



HAL
open science

The diversity and abundance of soft rot Pectobacteriaceae along the Durance River stream in the south-east of France revealed by multiple seasonal surveys

Hajar Ben Moussa, Claire Bertrand, Emma Rochelle-Newall, Sarah Fiorini,
Jacques Pédrón, Marie-Anne Barny

► To cite this version:

Hajar Ben Moussa, Claire Bertrand, Emma Rochelle-Newall, Sarah Fiorini, Jacques Pédrón, et al.. The diversity and abundance of soft rot Pectobacteriaceae along the Durance River stream in the south-east of France revealed by multiple seasonal surveys. *Phytopathology*, In press, 112 (8), pp.1676-1685. 10.1094/phyto-12-21-0515-r . hal-03660971

HAL Id: hal-03660971

<https://hal.sorbonne-universite.fr/hal-03660971v1>

Submitted on 6 May 2022

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
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19

The diversity of soft rot *Pectobacteriaceae* along the Durance River stream in the south-east of France revealed by multiple seasonal surveys

Hajar Ben Moussa^{1*}, Claire Bertrand^{1*}, Emma Rochelle-Newall¹, Sarah Fiorini², Jacques Pédrón¹, Marie-Anne Barny^{1#}

¹Sorbonne Université, INRAE, IRD, CNRS, UPEC, Institute of Ecology and Environmental Sciences-Paris (iEES-Paris), 4 place Jussieu, F-75252 Paris, France.

²Ecole Normale Supérieure, PSL Research University, CNRS, Centre de recherche en écologie expérimentale et prédictive (CEREAP-Ecotron IleDeFrance), UMS 3194, 11 chemin de Busseau, 77140 Saint-Pierre-lès-Nemours, France

#Corresponding author: MA Barny; e-mail : marie-anne.barny@sorbonne-universite.fr

* equal contribution to this work

Key words: Ecology, *Pectobacterium*, *Dickeya*, water

Funding: Agence Nationale de la Recherche grant number ANR-17-CE32-0004

20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40

ABSTRACT

Although irrigation water is frequently assessed for the presence of plant pathogens, large spatial and temporal surveys that provide clues on the diversity and circulation of pathogens is missing. We evaluate the diversity of soft rot *Pectobacteriaceae* (SRP) of the genera *Dickeya* and *Pectobacterium* over two years in a temperate, mixed use watershed. The abundance of isolated strains correlates with the agricultural gradient along the watershed with a positive correlation found with temperature, nitrate and dissolved organic carbon water concentration. We characterized 582 strains by amplification and sequencing of the *gapA* gene. MLSA analysis performed with 3 housekeeping genes for 99 strains and core genome analysis of 38 sequenced strains confirmed for all the strains but one the taxonomic assignment obtained with the sole *gapA* sequence. *Pectobacterium* spp. (549 isolates) were far more abundant than *Dickeya* spp. (33 isolates). *Dickeya* spp. were only observed in the lower part of the river when water temperature was above 19°C and we experimentally confirmed a decreased fitness of several *Dickeya* spp. at 8°C in river water. *D. oryzae* dominates the *Dickeya* spp. *P. versatile* and *P. aquaticum* dominate the *Pectobacterium* spp. but their repartition along the watershed was different, *P. versatile* being the only species regularly recovered all along the watershed. Excepting *P. versatile*, *Dickeya* and *Pectobacterium* spp. responsible for disease outbreak on crops were less abundant or rarely detected. This work sheds light on the various ecological behaviours of different SRP in stream water and indicates that SRP occupation is geographically structured.

41 **INTRODUCTION**

42 Soft rot *Pectobacteriaceae* (SRP) belonging to the *Dickeya*, *Pectobacterium* and *Musicola* genera
43 are plant pathogenic bacteria that collectively infect a wide range of plant species, infecting at least
44 35 % of angiosperm plant orders all over the world (Ma et al. 2007; Charkowski 2018; Portier et al.
45 2020; Hugouvieux-Cotte-Pattat et al. 2021). The virulence of SRP relies mainly on the secretion of
46 cell wall degrading enzymes (PCWDE) provoking maceration symptoms (Charkowski 2018;
47 Hugouvieux-Cotte-Pattat et al. 2014). Latent infections, where the pathogen is present on the plant
48 in the absence of symptoms, are common, symptoms being expressed only when environmental
49 conditions are conducive (Toth et al. 2021a) and the main route of infection and dissemination
50 occurs via latently contaminated plant material. However, environmental sources of contamination
51 also play an important role and it has been demonstrated on the potato host that axenically grown
52 seed stocks, when planted on the field could rapidly become contaminated when the environmental
53 conditions are conducive (Van Gijsegem et al. 2021).

54 Plants can become contaminated with SRP from a variety of environmental sources including
55 insects, soil, aerosols, water or rainwater (Toth et al. 2021a). Identifying the major source(s) of these
56 primary infections is complex and has still not been fully achieved. Particular attention has been
57 paid to surface water that could serve for irrigation purposes. SRP are rare in surface water and
58 neither *Pectobacterium* genus or *Dickeya* genus are detected in 16S metabarcoding studies
59 performed on fresh water (Pédrón et al. 2020). However, the development of an efficient semi
60 selective medium (Burr and Schroth 1977; Hélias et al. 2012) means SRP can be isolated from fresh
61 water and early reports observed the frequent contamination of surface water by SRP and the
62 potential contamination of plants via water reservoirs (Cappaert et al. 1988; Gudmestad and Secor
63 1983; McCarter-Zorner et al. 1984; Peltzer and Sivasithamparam 1988; Franc G.D. and Harrison
64 M.D. 1987). Serological analysis performed in these early works identified up to 21 serogroups, and
65 a significant proportion of the isolates did not belong to known serogroups, pointing out the wide

66 diversity of SRP water isolates (Cappaert et al. 1988; Peltzer and Sivasithamparam 1988; Powelson,
67 M.L. and Apple J.D. 1984). Unfortunately, the strains isolated during these early studies were not
68 deposited in international collections and therefore their taxonomic status remains unclear. Indeed,
69 during these early samplings only 3 different groups were recognized within SRP while today the
70 taxonomy of the SRP group, clarified through genomic studies, encompasses 20 *Pectobacterium*
71 spp. 12 *Dickeya* spp. and 2 *Musicola* spp. (Toth et al. 2021b; Ben Moussa et al. 2021; Hugouvieux-
72 Cotte-Pattat et al. 2021; Hugouvieux-Cotte-Pattat and Van Gijsegem 2021) and detailed up-to-date
73 taxonomy and known host range of various species could be found in Table S1. Therefore, from
74 these early studies, it is difficult to understand which particular species are circulating in stream
75 water. Furthermore, water was poorly characterized in these early studies and it is unclear whether
76 better characterization of water quality could help to identify the risk of SRP presence in stream
77 water. Recently, water environment has regained attention and 7 SRP species isolated from stream
78 water were described. *Pectobacterium fontis* was isolated from a waterfall in Malaysia (Oulghazi et
79 al. 2019a), *Pectobacterium quasiaquaticum* from streams in France (Pédron et al. 2019; Ben Moussa
80 et al. 2021), *Pectobacterium aquaticum* from streams in France and from a lake in Poland (Pédron
81 et al. 2019; Babinska et al. 2021), *Pectobacterium polonicum* from groundwater within vegetable
82 fields in Poland (Waleron et al. 2019) and *Pectobacterium versatile* from water and a wide range of
83 diverse plants (Portier et al. 2019, 2020). As well, in the *Dickeya* genus, *Dickeya lacustris* was
84 observed in small eutrophic lakes surrounded by wetlands in the French region of La Dombes and
85 from the rhizosphere of pond-dwelling plants around these lakes (Hugouvieux-Cotte-Pattat et al.
86 2019), *Dickeya aquatica* was isolated from rivers in Finland and Scotland (Parkinson et al. 2014)
87 and further reported on rotted carrot plants (Zaczek-Moczydłowska et al. 2019), *Dickeya undicola*
88 was isolated from freshwater sampled both in Asia and Europe (Oulghazi et al. 2019b) and *Dickeya*
89 *zeae* and *Dickeya chrysanthemi* from river water in Poland (Potrykus et al. 2016). It is currently
90 unclear whether the recently described « aquatic » species are the main species circulating in

91 freshwater or if freshwaters frequently harbour aggressive plant species regularly responsible for
92 disease outbreaks.

93 The aim of the present paper was to obtain a comprehensive and holistic view of the various species
94 circulating in an open freshwater system. To do so, we performed a seasonal survey along the
95 Durance River. This river runs along 323 km from the pristine Alpine source to the stream mouth
96 in the agricultural plain of Avignon in the south-east of France. The isolated strains were
97 characterized through amplification and sequencing of the *gapA* house keeping gene, routinely used
98 to characterize SPR at the species level (Cigna et al. 2017). This was further completed by MLSA
99 analysis with *gapA*, *recA* and *dnaX* for 99 strains and genomic analysis of a 38 strains. The obtained
100 diversity was analysed in regards to the species abundance, the isolation sites, the season and the
101 water physico-chemical properties in order to tease out the ecological behaviours of the different
102 SRP species.

