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, Claire Bertrand, Emma Rochelle-Newall, Sarah Fiorini, Jacques Pédron, Marie-Anne Barny

Hajar Ben Moussa
Phytopathology. 2022 Feb 27. doi: 10.1094/PHYTO-12-21-0515-R.

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-Newall[1], Sarah Fiorini[2], Jacques Pédron[1], Marie-Anne Barny[1#]

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

[2]Ecole Normale Supérieure, PSL Research University, CNRS, Centre de recherche en écologie expérimentale et prédictive (CEREEP-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

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 *gap*A 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 assignation obtained with the sole *gap*A 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.

INTRODUCTION

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

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

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

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

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

MATERIALS AND METHODS

Description of sampled sites and analysis of water quality

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

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

Statistical analysis

Pairwise Spearman correlation were calculated between variables. For each pairwise analysis missing data were first removed before calculation that were performed with the following web site: https://biostatgv.sentiweb.fr/?module=tests/spearman. Pairwise Spearman correlations between sampling sites altitudes and water quality parameters (temperature, conductivity, pH, turbidity, Dissoved Organic carbon, PO4-, NH4+, $NO2^{--}$, NO3-) are presented in Table 2. Pairwise correlation between SRP, genus or species strains occurrences and water quality variables are presented in Table 3. Correlations were considered as non-significant when the p-value was superior to 0.01.

Bacterial strains isolation

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

modified medium (per liter 1.02 g $CaCl_2$, 5 g tri-sodium citrate, 2.0 g NH_4NO_3, 2 ml crystal violet 0,075%, 4 g agar, 2.8 ml NaOH 5M, 18 g pectine Dipecta (ref AG366, Agdia biofords, USA; hereafter CVPm) prepared as described by Hélias et al. 2012 for single layer CVP. A ten times dilution (100 µl) was also spread on 2 or 3 plates of CVPm. The plates were incubated at 28°C for 2 days. Each pit-forming colony chosen for isolation was assigned with a number, collected with a toothpick and further diluted into 1ml of sterile water. The obtained dilution (100 µl) was further spread again on CVPm to check and isolate the pit-forming activity. When pit-forming activity was confirmed, a well-isolated colony was further spread on LB medium and incubated overnight at 28°C. One isolated colony formed on this LB plate was both spread again on LB plate and checked again for its pit-forming activity on CVPm plate. When the pit-forming activity was confirmed, bacteria were scratched out from the lawn obtained on LB, suspended LB liquid medium and the same volume of sterile glycerol 80% was added. The prepared bacterial suspension was conserved at -80°C.

Bacterial re-growth in river water

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

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

Species delineation of isolated strains

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

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

RESULTS

Analysis of water quality along the watershed

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

Strains isolation and assignation to genera

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

Strains belonging to the *Dickeya* genus

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

19.50°C (mean 20.71°C + /- 0.575°C). This prompted us to compare the viability and growth of *Dickeya* spp. and *Pectobacterium* spp. in river water at different temperatures (Figure 2). At 20°C, both *Dickeya* spp. and *Pectobacterium* spp. bacteria inoculated at 10^3 CFU/ml were able to grow and reached at least 10^5 CFU/ml at 10 days. At 8°C, *Pectobacterium* spp. grew slowly, reaching 5.10^3 to 2.10^4 CFU/ml at 10 days but *Dickeya* spp. did not grow and some species such as *D. fangzhongdai* declined rapidly.

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

Strains belonging to the *Pectobacterium* genus

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

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

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

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

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

DISCUSSION

This work represents a comprehensive view of SRP dissemination and diversity at the scale of a large watershed covering a surface of 14 280 km². This watershed was interesting to follow 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 Pédron et al (Pédron
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 *gap*A 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 *gap*A gene. This indicates that sequencing the housekeeping gene
336 *gap*A 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

