@univ-brest.fr

17 **1.4 Keywords**

18 *Porphyromonas*; *Porphyromonas catoniae*; whole genome sequencing; lung; antimicrobial
19 resistance

20 **1.5 Abbreviations**

21 Anaerobe Basal Broth (ABB)
22 Anaerobe Basal Agar supplemented with 5% sheep blood (ABA-SB)
23 Antibiotic resistance genes (ARGs)
24 Committee for Antibiogram of the French Society of Microbiology (CA-SFM)
25 Cystic fibrosis (CF)
26 European Committee on Antimicrobial Susceptibility Testing (EUCAST)
27 Hidden Markov Model (HMM)
28 Markov Cluster Algorithm (MCL)
29 Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF
30 MS)
31 MicroScope Microbial Genome Annotation and Analysis Platform (MaGe)
32 Peptone Yeast (PY)
33 *Porphyromonas* other than *gingivalis* (POTG)
34 Rheumatoid arthritis (RA)
35 Whole genome sequencing (WGS)

36 **2. Abstract**

37 Data on the genomics of *Porphyromonas* species other than *Porphyromonas gingivalis*
38 (POTG), particularly within pulmonary environments, are scarce. In this study, we conducted
39 whole genome sequencing (WGS) on pulmonary isolates of POTG, specifically *P. catoniae*
40 (n=3), *P. pasteri* (n=1), and *P. uenonis* (n=2), from people with cystic fibrosis. These genomic
41 analyses were complemented with antimicrobial susceptibility tests for these strains.

42 We compared the genomic sequences of these pulmonary isolates with those of previously
43 characterized *Porphyromonas* species. Our study revealed a distinct clade differentiation
44 between non-pigmented and pigmented *Porphyromonas* species. Interestingly, unlike
45 *P. gingivalis*, the pulmonary POTG strains lacked known virulence genes, except a putative
46 hemolysin gene. Regarding antibiotic resistance, notable resistances were limited to
47 vancomycin in *P. catoniae* and clindamycin in *P. uenonis*.

48 These findings support the hypothesis that POTG species may predominantly behave as
49 commensals in the lung environment rather than as pathogens.

50 **3. Impact statement**

51 *Porphyromonas* constitutes a prevalent genus of anaerobic bacteria within the respiratory
52 tract. Despite its ubiquity, the precise taxonomic delineation of *Porphyromonas* species
53 implicated in respiratory conditions needs to be better defined, as most microbiome analyses
54 report results only at the genus level. Consequently, data about these pulmonary species often
55 default to associations with *Porphyromonas gingivalis*, a well-characterized oral pathogen. In
56 this study, we sequenced the complete genomes of six *Porphyromonas* strains isolated from
57 the airway microbiota of people with cystic fibrosis (CF) to enhance the representation and
58 identification of *Porphyromonas* other than *gingivalis* (POTG) species. Our phylogenomic
59 analysis aimed to elucidate the diversity within pulmonary POTG. Notably, genomes of the
60 commensal species *Porphyromonas catoniae* isolated from CF patients did not harbor
61 virulence genes typically associated with *P. gingivalis*. Additionally, phenotypic resistance
62 profiling against nine clinically relevant antibiotics revealed low resistance levels,
63 notwithstanding the frequent antibiotic treatments administered to CF patients. Our findings
64 provide compelling evidence for the non-pathogenic role of these POTG species in the
65 pulmonary environment.

66 **4. Data summary**

67 Six *Porphyromonas* genomes generated in this study are available in the Sequence Read
68 Archive and GenBank databases under BioProject accession PRJEB75658.



69 The raw data set generated during the current study is available in the European Nucleotide
70 Archive (repository with the project accession number.

71 The authors confirm that all supporting data, code, and protocols have been provided within
72 the article or through supplementary data files.

73

74 **5. Introduction**

75 In recent years, the composition of the pulmonary microbiota has garnered significant interest,
76 particularly regarding its role in chronic respiratory diseases. The complex ecology of the lung
77 microbiota, rich even in anaerobes among healthy populations, provides a pivotal lens through
78 which disease progression can be viewed (1-15). *Porphyromonas*, ranking among the top ten
79 anaerobic taxa in healthy lungs, exemplifies the metabolic diversity of lung anaerobes,
80 spanning both commensal and pathogenic roles (16, 17). Notably, *Porphyromonas gingivalis*,
81 a key pathogenic species, has been extensively studied for its virulence and adaptation
82 strategies. However, other *Porphyromonas* species, such as *Porphyromonas catoniae* and
83 *Porphyromonas pasteri*, typically associated with the healthy oral microbiota, have yet to be
84 investigated, particularly from pulmonary sources (18, 19).

85 Our study aims to fill this gap by performing whole-genome sequencing (WGS) on pulmonary
86 *Porphyromonas* isolates previously collected (17). We explore the genomic distinctions
87 between oral and pulmonary strains and identify potential genes of interest within the genomes
88 of pulmonary *Porphyromonas* other than *gingivalis* (POTG). Additionally, we examine the
89 antibiotic resistance genes (ARGs) in conjunction with phenotypic resistance profiles to
90 assess whether POTG could act as a commensal species within the respiratory tract. Our
91 findings could significantly impact our understanding of the potential commensal or pathogenic
92 roles of less-studied *Porphyromonas* species.

93 **6. Methods**

94 **Isolation and Culture of Strains**

95 The clinical strains of POTG employed in this study comprised two strains of *P. catoniae* (PC),
96 one strain of *P. pasteri* (PP), and two strains of *P. uenonis* (PU), all isolated from the sputum
97 of people with cystic fibrosis (CF). Sample selection adhered to the Murray-Washington quality
98 score to exclude contamination from saliva (20). Collection was conducted using a device
99 specifically designed to preserve an anaerobic environment as previously described (17). This

100 involved plating the samples within an anaerobic chamber (Bactron 300, Sheldon
101 Manufacturing) that maintained a gas composition of 5% CO₂, 5% H₂, and 90% N₂.

102 The bacterial colonies were initially cultivated on Anaerobe Basal Agar supplemented with 5%
103 sheep blood (ABA-SB) from ThermoFisher. ABA-SB media containing kanamycin-vancomycin
104 and colistin-nalidixic acid were also employed for selective cultivation. The strains were
105 identified using matrix-assisted laser desorption/ionization time-of-flight mass spectrometry
106 (MALDI-TOF MS) with the Biotyper MBT system (Bruker). Additional confirmation for some
107 colonies was achieved by 16S rRNA sequencing performed on an Applied Biosystems 3130xl
108 instrument (ThermoFisher) using the BigDye Terminator Cycle Sequencing Kit Ready
109 Reaction v3.1 (Applied Biosystems) as outlined by Lamoureux et al. (17).

110 *Porphyromonas* cultures were grown on ABA-SB plates for 4 to 7 days, followed by
111 subculturing in a homemade Peptone Yeast (PY) liquid medium for genomic DNA extraction.
112 The cultures were incubated for an additional three days at 37°C within the anaerobic
113 chamber, ensuring optimal growth conditions for subsequent analyses.