103

104 **MATERIALS AND METHODS**

105 **Description of sampled sites and analysis of water quality**

106 Sampling of surface water was performed along the Durance River in the south-east of France. To
107 cover the Durance watershed, 8 sites were selected along the main river, 11 sites on tributaries and
108 2 sites on the lower part were selected on an irrigation canal in that diverts from the river. Eleven of
109 the sampled sites (1 to 11) were located in the Alpine pristine upstream of the Serre-Ponçon lake.
110 This upper Alpine watershed is mainly devoted to pastoralism. The remaining 10 downstream
111 sampled sites (12 to 21) were located in the agricultural part of the Durance watershed. Precise
112 description of the sampling sites is shown in Figure 1 and Table S2. Samplings were performed at
113 2 sites in may 2015, 20 sites at fall 2015 and 21 sites during winter, spring, summer and fall in 2016
114 and 2017. Surface water was recovered at 5 meters distant from the bank with a bucket secured with
115 a rope. The bucket was first rinsed with river water and water recovered the second time was kept

116 for analysis. Particular attention was paid to avoid any sediment or plant debris inside the bucket.
117 One liter was maintained in a cool box before treatment that occurred within 24 h. Each sample (500
118 ml) was filtered through 0.22 μm cellulose acetate filters (Sartorius, Germany). Water in situ
119 temperature and electrical conductivity were measured using a Multi Probe System (YSI 556 MPS)
120 and water turbidity was measured using a EUTECH Instruments (TN-100) turbidity meter.
121 Acidified (85% H_3PO_4), filtered river water samples (0.2 μm) were used to determine the dissolved
122 organic carbon (DOC) concentration with a Shimadzu TOCVcph, as described in (Rochelle-Newall
123 et al. 2014). The concentration of nitrates, nitrites, ammonium, ortho-phosphates and total dissolved
124 nitrogen and phosphorus was determined by colorimetry (Bran and Luebbe 2013a,b,c,d) in the
125 laboratory with a segmented continuous flow analyzer (AA3, Seal Analytical, UK). The samples
126 (15 ml) were filtered in situ on 0.2 μm for dissolved nutrients and on 50 μm for total nitrogen and
127 phosphorus and frozen (-20°C) before analysis. Details of the water quality parameters are presented
128 in Table S3.

129 **Statistical analysis**

130 Pairwise Spearman correlation were calculated between variables. For each pairwise analysis
131 missing data were first removed before calculation that were performed with the following web site:
132 <https://biostatv.sentiweb.fr/?module=tests/spearman>. Pairwise Spearman correlations between
133 sampling sites altitudes and water quality parameters (temperature, conductivity, pH, turbidity,
134 Dissolved Organic carbon, PO_4^- , NH_4^+ , NO_2^- , NO_3^-) are presented in Table 2. Pairwise correlation
135 between SRP, genus or species strains occurrences and water quality variables are presented in
136 Table 3. Correlations were considered as non-significant when the p-value was superior to 0.01.

137

138 **Bacterial strains isolation**

139 The fraction retained on the filter following the 500 ml sampled water filtration (0.2 μm) was
140 suspended in 1 ml of sterile distilled water and 100 μl were used to inoculate 2 plates of a CVP

141 modified medium (per liter 1.02 g CaCl₂, 5 g tri-sodium citrate, 2.0 g NH₄NO₃, 2 ml crystal violet
142 0,075%, 4 g agar, 2.8 ml NaOH 5M, 18 g pectine Dipecta (ref AG366, Agdia biofords, USA;
143 hereafter CVPm) prepared as described by Hélias et al. 2012 for single layer CVP. A ten times
144 dilution (100 µl) was also spread on 2 or 3 plates of CVPm. The plates were incubated at 28°C for
145 2 days. Each pit-forming colony chosen for isolation was assigned with a number, collected with a
146 toothpick and further diluted into 1ml of sterile water. The obtained dilution (100 µl) was further
147 spread again on CVPm to check and isolate the pit-forming activity. When pit-forming activity was
148 confirmed, a well-isolated colony was further spread on LB medium and incubated overnight at
149 28°C. One isolated colony formed on this LB plate was both spread again on LB plate and checked
150 again for its pit-forming activity on CVPm plate. When the pit-forming activity was confirmed,
151 bacteria were scratched out from the lawn obtained on LB, suspended LB liquid medium and the
152 same volume of sterile glycerol 80% was added. The prepared bacterial suspension was conserved
153 at -80°C.

154 **Bacterial re-growth in river water**

155 River water, collected at site 18 was used for the experiment. Just after collection, the water was
156 filtered and the filtrate was autoclaved and kept in plastic bottle in the dark at room temperature
157 until use. Just before use, water was filtered again through a 0.22 µm filter to eliminate potential
158 salt precipitates. The bacterial re-growth was followed for species belonging to the *Pectobacterium*
159 genus: *P. carotovorum*, *P. versatile*, *P. aquaticum* and *P. atrosepticum*, and species of the *Dickeya*
160 genus: *D. zea/D. orizae*, *D. chrysanthemi*, *D. fangzhongdai* and *D. solani*. The strains used are
161 described in Table S4. For most species, the bacterial growth of 3 to 4 different strains was followed.
162 Bacterial strains were inoculated at 10³ CFU/ml and grown at 20°C or 8°C. Solid 10% TSA medium
163 (tryptic Soy Agar: 14 g agar, 1.5 g pancreatic digest of casein, 0.5 g peptic digest of soybean meal,
164 0.5 g sodium chloride per liter) was used to calculate the water culture's viable cell count by
165 spreading a diluted sample over the plate's surface and placing the plate at 28°C for 48 h. Results

166 shown in Figure 2 are the mean for each species of 4 independent growth curves.

167 **Species delineation of isolated strains**

168 The genus and species of each conserved bacteria was determined following amplification and
169 sequencing of the *gapA* housekeeping gene, as previously described (Cigna et al. 2017). Briefly,
170 bacteria were grown overnight on LB, diluted ten times in sterile water, boiled for 5 minutes and
171 place in a -20°C freezer for further use. Amplifications were performed with 5 µl of the boiled bacteria
172 and the *gapA* primers as previously described (Cigna et al. 2017) for 598 strains, and for a subset of
173 99 strains amplifications with the *dnaX* primers (Sławiak et al. 2009) and the *recA* primers (Lee et al.
174 2014). were also performed. The amplified products were Sanger-sequenced by EUROFIN. The fasta
175 files of the analyzed sequenced are available at <https://doi.org/10.5281/zenodo.5779227>. The *gapA*
176 sequences were aligned with reference sequences extracted from complete genome sequences using
177 the MUSCLE software (Edgar 2004) and the alignments were filtered with the program GBLOCKS
178 (Castresana 2000). Tree was computed using the PHYML algorithm (Guindon and Gascuel 2003)
179 implemented in the sea view software (Gouy et al. 2010) under default settings using the GTR model
180 (Tavaré 1986). *GapA* sequences shorter than 800 pb were not included in the phylogenetic tree and
181 species assignation was performed following blast analysis on NCBI. These strains were assigned to
182 a given species when at least 99% of identity was unambiguously observed along 90% of the sequence
183 with a well-defined species. In addition, for a subset of 99 strains, a MLSA tree was constructed from
184 concatenated nucleotide sequences of 3 housekeeping genes, *gapA*, *dnaX* and *recA*. Each gene was
185 aligned using the MUSCLE software and then concatenated. The alignments were filtered using the
186 GBLOCK tool, the tree computed by the PhyML algorithm, implemented in the SeaView software,
187 under default settings using the GTR model. Furthermore, 38 strains isolated in the course of this
188 study were sequenced, out of which 17 were previously released in the NCBI database and 21 were
189 new genomes analysed in the course of this study (Table S5). For preparation of genomic DNA, the
190 strains were grown overnight at 28°C on solid LB medium. A single colony was then picked up and

191 grown overnight in 2 ml of liquid LB medium at 28°C agitated at 20 rpm. After centrifugation of the
192 culture broth (5 min at 12000 rpm), DNA was extracted with the wizard genomic DNA extraction kit
193 (Promega) following the supplier's instructions. Genome sequencing was performed at the next
194 generation sequencing core facilities of the Institute for Integrative Biology of the Cell (Avenue de
195 la Terrasse 91190 Gif-sur-Yvette France) or at Genoscreen (Lille, France). Nextera DNA libraries
196 were prepared and Paired end 2x75 pb or 2X150 pb sequencing was performed on an Illumina
197 NextSeq500 apparatus, with a High Output 150 cycle kit. CLC Genomics Workbench (Version 9.5.2,
198 Qiagen Bioinformatics) was used to assemble reads. Final sequencing coverage was between 60 and
199 180. Coding sequences were predicted using the RAST server (19) with the Glimmer 3 prediction
200 tool (20). Statistics of the 21 newly sequenced draft genomes are presented in Table S5.

201

202 **RESULTS**

203 **Analysis of water quality along the watershed**

204 Analysis of water quality characteristics included : temperature, pH, conductivity, turbidity,
205 Dissolved Organic Carbon (DOC) in 2016 and 2017 and nitrate, nitrite, ammonium and phosphate
206 in 2017. Minimal, maximum and mean values observed for each considered variables are indicated
207 Table 1 and complete results are provided Table S3. Pairwise Spearman correlations between
208 altitude and water quality variables were calculated (Table 2). No significant correlation was
209 observed between altitude and phosphate or ammonium water content. However, significant
210 negative correlations were observed between altitude and nitrite content, altitude and turbidity,
211 altitude and conductivity while a positive correlation was observed between altitude and pH. As
212 expected, a strong negative correlation was observed between altitude and temperature. Strong
213 negative correlations were also observed between altitude and nitrate or altitude and DOC
214 concentration. Overall, this analysis confirmed the increasing importance of agriculture along the
215 watershed from its top to its bottom.

216 **Strains isolation and assignation to genera**

217 Depending on the sites and seasons, various numbers of pit-forming colonies were observed on the
218 CVPm plates. When the number of pit-forming colonies observed in a given sample was less than
219 20, we attempted to isolate all of them, when the number of observed pit-forming colonies was
220 superior to 20, we attempted to isolate 20 colonies. Overall, this survey led to isolation of 657 pit-
221 forming colonies on CVPm. Successful isolation varied between sampling years and month. Out of
222 these 657 isolated strains, the *gapA* amplicon was successfully amplified and sequenced for 598
223 strains and 16 sequences (2.7%) were neither assigned to *Dickeya*, *Pectobacterium* or *Musicola*
224 genera but related to other species of the *Enterobacterale* order such as *Enterobacter* sp. (8) *Serratia*
225 sp. (2), *Kluyvera* sp. (1), *Klebsiella* sp.(1), *Pantoea* sp. (1), *Rahnella* sp. (1), *Buttiauxella* sp. (1) and
226 *Citrobacter* sp. (1). The 582 remaining sequences were all assigned to SRP (97.3%) with none
227 assigned to the newly described *Musicola* genus, 33 (5.67%) assigned to the *Dickeya* genus and 549
228 (94.32%) assigned to the *Pectobacterium* genus. Spearman correlations with environmental
229 variables indicated SRP isolation correlates negatively with altitude and positively with
230 temperature, nitrate, nitrite and DOC content (Table 3). Weak but significant correlations were also
231 observed with sampled water conductivity and turbidity while no correlation was observed with
232 sampled water pH (Table 3).