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

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

The isolation of *Dickeya* spp., *D. undicola*, *D. aquatica* and *D. lacustris* from surface water of rivers, lakes and irrigation canals have recently been described (Hugouvieux-Cotte-Pattat et al. 2019; Oulghazi et al. 2019b; Parkinson et al. 2014). Surprisingly, none of these species were isolated during our 2 year survey suggesting that these species are sporadically associated with water. In 2016 and 2017 *D. undicola* strains were isolated from a small irrigation canal at Monfavet (Oulghazi et al. 2019b). This sampling point is close to the points 20 and 21 sampled in this study, suggesting that extensive sampling in the lower Durance sites during spring and summer when water temperature is high, will likely extend the number of *Dickeya* spp. recovered. Indeed, we also isolated strains of *D. fangzhondai*, a highly aggressive *Dickeya* spp. (Ališ et al. 2017), from additional sampling from a small irrigation canal not included in the present study in the lower 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. zeae* clade has been split in two
370 species *D. zeae* and *D. oryzae* (Wang et al. 2020). In our work, we found that *D. orizae* was more
371 abundant than *D. zeae*. Interestingly, neither *D. zeae/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 (Tsror 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

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

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

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

The large majority of strains isolated during this survey belonged to two species, *P. aquaticum* and *P. versatile*, that are not known to be associated with severe outbreaks for crop 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. peruviense*. The *P. peruviense* 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. peruviense*
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**

Ben Moussa, H., Pédron, J., Bertrand, C., Hecquet, A., and Barny, M.-A. 2021. *Pectobacterium quasiaquaticum* sp. nov., isolated from waterways. Int. J. Syst. Evol. Microbiol. 71

Bran, and Luebbe. 2013a. Nitrite in water and seawater. Method No. G-173-96 Rev. 10

Bran, and Luebbe. 2013b. Phosphate in water and seawater. Method No. G-175-96 Rev. 15

Bran, and Luebbe. 2013c. Total nitrogen in water. Method No. G-392-08 Rev. 5.

Bran, and Luebbe. Total phosphorous in water. Method No. G-393-08 Rev.

Burr, T. J., and Schroth, M. N. 1977. Occurrence of soft-rot *Erwinia* spp. in soil and plant material. Phytopathology. 67:1382–1387.

Cappaert, M. R., Powelson, M. L., Franc, D. G., and Harrison, M. D. 1988. Irrigation water as a source of inoculum of soft rot *Erwinias* for aerial stem rot of potatoes. Phytopathology. 78:1668–1672.

Castresana, J. 2000. Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis. Mol. Biol. Evol. 17:540–552.

Charkowski, A. O. 2018. The changing face of bacterial soft-rot diseases. Annu. Rev. Phytopathol. 56:269–288.

Cigna, J., Dewaegeneire, P., Beury, A., Gobert, V., and Faure, D. 2017. A gapA PCR-sequencing assay for identifying the *Dickeya* and *Pectobacterium* potato pathogens. Plant Dis. 101:1278–1282.

Cother, E. J., and Gilbert, R. L. 1990. Presence of *Erwinia chrysanthemi* in two major river systems and their alpine sources in Australia. J. Appl. Bacteriol. 69:729–738.

Edgar, R. C. 2004. MUSCLE: a multiple sequence alignment method with reduced time and space complexity. BMC Bioinformatics. 5:113.

Failor, K. C., Schmale, D. G., Vinatzer, B. A., and Monteil, C. L. 2017. Ice nucleation active bacteria in precipitation are genetically diverse and nucleate ice by employing different mechanisms. ISME J. 11:2740–2753.

Faye, P., Bertrand, C., Pédron, J., and Barny, M.-A. 2018. Draft genomes of "*Pectobacterium*

peruviense" strains isolated from fresh water in France. Stand. Genomic Sci. 13

Fikowicz-Krosko, J., Wszalek-Rozek, K., Smolarska, A., & Czajkowski, R. 2017. First report on isolation of soft rot *Pectobacterium carotovorum* subsp. *carotovorum* from symptomless bittersweet nightshade occuring in rural area in Poland. J. Plant Pathol. *99*(1), 294.

Franc G.D. and Harrison M.D. 1987. The role of contaminated irrigation water in the recontamination of *Erwinia* free seed potatoes. Am. Potato J. 64:438.