114

115 **DNA Extraction and Sequencing**

116 DNA was extracted using the QIAamp DNA Mini Kit (Qiagen), following the manufacturer's
117 instructions, with elution in 100 µL. We employed two distinct sequencing platforms to ensure
118 comprehensive genomic data coverage:

119 1. *PacBio Sequel System*. Library preparation was carried out using the Template Prep Kit V2
120 per the manufacturer's specifications. Sequencing was conducted in a single run at the Faculty
121 of Medicine, Brest, France, using the Sequel Sequencing Kit 3.0 and the Sequel SMRT Cell
122 1M v3 LR Tray. This system is noted for its long-read capabilities, crucial for assembling
123 complex genomic regions.

124 2. *Illumina MiSeq System*. Libraries were prepared using the TruSeq Nano kit, following the
125 manufacturer's guidelines. Whole genome sequencing was performed in a single run using
126 V3 chemistry (2x250 bp) at the Genotoul platform, Castanet-Tolosan-Auzeville, France. This
127 system is favored for its high-precision short-read sequencing.

128 This hybrid sequencing strategy was designed to maximize the coverage and accuracy of the
129 genomic data, facilitating a deeper understanding of the genetic intricacies of the
130 *Porphyromonas* species under study.

131

132 **Genome Assembly**

133 *Quality Control*. Raw sequence reads underwent rigorous quality control to assess and filter
134 out low-quality sequences. Adapter sequences and substandard reads were removed using



135 bbduk from the bmap suite v38.92-0, setting a quality base threshold of 20 and a minimum
136 length requirement of 200 bp. Further, we assessed quality and potential contamination using
137 kat v2.4.2 and kraken2 v2.1.2 to identify human and other non-target sequences.

138 *Genome Assembly and Annotation.* Hybrid genome assemblies were conducted using the
139 Unicycler pipeline v0.4.8-beta, incorporating a suite of tools including spades.py v3.13.0,
140 racon v1.3.3, makeblastdb v2.9.0+, tblastn v2.9.0+, bowtie2-build v2.3.5, bowtie2 v2.3.5,
141 samtools v1.9, java v11.0.1, and pilon v1.23 (21). Assembly quality statistics were generated
142 using Quast v5.02, and genome completeness and contamination levels were evaluated with
143 CheckM v1.0.18.

144 *Scaffold Profiling and Gene Identification.* We utilized Anvi'o v7 for scaffold profiling,
145 employing Prodigal v2.6.354 with default settings for gene identification and HMMER v3.355
146 to detect bacterial single-copy genes and ribosomal RNA-based Hidden Markov Model
147 (HMM). Short reads from each strain were mapped back to their respective scaffolds using
148 Bowtie2 v2.4.4-0, with a minimum identity threshold of 95%. The mapped reads were then
149 stored as BAM files using samtools. Anvi'o was subsequently used to profile each BAM file,
150 allowing for precise estimation of coverage and detection statistics for each scaffold.

151

152 **Phylogenomics and Pan-genome Analysis**

153 To determine the phylogenetic position of our six pulmonary *Porphyromonas* strains, we
154 accessed 99 assigned *Porphyromonas* genomes and 23 non-redundant *Alloprevotella*
155 genomes (serving as outgroup) from the NCBI database using the ncbi-genome-download
156 tool (<https://github.com/kbclin/ncbi-genome-download>). We assessed whole-genome similarity
157 using the Average Nucleotide Identity (ANI), computed with FastANI v1.32. For phylogenomic
158 and pangenomic analyses, only genome pairs from NCBI exhibiting FastANI values greater
159 than 98% were retained, narrowing the dataset to 59 *Porphyromonas* genomes and 15
160 outgroup genomes. Using HMMER v3.3, we extracted HMM hits for each genome, focusing
161 on 36 ribosomal single-copy core genes. These genes were aligned, and gaps affecting more
162 than 50% of nucleotide positions were removed using trimAl v1.4.1, ensuring alignment quality
163 for subsequent analyses. The phylogenetic analysis of these high-quality ribosomal gene
164 alignments was conducted on a concatenated dataset. Substitution models for each gene
165 were automatically selected from the GTR + F + Γ + I + R family using model-testing algorithms
166 to find the best-fitting partition scheme. Phylogenetic trees were inferred using maximum
167 likelihood estimation with proportional variation in branch lengths across partition subsets,
168 implemented in IQ-TREE2 v1.7, using the LG + F + R4 model. The robustness of the
169 phylogenetic tree was supported by 5,000 bootstrap replicates, providing statistical confidence
170 in the branching patterns observed.

171 To explore the genetic diversity and evolutionary relationships within the *Porphyromonas*
172 genus, we employed the Anvi'o platform to compute and visualize the pan-genome of
173 *Porphyromonas*, encompassing both our strains and reference genomes. We stored 65
174 genomes (59 reference genomes from NCBI and our 6 pulmonary strains) in an Anvi'o
175 database using the `anvi-gen-genomes-storage` command. For the pan-genomic analysis, we
176 focused on closely related genomes available from NCBI. Creating the pan-genome involved
177 passing the stored database to the `anvi-pan-genome` command, which utilized NCBI's
178 BLAST to assess gene similarity within and between the genomes. To cluster these genes
179 into functionally coherent groups, we applied the Markov Cluster Algorithm (MCL), facilitating
180 the identification of core and accessory genomic components across the studied
181 *Porphyromonas* species. This comprehensive analysis allows us to better understand the
182 genomic architecture and variability among species, which is crucial for elucidating their
183 evolutionary trajectories and potential functional roles in different environmental niches.

184

185 **Genome Annotation and Characterization of Pulmonary *P. catoniae* Strains**

186 The assembled genomes of *P. catoniae* isolated from pulmonary sources were
187 comprehensively annotated and analyzed using several advanced bioinformatics tools. We
188 utilized Prokka v1.13 (22) for rapid annotation, the MicroScope Microbial Genome Annotation
189 and Analysis Platform (MaGe) (23), the NCBI integrated Prokaryotic Genome Annotation
190 Pipeline (PGAP), and the RAST server v2.0 (Overbeck, 2014). These tools were employed
191 with default parameters to ensure a standardized approach across all annotations.

192 To delve deeper into the metabolic capabilities of our lung *P. catoniae* strains, we compared
193 their profiles with those of *P. catoniae* ATCC 51270 and several strains of *P. gingivalis* (ATCC
194 33277, TDC60 and W83), using the KEGG and BioCyc databases. This comparison was
195 facilitated by the analytical tools available on the MicroScope platform, which also enabled the
196 detection of macromolecular systems via MacSyFinder v1.0.2, prediction of CRISPR arrays
197 using CRISPRCasFinder v4.2.19, identification of antibiotic resistance genes through the
198 CARD database v3.0.2, and exploration of secondary metabolite regions using antiSMASH
199 v5.0.0. Additionally, we investigated the presence of mobile genetic elements in the genomes
200 using Phaster, viralVerify (24), and VirSorter2 (25). These analyses provided insights into the
201 genomic flexibility and potential adaptative mechanisms of *P. catoniae* in the pulmonary
202 environment.

203

204 **Phenotypic Tolerance and Antibiotic Resistance Testing**



205 Given the genomic identity established among the strains PC103, PC123, and PC135,
206 phenotypic tests were exclusively conducted on the PC103 strain to evaluate its environmental
207 tolerance capabilities.