233 **Strains belonging to the *Dickeya* genus**

234 Out of the 582 SRP strains characterized, 33 (5.67%) were assigned to the *Dickeya* genus following
235 *gapA* sequencing (Table S6). Due to the small number of recovered *Dickeya* strains, we did not
236 calculate Spearman correlation with environmental variables. However, we noticed that these 33
237 *Dickeya* strains were only recovered during spring and summer (Table 4). Furthermore, these 33
238 *Dickeya* strains were isolated from 6 sites that were all located in the lower part of the watershed
239 below an altitude of 500 m and many strains (19/33) were collected from irrigation canals at site 16
240 and 20 (Table 5). All these *Dickeya* strains were isolated when water temperature was superior to

241 19.50°C (mean 20.71°C +/- 0.575°C). This prompted us to compare the viability and growth of
242 *Dickeya* spp. and *Pectobacterium* spp. in river water at different temperatures (Figure 2). At 20°C,
243 both *Dickeya* spp. and *Pectobacterium* spp. bacteria inoculated at 10³ CFU/ml were able to grow
244 and reached at least 10⁵ CFU/ml at 10 days. At 8°C, *Pectobacterium* spp. grew slowly, reaching
245 5.10³ to 2.10⁴ CFU/ml at 10 days but *Dickeya* spp. did not grow and some species such as *D.*
246 *fangzhongdai* declined rapidly.

247 The *gapA* sequences allowed assignment of all but one of the isolated strains to known species
248 (Table S6). The last strain could be assigned to the *Dickeya* genus but the *gapA* sequence was too
249 short to decipher to which species it belonged. *GapA* species assignment was further confirmed for
250 8 strains belonging to each identified species with MLSA analysis performed with 3 housekeeping
251 genes *recA*, *dnaX* and *gapA* (Figure 3) and with MLSA based core genome analysis (Figure 4, Table
252 S7). Particularly, the phylogenetic tree performed with the sole *gapA* gene, the three housekeeping
253 genes, or the core genome, allowed to distinguish strains assigned to *D. zea* from those assigned to
254 the recently described and closely related *D. orizae* species, the only difference being that strains
255 belonging to the species *D. zea* were splitted in two clades within the *gapA* phylogenetic tree and
256 grouped a single clade following the *recA*, *dnaX* and *gapA* MLSA analysis (Figure 3). This analysis
257 was completed with average nucleotide identities (ANI) calculation performed on 4 sequenced
258 strains that clearly differentiate strains belonging to *D. zea* from strains belonging to the closely
259 related species *D. oryzae* (Table 6). Overall, out of the 13 currently described *Dickeya* spp. and
260 subsp., the 32 assigned strains belonged to 6 species, *D. orizae*, *D. zea*, *D. chrysanthemi*, *D. solani*,
261 *D. dianthicola* and *D. dadantii*. A strong domination of *D. orizae* was observed as *D. orizae* strains
262 represented 23 out of the 32 strains assigned to *Dickeya* spp.

263 **Strains belonging to the *Pectobacterium* genus**

264 Out of the 582 SRP strains characterized, strains belonging to the *Pectobacterium* genus dominated
265 with 549 (94.32%) of the characterized SRP strains assigned to this genus on the basis of the *gapA*

266 sequence. It is not surprising therefore that Spearman correlations with environmental variables
267 follow the same trend as the that observed with the full set of SRP isolated strains (Table 3). The
268 strongest correlation found was with altitude of the sampling site, followed by nitrate content and
269 temperature of the sampled water. However, in contrast to what was observed with *Dickeya* spp.,
270 *Pectobacterium* spp. were isolated at all seasons (Table 4) and at all sites but 2 (Table 5) in a large
271 range of temperature (from 3.6°C to 22.6°C).

272 MLSA analysis of a subset of 91 strains, performed with the 3 housekeeping genes *gapA*, *dnaX* and
273 *recA*, was compared with the one obtained with the sole *gapA* gene (Figure 3 and Figure S1, S2 for
274 extended trees). Out of the 91 analysed strains, 90 were similarly assigned at the species level
275 between the two phylogenetic trees (Figure 3). The main differences observed were that the *P.*
276 *carotovorum* strains were split in two different clades within the *gapA* phylogenetic tree and grouped
277 in a single clade within the MLSA phylogenetic tree while the reverse was observed for the *P.*
278 *aquaticum* strains that were split in two clades within the MLSA analysis and grouped in a single
279 clade in the *gapA* phylogenetic tree (Figure 3). Strain A519-S5-A17 was the only strain
280 differentially assigned at the species level between the two phylogenetic tree. This strain grouped
281 within the *P. aquaticum* clade following the *gapA* analysis and was attached to the base of the *P.*
282 *versatile* clade with the MLSA analysis (Figures 3, S2, S3). We then compared the species
283 assignment obtained between the *gapA* phylogenetic trees with the one based on MLSA performed
284 on core genome analysis of 30 *Pectobacterium* strains isolated during our survey whose full genome
285 sequences were previously published or released in the NCBI data base (Pédron et al. 2019; Portier
286 et al. 2019, 2020; Faye et al. 2018; Ben Moussa et al. 2021) or sequenced in the course of the present
287 study (Table S5). These 30 strains were assigned to the same species following *gapA* analysis or
288 full genome sequences. Therefore the *gapA* sequence was used to classify the whole set of 549
289 *Pectobacterium* isolates within species. All the isolated *Pectobacterium* strains but one could be
290 assigned to 9 *Pectobacterium* spp. (Table S6). A strong dominance of two species, *P. versatile* and

291 *P. aquaticum* was observed. *P. versatile* with 256 isolates, accounted for 46,6% of the
292 *Pectobacterium* strains isolated while *P. aquaticum* with 219 isolates account for 39,9% of the
293 *Pectobacterium* isolated strains. Other species such as *P. carotovorum* (36 isolates), *P.*
294 *quasiaquaticum* (13 isolates) and *P. odoriferum* (11 isolates) represented each 2% to 6% of the
295 isolated strains. The repartition of these latter species was variable: *P. carotovorum* strains were
296 isolated from 9 sites at various seasons, *P. quasiaquaticum* were recovered at all seasons on 5 sites
297 located in lower part of the watershed and all the *P. odoriferum* strains but one were isolated from
298 a single site on a single date (Table 4 and 5). Finally, *P. atrosepeticum* (1 isolate), *P. peruvienne* (2
299 isolates) *P. polaris* (5 isolates) and *P. brasiliense* (4 isolates) represented each less than 1% of the
300 isolated strains.

301 For the two abundant species, *P. versatile* and *P. aquaticum*, we compared when and where strains
302 of each species were isolated. Both *P. versatile* and *P. aquaticum* were isolated at all seasons but *P.*
303 *versatile* was most preferentially isolated at fall and spring while *P. aquaticum* was preferentially
304 isolated at summer and fall (Table 4). The sites of isolation varied between these two species. While
305 *P. versatile* was isolated from 19 sites all along the watershed, *P. aquaticum* was isolated from 12
306 sites that were mostly located in the lower part of the watershed (Table 5). Furthermore, while the
307 number of strains isolated at each site was roughly equilibrated for *P. versatile*, more than 50% of
308 the *P. aquaticum* strains were isolated at 2 sites (Table 5). Interestingly, we also noted that 83%
309 (81/97) of the *Pectobacterium* spp. isolated in the upper part of the Durance watershed belonged to
310 the *P. versatile* species (Table 5).

311

312 **DISCUSSION**

313 This work represents a comprehensive view of SRP dissemination and diversity at the scale
314 of a large watershed covering a surface of 14 280 km². This watershed was interesting to follow
315 bacterial plant pathogens such as SRP because its runs along 323 km from an alpine area devoted

316 to pastoralism and hiking to the agricultural plain of Avignon where various crop species are
317 cultivated. Therefore, following distribution and diversity of SRP along this watershed over more
318 than two years along the four seasons helped to differentiate species that could be isolated across
319 the four seasons, those that are found in pristine areas and those that are more associated with crop
320 culture. Characterization of the sampled water confirmed the increasing importance of agriculture
321 along the watershed as nitrite, nitrate and DOC negatively correlates with altitude of the sampled
322 water. Water temperature from the top to the bottom was also an interesting parameter to follow as
323 it varied from 0.3°C to 25.6°C with a mean of 10.8°C.

324 Our study showed that SRP were rare in water. This is in agreement with Pédrón et al (Pédrón
325 et al, 2020) that previously showed that SRP are not detected through metabarcoding in river water
326 although they are detected when an efficient semi-selective medium is used. This indicates that SRP
327 are not indigenous planktonic species in surface water and are only sporadically passing by in water.
328 The biological continuity between soil and stream microbial communities via surface water runoff
329 has been shown elsewhere (Le et al. 2022) and this link probably explains the low and sporadic
330 prevalence.

331 Our phylogenetic analysis, through the sequencing of the single *gapA* housekeeping gene
332 proved efficient for roughly classifying 582 of the recovered strains within 15 SRP species. This
333 classification was congruent with that performed with 3 housekeeping genes for 99 strains for all
334 the analysed strains but one. Full genome analysis of 38 strains also confirmed the species
335 assignation based on the sole *gapA* gene. This indicates that sequencing the housekeeping gene
336 *gapA* gene is an efficient tool to rapidly classify large sets of strains within *Pectobacterium* and
337 *Dickeya* genera.

338 No strain of the recently described *Musicola* genus were observed during our survey
339 (Hugouvieux-Cotte-Pattat et al. 2021). This new genus was described following an in depth genomic
340 analysis showing that genomes of the formerly *Dickeya paradisiaca* species aligned with genomes

341 of other *Dickeya* species on less than one third of their genomes. The newly created genus was
342 named *Musicola* with reference to the isolation of most strains from *Musa paradisiaca*. The fact
343 that members of *Musicola paradisiaca* species were mostly isolated from tropical and subtropical
344 samples may explain their absence in a temperate watershed such as the one studied here.