Gascuel, O. 1997. BIONJ: an improved version of the NJ algorithm based on a simple model of sequence data. Mol. Biol. Evol. 14:685–695.

Gouy, M., Guindon, S., and Gascuel, O. 2010. SeaView Version 4: A Multiplatform Graphical User Interface for Sequence Alignment and Phylogenetic Tree Building. Mol. Biol. Evol. 27:221–224.

Gudmestad, N. C., and Secor, G. A. 1983. The bionomics of *Erwinia carotovora* in north dakota. Am. Potato J. 60:759–771.

Guindon, S., and Gascuel, O. 2003. A simple, fast, and accurate algorithm to estimate large phylogenies by maximum likelihood. Syst. Biol. 52:696–704.

Hélias, V., Hamon, P., Huchet, E., Wolf, J. V. D., and Andrivon, D. 2012. Two new effective semiselective crystal violet pectate media for isolation of *Pectobacterium* and *Dickeya*: Isolating pectolytic bacteria on CVP. Plant Pathol. 61:339–345.

Hugouvieux-Cotte-Pattat, N., des-Combes, C. J., Briolay, J., and Pritchard, L. 2021. Proposal for the creation of a new genus *Musicola* gen. nov., reclassification of *Dickeya paradisiaca* (Samson et al. 2005) as *Musicola paradisiaca* comb. nov. and description of a new species *Musicola keenii* sp. nov. Int. J. Syst. Evol. Microbiol. 71

Hugouvieux-Cotte-Pattat, N., Condemine, G., and Shevchik, V. E. 2014. Bacterial pectate lyases, structural and functional diversity. Environ. Microbiol. Rep. 6:427–440.

Hugouvieux-Cotte-Pattat, N., Jacot-des-Combes, C., and Briolay, J. 2019. 2019. *Dickeya lacustris* 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., Pédron, 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.,

isolated from waterways. Int. J. Syst. Evol. Microbiol. 64:2264–2266.

Pédron, J., Bertrand, C., Taghouti, G., Portier, P., and Barny, M.-A. 2019. *Pectobacterium aquaticum* sp. nov., isolated from waterways. Int. J. Syst. Evol. Microbiol. 69:745–751.

Pédron, J., Guyon, L., Lecomte, A., Blottière, L., Chandeysson, C., Rochelle-Newall, E., et al. 2020. Comparison of environmental and culture-derived bacterial communities through 16S metabarcoding: A powerful tool to assess media selectivity and detect rare taxa. Microorganisms. 8:1129.

Peltzer, S., and Sivasithamparam, K. 1988. Sero-groups of *Erwinia carotovora* associated with water, soil, tuber, and stems of potato plants in Western Australia. N. Z. J. Exp. Agric. 16:265–270.

Perombelon, M. C. M., and Hyman, L. J. 1989. Survival of soft rot coliforms, *Erwinia carotovora* subsp. *carotovora* and *E. carotovora* subsp. *atroseptica* in soil in Scotland. J. Appl. Bacteriol. 66:95–106.

Portier, P., Pédron, J., Taghouti, G., Dutrieux, C., and Barny, M.-A. 2020. Updated Taxonomy of *Pectobacterium* genus in the CIRM-CFBP bacterial collection: when newly described species reveal "old" endemic population. Microorganisms. 8:1441.

Portier, P., Pédron, J., Taghouti, G., Fischer-Le Saux, M., Caullireau, E., Bertrand, C., et al. 2019. Elevation of *Pectobacterium carotovorum* subsp. *odoriferum* to species level as *Pectobacterium odoriferum* sp. nov., proposal of *Pectobacterium brasiliense* sp. nov. and *Pectobacterium actinidiae* sp. nov., emended description of *Pectobacterium carotovorum* and description of *Pectobacterium versatile* sp. nov., isolated from streams and symptoms on diverse plants. Int. J. Syst. Evol. Microbiol. 69:3207–3216.

Potrykus, M., Golanowska, M., Sledz, W., Zoledowska, S., Motyka, A., Kolodziejska, A., et al. 2016. Biodiversity of *Dickeya* spp. isolated from potato plants and water sources in temperate climate. Plant Dis. 100:408–417.