208 *Oxygen Tolerance.* The PC103 strain's oxygen tolerance was assessed by suspending the
209 bacteria in non-oxygen-reduced physiological serum at two initial concentrations: a high
210 concentration of 10^7 CFU.mL⁻¹ and a lower concentration of 10^3 CFU.mL⁻¹. The suspensions
211 were exposed to ambient oxygen, and samples were periodically collected, plated, and
212 incubated for 7 days at 37°C in an anaerobic chamber (5% CO₂, 5% H₂, 90% N₂) to measure
213 the residual inoculum.

214 *Acid Tolerance.* PC103's acid tolerance was tested using ABA-SB medium acidified with 1M
215 HCl to achieve pH values of 6.86, 6.38, 6.0, 5.5, and 5.0. Actively growing cultures of
216 *P. catoniae* PC103 on standard ABA-SB were transferred to the pH-adjusted ABA-SB. These
217 plates were incubated under the same anaerobic conditions for 7 days, after which colony
218 growth was evaluated.

219 *Cadmium Tolerance.* Cadmium tolerance was examined using ABA-SB medium
220 supplemented with increasing cadmium concentrations, ranging from 0 to 80 mg.L⁻¹. Cultures
221 of PC103 grown on standard ABA-SB were subcultured onto cadmium-enriched ABA-SB and
222 incubated for 7 days in an anaerobic environment at 37°C. Growth assessment was performed
223 by examining the colonies' presence. These tests provide valuable insights into the robustness
224 of the PC103 strain under various stress conditions, which may contribute to understanding
225 its survivability and adaptability in the pulmonary niche.

226 *Antibiotic resistance.* The antibiotic resistance profiles of six *Porphyromonas* strains isolated
227 from pulmonary sources were assessed following the 2020 guidelines provided by the
228 European Committee on Antimicrobial Susceptibility Testing (EUCAST) and the Committee
229 for Antibiogram of the French Society of Microbiology (CA-SFM) (26). Freshly isolated cultures
230 from ABA-SB were suspended to a density equivalent to 1 McFarland standard in Anaerobe
231 Basal Broth (ABB) (ThermoFisher). These suspensions were then plated onto ABA-SB agar
232 to establish a uniform bacterial lawn. The antibiotics tested included amoxicillin (25 µg),
233 amoxicillin/clavulanic acid (20 µg/10 µg), clindamycin (2 µg), imipenem (10 µg), metronidazole
234 (5 µg), moxifloxacin (5 µg), piperacillin (30 µg), and piperacillin/tazobactam (30 µg/6 µg)
235 (ThermoFisher). Additionally, a vancomycin disk (30 µg) was included due to the noted
236 susceptibility of other *Porphyromonas* species, although *P. catoniae* typically shows
237 resistance (27). The inoculated agar plates were incubated in an anaerobic chamber
238 maintained at 37°C for at least 48 hours to allow adequate growth and interaction between the
239 bacterial isolates and the antibiotic agents.

240

241 7. Results

242 Whole genome sequencing (WGS) of POTG (*P. catoniae*, *P. pasteri*, and *P. uenonis*)

243 The genomic characteristics of pulmonary strains of the *Porphyromonas* genus, specifically
244 *P. catoniae*, *P. pasteri*, and *P. uenonis*, were investigated through WGS. The GC% content of
245 these strains fell within the range of 51% to 56%, consistent with prior findings for this genus.
246 Utilizing *de novo* hybrid assembly methods, the genomes of the pulmonary POTG were
247 effectively reconstructed, resulting in contig numbers ranging from 1 to 66, with N50 values
248 spanning from 216,372 bp to 2,401,504 bp. The largest contig length observed for
249 *Porphyromonas* ranged from 320,622 bp to 2,401,504 bp. Notably, the *de novo* assembly
250 lengths agreed with previously reported median total lengths for each species. Assessment of
251 genome completeness revealed values ranging from 95.05% to 99.92%. Detailed genomic
252 statistics are provided in Table 1.

253

254 **Table 1. Sequencing and assembly statistics of pulmonary strains of *Porphyromonas* other than**
255 ***P. gingivalis* (POTG)**

Strain	POTG species	Unicycler assembly	N50	Largest contig	GC%	Contigs number	Completeness	Contamination	Predicted CDS
PC103	<i>P. catoniae</i>	2,032,211	2,032,211	2,032,211	51.00	1	95.05	0.157	1619
PC123	<i>P. catoniae</i>	2,032,214	1,601,551	1,601,551	51.00	3	95.05	0.157	1622
PC135	<i>P. catoniae</i>	2,032,218	2,032,218	2,032,218	51.09	1	95.05	0.157	1620
PP88	<i>P. pasteri</i>	2,225,553	2,205,487	2,205,487	55.75	2	98.87	0	1703
PU103	<i>P. uenonis</i>	2,401,504	2,401,504	2,401,504	52.31	1	99.92	0	1882
PU61	<i>P. uenonis</i>	2,325,647	216,372	320,622	52.49	66	99.92	0	1876

256

257 Analysis of Average Nucleotide Identity (ANI) in Pulmonary *Porphyromonas* Strains

258 ANI analysis of pulmonary strains revealed two distinct genetic clusters among the
259 *Porphyromonas* species (Table 2). Non-pigmented strains of *P. catoniae* and *P. pasteri* formed
260 one group, while all *P. uenonis* strains comprised the second group. Specifically, strains
261 PC103, PC123, and PC135 were identified as belonging to the same species, adopting a 95%
262 similarity threshold for species delineation, as established by Konstantinidis and Tiedje (28)
263 and Konstantinidis *et al.* (29). These strains exhibited ANI values of 96.6% and 96.7% with
264 the reference strain of *P. catoniae* (F0037).



265 Conversely, strain PP88 showed only 78.7% to 79% similarity with strains PC103, PC123, and
266 PC135, indicating significant genomic divergence. However, PP88 displayed a higher
267 similarity of 96.4% with the reference strain of *P. pasteri* (JCM 30531), suggesting a closer
268 genetic relationship with this species. Furthermore, strains PU61 and PU103 showed a high
269 similarity of 97.6%. These strains and the reference strain of *P. uenonis* (DSM 23387), which
270 was initially isolated from a human sacral decubitus ulcer, shared a similarity of 97.4% ANI.
271 Notably, all analyzed POTG lung strains demonstrated less than 80% similarity with the
272 reference strain of *P. gingivalis* (ATCC 33277), further highlighting the genetic distinctiveness
273 of the pulmonary strains within this group.