345 We found that bacteria belonging to the *Dickeya* genus were restricted to the lower part of the
346 watershed and were only isolated during spring and summer when temperature was superior to 19.5°C
347 and we experimentally confirmed the difficulty of several *Dickeya* spp. to grow and survive at low
348 temperature in river water. This is in line with a previous survey from an Australian river that detected
349 *Dickeya* spp. after an enrichment procedure on water samples with temperature superior to 16.2°C
350 (Cothier and Gilbert 1990) and with the fact that *Dickeya* spp. were historically described as being
351 present mainly in tropical and subtropical regions although they now appear to be expanding their
352 global distribution. With the prospect of global warming, it is expected that *Dickeya* spp. population
353 will increase in river water in the future. Regular water monitoring when water temperature is high
354 could help to mitigate the risk of crop cultures infection by *Dickeya* spp. through irrigation.

355 The isolation of *Dickeya* spp., *D. undicola*, *D. aquatica* and *D. lacustris* from surface water
356 of rivers, lakes and irrigation canals have recently been described (Hugouvieux-Cotte-Pattat et al.
357 2019; Oulghazi et al. 2019b; Parkinson et al. 2014). Surprisingly, none of these species were isolated
358 during our 2 year survey suggesting that these species are sporadically associated with water. In
359 2016 and 2017 *D. undicola* strains were isolated from a small irrigation canal at Monfavet (Oulghazi
360 et al. 2019b). This sampling point is close to the points 20 and 21 sampled in this study, suggesting
361 that extensive sampling in the lower Durance sites during spring and summer when water
362 temperature is high, will likely extend the number of *Dickeya* spp. recovered. Indeed, we also
363 isolated strains of *D. fangzhondai*, a highly aggressive *Dickeya* spp. (Alič et al. 2017), from
364 additional sampling from a small irrigation canal not included in the present study in the lower
365 Durance watershed.

366 During our survey, the main isolated *Dickeya* spp. were *D. orizae* followed by *D.*
367 *chrysanthemi*. Similarly, a survey focusing on *Dickeya* spp. previously performed in Poland also
368 identified *D. zea* followed by *D. chrysanthemi* as the main *Dickeya* species circulating in surface
369 water (Potrykus et al. 2016). Since this 2016 publication, the *D. zea* clade has been split in two
370 species *D. zea* and *D. oryzae* (Wang et al. 2020). In our work, we found that *D. orizae* was more
371 abundant than *D. zea*. Interestingly, neither *D. zea*/*D. orizae* nor *D. chrysanthemi* have been
372 reported to cause crop disease outbreaks in Europe. *D. zea* and *D. oryzae* were, respectively,
373 described as infecting maize and rice, both of which are monocotyledon plants. Therefore, it is
374 plausible that their detection in river water could be due to their association with some common
375 herbaceous monocotyledons unrelated to crop culture found on the riverbanks. Indeed, while some
376 symptomless weeds in the vicinity of potato fields were found to harbour SRP (Tsrör et al. 2019;
377 Zoledowska et al. 2018), plants along riverbanks have not been investigated in detail, although *D.*
378 *lacustris* has been isolated from the rhizosphere the pond edge plant *Solanum dulcamara*
379 (Hugouvieux-Cotte-Pattat et al. 2019). As well, *P. carotovorum* has been isolated from *Solanum*
380 *dulcamara* in Poland (Fikowicz-Krosko et al. 2017). Enlarging sampling of common plants found
381 along river banks could help to decipher if these plants are important drivers of SRP circulation in
382 river water, particularly as the distance of plants from the stream course has been shown to be
383 important in structuring aquatic microbial communities (Le et al. 2018).

384 Among SRP, our survey indicated that *Pectobacterium* spp. were far more frequently isolated
385 than *Dickeya* spp. in river water. This is in line with previous surface water surveys performed along
386 the Colorado rivers in the USA where no *Dickeya* spp. were reported (Maddox and Harrison 1988).
387 Therefore, in temperate streams, both in Europe and USA, *Pectobacterium* spp. largely dominate.
388 This could be linked with the fact that *Pectobacterium* spp., in contrast with *Dickeya* spp., could be
389 recovered across a large range of water temperatures. *Pectobacterium* spp. are also known to survive
390 better in soils than *Dickeya* spp. (Perombelon and Hyman 1989) and this better survival could also

391 contribute to the higher detection of *Pectobacterium* spp. in river water through leaching and surface
392 water runoff during rain events.

393 Among the *Pectobacterium* recovered spp., two recently described species, *P. aquaticum*
394 and *P. versatile* dominated, representing respectively 46.6% and 39.9% of the recovered
395 *Pectobacterium* spp. A recent taxonomic update of 265 *Pectobacterium* strains hosted at the CIRM-
396 CFBP international collection that gathers strains isolated since 1944, showed that *P. versatile* is a
397 broad spectrum species regularly isolated from a large number of cultivated plants. In contrast, no
398 strain of *P. aquaticum* isolated from plants are hosted in this collection (Portier et al. 2020).
399 Therefore, the abundance of both species in river water might have different origins. Interestingly,
400 while *P. versatile* was recovered all along the stream course and was regularly isolated in the upper
401 part of the watershed in Alpine pristine water area, *P. aquaticum* was mostly found in the lower part
402 of the watershed and more than 50% of the strains were isolated from only two sampling sites. These
403 two sampling sites were characterized by abundant presence of aquatic plants on their banks,
404 suggesting that *P. aquaticum* could be associated with some of these plants on the riverbank. In
405 contrast, the large circulation of *P. versatile* all along the stream course, both in the pristine area
406 and the lower, agriculturally dominated part of the watershed, is reminiscent of what was observed
407 another plant bacterial pathogen *P. syringae*. Actually, bacteria belonging to the *P. syringae*
408 complex population were regularly isolated from alpine lakes and stream and their ecology was
409 proposed to be linked to the water cycle (Morris et al. 2013). However, strains of *P. versatile* are
410 far less abundant in Alpine pristine water than strains of the *P. syringae* complex (Pédron et al.
411 2020), and their presence in clouds or rain has not been reported (Failor et al. 2017). In addition, *P.*
412 *versatile* occurrence along the watershed strongly correlates with environmental variables such as
413 nitrate and DOC indicating a larger occurrence in the lower part of the watershed while previous
414 work indicated that strains of the *P. syringae* complex are abundant in the upper part of the
415 watershed (Monteil et al. 2014). Therefore, while *P. versatile* is the SRP species with the largest

416 observed prevalence both on plant and in river stream it behave quite differently from strains of the
417 *P. syringae* complex.

418 *P. versatile*, despite its large prevalence has only recently been described. This is principally
419 due to its close genomic proximity with *P. carotovorum* which explains the previous mix up of the
420 two species (Portier et al. 2019, 2020). This mix up was also favoured by the fact that *P.*
421 *carotovorum* and *P. versatile* both have a broad geographical distribution on plants (Portier et al.
422 2020; Ma et al. 2007). Our river survey however indicated that *P. carotovorum* was far less abundant
423 in stream water than *P. versatile* (36 isolates vs 246 isolates) and was isolated from a smaller number
424 of sites (9 vs 19 for *P. versatile*). The same was true for *P. quasiquaticum*, despite its close genomic
425 proximity to *P. aquaticum* (Ben Moussa et al. 2021), its prevalence in river water was also one order
426 of magnitude smaller (13 isolates vs 219 isolates) with less sampling sites positive (5 vs 12). This
427 highlights the fact that closely related species have different ecological behaviours and warns
428 against extensive generalisation without careful analysis of the studied bacterial populations.

429 SRP pathogens regularly isolated from crop disease outbreaks such as *P. atrosepticum* and
430 *P. brasiliense*, *D. solani* or *D. dianthicola* were rarely isolated along the watershed. While the mean
431 water temperature observed during this survey (10.9°C) could explain the rare occurrence of *Dickeya*
432 spp., *Pectobacterium* spp. were recovered from a large range of water temperatures and we
433 experimentally observed the ability of *Pectobacterium* spp. to grow in sterilized river water at low
434 temperature. *P. atrosepticum* is well known to be a specialized species mainly found on potato crops
435 which could explain its scarcity in water. *P. brasiliense* has a larger plant host spectrum and its rare
436 occurrence was more surprising although it may suggest a poor survival capacity on soil or non-
437 crop plants compared to species regularly observed in water such as *P. versatile*.

438 The large majority of strains isolated during this survey belonged to two species, *P.*
439 *aquaticum* and *P. versatile*, that are not known to be associated with severe outbreaks for crop
440 culture. This suggests that the risk of infection may be overestimated when surveys do not include

441 careful characterization of the SRP species involved. In that regard, taxonomical analysis with a
442 single housekeeping gene could help to rapidly analysed the isolated species. However, the risk of
443 infection with irrigation remains, as species responsible for disease outbreak such as *P.*
444 *atrosepticum*, *D. solani* or *D. dianthicola* were sporadically isolated, albeit at low frequency. We
445 also identified 2 strains of *P. peruvienne*. The *P. peruvienne* species was described following strains
446 isolated from diseased potato on the Peruvian altiplano but this species has not yet been described
447 on crop plants in Europe (Waleron et al. 2018; Faye et al. 2018). The occurrence of *P. peruvienne*
448 strains in European river water indicates surveys of microbial water quality could help to identify
449 new bacterial threats not yet reported on plant in a given geographic area. Overall, our survey also
450 revealed that the critical sites to be surveyed regularly to estimate the risk of SRP for crop culture
451 are the those located in the downstream section of the watershed where most SRP strains are
452 regularly found. Regular monitoring of well-chosen sites could help to prevent the risk of infection
453 for crops.

454

455 **Acknowledgements**

456 We thank Katherine Green for her help during the growth experiment in river water. We also thank
457 Odile Berge and Frédérique Van Gijsegem for their help during river sampling and Ariane Toussaint
458 and Antoine Pourbaix for hosting us during the field samplings. This work has benefited from the
459 expertise of the High-Throughput Sequencing Platform of I2BC, Gif sur Yvette, France.