Powelson, M.L. and Apple J.D. 1984. Soil and seed tubers as sources of inoculum of *Erwinia*

carotovora pv. *carotovora* for stem soft rot potatoes. Phytopathology. :429–432.

Rochelle-Newall, E., Hulot, F. D., Janeau, J. L., and Merroune, A. 2014. CDOM fluorescence as a proxy of DOC concentration in natural waters: a comparison of four contrasting tropical systems. Environ. Monit. Assess. 186:589–596.

Tavaré, S. 1986. Some probabilistic and statistical problems in the analysis of DNA sequences. In *Lectures on Mathematics in the Life Sciences*, American Mathematical Society, p. 57–86.

Toth, I. K., Barny, M., Brurberg, M. B., Condemine, G., Czajkowski, R., Elphinstone, J. G., et al. 2021a. *Pectobacterium* and *Dickeya*: environment to disease development. In *Plant Diseases Caused by Dickeya and Pectobacterium Species*, eds. Frédérique Van Gijsegem, Jan M. van der Wolf, and Ian K. Toth. Cham: Springer International Publishing, p. 39–84.

Toth, I. K., Barny, M., Czajkowski, R., Elphinstone, J. G., Li, X. (Sean), Pédron, J., et al. 2021b. *Pectobacterium* and *Dickeya*: taxonomy and evolution. In *Plant Diseases Caused by Dickeya and Pectobacterium Species*, eds. Frédérique Van Gijsegem, Jan M. van der Wolf, and Ian K. Toth. Cham: Springer International Publishing, p. 13–37.

Tsror, L., Lebiush, S., Erlich, O., Galilov, I., Chalupowicz, L., Reuven, M., et al. 2019. First report of latent infection of *Malva nicaeensis* caused by *Pectobacterium carotovorum* subsp. *brasiliense* in Israel. New Dis. Rep. 39:4–4.

Van Gijsegem, F., Toth, I. K., and van der Wolf, J. M. 2021. Soft Rot *Pectobacteriaceae*: A brief overview. In *Plant Diseases Caused by Dickeya and Pectobacterium Species*, eds. Frédérique Van Gijsegem, Jan M. van der Wolf, and Ian K. Toth. Cham: Springer International Publishing, p. 1–11.

Waleron, M., Misztak, A., Waleron, M., Franczuk, M., Wielgomas, B., and Waleron, K. 2018. Transfer of *Pectobacterium carotovorum* subsp. *carotovorum* strains isolated from potatoes grown at high altitudes to Pectobacterium peruviense sp. nov. Syst. Appl. Microbiol. 41:85–93.

Waleron, M., Misztak, A., Waleron, M., Jonca, J., Furmaniak, M., and Waleron, K. 2019. *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

FIGURE LEGEND

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

Figure 2: Growth curves of *Pectobacterium* spp. and *Dickeya* spp. in Durance river water. Water was collected at site 18 in December 2018, filtered on a 0.22 µm acetate cellulose membrane and autoclaved prior use. Bacteria were inoculated at 10^3 CFU/ml in 9 ml of river water. The 9 ml were then split evenly between two plastic tubes, containing a 4.5 ml culture each, one being incubated at 20°C and the other being put into an incubator at 8°C. The tested bacterial species are indicated below the graph. The graph represents the mean of 3 or 4 independent growth curves observed with up to 4 different strains of each species. *D. zeae* and *D. orizae* are presented together as they were included in the same species at the time of the experiment and were only recently split. All the strains used in this experiment are described in Table S4.