274

275 **Table 2. Average nucleotide identity between *Porphyromonas* genomes**

Strain*	PC103	PC123	PC	PC135	PP88	PP	PU103	PU61	PU	PA	PB	PE	PS	PG
PC103	100	99.9	96.6	99.9	79	78.5	<80	<80	<80	<80	<80	<80	<80	<80
PC123	99.9	100	96.7	99.9	78.8	78.4	<80	<80	<80	<80	<80	<80	<80	<80
PC135	99.9	99.9	96.6	100	79	78.6	<80	<80	<80	<80	<80	<80	<80	<80
PP88	79	78.8	78.7	79	100	96.4	<80	<80	<80	<80	<80	<80	<80	<80
PU103	<80	<80	<80	<80	<80	<80	100	97.6	97.4	88.8	<80	<80	<80	<80
PU61	<80	<80	<80	<80	<80	<80	97.6	100	97.4	88.9	<80	<80	<80	<80

276 *PC, *Porphyromonas catoniae* F0037; PP, *Porphyromonas pasteri* JCM 30531; PU, *Porphyromonas uenonis* DSM
277 23387; PA, *Porphyromonas asaccharolytica* DSM 20707; PB, *Porphyromonas bennonis* DSM 23058; PE,
278 *Porphyromonas endodontalis* 52278C01; PS, *Porphyromonas somerae* DSM 23386; PG, *Porphyromonas*
279 *gingivalis* ATCC 33277

280

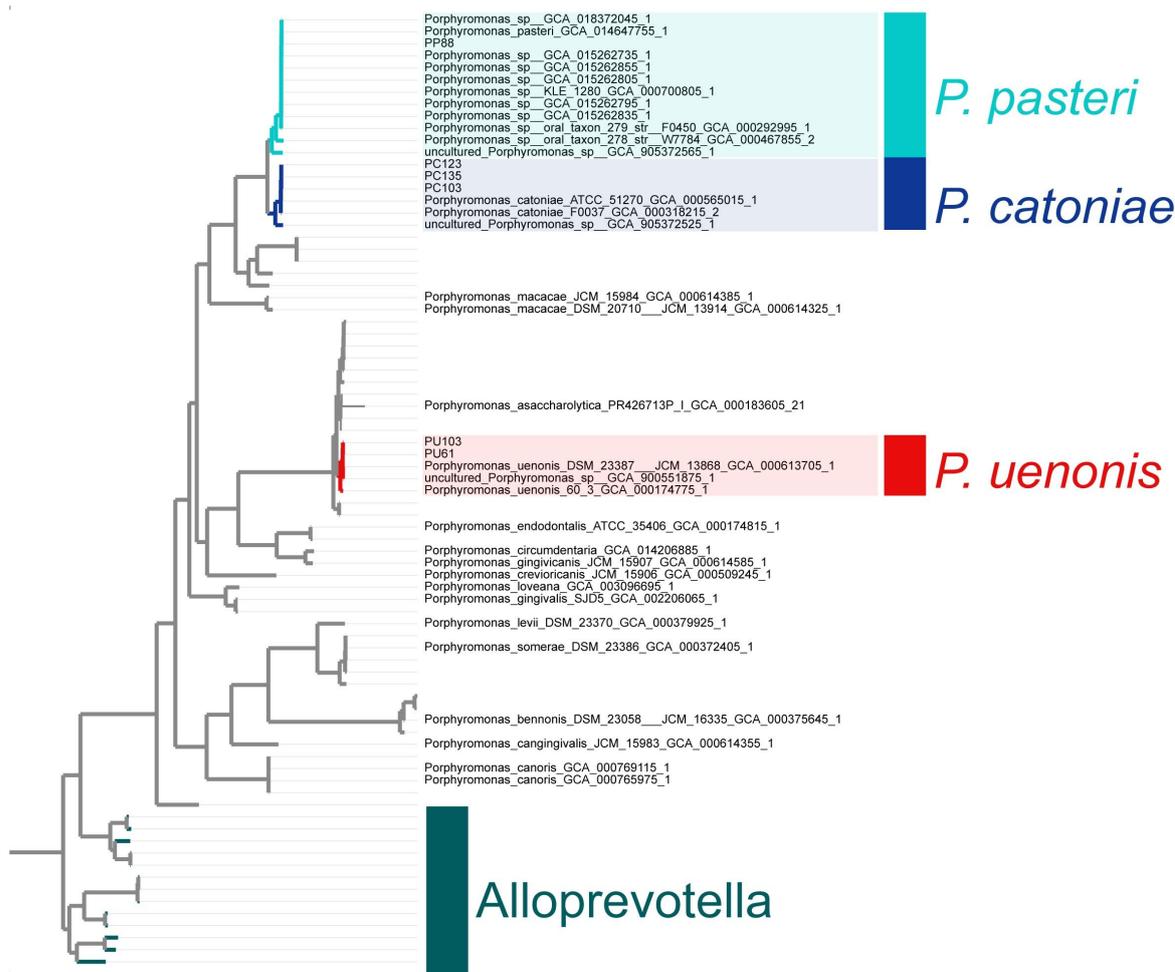
281 **Phylogenomic Analysis**

282 In the phylogenomic analysis, ribosomal sequences from all publicly available *Porphyromonas*
283 strains were concatenated to construct a comprehensive phylogenomic tree (Figure 1). This
284 analysis enabled the clear differentiation of species within the *Porphyromonas* genus. The
285 non-pigmented strains of the POTG, including *P. pasteri* and *P. catoniae*, formed a distinct
286 clade, demonstrating their phylogenetic separation from other pigmented *Porphyromonas*
287 species.

288 In detail, lung strains PU61 and PU103 closely aligned with all reference genomes of
289 *P. uenonis*, indicating strong phylogenetic consistency within this species. Similarly, PC103,
290 PC123, and PC135 clustered closely with all reference genomes of *P. catoniae*, confirming
291 their classification within this species. Additionally, PP88 was phylogenetically aligned with all
292 reference genomes of *P. pasteri*, underscoring its species-specific grouping.

293





294

295 **Fig. 1. Phylogenomic tree of *Porphyromonas* pulmonary species.** Tree based on
296 concatenating ribosomal sequences using IQ-TREE natively with ModelFinder Plus (30).
297

298 **Pan-Genomic Analysis of *Porphyromonas* Strains**

299 A comprehensive pan-genomic analysis was conducted on lung strains of *Porphyromonas*
300 from our collection alongside reference genomes from the POTG group. This analysis
301 revealed a core genome consistent across all *Porphyromonas* species, comprising 357 gene
302 clusters encompassing 8,587 open reading frames (ORFs). A species-specific accessory
303 genome was identified exclusively in *P. catoniae*, containing 255 gene clusters (419 ORFs).
304 The study aimed to elucidate potential virulence factors among the POTG, mainly focusing on
305 non-pigmented *Porphyromonas* strains. The virulence genes examined were primarily those
306 previously identified in *P. gingivalis* ATCC 33277 by Mendez *et al.* (31), who extensively
307 searched for such genes in that strain. However, our current analysis did not detect genes
308 associated with gingipains or hemagglutinin in the non-pigmented *Porphyromonas* genomes.
309 Only one putative hemolysin gene was found, and an additional gene was identified as
310 hyaluronidase/collagenase (K01197; EC: 3.2.1.35), indicating a reduced virulence gene profile

311 in these strains. Regarding secretion systems, a type IX secretion system was detected in
312 both PC103 and *P. catoniae* ATCC 51270 genomes, albeit incomplete, and a type IV secretion
313 system was noted in *P. catoniae* ATCC 51270, also incomplete. Interestingly, specific
314 metabolic pathways and transport mechanisms, such as the pentose phosphate pathway and
315 iron transport genes, were explicitly identified in the lung strains of *P. catoniae*, suggesting
316 adaptations to their niche. These screened genes and their potential implications are
317 summarized in Table in Supplementary Information.

318

319 **Genome comparison of *Porphyromonas* strains**

320 The comprehensive genomic analysis of the lung-derived *P. catoniae* strain PC103 reveals a
321 single circular chromosome measuring 2,032,211 base pairs with a genomic DNA G+C
322 content of 51.09%. Genome annotation, performed using the MaGe platform, predicted 1,669
323 genes, including 1609 protein-coding sequences (CDSs), as illustrated in Figure 3. The
324 analysis also identified one CRISPR locus between 805,723 and 806,843 bp downstream of
325 a CAS-type II cluster spanning from 800,799 to 805,615 bp. Additionally, two incomplete
326 prophage regions were detected, each approximately 27 kb and 26 kb in size, showing 40%
327 completeness.