460 **Ethical statement**

461 not applicable

462 **Conflicts of interest**

463 The authors declare that there are no conflicts of interest.

464

465 **LITERATURE CITED**

- 466 Ben Moussa, H., Pédrón, J., Bertrand, C., Hecquet, A., and Barny, M.-A. 2021. *Pectobacterium*
467 *quasiaquaticum* sp. nov., isolated from waterways. Int. J. Syst. Evol. Microbiol. 71
- 468 Bran, and Luebbe. 2013a. Nitrite in water and seawater. Method No. G-173-96 Rev. 10
- 469 Bran, and Luebbe. 2013b. Phosphate in water and seawater. Method No. G-175-96 Rev. 15
- 470 Bran, and Luebbe. 2013c. Total nitrogen in water. Method No. G-392-08 Rev. 5.
- 471 Bran, and Luebbe. Total phosphorous in water. Method No. G-393-08 Rev.
- 472 Burr, T. J., and Schroth, M. N. 1977. Occurrence of soft-rot *Erwinia* spp. in soil and plant material.
473 Phytopathology. 67:1382–1387.
- 474 Cappaert, M. R., Powelson, M. L., Franc, D. G., and Harrison, M. D. 1988. Irrigation water as a
475 source of inoculum of soft rot *Erwinias* for aerial stem rot of potatoes. Phytopathology. 78:1668–
476 1672.
- 477 Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in
478 phylogenetic analysis. Mol. Biol. Evol. 17:540–552.
- 479 Charkowski, A. O. 2018. The changing face of bacterial soft-rot diseases. Annu. Rev. Phytopathol.
480 56:269–288.
- 481 Cigna, J., Dewaegeneire, P., Beury, A., Gobert, V., and Faure, D. 2017. A *gapA* PCR-sequencing
482 assay for identifying the *Dickeya* and *Pectobacterium* potato pathogens. Plant Dis. 101:1278–1282.
- 483 Cother, E. J., and Gilbert, R. L. 1990. Presence of *Erwinia chrysanthemi* in two major river systems
484 and their alpine sources in Australia. J. Appl. Bacteriol. 69:729–738.
- 485 Edgar, R. C. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space
486 complexity. BMC Bioinformatics. 5:113.
- 487 Failor, K. C., Schmale, D. G., Vinatzer, B. A., and Monteil, C. L. 2017. Ice nucleation active bacteria
488 in precipitation are genetically diverse and nucleate ice by employing different mechanisms. ISME J.
489 11:2740–2753.
- 490 Faye, P., Bertrand, C., Pédrón, J., and Barny, M.-A. 2018. Draft genomes of “*Pectobacterium*

- 491 *peruviense*” strains isolated from fresh water in France. Stand. Genomic Sci. 13
- 492 Fikowicz-Krosko, J., Wszalek-Rozek, K., Smolarska, A., & Czajkowski, R. 2017. First report on
493 isolation of soft rot *Pectobacterium carotovorum* subsp. *carotovorum* from symptomless bittersweet
494 nightshade occurring in rural area in Poland. J. Plant Pathol. 99(1), 294.
- 495 Franc G.D. and Harrison M.D. 1987. The role of contaminated irrigation water in the recontamination
496 of *Erwinia* free seed potatoes. Am. Potato J. 64:438.
- 497 Gascuel, O. 1997. BIONJ: an improved version of the NJ algorithm based on a simple model of
498 sequence data. Mol. Biol. Evol. 14:685–695.
- 499 Gouy, M., Guindon, S., and Gascuel, O. 2010. SeaView Version 4: A Multiplatform Graphical User
500 Interface for Sequence Alignment and Phylogenetic Tree Building. Mol. Biol. Evol. 27:221–224.
- 501 Gudmestad, N. C., and Secor, G. A. 1983. The bionomics of *Erwinia carotovora* in north dakota.
502 Am. Potato J. 60:759–771.
- 503 Guindon, S., and Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large
504 phylogenies by maximum likelihood. Syst. Biol. 52:696–704.
- 505 Hélias, V., Hamon, P., Huchet, E., Wolf, J. V. D., and Andrivon, D. 2012. Two new effective
506 semiselective crystal violet pectate media for isolation of *Pectobacterium* and *Dickeya*: Isolating
507 pectolytic bacteria on CVP. Plant Pathol. 61:339–345.
- 508 Hugouvieux-Cotte-Pattat, N., des-Combes, C. J., Briolay, J., and Pritchard, L. 2021. Proposal for the
509 creation of a new genus *Musicola* gen. nov., reclassification of *Dickeya paradisiaca* (Samson et al.
510 2005) as *Musicola paradisiaca* comb. nov. and description of a new species *Musicola keenii* sp. nov.
511 Int. J. Syst. Evol. Microbiol. 71
- 512 Hugouvieux-Cotte-Pattat, N., Condemine, G., and Shevchik, V. E. 2014. Bacterial pectate lyases,
513 structural and functional diversity. Environ. Microbiol. Rep. 6:427–440.
- 514 Hugouvieux-Cotte-Pattat, N., Jacot-des-Combes, C., and Briolay, J. 2019. 2019. *Dickeya lacustris*
515 sp. nov., a water-living pectinolytic bacterium isolated from lakes in France. Int. J. Syst. Evol.

- 516 Microbiol. 69:721–726.
- 517 Le, H. T., Pommier, T., Ribolzi, O., Soulileuth, B., Huon, S., Silvera, N., et al. 2022. Overland flow
518 during a storm event strongly affects stream water chemistry and bacterial community structure.
519 Aquat. Sci. 84:7.
- 520 Le, H. T., Rochelle-Newall, E., Auda, Y., Ribolzi, O., Sengtaheuanghoung, O., Thébault, E., et al.
521 2018. Vicinal land use change strongly drives stream bacterial community in a tropical montane
522 catchment. FEMS Microbiol. Ecol. 94.
- 523 Ma, B., Hibbing, M. E., Kim, H.-S., Reedy, R. M., Yedidia, I., Breuer, J., et al. 2007. Host range and
524 molecular phylogenies of the soft rot *Enterobacterial* genera *Pectobacterium* and *Dickeya*.
525 Phytopathology. 97:1150–1163.
- 526 McCarter-Zorner, N. J., Franc, G. D., Harrison, M. D., Michaud, J. E., Quinn, C. E., Sells, I. A., et
527 al. 1984. Soft rot *Erwinia* bacteria in surface and underground waters in southern Scotland and in
528 Colorado, United States. J. Appl. Bacteriol. 57:95–105.
- 529 Monteil, C. L., Lafolie, F., Laurent, J., Clement, J.-C., Simler, R., Travi, Y., et al. 2014. Soil water
530 flow is a source of the plant pathogen *Pseudomonas syringae* in subalpine headwaters. Environ.
531 Microbiol. 16:2038–2052.
- 532 Morris, C. E., Monteil, C. L., and Berge, O. 2013. The life history of *Pseudomonas syringae* : Linking
533 agriculture to earth system processes. Annu. Rev. Phytopathol. 51:85–104.
- 534 Oulghazi, S., Cigna, J., Lau, Y. Y., Moumni, M., Chan, K. G., and Faure, D. 2019a. Transfer of the
535 waterfall source isolate *Pectobacterium carotovorum* M022 to *Pectobacterium fontis* sp. nov., a deep-
536 branching species within the genus *Pectobacterium*. Int. J. Syst. Evol. Microbiol. 69:470–475.
- 537 Oulghazi, S., Pédrón, J., Cigna, J., Lau, Y. Y., Moumni, M., Van Gijsegem, F., et al. 2019b. *Dickeya*
538 *undicola* sp. nov., a novel species for pectinolytic isolates from surface waters in Europe and Asia.
539 Int. J. Syst. Evol. Microbiol. 69:2440–2444.
- 540 Parkinson, N., DeVos, P., Pirhonen, M., and Elphinstone, J. 2014. 2014. *Dickeya aquatica* sp. nov.,

- 541 isolated from waterways. *Int. J. Syst. Evol. Microbiol.* 64:2264–2266.
- 542 Pédrón, J., Bertrand, C., Taghouti, G., Portier, P., and Barny, M.-A. 2019. *Pectobacterium aquaticum*
543 sp. nov., isolated from waterways. *Int. J. Syst. Evol. Microbiol.* 69:745–751.
- 544 Pédrón, J., Guyon, L., Lecomte, A., Blottière, L., Chandeysson, C., Rochelle-Newall, E., et al. 2020.
545 Comparison of environmental and culture-derived bacterial communities through 16S
546 metabarcoding: A powerful tool to assess media selectivity and detect rare taxa. *Microorganisms.*
547 8:1129.
- 548 Peltzer, S., and Sivasithamparam, K. 1988. Sero-groups of *Erwinia carotovora* associated with water,
549 soil, tuber, and stems of potato plants in Western Australia. *N. Z. J. Exp. Agric.* 16:265–270.
- 550 Perombelon, M. C. M., and Hyman, L. J. 1989. Survival of soft rot coliforms, *Erwinia carotovora*
551 subsp. *carotovora* and *E. carotovora* subsp. *atroseptica* in soil in Scotland. *J. Appl. Bacteriol.* 66:95–
552 106.
- 553 Portier, P., Pédrón, J., Taghouti, G., Dutrieux, C., and Barny, M.-A. 2020. Updated Taxonomy of
554 *Pectobacterium* genus in the CIRM-CFBP bacterial collection: when newly described species reveal
555 “old” endemic population. *Microorganisms.* 8:1441.
- 556 Portier, P., Pédrón, J., Taghouti, G., Fischer-Le Saux, M., Caullireau, E., Bertrand, C., et al. 2019.
557 Elevation of *Pectobacterium carotovorum* subsp. *odoriferum* to species level as *Pectobacterium*
558 *odoriferum* sp. nov., proposal of *Pectobacterium brasiliense* sp. nov. and *Pectobacterium actinidiae*
559 sp. nov., emended description of *Pectobacterium carotovorum* and description of *Pectobacterium*
560 *versatile* sp. nov., isolated from streams and symptoms on diverse plants. *Int. J. Syst. Evol. Microbiol.*
561 69:3207–3216.
- 562 Potrykus, M., Golanowska, M., Sledz, W., Zoledowska, S., Motyka, A., Kolodziejaska, A., et al. 2016.
563 Biodiversity of *Dickeya* spp. isolated from potato plants and water sources in temperate climate. *Plant*
564 *Dis.* 100:408–417.
- 565 Powelson, M.L. and Apple J.D. 1984. Soil and seed tubers as sources of inoculum of *Erwinia*