Figure 3: Comparative phylogenetic analysis

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

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

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

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

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

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

Table 3: Pairwise Spearman correlation of strains occurrence along the watershed with water quality 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. zeae* 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	2.32-05	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 Each water quality variables indicated in lane 2 were analyzed in regard to the altitude of the sampling points.
672

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)
Pectobacterium 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)
P. versatile	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)
P. aquaticum	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

	P. genus	*D.* genus	*Pve*	*Paq*	*Pcar*	*Pqu*	*Pod*	*Dor*	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	*P.* genus	*D.* genus	*Pve*	*Paq*	*Pca*	*Pqu*	*Pod*	*Dor*	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

*P.: Pectobacterium; D.: Dickeya; Pve: P. versatile; Paq: P. aquaticum; Pca: P. carotovorum; Pqu: P. quasiaquaticum; Pod: P. odoriferum; Dor: D. oryzae. The double line indicates the limit between the Alpine upper watershed and the downstream agricultural part of the watershed. *Total is the sum of Pectobacterium and Dickeya genera and encompasses also rare species with occurrence <10 not displayed in this table.*

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

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

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

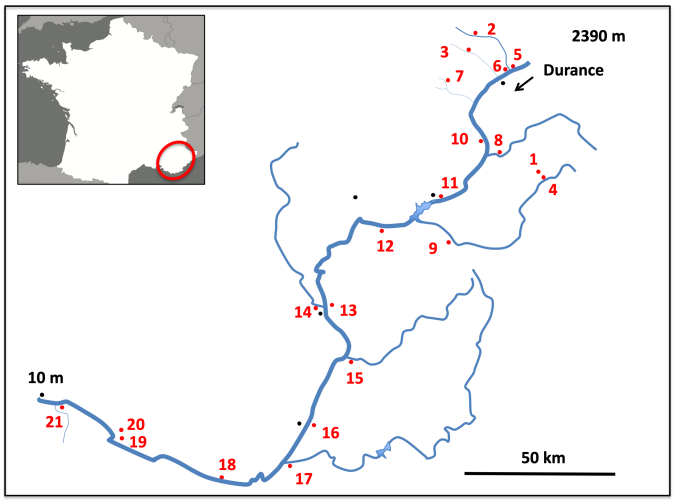

Figure 1: Schematic representation of sampling points along the river Durance.

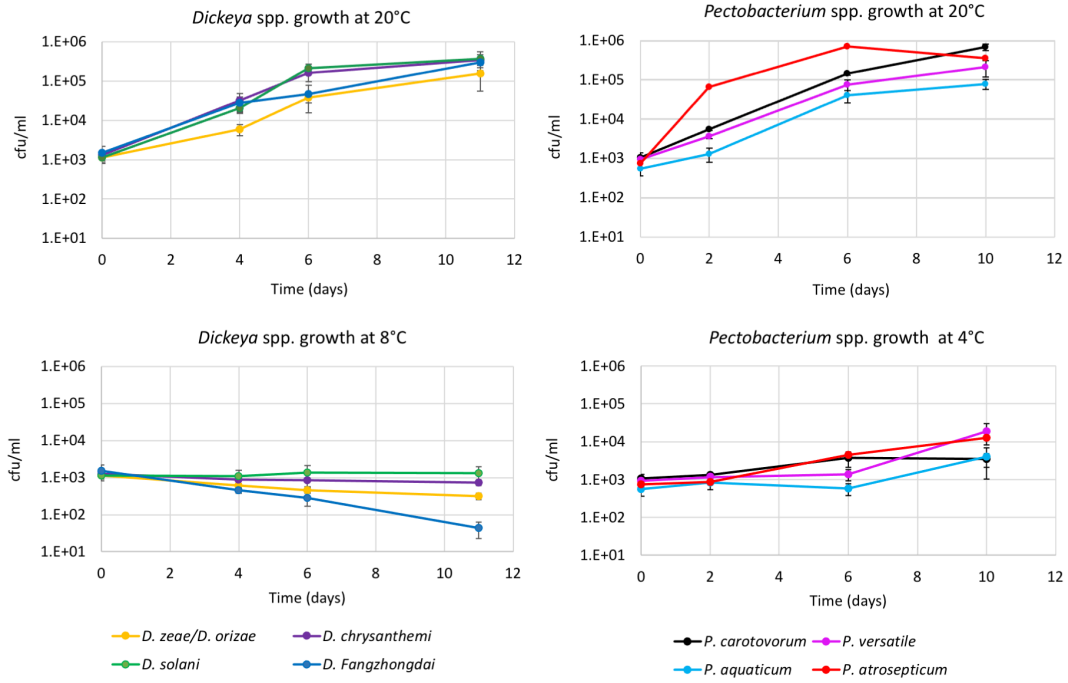

Figure 2: Growth curves of Pectobacterium spp. and Dickeya spp. in Durance river water.