328 The metabolic profile of *P. catoniae* PC103 was compared against *P. catoniae* ATCC 51270
329 and several *P. gingivalis* strains (ATCC 33277, TDC60, and W83) using the MicroScope
330 platform. This comparison highlighted significant metabolic distinctions between the
331 *P. catoniae* and *P. gingivalis* strains, irrespective of their source (oral or pulmonary). Key
332 pathway differences are summarized in Table 3, including a cadmium transport system unique
333 to *P. catoniae* strains and absent in *P. gingivalis* strains. Conversely, genes involved in protein
334 citrullination were exclusive to *P. gingivalis* strains.

335 *P. catoniae* strains exhibited unique sulfur metabolism pathways (sulfide oxidation and sulfur
336 reduction) and an arginine-dependent acid resistance gene. The genomes also showed a
337 partial presence of pathways for detoxifying oxidized GTP and dGTP and the degradation of
338 superoxide radicals. Furthermore, the fatty acid biosynthesis pathways showed 60%
339 completion for initiation and 80% completion for elongation processes. Additionally, various
340 vitamin biosynthesis pathways were identified in *P. catoniae*, including those for folate, lipoic
341 acid, pantothenate, riboflavin, and pyridoxal, contrasting with the degradation pathways for
342 amino acids and sugars found in *P. gingivalis* genomes.

343

344

345

346 **Table 3. Metabolic comparison of *P. catoniae* PC103, *P. catoniae* ATCC 51270, *P. gingivalis* ATCC**
 347 **33277, *P. gingivalis* TDC60, and *P. gingivalis* W83 with the MicroScope platform using KEGG and**
 348 **MicroCyc databases.**
 349

Reaction	Database	Number of reactions	Completeness of pathway				
			<i>P. catoniae</i> PC103	<i>P. catoniae</i> ATCC51270	<i>P. gingivalis</i> ATCC 33277	<i>P. gingivalis</i> TDC60	<i>P. gingivalis</i> W83
Cadmium transport	MicroCyc	1	1	1	0	0	0
Protein citrullination	MicroCyc	1	0	0	1	1	1
Sulfide oxidation II	MicroCyc	1	1	1	0	0	0
Sulfure reduction II	MicroCyc	1	1	1	0	0	0
Oxidized GTP and dGTP detoxification	MicroCyc	2	0,5	0,5	0	0	0
Superoxide radical degradation	MicroCyc	2	0,5	0,5	0	0	0
Arginine degradation IV	MicroCyc	3	1	1	0,67	0,67	0,67
Arginine dependent acid resistance	MicroCyc	1	1	1	0	0	0
Glutamate degradation	MicroCyc	1	0	0	1	1	1
Glutamine degradation I	MicroCyc	1	0	0	1	1	1
Glutamine degradation II	MicroCyc	1	0	0	1	1	1
Histidine degradation I	MicroCyc	4	0	0	0,75	0,75	0,75
L-Serine degradation	MicroCyc	1	0	0	1	1	1
Methionine degradation II	MicroCyc	1	0	0	1	1	1
Tryptophane degradation II	MicroCyc	1	0	0	1	1	1
Melibiose degradation	MicroCyc	1	0	0	1	1	1
Fatty acid biosynthesis, elongation	KEGG	35	1	1	0,8	0,8	0,8
Fatty acid biosynthesis, initiation	KEGG	5	0,6	0,6	0,6	0,6	0,6
L-threo-Tetrahydrobiopterin biosynthesis, GTP to L-threo-BH4 (Folate biosynthesis)	KEGG	4	0,75	0,75	0,75	0,75	0,75
Tetrahydrobiopterine biosynthesis, GTP to BH4 (Folate biosynthesis)	KEGG	12	0,75	0,75	0,83	0,83	0,83
Tetrahydrofolate biosynthesis, GTP to THF (Folate Biosynthesis)	KEGG	9	0,89	0,89	0,89	0,89	0,89
Tetrahydrofolate biosynthesis, mediated by PTPS, GTP to THF (Folate biosynthesis)	KEGG	10	0,8	0,8	0,8	0,8	0,8
Tetrahydrofolate biosynthesis, mediated by ribA and trpF, GTP to THF (Folate biosynthesis)	KEGG	3	0,67	0,67	0,67	0,67	0,67
Lipoic acid biosynthesis, animals and bacteria, octanoyl-ACP to dihydrolipoyl-H to dihydrolipoyl-E2 (Lipoic acid biosynthesis)	KEGG	2	1	1	1	1	1
Lipoic acid biosynthesis, eukaryotes, octanoyl-ACP to dihydrolipoyl-H (Lipoic acid biosynthesis)	KEGG	2	0,5	0,5	0,5	0,5	0,5
Lipoic acid biosynthesis, octanoyl-CoA to dihydrolipoyl-E2 (Lipoic acid biosynthesis)	KEGG	4	1	1	0	0	0
Lipoic acid biosynthesis, plants and bacteria, octanoyl-ACP to dihydrolipoyl-E2/H (Lipoic acid biosynthesis)	KEGG	5	1	1	1	1	1
Pantothenate biosynthesis, valine/L-aspartate to pantothenate	KEGG	9	1	1	1	1	1
Riboflavin biosynthesis, plants and bacteria, GTP to riboflavin/FMN/FAD	KEGG	6	0,83	0,83	0,83	0,83	0,83
Pyridoxal-P biosynthesis, erythrose-4P to pyridoxal-P (Vit B6 metabolism)	KEGG						

350 *The pathway's completeness is indicated as 1 = 100% and 0 = 0%. Reactions are named as they are in the KEGG*
 351 *or MicroCyc databases.*

352 ***In Vitro* Confirmation of Genotypic Characteristics of *Porphyromonas* Strains**

353 *Oxygen Tolerance.* The presence of genes related to the detoxification of oxidized GTP and
354 dGTP and the degradation of superoxide radicals in *P. catoniae* PC103 suggested potential
355 oxygen tolerance. To verify this, we exposed a suspension of PC103 to atmospheric oxygen
356 concentrations, monitoring the bacterial survival over time. The results, depicted in Figures 3A
357 and B, showed only a 1.86 log decrease in CFU per milliliter after 8 hours at an initial
358 concentration of 10^7 CFU.mL⁻¹ and a 1.25 log decrease after 6 hours at 10^3 CFU.mL⁻¹.

359 *pH Tolerance.* The *in silico-identified* arginine-dependent acid resistance gene prompted an
360 *in vitro* assessment of PC103's acid tolerance. Cultures were grown on ABA-SB medium at
361 various pH levels (6.86, 6.38, 6.0, 5.5, and 5.0) and incubated for 7 days. PC103 was able to
362 grow at pH levels down to 5.5, but no growth was observed at pH 5.0, even with extended
363 incubation.

364 *Cadmium Resistance.* The WGS of PC103 had revealed a cadmium transport gene,
365 suggesting possible resistance to cadmium. This was tested by culturing PC103 on an ABA-
366 SB medium supplemented with increasing cadmium concentrations (0 to 80 mg.L⁻¹). The
367 results demonstrated that PC103 could tolerate cadmium concentrations up to 20 mg.L⁻¹, but
368 growth was inhibited at 40 mg.L⁻¹ and above.