- 566 *carotovora* pv. *carotovora* for stem soft rot potatoes. *Phytopathology*. :429–432.
- 567 Rochelle-Newall, E., Hulot, F. D., Janeau, J. L., and Merroune, A. 2014. CDOM fluorescence as a
568 proxy of DOC concentration in natural waters: a comparison of four contrasting tropical systems.
569 *Environ. Monit. Assess.* 186:589–596.
- 570 Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. In
571 *Lectures on Mathematics in the Life Sciences*, American Mathematical Society, p. 57–86.
- 572 Toth, I. K., Barny, M., Brurberg, M. B., Condemine, G., Czajkowski, R., Elphinstone, J. G., et al.
573 2021a. *Pectobacterium* and *Dickeya*: environment to disease development. In *Plant Diseases Caused*
574 *by Dickeya and Pectobacterium Species*, eds. Frédérique Van Gijsegem, Jan M. van der Wolf, and
575 Ian K. Toth. Cham: Springer International Publishing, p. 39–84.
- 576 Toth, I. K., Barny, M., Czajkowski, R., Elphinstone, J. G., Li, X. (Sean), Pédrón, J., et al. 2021b.
577 *Pectobacterium* and *Dickeya*: taxonomy and evolution. In *Plant Diseases Caused by Dickeya and*
578 *Pectobacterium Species*, eds. Frédérique Van Gijsegem, Jan M. van der Wolf, and Ian K. Toth. Cham:
579 Springer International Publishing, p. 13–37.
- 580 Tsrör, L., Lebiush, S., Erlich, O., Galilov, I., Chalupowicz, L., Reuven, M., et al. 2019. First report
581 of latent infection of *Malva nicaeensis* caused by *Pectobacterium carotovorum* subsp. *brasiliense* in
582 Israel. *New Dis. Rep.* 39:4–4.
- 583 Van Gijsegem, F., Toth, I. K., and van der Wolf, J. M. 2021. Soft Rot *Pectobacteriaceae*: A brief
584 overview. In *Plant Diseases Caused by Dickeya and Pectobacterium Species*, eds. Frédérique Van
585 Gijsegem, Jan M. van der Wolf, and Ian K. Toth. Cham: Springer International Publishing, p. 1–11.
- 586 Waleron, M., Misztak, A., Waleron, M., Franczuk, M., Wielgomas, B., and Waleron, K. 2018.
587 Transfer of *Pectobacterium carotovorum* subsp. *carotovorum* strains isolated from potatoes grown at
588 high altitudes to *Pectobacterium peruviense* sp. nov. *Syst. Appl. Microbiol.* 41:85–93.
- 589 Waleron, M., Misztak, A., Waleron, M., Jonca, J., Furmaniak, M., and Waleron, K. 2019.
590 *Pectobacterium polonicum* sp. nov. isolated from vegetable fields. *Int. J. Syst. Evol. Microbiol.*

591 69:1751–1759.

592 Wang, X., He, S.-W., Guo, H.-B., Han, J.-G., Thin, kyu kyu, Gao, J., et al. 2020. *Dickeya oryzae* sp.

593 nov., isolated from the roots of rice. *Int. J. Syst. Evol. Microbiol.* 70:4171–4178.

594 Zaczek-Moczydłowska, M. A., Fleming, C. C., Young, G. K., Campbell, K., and O’Hanlon, R. 2019.

595 *Pectobacterium* and *Dickeya* species detected in vegetables in Northern Ireland. *Eur. J. Plant Pathol.*

596 154:635–647.

597 Zoledowska, S., Motyka, A., Zukowska, D., Sledz, W., and Lojkowska, E. 2018. Population structure

598 and biodiversity of *Pectobacterium parmentieri* isolated from potato fields in temperate climate. *Plant*

599 *Dis.* 102:154–164.

600

601

602 **FIGURE LEGEND**

603 **Figure 1:** Schematic representation of sampling points along the river Durance. The Durance river
604 is located in south-east of France. Its headwaters rise in the Alp mountains at an altitude of 2390 m.
605 The Durance river runs for 323 km until it flows into the Rhône river at an altitude of 10 m. The 21
606 sampling points are indicated by red dots together with their assigned number from the highest to
607 lowest altitude. Points 1 to 11 are in the Alpine part of the river above the Serre Ponçon lake, points
608 12 to 21 are in the lower part of the rivers where agriculture is important. The sampled tributaries
609 are depicted from top to bottom: Clarée, Guisane, Gyronde, Guil, Ubaye, Buëch, Bléone, Verdon
610 and Grand Anguillon. Sampling points 1 and 20 are on irrigation canals that derivate from the River
611 Durance. The lowest tributary, the Grand-Anguillon River runs along 20.1 km as an irrigation canal
612 from its spring to its confluence with the Durance River. The black dots indicate the main cities
613 along the river, from top to bottom: Briançon, Embrun, Gap, Sisteron, Manosque and Avignon.
614 Detailed GPS coordinates and altitude of sampled points are indicated in Table S2.

615 **Figure 2:** Growth curves of *Pectobacterium* spp. and *Dickeya* spp. in Durance river water. Water
616 was collected at site 18 in December 2018, filtered on a 0.22 µm acetate cellulose membrane and
617 autoclaved prior use. Bacteria were inoculated at 10³ CFU/ml in 9 ml of river water. The 9 ml were
618 then split evenly between two plastic tubes, containing a 4.5 ml culture each, one being incubated
619 at 20°C and the other being put into an incubator at 8°C. The tested bacterial species are indicated
620 below the graph. The graph represents the mean of 3 or 4 independent growth curves observed with
621 up to 4 different strains of each species. *D. zae* and *D. orizae* are presented together as they were
622 included in the same species at the time of the experiment and were only recently split. All the
623 strains used in this experiment are described in Table S4.

624 **Figure 3:** Comparative phylogenetic analysis

625 a) Phylogenetic tree constructed on the basis of the partial *gapA* gene sequence; b) Phylogenetic tree
626 constructed on the basis of concatenated partial gene sequences of *gapA*, *dnaX*, and *recA*.

627 99 strains (91 *Pectobacterium* spp. and 8 *Dickeya* spp.) isolated in the course of this study and 29
628 reference strains representatives of *Pectobacterium* and *Dickeya* spp. were included in this analysis.
629 The number in brackets indicates the number of isolated strains present in each clade. The position
630 of strain A519-S5-A17, the only strain out of 99 that grouped with different species following each
631 analysis, is indicated with an asterisk. In both phylogenetic analyses bootstrap percentages were
632 calculated based on 100 replicates and bootstrap support values are indicated if less than 70%. Bar,
633 0.07 changes per nucleotide position. Fasta sequences used to construct these phylogenetics analysis
634 are provided at <https://doi.org/10.5281/zenodo.5779227>. Extended phylogenetic trees are provided in
635 Figures S2 and S3.

636 **Figure 4:** MLSA phylogenetic tree reconstructed from concatenated nucleotide sequences of 601
637 homologous gene sequences.

638 The 38 sequenced strains, with strain names starting with an “A”, are included in the phylogenetic
639 tree together with 29 reference strains for each species. Clustering of homologous nucleotide
640 sequences was performed with SiLix software with a 80% identity threshold. Homologous sequences
641 of each gene were aligned using MUSCLE (Edgar 2004) software then concatenated. The alignments
642 were filtered using the GBLOCK tool (Castresana 2000) resulting in a data set of 627806 sites (of
643 which 211221 are informative). The tree was computed with the SeaView software (Gouy et al. 2010)
644 using the BioNJ method (Gascuel 1997). Bootstrap percentages were calculated based on 100
645 replicates and only bootstrap values <100 are represented. Bar, 0.02 changes per nucleotide position.
646 The NCBI accession numbers for the genomes used in this analysis are available in Table S7.

647
648 **Table 1:** Observed minimal, maximal and mean values for each water quality variables
649 **Table 2 :** Pairwise Spearman correlations of water quality variables with altitude
650 **Table 3:** Pairwise Spearman correlation of strains occurrence along the watershed with water quality
651 variables for the all SRP, the *Pectobacterium* genus or the indicated species

- 652 **Table 4** : Number of isolated strains for each season for the indicated genera or species
 653 **Table 5** : Number of strains isolated at each site for the indicated genera or species
 654 **Table 6** : Pairwise ANI values between *D. zea* and *D. oryzae* genomes.
 655 **Fig S1**: Extended *gapA* phylogenetic tree corresponding to Fig 3A
 656 **Fig S2**: Extended *gapA-dnaX-recA* phylogenetic tree corresponding to Fig 3B
 657 Provided as supplementary exel file:
 658 **Table S1**: Description of SRP species with recorded host range
 659 **Table S2**: Description of the sampled sites
 660 **Table S3**: Water physico-chemical data
 661 **Table S4**: Strains used for growth in river Durance water
 662 **Table S5** : Characteristics and accession numbers of the 21 genomes sequenced in the course of this
 663 study
 664 **Table S6**: *gapA* assignation of the 582 SRP isolated strains
 665 **Table S7**: NCBI accession number for genome presented in the phylogenetic tree Figure 4
 666

667

668 **Table 1** : Observed minimal, maximal and mean values for each water quality variables

	pH	Temp (°C)	Conductivity (µS)	Turbidity (NTU)	DOC (µMC)	PO4- (µg/L)	NH4+ (µg/L)	NO2-- (µg/L)	NO3- (µg/L)
min	7,80	0,30	81,00	0,00	16,66	1,05	0,56	0,11	2,54
max	9,03	25,60	1175,00	494,00	235,83	61,45	107,13	18,05	930,42
mean	8,55	10,9	493,01	20,96	84,69	7,45	18,05	2,91	225,94

669

670 **Table 2**: Pairwise Spearman correlations of water quality variables with altitude

	Spearman correlation with altitude								
	temp	conductivity	pH	turbidity	DOC	PO4-	NH4+	NO2--	NO3-
Spearman correlation	-0.6351	-0.386	0.369	-0.4599	-0.5325	0.09	-0.0908	-0.4573	-0.6029
p-value	7.47E-23	3.80E-08	2.32E-05	3.63E-09	1.42E-14	0.42	0.41	1.38E-05	1.63E-9
nb observations	190	190	125	149	180	83	83	83	83

671

672

Each water quality variables indicated in lane 2 were analyzed in regard to the altitude of the sampling points.