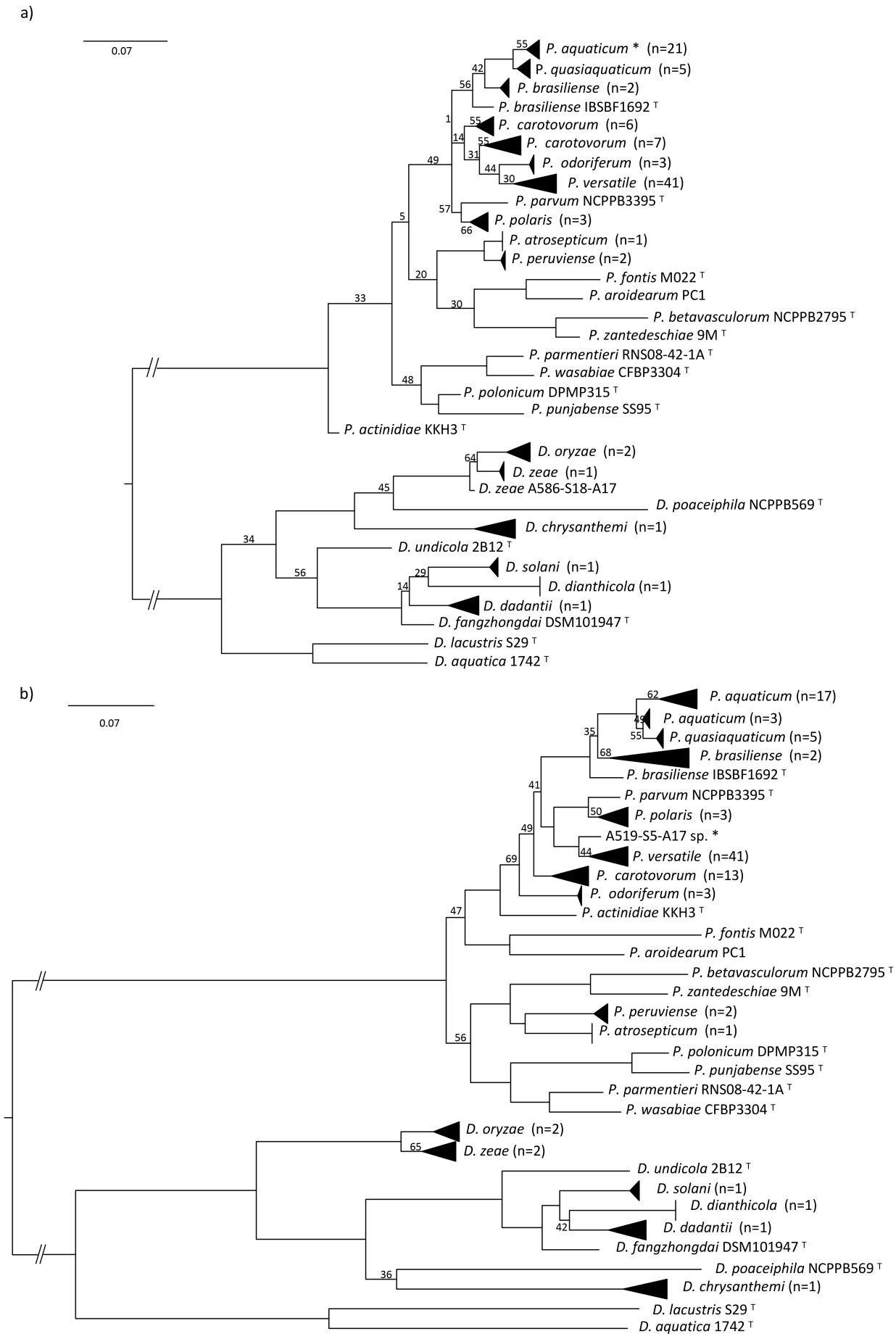

Figure 3: Comparative phylogenetic analysis

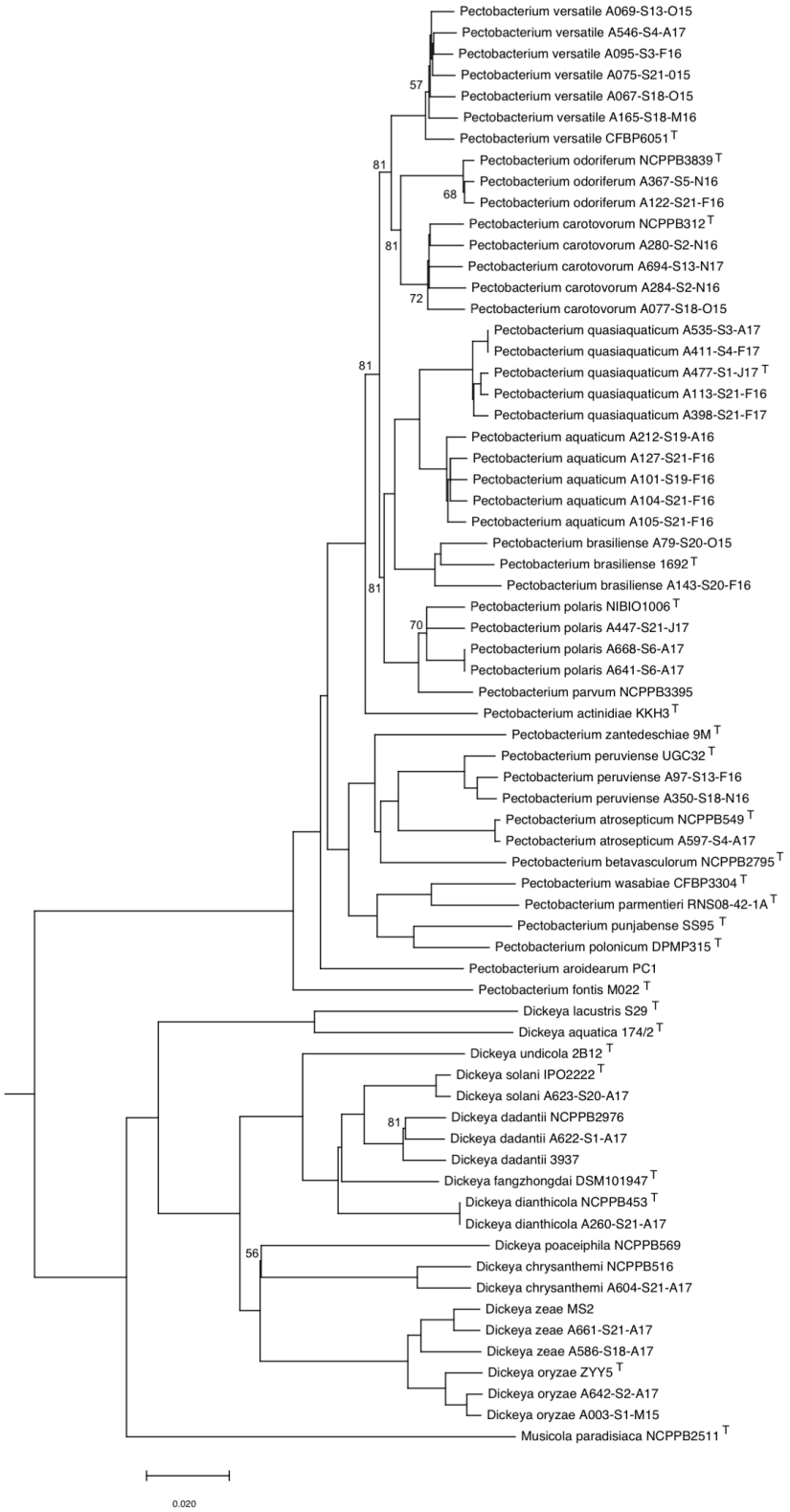

Figure 4: MLSA phylogenetic tree reconstructed from concatenated nucleotide sequences of 601 homologous gene sequences.