369 *Antibiotic Resistance.* Investigations into the antibiotic susceptibility of pulmonary
370 *Porphyromonas* species revealed sensitivity to a broad spectrum of antibiotics, with consistent
371 antibiogram profiles across the species, except for specific resistances. Strains of *P. uenonis*
372 (n=2) displayed resistance to clindamycin. Strains of *P. catoniae* (n=2), including PC103, were
373 resistant to vancomycin. Intriguingly, no vancomycin resistance genes were identified in the
374 reconstructed genomes of these strains using the CARD database, indicating a potential non-
375 genomic resistance mechanism or an unidentified resistance gene.

376

377 **8. Discussion**

378 In this study, we conducted WGS of pulmonary strains of the POTG group, specifically
379 *P. catoniae*, *P. pasteri*, and *P. uenonis*, isolated from the sputum of people with CF. Our
380 genomic analysis revealed high nucleotide similarity among these strains and identified a
381 distinct clade of non-pigmented *Porphyromonas* in the phylogenomic tree. Notably,
382 *P. catoniae* was found to lack the virulence genes typical of *P. gingivalis*, except for a putative
383 hemolysin gene.

384

385 As reported in several studies, *P. gingivalis* is a well-documented pathogen extensively
386 associated with periodontitis (18, 32-34). Recent research has also implicated this bacterium
387 in systemic diseases, including lung cancer (35), rheumatoid arthritis (RA) (36), and
388 Alzheimer's disease (37), highlighting its potential role beyond oral health. In contrast, the
389 genus *Porphyromonas* has been detected in both diseased and healthy pulmonary niches
390 using molecular and cultural methods (2, 5, 13-15, 17, 38). Specifically, targeted metagenomic
391 analyses have frequently identified POTG strains, predominantly non-pigmented
392 *Porphyromonas*, in pulmonary samples (2, 16, 17). Unlike *P. gingivalis*, species such as
393 *P. catoniae* and *P. pasteri* are primarily associated with healthy individuals, suggesting
394 potentially different roles in the microbial ecology of the lung (18, 39-41).

395

396 Historically, genomic data on non-pigmented *Porphyromonas* species have been sparse. A
397 notable study by O'Flynn *et al.* (42) included a genome of *P. catoniae* isolated from a human
398 gingival crevice, revealing significant genomic distinctions between *P. catoniae* and
399 *P. gingivalis*. Key differences highlighted were the absence of hemagglutinin and gingipains,
400 virulence factors in *P. gingivalis*, whereas genes for hemolysins and iron transport were
401 detected in *P. catoniae*. Consistent with these findings, our current analysis of pulmonary non-
402 pigmented *Porphyromonas* strains identified similar absences and noted the presence of a
403 putative hemolysin (cds 1638831 1640078 - HANLCC_06880 hemolysin SO:0001217,
404 UniRef: UniRef50_A0A0A2G0A3, UniRef: UniRef90_L1N9Y2).

405 Moreover, O'Flynn *et al.* reported the absence of glutamate enzymes in *P. catoniae* (42),
406 except for a NADP-specific glutamate dehydrogenase (EC 1.4.1.4), which we also identified
407 in our study, indicating its conservation across different isolates of this species. Interestingly,
408 our research further established the presence of genes encoding pentose phosphate pathway
409 enzymes in *P. catoniae*, contrasting with their absence in *P. gingivalis* genomes. This study
410 also uncovered genes involved in synthesizing glutamate from ammonia and maintaining the
411 glutamate pool in *P. catoniae*, elements missing in *P. gingivalis*. Conversely, genes related to
412 the synthesis of ammonia from nitrate, the synthesis of vitamin B12 from uroporphyrinogen,
413 and the release of iron from blood cells, typically found in *P. gingivalis*, were absent in
414 *P. catoniae* (42).

415 Significantly, none of the *P. catoniae* genomes, including lung clinical strains and *P. catoniae*
416 ATCC 51270, exhibited genes involved in protein citrullination, a process linked to the
417 production of citrullinated protein antigens, which are specific markers for RA (43). This
418 absence is crucial because *P. gingivalis* has been implicated in RA due to its ability to produce
419 such antigens (44). The lack of citrullination genes in non-pigmented *Porphyromonas*
420 suggests that these species play no role in RA and possibly other autoimmune diseases. This

421 genetic profile supports the hypothesis that non-pigmented *Porphyromonas*, particularly
422 *P. catoniae*, are less likely to contribute to autoimmune pathogenesis than *P. gingivalis*.

423

424 In addition to the absence of specific *P. gingivalis* virulence factors such as hemagglutinin and
425 gingipains, our genomic analysis of *P. catoniae* strains revealed a notable lack of amino acid
426 degradation pathways for glutamate, glutamine, histidine, serine, methionine, tryptophan, and
427 melibiose. Unlike *P. gingivalis*, which utilizes these pathways as carbon sources to support its
428 pathogenic lifestyle, the absence of such metabolic capabilities in *P. catoniae* may suggest a
429 non-pathogenic nature. This potential for commensalism is further supported by the lack of
430 other virulence genes commonly associated with *P. gingivalis*, aligning with findings that
431 characterize non-pigmented *Porphyromonas* species, notably *P. catoniae*, as less likely to
432 contribute to disease states (16, 45). Future studies should incorporate *in vivo* and *in vitro*
433 testing to comprehensively assess these bacteria's biological roles and impacts.

434 Moreover, our study identified genes associated with acid resistance in *P. catoniae* strains.
435 Experimentally, these strains demonstrated the ability to grow at a low pH of 5.5 on ABA-SB
436 agar, indicating a robust capacity for acid tolerance. This finding aligns with previous
437 descriptions of *P. catoniae* as a lactic acid producer. This attribute contributes to its survival
438 in acidic environments, much like other lactic acid bacteria, including species of *Lactobacillus*
439 (27). This acid resistance capability not only underscores the adaptive strategies of *P. catoniae*
440 to various environmental stresses but also suggests its potential role in maintaining microbial
441 balance within its niche. The ability to produce and withstand acidic conditions may facilitate
442 the colonization and persistence of *P. catoniae* in the human microbiome, particularly in acid-
443 prone sites, which is a characteristic typical of commensal bacteria.

444

445 Our investigation into the antibiotic resistance profiles of pulmonary *Porphyromonas* strains
446 highlights key findings. Consistent with prior studies (27, 46), *Porphyromonas* generally
447 exhibits high susceptibility to antibiotics targeting anaerobes, including β -lactams. However,
448 notable exceptions were observed: *P. uenonis* demonstrated resistance to clindamycin, a
449 relatively new observation given that this species was only recently described (47), and
450 *P. catoniae* displayed resistance to vancomycin, as previously reported (27). In the 2 strains
451 of *P. catoniae*, the search for glycopeptide resistance genes in their genome was negative.
452 Gram-negative bacilli are therefore intrinsically resistant to vancomycin, due to the inability of
453 this glycopeptide to cross the outer membrane. While the present dataset includes only six
454 pulmonary strains, expanding this dataset could provide deeper insights into glycopeptide
455 resistance mechanisms within the POTG group. The isolation of these strains from lung

456 samples remains challenging due to their sensitivity to most antibiotics, oxygen and their slow
457 growth rate.