673 Table 3: Pairwise Spearman correlations of strains occurrence along the watershed with
674 environmental variables for the all SRP, the *Pectobacterium* genus or the indicated species

	Nb of collected strains	Altitude	Temperature	DOC	Conductivity	pH	turbidity	NO3 ⁻	NO2 ⁻
SRP	657	-0.5881 (4.60E-19)	0.49 (7.25E-13)	0.3775 (2.22E-05)	0.27 (1.64E-04)	-0.1388 (1.23E-01)	0.2483 (1.49E-03)	0.44 (2.22E-05)	0.4122 (1.08E-04)
<i>Pectobacterium</i> genus	549	-0.5689 (1.10E-17)	0.4524 (5.66E-11)	0.3572 (8.58E-07)	0.2814 (8.38E-05)	-0.1158 (1.98E-01)	0.2384 (2.33E-03)	0.4601 (1.07E-05)	0.4046 (1.48E-04)
<i>P. versatile</i>	256	-0.4165 (2.27E-09)	0.3228 (5.58E-06)	0.266 (4.02E-03)	0.2153 (2.86E-03)	0.0181 (8.41E-01)	0.1584 (4.48E-02)	0.3107 (4.02E-03)	0.2532 (2.09E-02)
<i>P. aquaticum</i>	219	-0.5112 (4.82E-14)	0.3758 (9.15E-18)	0.3657 (3.90E-04)	0.2176 (2.56E-03)	-0.177 (4.83E-02)	0.1017 (1.99E-01)	0.3782 (3.90E-04)	0.36 (8.03E-04)

675 Pairwise Spearman correlations between the bacterial groups indicated on the first column and the water quality
676 variables indicated on the first line. The p-value are indicated in bracket
677

678 Table 4: Number of isolated strains for each seasons for the
679 indicated genera or species

	<i>P. genus</i>	<i>D. genus</i>	<i>Pve</i>	<i>Paq</i>	<i>Pcar</i>	<i>Pqu</i>	<i>Pod</i>	<i>Dor</i>	Total*
fall	197	0	73	84	22	4	11	0	197
winter	81	0	45	28	3	3	1	0	81
spring	134	12	87	36	3	4	0	10	146
summer	137	21	51	71	8	2	0	13	158
total	549	33	256	219	36	13	12	23	582

680 *P.*: *Pectobacterium*; *D.*: *Dickeya*; *Pve*: *P. versatile*; *Paq*: *P. aquaticum*;
681 *Pca*: *P. carotovorum*; *Pqu*: *P. quasiaquaticum*; *Pod*: *P. odoriferum*; *Dor*: *D. oryzae*.
682 *Total is the sum of *Pectobacterium* and *Dickeya* genera and encompassed also
683 rare species with occurrence <10 not displayed in this table.
684
685

686 Table 5 : number of strains isolated at each site for the indicated genera or species

site	altitude	<i>P. genus</i>	<i>D. genus</i>	<i>Pve</i>	<i>Paq</i>	<i>Pca</i>	<i>Pqu</i>	<i>Pod</i>	<i>Dor</i>	Total*
1	2090	0	0	0	0	0	0	0	0	0
2	1813	5	0	5	0	0	0	0	0	5
3	1659	4	0	2	0	2	0	0	0	4
4	1443	4	0	4	0	0	0	0	0	4
5	1363	4	0	4	0	0	0	0	0	4
6	1363	15	0	13	2	0	0	0	0	15
7	1294	0	0	0	0	0	0	0	0	0
8	1066	7	0	7	0	0	0	0	0	7
9	968	6	0	6	0	0	0	0	0	6
10	907	22	0	15	0	6	0	0	0	22
11	790	30	0	25	1	2	0	0	0	30
12	620	46	0	6	27	2	0	11	0	46
13	459	41	0	15	20	4	1	0	0	41

14	459	11	1	10	1	0	0	0	1	12
15	438	27	1	18	4	3	0	0	0	28
16	291	38	0	17	16	0	4	0	0	38
17	274	19	0	21	71	0	0	0	0	92
18	188	38	9	19	6	12	0	0	8	47
19	106	92	13	11	5	1	2	1	11	32
20	98	34	6	18	11	0	2	0	3	40
21	39	106	3	40	55	4	4	0	0	109
total		549	33	256	219	36	13	12	23	582

687 *P.*: *Pectobacterium*; *D.*: *Dickeya*; *Pve*: *P. versatile*; *Paq*: *P. aquaticum*; *Pca*: *P. carotovorum*;

688 *Pqu*: *P. quasiaquaticum*; *Pod*: *P. odoriferum*; *Dor*: *D. oryzae*.

689 The double line indicates the limit between the Alpine upper watershed and the downstream agricultural part

690 of the watershed. *Total is the sum of *Pectobacterium* and *Dickeya* genera and encompasses also rare species

691 with occurrence <10 not displayed in this table.

692

693

694

Table 6: Pairwise ANI values between *D. zea* and *D. oryzae* genomes

	1	2	3	4	5	6	7	8
1: <i>D. oryzae</i> A003-S1-M15	1.00	0.99	0.97	0.97	0.95	0.95	0.95	0.94
2: <i>D. oryzae</i> A642-S2-A17	0.99	1.00	0.97	0.97	0.95	0.95	0.95	0.94
3: <i>D. oryzae</i> ZYY5	0.97	0.97	1.00	0.99	0.95	0.95	0.95	0.94
4: <i>D. oryzae</i> EC1	0.97	0.97	0.99	1.00	0.95	0.95	0.95	0.94
5: <i>D. zea</i> MS2	0.95	0.95	0.95	0.95	1.00	0.98	0.98	0.96
6: <i>D. zea</i> A661-S21-A17	0.95	0.95	0.95	0.95	0.98	1.00	0.98	0.96
7: <i>D. zea</i> NCPPB3532	0.95	0.95	0.95	0.95	0.98	0.98	1.00	0.96
8: <i>D. zea</i> A586-S18-A17	0.94	0.94	0.94	0.94	0.96	0.96	0.96	1.00

695 ANI values above 96% are shown in orange, those below 96% in blue

696

697

698

699

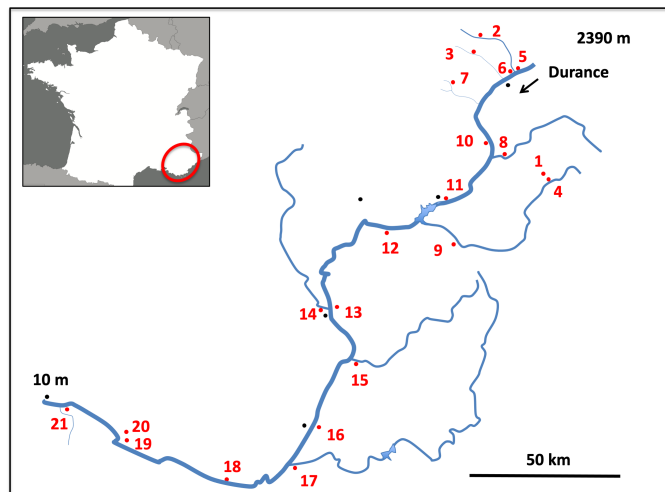


Figure 1: Schematic representation of sampling points along the river Durance.

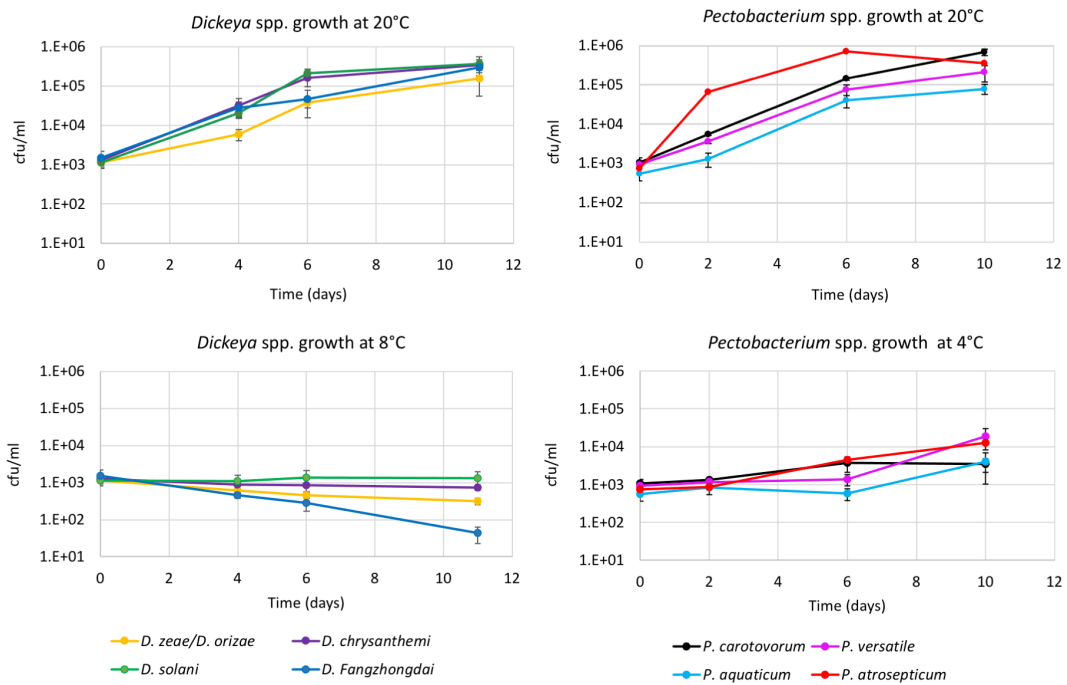
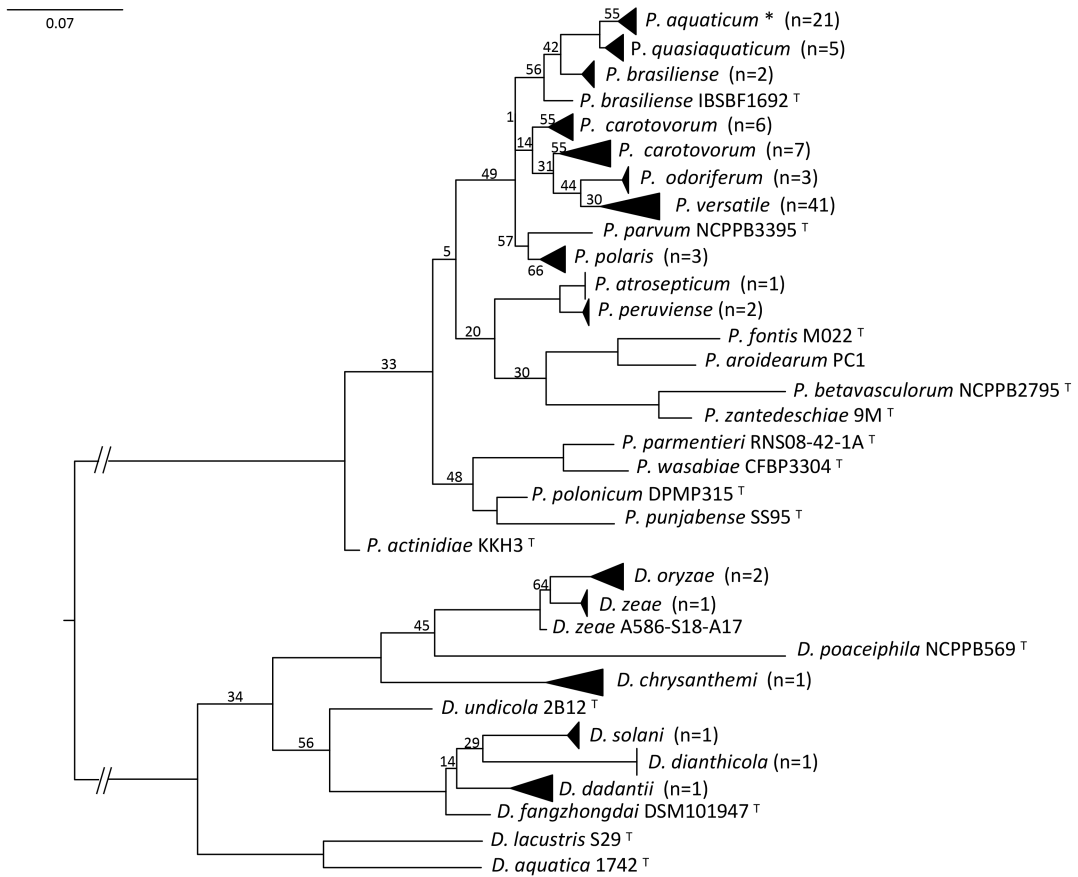