458

459 In conclusion, our study did not detect virulence genes typically associated with *P. gingivalis*
460 in non-pigmented *Porphyromonas* strains, except for a putative hemolysin. These findings
461 suggest that non-pigmented POTG, particularly *P. catoniae*, and *P. pasteri*, are endogenous
462 inhabitants of the lung, contributing to homeostasis rather than disease. This distinction
463 between POTG and the pathogenic *P. gingivalis* is crucial for understanding the lung niche's
464 microbial dynamics and guiding clinical approaches to lung health and disease.

465

466 **9. Author statements**

467 **9.1 Conflicts of interest**

468 The authors declare no conflicts of interest.

469 **9.2 Funding information**

470 This study was supported by a grant from the French Cystic Fibrosis Association 'Vaincre la
471 Mucoviscidose' (Contract No. RC20180502218).

472 **9.3 Ethical approval**

473 The study was approved by the French Ethical Research Committee in March 2018 (2018-
474 A00624- 51). All experiments were performed following relevant guidelines and regulations.
475 Informed consent was obtained from all participants and/or their legal guardians.

476 **9.4 Acknowledgements**

477 We gratefully acknowledge the support from LABGeM (CEA/Genoscope & CNRS UMR8030),
478 France Génomique, and the French Bioinformatics Institute. These institutions have provided
479 significant assistance as part of the Investissement d'Avenir program, managed by the Agence
480 Nationale pour la Recherche under contracts ANR-10-INBS-09 and ANR-11-INBS-0013. Their
481 contributions were instrumental in facilitating our use of the MicroScope annotation platform,
482 which was essential for the genomic analyses conducted in this study.

483

484 10. References

- 485 1. Carmody LA, Caverly LJ, Foster BK, Rogers MAM, Kalikin LM, Simon RH, et al.
486 Fluctuations in airway bacterial communities associated with clinical states and disease
487 stages in cystic fibrosis. *PLoS One*. 2018;13(3):e0194060.
- 488 2. Muhlebach MS, Hatch JE, Einarsson GG, McGrath SJ, Gilipin DF, Lavelle G, et al.
489 Anaerobic bacteria cultured from cystic fibrosis airways correlate to milder disease: a
490 multisite study. *Eur Respir J*. 2018;52(1).
- 491 3. Zemanick ET, Harris JK, Wagner BD, Robertson CE, Sagel SD, Stevens MJ, et al.
492 Inflammation and airway microbiota during cystic fibrosis pulmonary exacerbations. *PLoS*
493 *One*. 2013;8(4):e62917.
- 494 4. Charlson ES, Bittinger K, Haas AR, Fitzgerald AS, Frank I, Yadav A, et al. Topographical
495 continuity of bacterial populations in the healthy human respiratory tract. *Am J Respir Crit*
496 *Care Med*. 2011;184(8):957-63.
- 497 5. Erb-Downward JR, Thompson DL, Han MK, Freeman CM, McCloskey L, Schmidt LA, et
498 al. Analysis of the lung microbiome in the "healthy" smoker and in COPD. *PLoS One*.
499 2011;6(2):e16384.
- 500 6. Whelan FJ, Waddell B, Syed SA, Shekarriz S, Rabin HR, Parkins MD, et al. Culture-
501 enriched metagenomic sequencing enables in-depth profiling of the cystic fibrosis lung
502 microbiota. *Nat Microbiol*. 2020;5(2):379-90.
- 503 7. Dickson RP, Erb-Downward JR, Freeman CM, McCloskey L, Falkowski NR, Huffnagle
504 GB, et al. Bacterial Topography of the Healthy Human Lower Respiratory Tract. *mBio*.
505 2017;8(1).
- 506 8. Lee SY, Mac Aogain M, Fam KD, Chia KL, Binte Mohamed Ali NA, Yap MMC, et al. Airway
507 microbiome composition correlates with lung function and arterial stiffness in an age-
508 dependent manner. *PLoS One*. 2019;14(11):e0225636.
- 509 9. Weinberg F, Dickson RP, Nagrath D, Ramnath N. The Lung Microbiome: A Central
510 Mediator of Host Inflammation and Metabolism in Lung Cancer Patients? *Cancers*
511 (Basel). 2020;13(1).
- 512 10. Chen J, Li T, Ye C, Zhong J, Huang JD, Ke Y, et al. The Lung Microbiome: A New Frontier
513 for Lung and Brain Disease. *Int J Mol Sci*. 2023;24(3).
- 514 11. Beck JM, Young VB, Huffnagle GB. The microbiome of the lung. *Transl Res*.
515 2012;160(4):258-66.
- 516 12. Morris A, Beck JM, Schloss PD, Campbell TB, Crothers K, Curtis JL, et al. Comparison
517 of the respiratory microbiome in healthy nonsmokers and smokers. *Am J Respir Crit Care*
518 *Med*. 2013;187(10):1067-75.

- 519 13. Park H, Shin JW, Park SG, Kim W. Microbial communities in the upper respiratory tract
520 of patients with asthma and chronic obstructive pulmonary disease. *PLoS One*.
521 2014;9(10):e109710.
- 522 14. Lim MY, Yoon HS, Rho M, Sung J, Song YM, Lee K, et al. Analysis of the association
523 between host genetics, smoking, and sputum microbiota in healthy humans. *Sci Rep*.
524 2016;6:23745.
- 525 15. Bassis CM, Erb-Downward JR, Dickson RP, Freeman CM, Schmidt TM, Young VB, et al.
526 Analysis of the upper respiratory tract microbiotas as the source of the lung and gastric
527 microbiotas in healthy individuals. *mBio*. 2015;6(2):e00037.
- 528 16. Guilloux CA, Lamoureux C, Beauruelle C, Hery-Arnaud G. *Porphyromonas*: A neglected
529 potential key genus in human microbiomes. *Anaerobe*. 2021;68:102230.
- 530 17. Lamoureux C, Guilloux CA, Beauruelle C, Gouriou S, Ramel S, Dirou A, et al. An
531 observational study of anaerobic bacteria in cystic fibrosis lung using culture dependant
532 and independent approaches. *Sci Rep*. 2021;11(1):6845.
- 533 18. Abusleme L, Dupuy AK, Dutzan N, Silva N, Burleson JA, Strausbaugh LD, et al. The
534 subgingival microbiome in health and periodontitis and its relationship with community
535 biomass and inflammation. *ISME J*. 2013;7(5):1016-25.
- 536 19. Camelo-Castillo AJ, Mira A, Pico A, Nibali L, Henderson B, Donos N, et al. Subgingival
537 microbiota in health compared to periodontitis and the influence of smoking. *Front*
538 *Microbiol*. 2015;6:119.
- 539 20. Murray PR, Washington JA. Microscopic and bacteriologic analysis of expectorated
540 sputum. *Mayo Clin Proc*. 1975;50(6):339-44.
- 541 21. Wick RR, Judd LM, Gorrie CL, Holt KE. Completing bacterial genome assemblies with
542 multiplex MinION sequencing. *Microb Genom*. 2017;3(10):e000132.
- 543 22. Seemann T. Prokka: rapid prokaryotic genome annotation. *Bioinformatics*.
544 2014;30(14):2068-9.
- 545 23. Vallenet D, Calteau A, Dubois M, Amours P, Bazin A, Beuvin M, et al. MicroScope: an
546 integrated platform for the annotation and exploration of microbial gene functions through
547 genomic, pangenomic and metabolic comparative analysis. *Nucleic Acids Res*.
548 2020;48(D1):D579-D89.
- 549 24. Antipov D, Raiko M, Lapidus A, Pevzner PA. Metaviral SPAdes: assembly of viruses from
550 metagenomic data. *Bioinformatics*. 2020;36(14):4126-9.
- 551 25. Guo J, Bolduc B, Zayed AA, Varsani A, Dominguez-Huerta G, Delmont TO, et al.
552 VirSorter2: a multi-classifier, expert-guided approach to detect diverse DNA and RNA
553 viruses. *Microbiome*. 2021;9(1):37.