Figure 2: Growth curves of Pectobacterium spp. and Dickeya spp. in Durance river water.

a)



b)

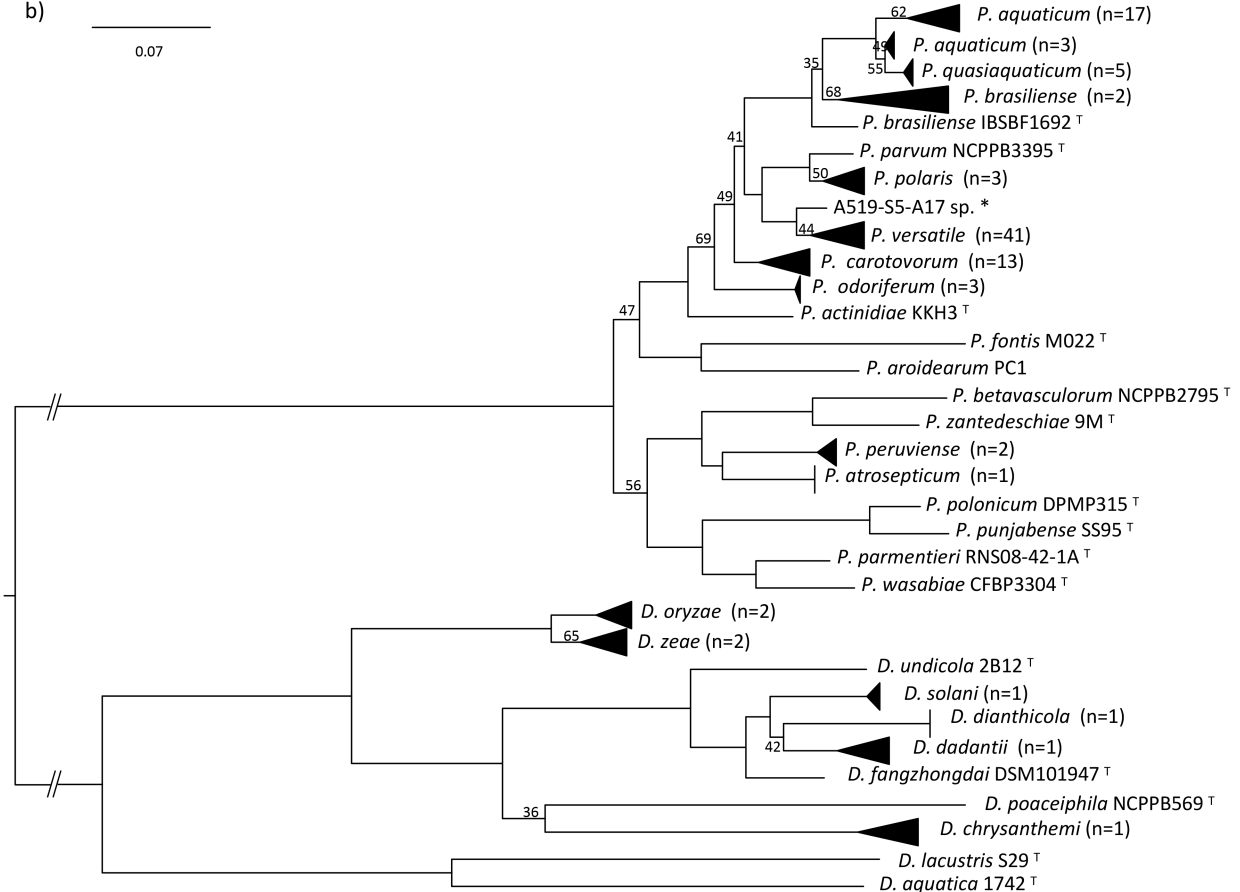


Figure 3: Comparative phylogenetic analysis

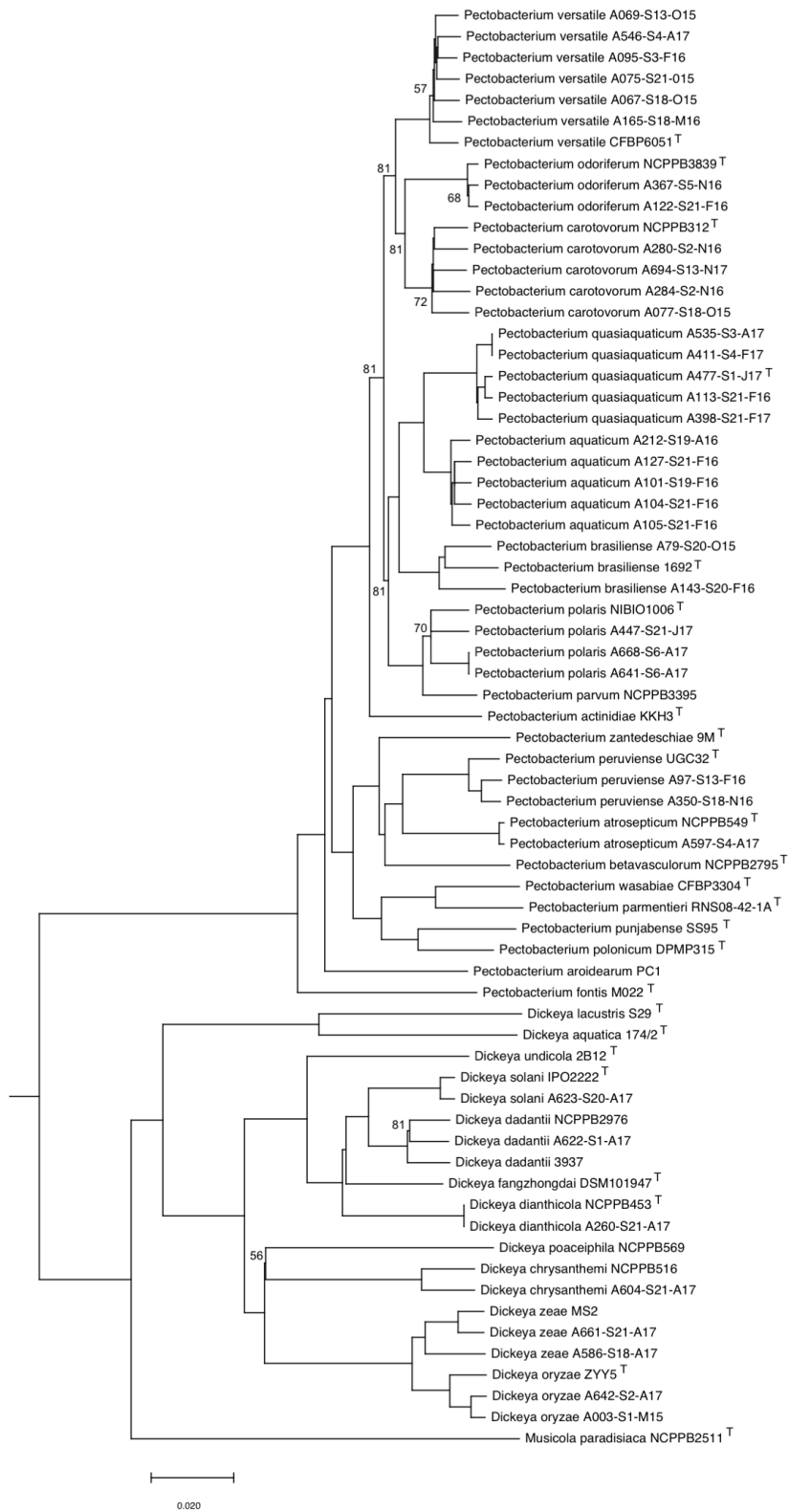


Figure 4: MLSA phylogenetic tree reconstructed from concatenated nucleotide sequences of 601 homologous gene sequences.