- 554 26. Microbiologie SFd. Anaérobies. CA-SFM / EUCAST. Société Française de Microbiologie
555 ed2022. p. 128-31.
- 556 27. Summanen P, Finegold S. *Porphyromonas*. Bergey's Manual of Systematics of Archaea
557 and Bacteria. Kindle ed2015. p. 1-14.
- 558 28. Konstantinidis KT, Tiedje JM. Towards a genome-based taxonomy for prokaryotes. J
559 Bacteriol. 2005;187(18):6258-64.
- 560 29. Konstantinidis KT, Ramette A, Tiedje JM. The bacterial species definition in the genomic
561 era. Philos Trans R Soc Lond B Biol Sci. 2006;361(1475):1929-40.
- 562 30. Kalyaanamoorthy S, Minh BQ, Wong TKF, von Haeseler A, Jermiin LS. ModelFinder: fast
563 model selection for accurate phylogenetic estimates. Nat Methods. 2017;14(6):587-9.
- 564 31. Mendez KN, Hoare A, Soto C, Bugueno I, Olivera M, Meneses C, et al. Variability in
565 Genomic and Virulent Properties of *Porphyromonas gingivalis* Strains Isolated From
566 Healthy and Severe Chronic Periodontitis Individuals. Front Cell Infect Microbiol.
567 2019;9:246.
- 568 32. Hajishengallis G. Periodontitis: from microbial immune subversion to systemic
569 inflammation. Nat Rev Immunol. 2015;15(1):30-44.
- 570 33. Payungporn S, Arirachakaran P, Poomipak W, Praianantathavorn K, Charalampakis G,
571 Poovorawan Y. Identification of Bacteria Associated with a Periodontal Disease in Thai
572 Patients Based on Next-Generation Sequencing. Jundishapur Journal of Microbiology.
573 2017;10(6).
- 574 34. Yang HW, Huang YF, Chou MY. Occurrence of *Porphyromonas gingivalis* and *Tannerella*
575 *forsythensis* in periodontally diseased and healthy subjects. J Periodontol.
576 2004;75(8):1077-83.
- 577 35. Liu Y, Yuan X, Chen K, Zhou F, Yang H, Yang H, et al. Clinical significance and prognostic
578 value of *Porphyromonas gingivalis* infection in lung cancer. Transl Oncol.
579 2021;14(1):100972.
- 580 36. Scher JU, Abramson SB. Periodontal disease, *Porphyromonas gingivalis*, and rheumatoid
581 arthritis: what triggers autoimmunity and clinical disease? Arthritis Res Ther.
582 2013;15(5):122.
- 583 37. Dominy SS, Lynch C, Ermini F, Benedyk M, Marczyk A, Konradi A, et al. *Porphyromonas*
584 *gingivalis* in Alzheimer's disease brains: Evidence for disease causation and treatment
585 with small-molecule inhibitors. Sci Adv. 2019;5(1):eaau3333.
- 586 38. Pattaroni C, Watzenboeck ML, Schneidegger S, Kieser S, Wong NC, Bernasconi E, et al.
587 Early-Life Formation of the Microbial and Immunological Environment of the Human
588 Airways. Cell Host Microbe. 2018;24(6):857-65 e4.

- 589 39. Crielaard W, Zaura E, Schuller AA, Huse SM, Montijn RC, Keijser BJ. Exploring the oral
590 microbiota of children at various developmental stages of their dentition in the relation to
591 their oral health. *BMC Med Genomics*. 2011;4:22.
- 592 40. Könönen E, Väisänen M, Finegold S, Heine R, Jousimies-Somer H. Cellular fatty acid
593 analysis and enzyme profiles of *Porphyromonas catoniae*— a frequent colonizer of the
594 oral cavity in children. *Anaerobe*. 1996(2):329-35.
- 595 41. Yasunaga H, Takeshita T, Shibata Y, Furuta M, Shimazaki Y, Akifusa S, et al. Exploration
596 of bacterial species associated with the salivary microbiome of individuals with a low
597 susceptibility to dental caries. *Clin Oral Investig*. 2017;21(8):2399-406.
- 598 42. O'Flynn C, Deusch O, Darling AE, Eisen JA, Wallis C, Davis IJ, et al. Comparative
599 Genomics of the Genus *Porphyromonas* Identifies Adaptations for Heme Synthesis within
600 the Prevalent Canine Oral Species *Porphyromonas cangingivalis*. *Genome Biol Evol*.
601 2015;7(12):3397-413.
- 602 43. Romero V, Fert-Bober J, Nigrovic PA, Darrah E, Haque UJ, Lee DM, et al. Immune-
603 mediated pore-forming pathways induce cellular hypercitrullination and generate
604 citrullinated autoantigens in rheumatoid arthritis. *Sci Transl Med*. 2013;5(209):209ra150.
- 605 44. Gomez-Banuelos E, Mukherjee A, Darrah E, Andrade F. Rheumatoid Arthritis-Associated
606 Mechanisms of *Porphyromonas gingivalis* and *Aggregatibacter actinomycetemcomitans*.
607 *J Clin Med*. 2019;8(9).
- 608 45. Nemoto TK, Ohara-Nemoto Y. Exopeptidases and gingipains in *Porphyromonas*
609 *gingivalis* as prerequisites for its amino acid metabolism. *Jpn Dent Sci Rev*.
610 2016;52(1):22-9.
- 611 46. Aldridge KE, Ashcraft D, Cambre K, Pierson CL, Jenkins SG, Rosenblatt JE. Multicenter
612 survey of the changing in vitro antimicrobial susceptibilities of clinical isolates of
613 *Bacteroides fragilis* group, *Prevotella*, *Fusobacterium*, *Porphyromonas*, and
614 *Peptostreptococcus* species. *Antimicrob Agents Chemother*. 2001;45(4):1238-43.
- 615 47. Finegold SM, Vaisanen ML, Rautio M, Eerola E, Summanen P, Molitoris D, et al.
616 *Porphyromonas uenonis* sp. nov., a pathogen for humans distinct from *P. asaccharolytica*
617 and *P. endodontalis*. *J Clin Microbiol*. 2004;42(11):5298-301.

618 11. Supplementary information

619 **Supplementary Table 1. Screened genes of *Porphyromonas* previously identified by Mendez *et al.***
620 **and O'Flynn *et al.* (31, 42).**

621

622 **Supplementary Figure 1. Pan-genome analysis of pulmonary strains of *Porphyromonas*.**

623 PC103, PC123, PC135, PP88, PU61, and PU103 were isolated from the lungs of people with cystic
624 fibrosis. Other genomes were downloaded from NCBI.

625

